Investigating the relationship between diabetes and frozen shoulder: longitudinal analysis of primary care medical records

Brett Paul Dyer

A thesis submitted for the degree of Doctor of Philosophy

June 2022

Keele University

Declaration

This thesis was proposed following discussions between Professor Danielle van der Windt, Dr Linda Chesterton, Dr Miliça Blagojevic-Bucknall, Dr Claire Burton.

The protocol and search strategy for the systematic reviews was constructed by myself with advice from the Keele University Primary Care Centre Versus Arthritis health information team. The systematic review study selection was conducted by myself and independently checked by Dr Miliça Blagojevic-Bucknall, Dr Claire Burton or Dr Trishna Rathod-Mistry. Risk of bias assessments were independently conducted by myself and by Dr Miliça Blagojevic-Bucknall, Dr Claire Burton or Dr Trishna Rathod-Mistry. Evidence synthesis was conducted by myself and the interpretation and discussion of results is my own.

The design of the Clinical Practice Research Datalink (CPRD) studies were my own. Dr Miliça Blagojevic-Bucknall provided advice on writing the Independent Scientific Advisory Committee (ISAC) protocol. Data were downloaded by James Bailey and cleaned/manipulated and analysed by myself. The interpretation and discussion of results is my own.

I have no competing interests to declare.

Abstract

Introduction

Frozen shoulder, a condition that can cause prolonged pain and disability, has previously been found to be common amongst people with diabetes. This thesis aimed to improve the understanding of the nature of the relationship between diabetes and frozen shoulder.

Methods

Evidence from existing longitudinal observational studies was summarised to assess whether diabetes is a risk factor for the onset of frozen shoulder, as well as whether it is a prognostic factor for poor outcomes of frozen shoulder. A series of cohort studies based on data from Clinical Practice Research Datalink (CPRD) were subsequently undertaken to establish these relationships more accurately and comprehensively, and to investigate whether the effect of diabetes on the risk of developing frozen shoulder is mediated through metabolic health. The association between newly diagnosed frozen shoulder and a subsequent diabetes diagnosis was also estimated.

Results

Most of the relevant studies identified in the literature search were at high risk of bias. Causal mediation analysis of 87,954 patients from CPRD suggested that diabetes does affect the development of frozen shoulder, but the effect is unlikely to be mediated by metabolic health. Following the 15.8-year follow-up, the probability for the frozen shoulder group to be diagnosed with diabetes was 5% versus 0.28% in those without frozen shoulder. Diabetes was a predictor of surgery in 40,644 patients with frozen shoulder.

Conclusion

Evidence suggests that diabetes is a potential cause of frozen shoulder, although our results did not support the hypothesis that metabolic health mediates the effect of diabetes on the development of frozen shoulder. People with frozen shoulder are more likely to have a subsequent diagnosis of diabetes; future research is required to determine whether testing all patients with frozen shoulder is an effective strategy to detect diabetes early in its course and reduce the like-

lihood of complications.

Publications and presentations associated with this PhD project

Peer-reviewed publications

B. P. Dyer, C. Burton, T. Rathod-Mistry, M. Blagojevic-Bucknall, D. A. van der Windt. Diabetes as a Prognostic Factor in Frozen Shoulder: A Systematic Review. *Archives of Rehabilitation Research and Clinical Translation*, 3(3):100141, 2021. https://doi.org/10.1016/j.arrct.2021.100141

Oral presentations

B. P. Dyer, C. Burton, T. Rathod-Mistry, M. Blagojevic-Bucknall, D. A. van der Windt. Investigating the relationship between Diabetes and Frozen Shoulder: A Plan of Study. *Keele University School of Medicine Postgraduate Symposium*, 16 May 2019.

B. P. Dyer, C. Burton, T. Rathod-Mistry, M. Blagojevic-Bucknall, D. A. van der Windt. Diabetes as a Prognostic Factor in Frozen Shoulder: A Systematic Review. *UK RiME Spring Series Webinars*, 1 March 2021.

B. P. Dyer, C. Burton, T. Rathod-Mistry, M. Blagojevic-Bucknall, D. A. van der Windt. Diabetes as a Prognostic Factor in Frozen Shoulder: A Systematic Review. *Keele University School of Medicine Postgraduate Symposium*, 6 May 2021.

Poster presentations

B. P. Dyer, C. Burton, T. Rathod-Mistry, M. Blagojevic-Bucknall, D. A. van der Windt. Diabetes as a Prognostic Factor in Frozen Shoulder: A Systematic Review. *Versus Arthritis Fellows Meeting*, 21 April 2021.

Acknowledgements

Many thanks to Versus Arthritis for funding this PhD project.

I would like to thank Elaine Willmore for checking the results of the systematic review searches and thank the Keele University Primary Care Centre Versus Arthritis health information team for their advice on constructing the search strategy and review protocol.

A huge thank you to all the teachers, lecturers, and supervisors that have helped me develop into the researcher I am today. A special thanks to my current supervisory team, Claire Burton, Trishna Rathod-Mistry, Miliça Blagojevic-Bucknall, and Danielle van der Windt, for their feedback and guidance on this PhD project.

Lastly, thank you to my friends and family for their support throughout my journey through academia.

Data statement

This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data is provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this study are those of the author/s alone.

Data may be obtained from a third party and are not publicly available. The data were obtained from the Clinical Practice Research Datalink (CPRD). CPRD data governance does not allow us to distribute patient data to other parties. Researchers may apply for data access at http://www.CPRD.com/.

Abbreviations

AGE	Advanced glycation end product
ASES	American Shoulder and Elbow Surgeons
BMI	Body mass index
CI	Confidence interval
CDE	Controlled direct effect
CPRD	Clinical Practice Research Datalink
DAG	Directed Acyclic Graph
DL	DerSimonian and Laird
DM	Diabetes Mellitus
GEE	Generalised estimating equations
GP	General practitioner
GPRD	General Practice Research Database
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HES	Hospital Episode Statistics
HKSJ	Hartung-Knapp-Sidik-Jonkman
HR	Hazard ratio
ICD	International Classification of Diseases
IMD	Index of Multiple Deprivation
IP	Inverse probability
IQR	Interquartile range
ISAC	Independent Scientific Advisory Committee
LASSO	Least Angle Selection and Shrinkage Operator
MeSH	Medical Subject Heading
MHRA	Medicines and Healthcare products Regulatory Agency
MLE	Maximum likelihood estimation
MUA	Manipulation Under Anaesthesia
NDE	Natural direct effect

NHS	National Health Service
NIE	Natural indirect effect
ONS	Office for National Statistics
OR	Odds ratio
OSS	Oxford Shoulder Score
PF	Prognostic factor
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analyses
PROSPERO	Prospective Register of Systematic Reviews
QOF	Quality Outcomes Framework
QUIPS	Quality In Prognosis Studies
RCT	Randomised controlled trial
REML	Restricted maximum likelihood
RMPW	Ratio-of-Mediator-Probability Weighting
ROM	Range of motion
SD	Standard deviation
SE	Standard error
SGLT2	Sodium-glucose cotransporter type 2
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
STT	Simple Shoulder Test
SUTVA	Stable-Unit-Treatment-Value-Assumption
VAMP	Value Added Medical Products
VAS	Visual Analogue Score
VQS	Visual Quality Scale

Contents

List of Figures

List of Tables

Chapter 1

Background, aims and objectives

1.1 Frozen shoulder

1.1.1 What is frozen shoulder?

Frozen shoulder, also known as adhesive capsulitis, is a common condition that often arises spontaneously [1] and can cause prolonged pain and disability [2]. People with frozen shoulder may struggle to carry out basic everyday tasks and experience disturbed sleep [3]. Frozen shoulder arises from the contraction of the glenohumeral joint capsule [4] (see Figure 1.1 for a cross-sectional diagram of the glenohumeral joint), which leads to a reduction in both active and passive range of motion (ROM) [5].

Codman introduced the term "frozen shoulder" in 1934 [6]. They described some common features of frozen shoulder as: feeling pain near the insertion of the deltoid; being unable to sleep on the affected side; painful and restricted flexion and external rotation (both active and passive); and a normal radiological appearance [6]. (Note that the different shoulder movements that will be described in this thesis are illustrated in Figure 1.2.) In 1945, Neviaser used the term adhesive capsulitis to describe the condition since they found the capsule to be adherent to the humeral head [7]. However, since the capsule has no adhesions and is not adhesive, authors have criticised the name adhesive capsulitis [8–10]. The name frozen shoulder will be used in this thesis because, traditionally, it has been the most commonly used term [1]. Although, the

name frozen shoulder has also been criticised for being non-specific and for encouraging its use as a 'waste-bin diagnosis' for any stiff and painful shoulder [1, 9, 10].

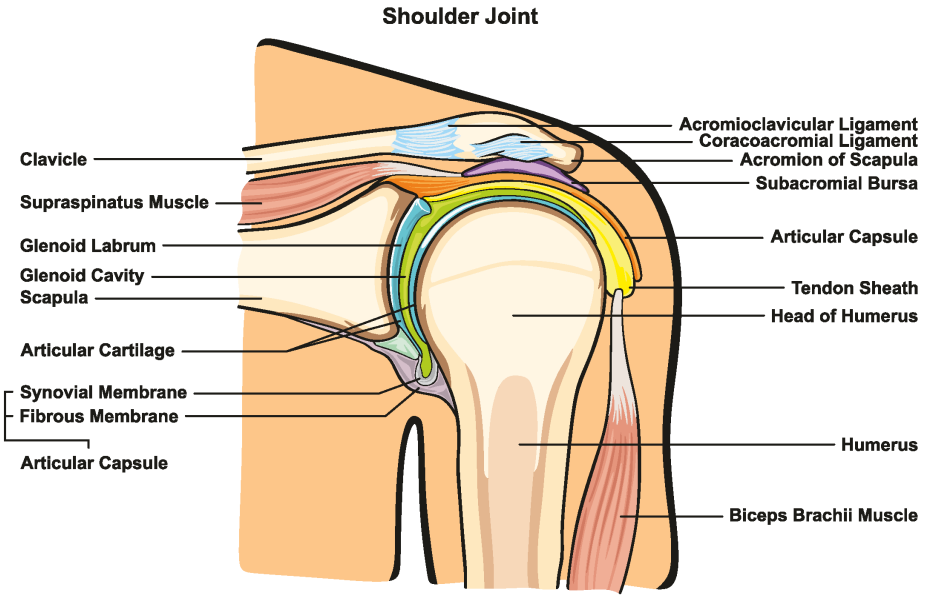

Figure 1.1: Cross-sectional diagram of the glenohumeral joint [11]

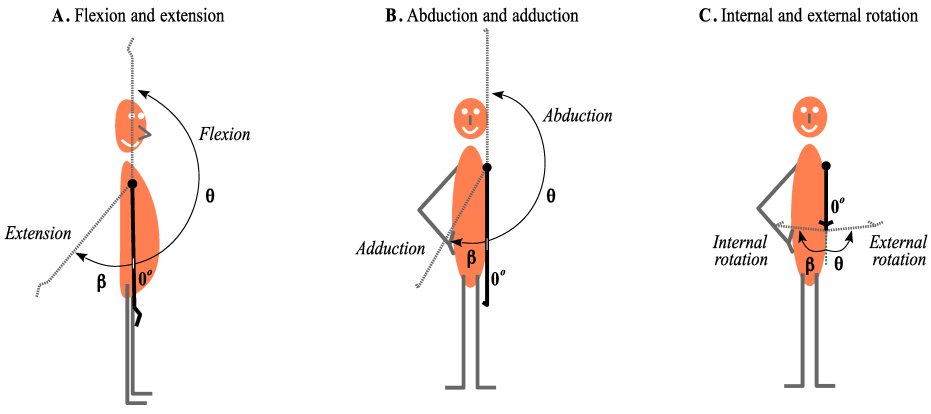

Figure 1.2: Diagram to illustrate different shoulder movements [12]

To avoid the inclination for clinicians to diagnose any painful, stiff shoulder as frozen shoulder, some authors have called for a standard set of criteria to define frozen shoulder [13, 14]. In 2011, Zuckerman et al. attempted to formally define frozen shoulder and surveyed 211 clinician members of the American Shoulder and Elbow Surgeons (ASES) to ask them whether they agree with the following proposed definition:

"Frozen shoulder is a condition characterised by functional restriction of both active and passive shoulder motion for which radiographs of the glenohumeral joint are essentially unremarkable except for the possible presence of osteopenia or calcific tendonitis" [15].

One hundred and ninety of the 211 ASES members responded to the survey. When asked *"Do you agree with the proposed definition of frozen shoulder?"*, 82% either agreed or strongly agreed with the definition; 13% either disagreed or strongly disagreed, although reasons for disagreeing were not reported in the article [15].

Some authors further classify frozen shoulder as either primary or secondary [1, 16–18]. A frozen shoulder is classed as primary if it is not associated with any underlying aetiology or associated condition and classed as secondary if a patient has an associated condition (e.g. diabetes) or an event that may have caused frozen shoulder (e.g. shoulder trauma) [15]. However, some authors classify patients with frozen shoulder and associated diabetes as having a primary frozen shoulder [17].

1.1.2 Epidemiology

The lifetime prevalence of frozen shoulder in the UK general working-age population (aged 25–66) has been estimated to be 8.2% in men and 10.1% in women [19]. A study in Dutch general practice estimated a 95% confidence interval for the cumulative incidence of frozen shoulder to be 1.9–2.9 cases per 1000 registered patients per year [20]. Variability in diagnostic criteria means that incidence and prevalence estimates will vary [20]. Furthermore, since frozen shoulder is often overdiagnosed due to other stiff and painful shoulders being wrongly labelled

as frozen shoulder, the true prevalence of frozen shoulder may be lower than the estimates reported above [10].

Frozen shoulder generally presents between 40 and 70 years of age, and the mean age of onset is 55 [21]. The shoulder of the less-dominant arm (61%) is affected more than the dominant arm (39%) [21], and it has been estimated that 12% of patients develop a second frozen shoulder on the contralateral side of the body [22].

The most common condition that is known to be associated with frozen shoulder is diabetes (type 1 and type 2) [8, 23, 24]. A meta-analysis estimated the prevalence of diabetes (all types) in frozen shoulder to be 30% (95% CI: 24 – 37%) [25]. Other markers of metabolic health have also been shown to be associated with frozen shoulder. Lipid dysfunction has been shown to be associated with frozen shoulder in two case-control studies and one cohort study [26, 27]. Hypertension has also been shown to be associated with frozen shoulder in a cross-sectional study [28]. Thyroid dysfunction has also been reported to be a risk factor for frozen shoulder [29–31]. Other factors associated with frozen shoulder include shoulder trauma [32], cardiovascular disease [8] (including stroke [23]), Dupuytren's contracture [33] and Parkinson's disease [24].

1.1.3 Pathophysiology

Frozen shoulder has long been understood to be a fibrotic disorder, which some have likened to Dupuytren's contracture of the hand [34]. Histological studies have observed an abundance of fibroblasts and then later myofibroblasts in a dense Type I and Type III collagen matrix [34–37]. The shift from fibroblasts to myofibroblasts is thought to be a key stage in the fibrosis and subsequent capsular contracture of frozen shoulder [24, 38].

More recently it has been shown that inflammation may play a role in the development of frozen shoulder [35, 39]. Patients with frozen shoulder have been shown to have an abnormal level of inflammatory cytokines [40, 41] and growth factors [41–44]. These inflammatory medi-

ators may play a role in activating fibroblasts to becomes myofibroblasts [24, 45]. Thus, frozen shoulder may initially start with an inflammatory process which then leads to the hypothesised fibrotic process described in previous paragraph.

Further, studies have identified the presence of B- and T-lymphocytes, macrophages and mast cells, which may suggest that frozen shoulder starts with an immune response which leads to inflammation and subsequent fibrosis [39, 46].

1.1.4 Clinical presentation and diagnosis

Primarily, the diagnosis of frozen shoulder is based on clinical history and examination. A patient with frozen shoulder will have restricted active and passive ROM [47–49]; loss of both active and passive external rotation is characteristic of frozen shoulder [8, 50]. If passive external rotation is preserved, then this may suggest another type of shoulder condition, such as a rotator cuff tear [4, 50]. Patients may experience most discomfort at the extremes of ROM movements [48].

A glenohumeral radiograph may be used to exclude other shoulder pathologies, especially osteoarthritis of the glenohumeral joint. A radiograph of a frozen shoulder should not show anything unusual, except for potential osteopenia of the humeral head and calcific tendonitis [15, 49]. Although not commonly used, an MRI may also help to rule out other stiff and painful shoulder conditions [50]. Thickening of the glenohumeral joint capsule may be observable in the MRI of a frozen shoulder [50].

1.1.5 Natural history

It has long been said that frozen shoulder is a self-limiting condition that progresses through a painful phase to a stiff phase to a recovery phase (although the names of the phases vary) [8, 50–52]. However, there is a lack of evidence to support these claims [51, 53]. A systematic review conducted by Wong et al. states that the theory that frozen shoulder recovers in ‘phases’ may

have originated by mistakenly combining two articles by Neviaser – one article about partial rotator cuff tears and one article about adhesive capsulitis [51].

The claim that frozen shoulder is self-limiting may have originated based on low quality evidence. Grey et al. produced a brief[1] case series report describing how they had observed that *"in the great majority of patients idiopathic frozen shoulder is a self-limiting condition, in which symptoms subside and full shoulder movement returns within a maximum of two years from the onset of symptoms"* [54]. Since this brief report was published, many studies have found evidence to suggest that for many patients frozen shoulder is not a self-limiting condition and that a reduced range of motion, lack of function, and pain may persist long-term [48, 49, 55–61].

A systematic review summarised evidence regarding the claims that frozen shoulder is a self-limiting condition and that frozen shoulder progresses through painful, stiff and recovery phases. The systematic review concluded that the claims were not supported by evidence [51]. The evidence from the review suggested that most improvement in range of motion and function does occur early, but improvement slows with time and a reduced range of motion and lack of function may persist for many years [51].

1.1.6 Management

A variety of treatment options for frozen shoulder are available. In the early course of frozen shoulder, a 'watch and wait approach' and physical therapy are the most frequently used treatment types [62]. Physical therapy may be coupled with an intra-articular glucocorticoid injection to reduce short-term pain and inflammation [48, 63]. Glucocorticoid injections have been shown to be a beneficial accompaniment to physical therapy to help improve ROM, function and pain in the short-term [64–67], but they may produce short-term (<12 weeks) side effects [48, 65].

[1]The brief half page report did not include methods or statistical analysis, meaning that the study was at a high risk of bias.

If patients fail to recover following watchful waiting, corticosteroid injections, and/or physical therapy, then surgical treatment may be considered [63]. The two most common surgical options are arthroscopic capsular release and Manipulation Under Anaesthesia (MUA) [62]. In previous decades, the popularity of arthroscopic capsular release has grown [48, 68]. In 1987, Neviaser stated that *"arthroscopy is not useful for either diagnosis or treatment of* [*frozen shoulder*]*"* [69]. However, today arthroscopic capsular release is the most common type of frozen shoulder surgery [62]. The procedure is a keyhole (arthroscopic) surgery, performed under a general anaesthesia where damaged tissue is removed and ligaments are split to 'release' the glenohumeral joint capsule [70]. The procedure may also include shoulder manipulation [49]. Patients may need at least one week off work post-surgery [49].

MUA is used in a similar frequency to arthroscopic capsular release. The patient lays supine under general anaesthesia whilst the surgeon manipulates the shoulder with flexion, abduction, and adduction movements to gently stretch the glenohumeral joint capsule [49]. Sometimes external rotation movements are also included [68]. Whilst performing the manipulation movements, the surgeon ensures the scapula is stable.

Hydrodilatation (also known as arthrographic distension), whilst not as common as arthroscopic capsular release and MUA, is growing in popularity [62, 71]. The procedure is performed by injecting local anaesthetic into the glenohumeral joint capsule at a high pressure to stretch the capsule; however, this can be an uncomfortable procedure [48, 68].

There has been debate about which treatment option is the most efficacious and cost-effective; however, recent systematic reviews have been unable to arrive at an answer due to the limited amount of evidence of a suitable quality [64, 68]. Since the aforementioned systematic reviews were published, a multicentre, pragmatic, three-arm, randomised superiority trial has been conducted to compare the effectiveness and cost-effectiveness of MUA, arthroscopic capsular release and physical therapy plus intra-articular glucocorticoid injection; however, no treatment was shown to be clinically superior [71]. MUA was the most cost-effective treatment and arthroscopic capsular release carried the highest risk of adverse events (although serious complications

were rare) [71].

1.1.7 Patient perceptions

Jones et al. interviewed 12 patients diagnosed with frozen shoulder [72]. It was reported that the primary concern of the patients was to recover from functional disability and, to a lesser extent, pain [72]. The pain was described as "severe" and "inexplicable" [72]. Restriction in movement meant that patients struggled doing everyday tasks, disrupted work, and in one case meant that the patient had to resign from work [72]. Anxiety regarding the uncertainty in diagnosis from non-specialist clinicians was also identified as a key theme amongst patients [72].

Recovery from pain and reduced shoulder function was also found to be a priority amongst the 44 patients interviewed by Srikesavan et al. in a qualitative study which sought to understand patient perceptions of treatment in the UK FROST trial [73]. The study attempted to summarise patient treatment preferences by stating, *"Trial participants had mixed treatment preferences before the trial. Some considered physiotherapy to be ineffective, while a few wanted to avoid the risks of surgery. Some preferred the less invasive MUA while some thought* [*arthroscopic capsular release*] *as the final solution. Some didn't have any preference at all."* [73]

1.2 Diabetes

1.2.1 What is diabetes?

Diabetes mellitus, more commonly referred to as diabetes, is a group of chronic diseases characterised by hyperglycaemia. The different types of diabetes can be classified according to their aetiology as type 1 diabetes, type 2 diabetes, gestational diabetes, and other specific types of diabetes [74, 75]. The two most prevalent types of diabetes are types 1 and 2 [76, 77].

Individuals that have blood sugar levels which are elevated but that are not high enough to meet the threshold for a diagnosis of diabetes are diagnosed as having pre-diabetes [78, 79].

People with pre-diabetes are at an increased risk of developing type 2 diabetes and are at an increased risk of having some of the complications associated with type 2 diabetes, such as developing cardiovascular diseases [80, 81]. To prevent patients progressing to develop type 2 diabetes, the NHS offers people diagnosed with pre-diabetes the opportunity to join the diabetes prevention programme which includes a behaviour change programme and annual glucose monitoring [82].

1.2.2 Pathogenesis and pathophysiology

Type 1 diabetes is thought to arise due to an autoimmune process that destroys insulin-producing β-cells [83, 84]. Insulin is a hormone that aids cellular glucose uptake; thus, insufficient insulin production results in hyperglycaemia and glucose-starved mitochondria. People with type 1 diabetes are usually either completely or nearly completely insulin deficient and require exogenous insulin for survival [85].

In type 2 diabetes the pancreas does make insulin, but cells do not adequately respond to the insulin [86]. In a normally functioning cell, insulin receptors are activated by insulin, which results in glucose transporters moving to the cell membrane to facilitate glucose uptake. In type 2 diabetes there are a reduced number of fully functioning insulin receptors; therefore, less glucose is taken into the cell and more glucose remains in the blood, which results in hyperglycaemia [87, 88]. In response to the elevated blood glucose levels, β-cells secrete a greater amount of insulin to help regain healthy blood glucose levels and facilitate cellular glucose uptake [86]. However, this response is not sustainable and is believed to result in β-cell dysfunction and deficits in β-cell mass [89]. This subsequently causes a reduction in the amount of insulin and results in hyperglycaemia. Management of type 2 diabetes includes a combination of lifestyle changes, oral hypoglycaemics and exogenous insulin.

1.2.3 Epidemiology of diabetes

In the 2018 Health Survey for England, the prevalence of clinically diagnosed diabetes was estimated to be 7% [90]. The prevalence has risen from the 2.8% prevalence estimated two decades prior in 1998. The rise may be partly due to the rising standards of testing and screening programmes, but also due to a change in lifestyle behaviours [91]. Also in the UK, the 2018–19 National Diabetes Audit estimated that 90% of people with diabetes have type 2 diabetes, 8% have type 1 diabetes, and the remaining 2% have other types of diabetes [76].

Factors thought to be associated with type 1 diabetes include: age (greatest incidence in individuals younger than 15 years of age); geographical location (nationally); genetics/family history; environmental factors, such as exposure to viruses [77, 92, 93]. Factors associated with type 2 diabetes include: lifestyle factors (diet, physical activity, smoking, alcohol); obesity; genetics/ family history; age (most common in individuals over 35 years of age); other metabolic factors [77, 94–96].

1.2.4 Metabolic syndrome

The term metabolic syndrome is used to represent a cluster of interrelated factors that are associated with an elevated risk of developing atherosclerotic cardiovascular disease [97, 98]. Definitions of metabolic syndrome and its components vary, but generally the factors that make up metabolic syndrome include: glucose intolerance, obesity, dyslipidaemia, and high blood pressure [99]. A consensus definition, arrived at during a meeting between the International Diabetes Federation and the American Heart Association/National Heart, Lung, and Blood Institute, can be found in Table 1.1 [100].

As mentioned above, the components of metabolic syndrome are associated with an elevated risk of developing atherosclerotic cardiovascular disease. It is thought that an increase in the number of metabolic factors is associated with an increased risk of cardiovascular disease, over and above the additional risk associated with each additional individual component

Consensus definition of metabolic syndrome
A patient should be diagnosed with metabolic syndrome if they satisfy any 3 of the following 5 criteria: • Elevated waist circumference, with cut points defined according to population- and country-specific definitions. (See Alberti et al. 2009 for cut points.) • Triglyceride levels $\geq$ 8.3 mmol/L and/or if patient is receiving drug treatment for elevated triglycerides. • High-density lipoprotein cholesterol (HDL-C) levels $<$ 2 mmol/L in men or $<$ 2.8 mmol/L in women, and/or if patient is receiving drug treatment for reduced HDL-C. • Systolic blood pressure $\geq$ 130 mm Hg, and/or diastolic blood pressure $\geq$ 85 mm Hg, and/or if patient is receiving antihypertensive drug treatment and has a history of hypertension. • Fasting glucose $\geq$ 5.6 mmol/L, and/or if patient is receiving drug treatment for elevated glucose.

Table 1.1: Table including a consensus definition of metabolic syndrome, arrived at during a meeting between the International Diabetes Federation and the American Heart Association/National Heart, Lung, and Blood Institute

[98, 99]. However, evidence of this is unclear and is complicated by the variation in the definitions of metabolic syndrome [101].

Metabolic syndrome is also associated with musculoskeletal disorders [102]. The inflammatory state brought about by metabolic syndrome is thought to play a key role in the pathogenesis of many musculoskeletal disorders [102–104]. Some authors believe that understanding the role of metabolic syndrome in musculoskeletal disease pathology is critical, but often overlooked [102, 105]. It has also been argued that metabolic syndrome may be a key component in explaining the association between type 2 diabetes and frozen shoulder [24, 106]. Metabolic syndrome is associated with chronic inflammation which, as described in Section 1.1.3, has been hypothesised to be involved in the pathogenesis of frozen shoulder.

1.3 Diabetes and frozen shoulder

Many complications, such as diabetic retinopathy, neuropathy, nephropathy, and cardiovascular conditions, in people with diabetes are well-known [107, 108]. Musculoskeletal complications of diabetes are much less well-known [107]. Such complications may impact a patients quality of life by causing pain and dysfunction. Further, musculoskeletal disease may impede diabetes treatment by restricting the patients ability to exercise. Musculoskeletal disorders commonly observed in patients with diabetes include: joint disorders (e.g. rheumatoid arthritis, osteoarthritis, gout), muscle-related disorders (e.g. diabetic muscle infarction, diabetic amyotrophy), skeletal disorders (e.g. osteoporosis, diffuse idiopathic skeletal hyperostosis), and fibrotic tissue disorders (e.g. frozen shoulder, Dupuytren's contracture, carpal tunnel syndrome) [109].

The prevalence of frozen shoulder in patients with diabetes has been estimated to be 13.4% (95% CI: 10.2 – 17.2%) and estimated to be 5 times (95% CI: 3.2 – 7.7) the prevalence in people without diabetes [25]. Patients with type 1 diabetes and patients with type 2 diabetes have a similar probability of having frozen shoulder [25].

Epidemiological evidence of the temporal and causal relationship between diabetes and frozen shoulder is unclear. However, there are hypotheses about the reason why there is such a high prevalence of frozen shoulder in people with diabetes. A hyperglycaemic state may result in simple sugars, such as glucose, bonding non-enzymatically with the amine group in protein molecules in a process called glycation [110]. This process produces advanced glycation end products (AGE's) that cause collagen cross-linking [111, 112], which may explain the capsular fibrosis in frozen shoulder [113, 114].

Additionally, hyperglycaemia may induce inflammatory cytokines [115], which are overexpressed in the subacromial bursa of patients with frozen shoulder [40]. This may explain the role of diabetes in the hypothesised pathogenesis, explained in Section 1.1.3, in which a process of inflammation leads to capsular fibrosis and subsequently to the development of frozen shoulder.

Two reviews have summarised evidence from studies comparing the outcomes of frozen shoulder in people with diabetes versus people without diabetes. The reviews reported the results of studies but did not contain any evidence synthesis. After reporting the results of 23 studies, Whelton and Peach judged that people with frozen shoulder and coexisting diabetes have *"a more severe and intractable condition"* than people with frozen shoulder but no diabetes [116]. Boutefnouchet et al. reported the results of six studies and concluded that, following arthroscopic capsular release, patients with diabetes *"have more residual pain, reduced motion and inferior function"* than patients without diabetes [117].

1.4 Electronic health records

In the UK, general practices are the first point of contact in the healthcare system. Consultations in general practices are recorded in electronic records and may be used for research. The Clinical Practice Research Datalink (CPRD) is a UK electronic health record database that was first established in 1987, under the name Value Added Medical Products (VAMP) dataset, before becoming the General Practice Research Database (GPRD) and then CPRD [118]. As of January 2021, `www.cprd.com` reported that the data included patient records for 50 million people, of which 16 million were currently registered [119].

Until their recent phaseout, Read codes have been used to record CPRD data which include but are not limited to: demographic data, diagnoses, symptoms, prescriptions, tests, and referrals. CPRD data may be linked with other databases, such as Hospital Episode Statistics (HES), Index of Multiple Deprivation (IMD) and Townsend deprivation scores, and Office for National Statistics (ONS) mortality data [118].

CPRD contains two different databases called GOLD and Aurum. Aurum data are collected in contributing English practices using the EMIS clinical system, whereas GOLD data are collected from contributing UK practices using the Vision system [120].[2] This thesis will use data from the CPRD GOLD database. Recently, NHS healthcare providers have moved towards using SNOMED-CT codes which provide a single coding language to capture clinical data. The system will be used for all clinical coding across the NHS, including coding for diagnoses, symptoms, procedures, tests and medications [122]. The SNOMED-CT coding system will be used across the UK and in other countries around the world [122].

The size and length of follow-up is one main strength of CPRD data, but also that CPRD GOLD participant data are believed to be broadly representative of the general UK population in terms of age, gender and ethnicity [118]. In 2004, the Quality Outcomes Framework (QOF)

[2] Further details of the differences between GOLD and Aurum can be found at `https://www.cprd.com/sites/default/files/CPRD%20Aurum%20FAQs%20v2.2.pdf` [121].

was introduced to improve the quality of primary care and necessitated the improvement in the quality of coding [123, 124]. The QOF standardised treatment for people with diabetes and encouraged patient registers to be kept. The register and routine follow up of patients with diabetes is thought to have improved data collection and recording [125, 126].

Incomplete data may be a limitation of using CPRD data. Data on variables such as BMI are more likely to be complete for overweight individuals[3] [127]. Lifestyle habits such as smoking and drinking alcohol are also known to be missing not at random in primary care electronic health records [128]. Also, in CPRD it is often assumed that the absence of a Read code implies the absence of the disease. This may reduce sensitivity, but the extent of this will depend on the nature of the disease and the likelihood that the participant would not consult their GP [118]. Additionally, it has been shown that there exists variation in coding between GP practices and clinicians and that valuable information in 'free text' may be missed [118].

CPRD requires that data meet two sets of criteria: patient-level criteria and practice-level criteria. Patient-level data quality criteria are based on the completeness of recording of the patient's registration status, valid transfer dates, valid age and gender, and checking that events are not recorded as occurring prior to the patient's birth year [129]. Practice-level data quality criteria are based on a quality marker called the 'up-to-standard' date which indicates the continuity of data recording (reasons for any gaps in data recording are investigated) and death recording (to determine if the number of deaths recorded in the practice is within a range which may be expected for a practice of its size) [129].

To access CPRD data a protocol must be submitted to, and approved by, the MHRA ISAC. The ISAC protocol corresponding to the work submitted in this thesis (protocol number 19_219R) was accepted on 16th December 2020.

[3]It is worth noting that people with diabetes will routinely be weighed as part of their annual review.

1.5 Gaps in knowledge

It has been shown that diabetes is highly prevalent amongst people with frozen shoulder and it has been hypothesised that diabetes could be a cause of frozen shoulder, but epidemiological evidence of the longitudinal association between diabetes and the onset of frozen shoulder is lacking. Further, little research has been done to try and provide epidemiological evidence to support the hypotheses about inflammation and/or glycation processes being the mechanism(s) underlying the diabetes-frozen shoulder association.

Previous studies have suggested that the prevalence of type 2 diabetes amongst people with frozen shoulder is high, although sample sizes were small and it was unclear whether the samples were representative of the entire population of people with frozen shoulder [130, 131]. To judge the effectiveness of testing all patients with frozen shoulder upon their diagnosis, it is necessary to demonstrate in a large, representative sample that patients presenting with frozen shoulder are at a higher risk of having undiagnosed type 2 diabetes than people without frozen shoulder.

The two reviews mentioned in Section 1.3 provided some very preliminary evidence that suggests diabetes may be associated with the outcomes of frozen shoulder. However, the lack of any evidence synthesis greatly reduces the transparency of how the authors drew their conclusions and means that the reviews are at a high risk of bias. A systematic review with a transparent approach to evidence synthesis would help to confirm whether diabetes is a prognostic factor in people with frozen shoulder. The systematic review would also help to identify any other gaps in the literature that may lead to improving the management of frozen shoulder in people with diabetes.

1.6 Aims of the thesis

The overall aim of this thesis was to improve the understanding of the nature of the relationship between diabetes and frozen shoulder. Firstly, the studies within this thesis aimed to understand whether people with diabetes are more likely to develop frozen shoulder, whether the association could be causal, and to understand the extent to which the association can be explained by metabolic health. The PhD also aimed to understand whether patients with newly-diagnosed frozen shoulder are more likely to have a subsequent diagnosis of type 2 diabetes than people without frozen shoulder. Lastly, the studies in this thesis aimed to understand whether diabetes is associated with the outcomes of frozen shoulder.

The knowledge gained from the studies contained in this thesis may inform patients and clinicians of how to reduce the risk of developing frozen shoulder and how to improve the management of frozen shoulder in people with diabetes.

1.7 Research objectives

The objective of each study included in this thesis is given below.

Chapter 2: Diabetes as a risk factor for the onset of frozen shoulder: a systematic review and meta-analysis

To summarise evidence from longitudinal observational studies to investigate whether diabetes (types 1 and 2) is a risk factor for the onset of frozen shoulder.

Chapter 4: Type 2 diabetes and the risk of developing frozen shoulder: a cohort study

Objective i

To estimate the causal effect of type 2 diabetes on the development of frozen shoulder.

Objective ii

To estimate the proportion of the effect of type 2 diabetes on the risk of developing frozen shoulder that is mediated through metabolic health.

Chapter 5: Are patients with newly diagnosed frozen shoulder more likely to be diagnosed with type 2 diabetes? A cohort study

To determine the association of newly diagnosed frozen shoulder with a subsequent diagnosis of type 2 diabetes in primary care.

Chapter 6: Diabetes as a prognostic factor in frozen shoulder: a systematic review

To summarise evidence from longitudinal observational studies to investigate whether diabetes is a prognostic factor for the outcomes of frozen shoulder.

Chapter 7: Is diabetes a predictor of surgery in people with frozen shoulder? A cohort study

To investigate the association between diabetes (the candidate prognostic factor) and referral for surgical intervention (a proxy for poor outcome) in people presenting with frozen shoulder in primary care.

Chapter 2

Diabetes as a risk factor for the onset of frozen shoulder: a systematic review and meta-analysis

The protocol for this systematic review was registered on the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42019122963; available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019122963).

The review was conducted and reported using the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines [132]. A completed PRISMA checklist can be found in Appendix Section A.1.

2.1 Introduction

As described in Section 1.3, a systematic review and meta-analysis of cross-sectional studies has previously demonstrated that diabetes is highly prevalent amongst people with frozen shoulder, and vice versa [25]. Additionally, it has been hypothesised that diabetes may be a cause of frozen shoulder. To better understand the nature of the relationship between diabetes and frozen shoulder, this systematic review summarised evidence from longitudinal observational studies to investigate whether diabetes is associated with the development of frozen shoulder.

2.1.1 Systematic review objective

To summarise evidence from longitudinal observational studies to investigate whether diabetes (types 1 and 2) is a risk factor for the onset of frozen shoulder.

2.2 Methods

2.2.1 Defining the eligibility criteria

Studies were eligible if they had a longitudinal observational design (prospective or retrospective) and presented results about the association between diabetes (type 1 or type 2) and the onset of frozen shoulder. Detailed inclusion criteria are described in Table 2.1. Studies were excluded from the review if they met any of the exclusion criteria, listed in Table 2.2.

	Inclusion Criteria
Population	Cohort study population at baseline: People without frozen shoulder. Case-control study: Cases - People with frozen shoulder, Controls - People without frozen shoulder.
Exposure	Diabetes. (All types of diabetes were considered.)
Outcome of interest	The onset of frozen shoulder.
Setting	No restrictions to study setting; population based as well as clinical cohorts were eligible.
Study design	Longitudinal observational studies (prospective and retrospective).

Table 2.1: Table summarising inclusion criteria

Exclusion Criteria
• If the full text was not available then the study was excluded. (Authors were contacted in an attempt to access full-text documents.) • Non–English language papers were eligible dependent upon finding a translator within the research institute. • Cross–sectional studies and case series were excluded. • Randomised controlled trials were ineligible. • If the paper did not present an effect estimate (odds ratio, risk ratio, hazard ratio) or present sufficient data to estimate an effect estimate then the study was excluded.

Table 2.2: Table summarising exclusion criteria

2.2.2 Identification of suitable literature

Suitable literature for the review was identified through a systematic literature search of 11 bibliographic databases, checking reference lists of included studies and emailing a professional contact (a clinician with an interest in frozen shoulder).

2.2.2.1 Sources of literature

MEDLINE, the U.S. National Library of Medicine's premier bibliographic database, contains more than 25 million references to life science journal articles dating back to 1966 (and a selected coverage of pre–1966 literature) [133]. The database uses National Library of Medicine Medical Subject Heading (MeSH) terms to index articles by their key concepts. The MeSH terms are structured into a hierarchical tree so that MeSH terms can be 'exploded' to ensure the user has captured all of the more specific terms arranged beneath the broader term in the tree.

EMBASE contains over 30 million records of biomedical evidence covering from 1974 to the present [134]. Similar to MEDLINE's MeSH terms, EMBASE utilises the Emtree thesaurus to index studies by drug, disease, and medical device.

AMED, produced by the Health Care Information Service of the British Library, contains over 150 thousand records from professions allied to medicine, complementary medicine, and

palliative care. The database covers 1985–present and has its AMED thesaurus indexing system based on MEDLINE's MeSH terms to help facilitate searching [135].

PsycINFO contains over 4 million records in psychology and the behavioural and social sciences. PsycINFO contains peer–reviewed journal articles covering from 1806 to present, book chapters and full books published from 1987 to present, and dissertations and technical reports [136].

Web of Science Core Collection gives access to six databases, but for this review only the Science Citation Index Expanded (1970–present) and the Conference Proceedings Citation Index–Science (1990–present) were required [137].

CINAHL provides access to more than 6 million records of literature from nursing and allied health professions that date as far back as 1937 [138]. CINAHL indexes terms using subject headings based on MEDLINE's MeSH terms to help facilitate searching.

Epistemonikos is a collection of systematic review records in the field of health [139]. The database also includes details of the primary studies that are included in the systematic reviews.

Trip is a search engine designed for users to find high quality clinical research evidence [140].

PEDro contains citations and abstracts of physiotherapy literature, with their oldest record dating back to 1929 [141].

OpenGrey is a database containing 700 thousand citations of European grey literature from the Sciences, Technology, Economics, and Humanities. OpenGrey's content includes research reports, doctoral dissertations, conference papers, plus more [142].

The Grey Literature Report is a database of public health research citations that was pro-

duced by The New York Academy of Medicine between 1999–2016, but is now discontinued [143].

2.2.2.2 Additional sources of literature

Reference checking – Reference lists of all studies included in the full–text screening stage of selection were checked to identify any additional relevant studies.

Versus Arthritis were given the list of studies that were identified for the review, along with the inclusion criteria, and asked if they could identify any further studies that may be suitable.

A professional contact (a clinician) of Professor Danielle van der Windt with a specific interest in frozen shoulder was contacted in the same manner as Versus Arthritis.

2.2.2.3 Constructing a search strategy

The search strategy for this review was constructed with the support of a health information specialist. The search strategy for each database can be found in Appendix Section A.2. All studies classified under index/MeSH terms or including free-text words relating to both diabetes and shoulder pain were identified. Studies relating to shoulder pain in general, rather than specifically frozen shoulder, were retrieved to increase the sensitivity of the search.

2.2.2.4 The selection process

Citations were downloaded using Mendeley [144]. Excel was used to check for duplicates and was used for reviewers to explain their reasoning for excluding studies. Titles and abstracts were screened by reviewer BD and a random sample consisting of 20% of the studies were independently checked by reviewers MB-B or CB using the pre-defined inclusion and exclusion criteria in Tables 2.1 and 2.2. Disagreements were resolved through discussion with reviewer DvdW. Full-text documents were screened by BD and one of the reviewers MB-B, CB or TR-M

using the same inclusion and exclusion criteria. Again, disagreements were resolved through discussion with DvdW.

2.2.3 Data extraction

Data extraction was executed by BD and independently checked by MB-B and TR-M using a pre-defined data extraction sheet. The extraction sheet was piloted using three studies to identify data that were missing from the extraction sheet. Extracted data included details of study design, setting, sample characteristics, sample size, exposure/outcome/covariate measurement, inclusion and exclusion criteria, duration of follow-up, attrition, statistical analysis, association estimates (odds ratio, risk ratio or hazard ratio) and their corresponding raw data.

2.2.4 Risk of bias assessment

Two sources of uncertainty may be present when analysing epidemiologic data – random error and systematic error. Meta-analysis (described in Section 2.2.5) is a method that can be used to provide a more precise association estimate (precise meaning that there is little random error). However, even in the absence of random error, systematic errors may remain. Systematic errors in estimates may be called biases [145] (although definitions of bias do differ [146]).

It is important that systematic reviews must not only assess the impact of random error, but also assess the potential for bias in the current body of evidence. Risk of bias assessments help to grade the certainty in evidence in a systematic review, and can help to identify whether methodological differences in studies may be potential sources of heterogeneity. Additionally, if a reviewer identifies a potential source of bias within current studies then the reviewer may suggest how future research may be improved to attempt to avoid such biases.

The Quality In Prognosis Studies (QUIPS) tool was used to judge the risk of bias in primary studies for this review [147]. The six domains covered by the QUIPS tool are:

- **Study participation** – To judge if the risk factor–outcome relationship is likely to differ between participants and eligible non-participants (i.e. selection bias).
- **Study attrition** – To judge if the risk factor–outcome relationship is likely to differ between participants that failed to complete the study and participants that completed the study (i.e. attrition bias).
- **Risk factor measurement** – To judge if the measurement of the risk factor is likely to differ according to the value of the outcome.
- **Outcome measurement** – To judge if the measurement of the outcome is likely to differ according to the value of the risk factor.
- **Study confounding** – To judge if the risk factor–outcome relationship is likely distorted due to some other variable(s) that are related to both the risk factor and outcome.
- **Statistical analysis and reporting** – To judge if the reported results are likely to be biased or spurious due to the statistical analysis or reporting.

Each domain contains a number of prompting items which help guide the scoring of the domain risk of bias score. The scoring of the risk of bias for the six domains was used to guide judgement of the overall risk of bias for the study; this is scored as low, moderate or high risk of bias. The overall risk of bias score was based on reviewer judgement to avoid the use of a tallied score.[1] BD judged the risk of bias for all studies and MB-B or TR-M also independently judged the risk of bias for all studies. All disagreements were resolved through discussion.

The QUIPS tool was selected for use in this review since it covers the key types of biases (selection/participation, attrition, measurement/misclassification, confounding, statistical and reporting)[2] that may affect the validity of the studies in this review. The prompting ques-

[1] The reason a tallied score was avoided is because only one major flaw on one of the bias domains can mean a study is at a high risk of bias.

[2] Note that there are differing opinions about whether confounding is [146, 148, 149] or is not [150–152] a type of bias.

tions within each domain ensures that all key components of a study that may introduce bias are examined by the reviewer. QUIPS also encourages that reviewers leave comments to argue why bias may be introduced (summaries of these comments for each risk of bias domain are provided in Section 2.3.3). These comments helped to provide suggestions to improve future research; these suggestions are provided in Section 2.4.

2.2.5 Meta-analysis methods

Meta-analysis is a systematic, transparent and reproducible approach to data synthesis. Meta-analysis uses statistical modelling to combine results from multiple studies estimating comparable associations [153]. Combining association estimates from multiple studies can help to produce a more precise 'summary' or 'pooled' association estimate, and can help to identify the existence of any heterogeneity between study results [154, 155].

Two models are commonly used to pool effect estimates in meta-analysis - the fixed-effect model and the random-effects model. The fixed-effect model is introduced in Section 2.2.5.1 and the random-effects model is introduced in Section 2.2.5.3. A comparison of the two models is given in Section 2.2.5.5.

2.2.5.1 Fixed-effect model structure

Assume that the observed effect sizes for all primary studies are distributed about some true underlying effect size, θ, and that the only variation about θ is due to within-study sampling error. Then, for a given study, i, the observed effect, y_i, may be modelled as

$$\begin{aligned} y_i &= \theta + \epsilon_i, \\ \epsilon_i &\sim N\big(0, \sigma_i^2\big), \end{aligned} \tag{2.1}$$

where ϵ_i is the within-study sampling error for study i.

2.2.5.2 Fixed-effect model parameter estimation

The inverse-variance method may be used to estimate θ. The method uses a weighted mean of the observed effect estimates from each study. A weighted mean is preferable to an arithmetic mean for meta-analyses since smaller studies are more subject to chance and therefore should contribute less to the pooled effect estimate [154]. For each study i, assign the weight

$$w_i = \frac{1}{\sigma_i^2},$$

where σ_i^2 is the within-study variance for study i.

Then, the underlying effect size θ from equation 2.1 can be estimated as the weighted mean

$$\hat{\theta} = \frac{\sum_{i=1}^{k} y_i w_i}{\sum_{i=1}^{k} w_i},$$

where k is the number of observed effect estimates contributing to the meta-analysis.

Hence, the estimate $\hat{\theta}$ has variance

$$\text{var}(\hat{\theta}) = \frac{1}{\sum_{i=1}^{k} w_i}.$$

This variance estimate can then be used to obtain a 95% confidence interval about $\hat{\theta}$:

$$\hat{\theta} \pm \big(1.96 \times \text{SE}(\hat{\theta})\big),$$

where $\text{SE}(\hat{\theta}) = \sqrt{\text{var}(\hat{\theta})}$.

2.2.5.3 Random-effects model structure

Building on the model that was introduced in Equation 2.1, the random-effects model introduces random-effects to model between-study variation. The model becomes

$$\begin{aligned} y_i &= \theta_i + \epsilon_i \\ &= \theta + u_i + \epsilon_i, \\ u_i &\sim N(0, \tau^2), \\ \epsilon_i &\sim N\big(0, \sigma_i^2\big), \\ u_i &\perp\!\!\!\perp \epsilon_i, \end{aligned} \tag{2.2}$$

where y_i is the observed effect size for study i, θ_i is the true underlying effect size for study i, θ is the mean of the true underlying effect sizes, u_i is the random-effect for study i, $\tau^2 = \text{var}(\theta_i)$ is the variance of true effect sizes (also known as between-study variance or heterogeneity), ϵ_i is the within-study sampling error for study i, and σ_i^2 is the within-study variance for study i.

Random-effects represent between-study differences; more specifically, the random-effect u_i represents how much study i's true effect size, θ_i, deviates from the mean of all of the study's true effect sizes, θ. The random-effects are a second source of error and are assumed to be normally distributed with mean 0 and variance $\tau^2 = \text{var}(\theta_i)$. The true effect size for study i may be different from the mean of all of the study's true effect sizes due to differences between studies such as different population characteristics, study design, exposure/outcome measurement. It is worth noting that if there are no differences between studies then $u_i = 0$ for $i = 1, \ldots, k$ and $\tau^2 = 0$ and so equations 2.1 and 2.2 become equal. Therefore, if no heterogeneity is present, then the random-effects model becomes a fixed-effect model.

2.2.5.4 Random-effects model parameter estimation

The estimate for the random-effects pooled estimate is similar to that of the fixed-effect model, but the additional source of variation that is modelled by the random-effects is also incorporated.

So, $\hat{\theta}$ is estimated as

$$\hat{\theta} = \frac{\sum_{i=1}^{k} y_i w_i^*}{\sum_{i=1}^{k} w_i^*}, \tag{2.3}$$

as before but with different weights, w_i^*, because the weights are now the inverses of two sources of variation instead of one.[3] Hence,

$$w_i^* = \frac{1}{\sigma_i^2 + \tau^2}, \tag{2.4}$$

and it follows that

$$\text{var}(\hat{\theta}) = \frac{1}{\sum_{i=1}^{k} w_i^*}. \tag{2.5}$$

In equations 2.4 and 2.5 above, there is an unknown parameter, τ^2, that needs to be estimated. One method for estimating τ^2 is restricted maximum likelihood (REML) estimation. Usually maximum likelihood estimation (MLE) is used to estimate a single model parameter; however, when two parameters need estimating MLE can produce biased parameter estimates [156]. This issue arises due to the need to estimate one parameter using another parameter estimate, thus losing one degree of freedom. REML corrects this bias with a small change to the log-likelihood function. REML then uses a process of iteration to calculate τ^2. More information about the REML algorithm can be found in (Veroniki, et al., 2016), which also describes other methods that are available for parameter estimation, such as the DerSimonian and Laird method of moments.

The estimation of τ^2 also causes issues in the calculation of 95% confidence intervals for the pooled mean effect estimate, $\hat{\theta}$. Using the same approach as for the fixed-effect model,

$$\hat{\theta} \pm \big(1.96 \times \text{SE}(\hat{\theta})\big),$$

[3]Using equation 2.4, note that the weights for large studies in random-effects meta-analyses will be smaller than in fixed-effect meta-analyses. It follows, from Equation 2.3, that the large studies have less influence on the summary estimate in random-effects meta-analyses than in fixed-effect meta-analyses. Additionally, using equation 2.5, it can be seen that the variance will increase and the 95% confidence interval will become wider in random-effects meta-analyses than fixed-effect meta-analyses.

would result in the size of the confidence interval being too narrow since the formula does not account for the uncertainty in the estimation of τ^2 [157]. (Recall that τ^2 was used to estimate var$(\hat{\theta})$ in equations 2.4 and 2.5, and hence is used to estimate SE$(\hat{\theta})$ here.) Following REML estimation, one can use the Hartung-Knapp-Sidik-Jonkman (HKSJ) method to calculate confidence intervals. Traditionally, the DerSimonian and Laird (DL) method for random-effects meta-analysis has been used, but it has been shown to have inflated type I error rates unless there are a large number of studies and little to no heterogeneity [158]. The HKSJ method has been shown to outperform the DL method, especially when the number of studies is small [158].

The HKSJ 95% confidence interval is given by:

$$\hat{\theta} \pm \sqrt{q} \times \text{SE}(\hat{\theta}) \times t_{(k-1);0.975},$$

where

$$q = \frac{1}{k-1} \sum_{i=1}^{k} w_i (y_i - \hat{\theta})^2,$$

and where $t_{(k-1);0.975}$ is the 0.975^{th} quantile of the Student's t-distribution with $k-1$ degrees of freedom [159–161].

2.2.5.5 Fixed-effect vs. random-effects meta-analysis

Sections 2.2.5.1 and 2.2.5.3 have introduced the fixed-effect and random-effects meta-analysis methods to compute summary estimates. The key difference between the two modelling strategies lies in the assumptions that are made about the studies that are being pooled. Fixed-effect meta-analysis assumes that the true association size being estimated is the same in all studies (an assumption that is often not met [162]). Random-effects meta-analysis does not require this assumption. The model incorporates random-effects to model the possible differences in true association size between studies. Differences in true association size may be due to, for example, differences in study populations, study design, or management/treatment of disease. The distribution of the true underlying association sizes is usually assumed to be Gaussian.

Since fixed-effect and random-effects meta-analysis models make different assumptions about the true underlying association sizes, their summary estimates have different meanings. A fixed-effect model summary value is the 'best estimate' of the true underlying association size, whereas the random-effects model summary value is the 'best estimate' of the mean of the different true underlying association sizes in the different studies.

The decision of whether a fixed-effect model or random-effects model is best-suited to a particular research question should initially be guided by whether there is evidence to suggest that the assumption of all studies having a common effect size is true [163, 164]. If there is a strong belief that no heterogeneity of true association sizes exists, then a fixed-effect model may be appropriate. However, if differences in true underlying association sizes may exist, then a random-effects model is required. Further, if a fixed-effect model is used but heterogeneity is detected, then a random-effects model should be used to avoid underestimating the width of the summary estimate confidence interval [165].

2.2.6 Measures of heterogeneity

Previous sections have stated the importance of identifying study heterogeneity for model-selection and the interpretation of the pooled association estimate. The τ^2 statistic is one measure of heterogeneity that has been introduced in Section 2.2.5.3, but there are additional analytic tools to assess heterogeneity. Firstly, visual inspection of the overlap between the confidence intervals of estimates from each study in a forest plot can be used. Little overlap in confidence intervals could suggest that there is heterogeneity in study association sizes.

Cochran's Q statistic is a weighted sum of squared deviations of the association size estimates for each study from the pooled association estimate; that is,

$$Q = \sum_{i=1}^{k} w_i'(\hat{\theta}_i - \hat{\theta})^2,$$

where

$$w_i' = \begin{cases} w_i, & \text{for fixed-effect meta-analysis,} \\ w_i^*, & \text{for random-effects meta-analysis.} \end{cases}$$

A p-value can be obtained from a $\chi^2(k-1)$ distribution, however the test is underpowered when sample size is small and/or when there is an imbalance in weights across studies [166].

The I^2 statistic is an alternative measure of heterogeneity that quantifies the percentage of total variance that is due to variation in true underlying association sizes rather than within-study sampling error. I^2 is given by:

$$I^2 = \begin{cases} 100\% \times (Q - \text{df})/Q, & Q > \text{df}, \\ 0, & Q < \text{df}, \end{cases}$$

where Q is Cochran's Q statistic and df is its degrees of freedom [167].

The I^2 statistic, unlike the Q statistic, is largely unaffected by the sample size [168]. However, since I^2 describes heterogeneity as a proportion of total variance, it should be reported and interpreted alongside a forest plot and an absolute measure of heterogeneity such as τ^2 [168].

2.2.7 Influence analysis

Influence analysis is a process in which the original meta-analysis model is re-run k times (where k is the number of studies included in the meta-analysis), each time leaving out a single study [169]. In a meta-analysis, it may be the case that outliers or larger studies have a big influence on the pooled estimate. Re-running the meta-analysis without the potentially highly influential study and assessing the extent to which the pooled estimate and its confidence interval have changed will help to determine the influence of the study. This may also be of interest to check the extent to which including a study suspected of being biased may change the pooled estimate and its confidence interval. If the pooled estimate and its confidence interval are similar in each

of the k meta-analysis re-runs then the reviewer can be confident in the robustness of the pooled estimate.

2.2.8 Systematic review analysis

This section will now describe the analysis plan for this systematic review.

Case-control and cohort studies identified from the systematic search were analysed separately. Where more than four studies provided results for the association between diabetes and the odds of developing frozen shoulder, a random-effects meta-analysis model was used to estimate a pooled odds ratio. Random-effects meta-analysis was used since differences in study populations (e.g. the type of diabetes that participants had), study design (e.g. length of follow-up), and analysis (e.g. covariates conditioned on) were anticipated. Where a study reported both adjusted and unadjusted association estimates, the adjusted estimate was included in the analysis. τ^2 was estimated using REML and a HKSJ 95% confidence interval was estimated for the summary value.

Where four or less studies were identified, a narrative summary was used to summarise results. It has been shown that when there are less than five studies, it is far from probable that the power from a random-effects meta-analysis will be more than the power from the individual studies that contribute to the meta-analysis [170]. Since it is assumed for a random-effects meta-analysis that the true underlying association sizes for each study are random draws from a normal distribution, then the interpretation of a summary estimate from a meta-analysis with a small number of studies becomes difficult; this is especially the case in the presence of large between-study variance, τ^2, which itself is difficult to estimate with a small number of included studies [171, 172].

Prediction intervals[4] have been shown to be inaccurate when little heterogeneity is present ($I^2 < 0.3$) or when there is an imbalance in study sizes [174], as is the case for the studies

[4]Prediction intervals provide a range for what true association sizes can be expected in future settings [173].

being pooled in this analysis (see Section 2.3.4); therefore, prediction intervals were not estimated. Between-study heterogeneity was assessed using Cochran's Q statistic, the I^2 index, and through the inspection of forest plots. Appendix Section A.3 contains the Stata [175] code that was used to fit the meta-analysis model.

A funnel plot of log odds ratios versus their standard errors was used to assess evidence of small-study bias [176]. (Since less than 10 studies were included in the meta-analysis, a test for asymmetry[5] was not conducted [177].) An influence analysis was undertaken to assess the impact of each study on the pooled estimate.

2.3 Results

2.3.1 Summary of search results

The search strategy outlined in Section 2.2.2.3 returned a total of 2371 citations, of which 1681 were unique. Eleven full-text articles were assessed for eligibility, eight of which met the eligibility criteria. A PRISMA flow chart can be seen in Figure 2.1.

2.3.2 Study characteristics

Table 2.3 summarises the characteristics of the six case-control studies, containing a total of 5,388 people, included in the meta-analysis. Table 2.4 summarises the characteristics of the two cohort studies, containing a total of 340,890 people, that were summarised narratively. Both tables contain the QUIPS risk of bias score, study design, setting, percentage of female study participants, sample size, method of diagnosis for diabetes and frozen shoulder, and the variables that were conditioned on through matching or adjustment.

Four of the case-control studies were hospital-based, one was based in a physical therapy clinic and the other used electronic health records. Two case-control studies were conducted in the USA and the others were conducted in China, South Korea, Israel, and Australia. Presence

[5]Details of the various tests for funnel plot asymmetry can be found in section 10.4.3.1 of the Cochrane Handbook for Systematic Reviews of Interventions [177].

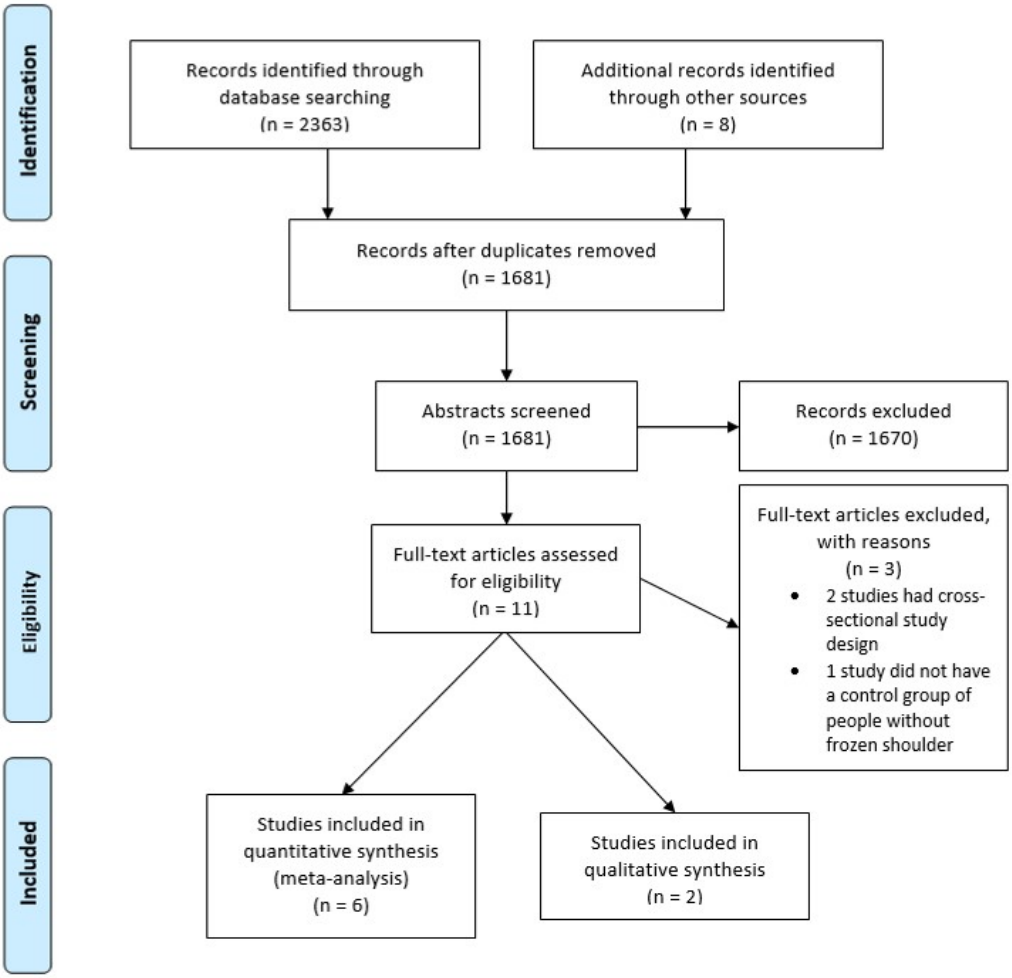

Figure 2.1: PRISMA flow diagram summarising the process of citation identification and study selection

of diabetes was self-reported in three of the case-control studies, identified using ICD-9 codes in one case-control study, another identified diabetes by a glucose test or if the patient was receiving diabetes drug treatment, and the other case-control study was unclear about how they identified diabetes. Frozen shoulder was identified clinically in five of the case-control studies and using ICD-9 codes in the other. In the case-control studies, the percentage of patients with frozen shoulder that were female ranged from 52% to 75%, and the mean age for the frozen shoulder group ranged from 52.8 years to 57.2 years. Two case-control studies matched on age and gender, two on gender only, one on age only, and one matched on time of hospitalisation and adjusted for history of minor shoulder trauma.

The two cohort studies were both based in Taiwan and were conducted using electronic health records; diabetes and frozen shoulder were diagnosed using ICD-9 codes. The percentage of female patients in the diabetes group was 47% in one cohort study and 52% in the other. The mean age for the diabetes group was not reported in one cohort study, but was 55.7 years in the other study. Both cohort studies included only newly diagnosed cases of diabetes. One cohort study matched on age and gender and adjusted for age, gender and dyslipidaemia in a Cox regression model; the other study adjusted for age, income, stroke, hypertension, hyperlipidaemia, obesity, and chronic obstructive pulmonary disease in a Cox regression model.

Author, Year (Country of Study)	Overall QUIPS Risk of Bias	Study Design and Setting	% Female	Mean age (years)	Sample Size	Method of Diabetes Diagnosis	Method of Frozen Shoulder Diagnosis	Variables conditioned on
K. L. Boyle-Walker, et al., 1997 (USA) [178]	High	Case-Control at a Physical Therapy Clinic	Case Group: 75%, Control Group: 68%	Not reported	Cases: 32, Controls: 31	Self-reported Questionnaire	Clinically diagnosed	Gender-matched
W. Li, et al., 2014 (China) [179]	High	Hospital based case-control	Case Group: 63%, Control Group: 55%	Cases: 57.2, Controls: 45.9	Cases: 182, Controls: 196	Face-to-face interview	Clinically diagnosed	Matched on time of hospitalisation. Adjusted for history of minor shoulder trauma
S-Y. Lee, et al., 2012 (South Korea) [180]	High	Hospital-based case-control	Case Group: 55%, Control Group: not reported	Cases: 52.8, Controls: not reported	Cases: 40, Controls: 40	Unclear	Clinically diagnosed	Age- and gender-matched

Continued on next page

Author, Year (Country of Study)	**Overall QUIPS Risk of Bias**	**Study Design and Setting**	**% Female**	**Mean age (years)**	**Sample Size**	**Method of Diabetes Diagnosis**	**Method of Frozen Shoulder Diagnosis**	**Variables conditioned on**
C. Milgrom, et al., 2008 (Israel) [181]	High	Hospital based case-control	Case Group: 60%, Control Group: 65%	Cases: 54.9, Controls: 55.4	Cases: 126, Controls: 98	If patient was receiving drug treatment for Diabetes or whose serum glucose was higher than 200 mg/dl	Clinically diagnosed	Age-matched
K. Wang, et al., 2013 (Australia) [182]	High	Hospital based case-control	Case Group: 64%, Control Group: 58%	Cases: 56, Controls: 55.3	Cases: 87, Controls: 176	Self-reported	Clinically diagnosed	Age- and gender-matched
K. Kingston, et al., 2018 (USA) [183]	High	Case-control using electronic health records	Case Group: 58%, Control Group: 58%	Cases: 56.4, Controls: Not Reported	Cases: 2190, Controls: 2190	ICD-9 Code	ICD-9 Code	Gender-matched

Table 2.3: Summary of study characteristics for case-control studies reporting results for the association between diabetes and the onset of frozen shoulder

Author, Year (Country of Study)	Overall QUIPS Risk of Bias	Study Design and Setting	% Female	Mean age (years)	Sample Size	Method of Diabetes Diagnosis	Method of Frozen Shoulder Diagnosis	Variables conditioned on
Y-P. Huang, et al., 2013 (Taiwan) [184]	High	Cohort with 3-year follow-up using electronic health records	Diabetes Group: 47%, Non-diabetes Group: 47%	Diabetes Group: 55.7, Non-Diabetes Group: 55.5	Diabetes Group: 78827, Non-Diabetes Group: 236481	ICD-9 Code	ICD-9 Code	Age- and gender-matched. Multivariable analysis adjusted for age, gender, dyslipidaemia
S-F. Lo, et al., 2013 (Taiwan) [31]	Moderate	Cohort with 8-year follow-up using electronic health records	Diabetes Group: 52%, Non-Diabetes Group: 51%	Not reported	Diabetes Group: 5109, Non-Diabetes Group: 20473	ICD-9 Code	ICD-9 Code	Multivariable analysis adjusted for age, income, stroke, hypertension, hyperlipidaemia, obesity, chronic obstructive pulmonary disease

Table 2.4: Summary of study characteristics for cohort studies reporting results for the association between diabetes and the onset of frozen shoulder

2.3.3 Risk of bias

QUIPS risk of bias scores for each study can be found in Table 2.5 and a bar graph of the risk of bias scores for the six bias domains can be found in Figure 2.2. Reviewers agreed on 75% of the domain risk of bias scores and on four out of the eight overall study risk of bias scores. After discussion, 100% agreement was achieved for all domains and overall risk of bias scores.

One cohort study, (Lo et al. 2014), was scored as being at a moderate risk of bias. All seven other studies were scored as being at a high risk of bias. All six QUIPS risk of bias domains contained studies scored at either a moderate or a high risk of bias. A brief summary of common[6] reasons for possible biases in each domain are given below.

Participation

One study was scored as being at a high risk of bias, three at a moderate risk of bias, and four at a low risk of bias. Common reasons for potential risk of bias were that the place of recruitment was unclear (four studies) and the recruitment rate was not given (three studies).

Attrition

No studies were deemed to be at a high risk of bias, three studies were scored as being at a moderate risk of bias, and five at a low risk of bias. The reasoning given for possible risk of bias was that little or no data was provided about how many people were lost to follow-up or did not respond to questionnaires, and no characteristics were given for these people (three studies).

Risk factor/diabetes measurement

One study was scored as being at a high risk of bias, three at a moderate risk of bias, and four at a low risk of bias. In the four studies scored as moderate or high risk of bias, it was unclear how diabetes was defined/established. In the studies that used questionnaires to identify the presence of diabetes, it was unclear what questions were asked and whether a blank response was treated as meaning the patient did not have diabetes. Code lists were not available for studies that used

[6]Here, 'common' refers to the reason being given for more than two studies.

ICD-9 codes to identify diabetes.

Outcome/frozen shoulder measurement

Two studies were judged to be at a high risk of bias, three at a moderate risk of bias, and three at a low risk of bias. The studies rated as high or moderate risk of bias did not give clear criteria for the diagnosis of frozen shoulder. Additionally, the duration of follow-up in the two cohort studies (three years and eight years) could have been too short to reliably estimate the association between diabetes and frozen shoulder.

Confounding

The standard of accounting for confounding was especially poor, thus all eight studies were deemed to be at a high risk of unaccounted confounding. In five studies there was no explanation of how the set of adjustment variables was decided upon. All studies missed potentially important confounders, often with only age and gender accounted for (Tables 2.3 and 2.4). (Li et al. 2014), (Huang, et al. 2013) and (Lo, et al. 2013) used univariable prefiltering and stepwise selection methods which are inappropriate for aetiologic models. Lo et al. erroneously adjusted for a potential mediator (i.e. a variable on the causal pathway between diabetes and frozen shoulder), stroke.

Statistics and reporting

Two studies were scored as being at a high risk of bias, three at a moderate risk of bias, and three at a low risk of bias. Reasons given for potential bias were that inappropriate methods for covariate selection were used (two studies), basic statistical tests that do not account for confounding were used (six studies), and studies were unclear about how well assumptions behind statistical models were met (two studies).

Author, Year	Participation	Attrition	Risk Factor Measurement	Outcome Measurement	Confounding	Statistical Analysis & Presentation	Overall Risk of Bias
Case–Control Studies							
K. L. Boyle-Walker, et al., 1997	High	Moderate	High	Moderate	High	Moderate	**High**
W. Li, et al., 2014	Moderate	Low	Moderate	High	High	High	**High**
S-Y. Lee, et al., 2012	Moderate	Low	Moderate	Moderate	High	Moderate	**High**
C. Milgrom, et al., 2008	Moderate	Low	Low	Low	High	Low	**High**
K. Wang, et al., 2013	Low	Low	Low	Low	High	Low	**High**
K. Kingston, et al., 2018	Low	Moderate	Moderate	Low	High	Moderate	**High**
Cohort Studies							
Y-P Huang, et al., 2013	Low	Moderate	Low	High	High	High	**High**
S-F. Lo, et al., 2013	Low	Low	Low	Moderate	High	Low	**Moderate**

Table 2.5: QUIPS risk of bias scores for the six QUIPS risk of bias domains

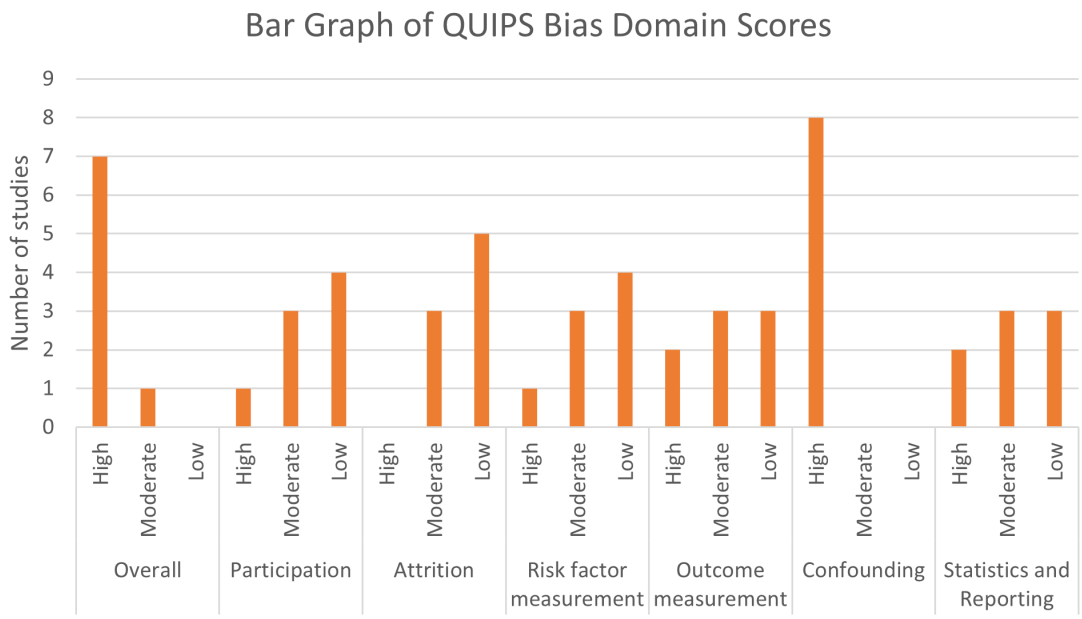

Figure 2.2: Bar graph of QUIPS risk of bias scores for overall risk of bias and for the six QUIPS risk of bias domains: study participation, attrition, risk factor measurement, outcome measurement, study confounding, statistical analysis and reporting

2.3.4 Results for the meta-analysis of case-control studies investigating the association between diabetes and the odds of developing frozen shoulder

The random-effects meta-analysis of six case-control studies contained 5388 people. The pooled odds ratio was estimated to be 3.69 (95% CI: 2.99 – 4.57). A forest plot of the meta-analysis results can be seen in Figure 2.3. The raw data that were used to calculate the odds ratios for each study can be found in Appendix A.4 Table A.1.

Between-study variance, τ^2, was estimated to be less than 0.01 (95% CI: <0.01 – 0.23). Estimated heterogeneity was small (Q=2.07, df=5, p=0.84; I^2 <0.01% (95% CI: <0.1% – 67.6%)), but the I^2 estimate was very imprecise. The forest plot in Figure 2.3 shows that the confidence intervals overlap well; although, this is partly due to the confidence intervals being wide, especially for the two smallest studies, (Lee et al. 2012) and (Boyle Walker et al. 1997).

The meta-analysis was robust to the exclusion of any individual study (Figure 2.4). One study, (Kingston et al. 2018), contained 4380 of the 5388 participants included in the meta-

analysis. Excluding this study only reduced the precision of the pooled estimate and did not considerably change the point estimate (Figure 2.4).

There was no clear evidence of small study bias. The two smallest case-control studies had the largest association estimates (Figure 2.5); however, since there are a small number of studies present in the plot, it is difficult to say that the slight asymmetrical appearance is clear evidence of small study bias.

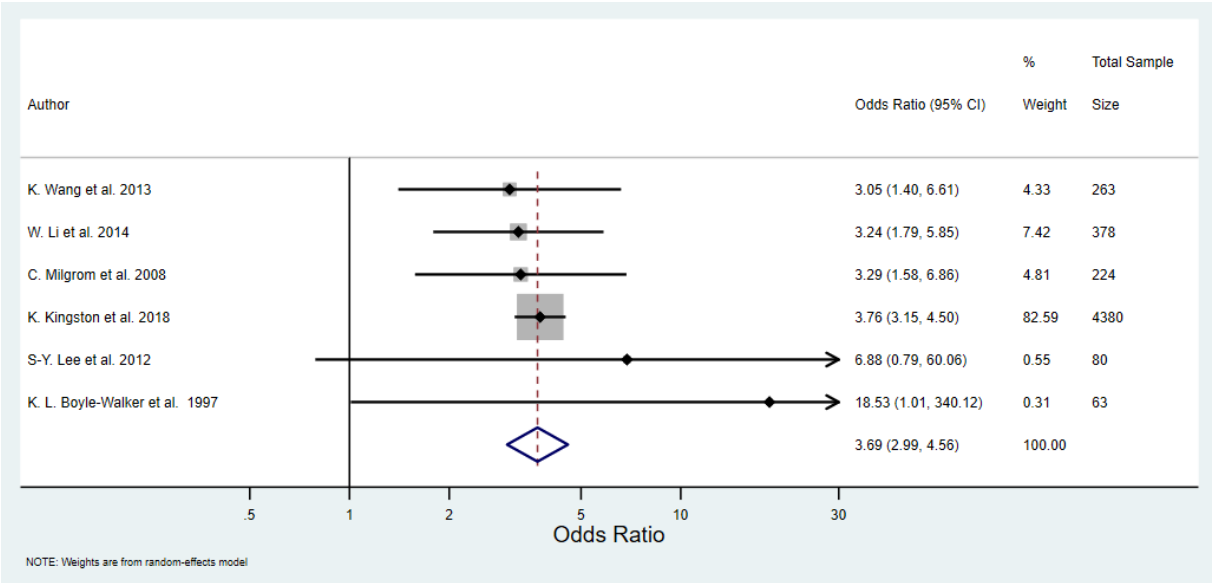

Figure 2.3: Random-effects meta-analysis of case-control studies estimating the association between diabetes and the odds of developing frozen shoulder

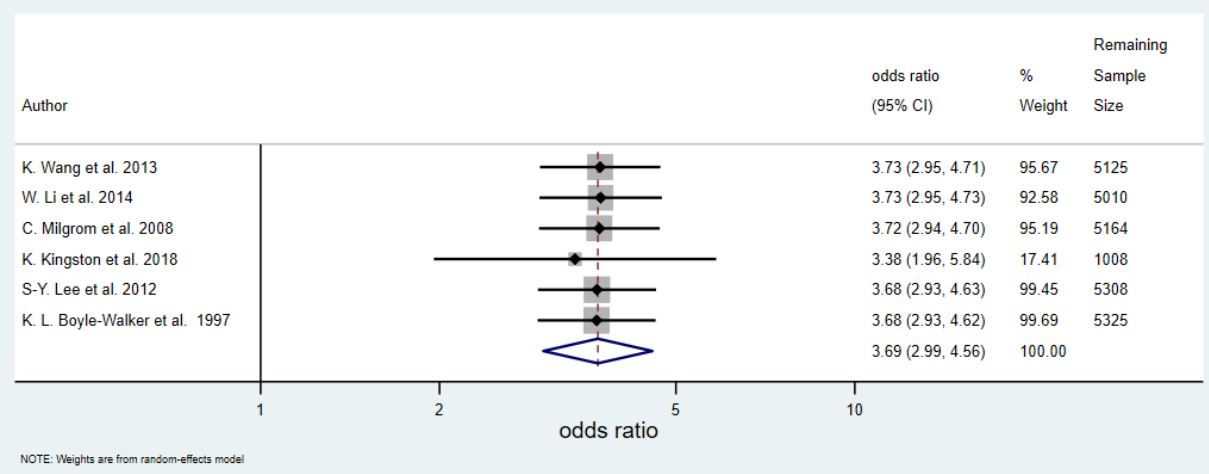

Figure 2.4: Influence analysis forest plot, showing the pooled effects estimated from repeating the original meta-analysis in Figure 2.3, each time leaving out a single primary study

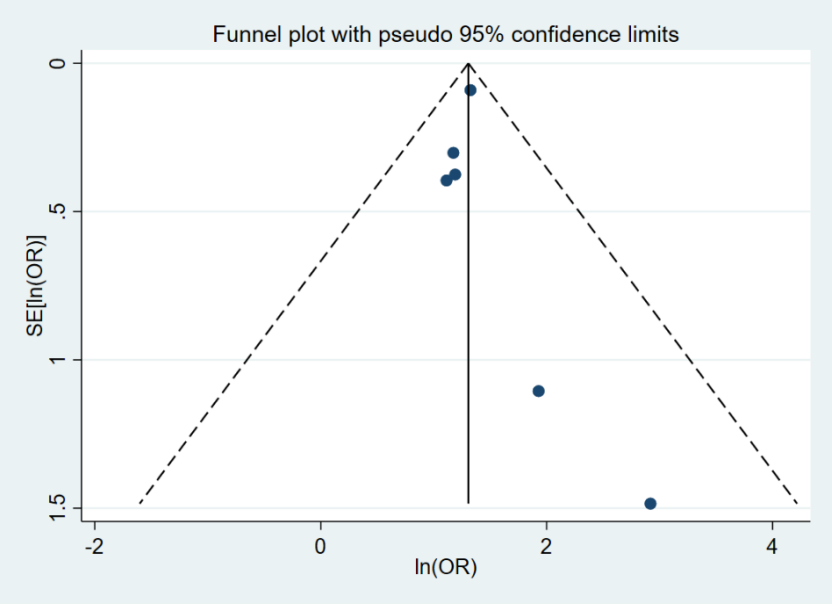

Figure 2.5: Funnel plot of log odds ratios vs standard errors for the studies included in the meta-analysis in Figure 2.3

2.3.5 Results for the narrative review of cohort studies investigating the association between diabetes and the onset of frozen shoulder

A meta-analysis of cohort studies reporting results about the association between diabetes and the risk of developing frozen shoulder was not possible since only two such studies were identified. The two cohort studies were summarised narratively.

(Lo et al. 2014) used a Cox proportional hazards model to estimate the hazard ratio for developing frozen shoulder for people with newly diagnosed diabetes compared to people without diabetes. Results suggested that people with newly diagnosed diabetes were more likely to develop frozen shoulder than people without diabetes; the age-, income-, stroke-, hypertension-, hyperlipidaemia-, obesity-, and chronic obstructive pulmonary disease-adjusted hazard ratio was estimated to be 1.67 (95% CI: 1.46 – 1.91).

(Huang et al. 2013) used Cox regression and found evidence to suggest that people with diabetes were more likely to develop frozen shoulder than people without diabetes; the age-, gender-, and dyslipidaemia-adjusted hazard ratio was estimated to be 1.32 (95% CI: 1.22 – 1.42).

2.4 Discussion

This systematic review is the first review to summarise evidence from longitudinal studies investigating the association between diabetes and the onset of frozen shoulder. All eight studies in this review provided evidence to suggest that people with diabetes are more likely to develop frozen shoulder than people without diabetes.

The meta-analysis of six control studies estimated the odds of developing frozen shoulder for people with diabetes to be 3.69 (95% CI: 2.99 – 4.56) times the odds for people without diabetes. The meta-analysis was robust to the exclusion of any single study. The two smallest studies included in the meta-analysis had the largest odds ratios; this made the funnel plot

appear slightly asymmetrical, but since only six studies contributed to the funnel plot, it is difficult to determine whether the asymmetrical appearance is evidence of small-study bias or due to chance. Since the influence analysis showed that the value of the pooled estimate did not substantially change after the omission of the smaller studies, the impact of small-study bias on the pooled estimate and its 95% confidence interval would have been small.

There was little heterogeneity detected among the studies included in the meta-analysis. Confidence intervals overlapped well, but this was partly due to the estimates of five of the six studies being imprecise. Further, whilst Cochran's Q test showed little evidence of heterogeneity, the test should be interpreted with caution when the sample size is small and when one study contributes a large proportion of the total sample size [166].

The systematic literature search identified two cohort studies that both suggested that people with diabetes were more likely to develop frozen shoulder than people without frozen shoulder. There was a difference in estimated hazard ratios in the two studies; one study estimated the hazard ratio to be 1.67 (95% CI: 1.46 – 1.91) and the other to be 1.32 (95% CI: 1.22 – 1.42). This heterogeneity could have been due in part to the differences in the covariates that were conditioned on in each study or due to the unequal follow-up durations, but equally could have been due to other factors.

One cohort study had a follow-up duration of three years [184] and the other of eight years [31], both of which could have been too short to reliably estimate the association between diabetes (both cohort studies only included patients with newly diagnosed diabetes) and frozen shoulder. Studies have found evidence to suggest that the duration of diabetes may be associated with the likelihood of developing frozen shoulder [185, 186]. One of the cohort studies included in this review, (Huang et al. 2013), also stated that their study *"suggests that the development of* [*frozen shoulder*] *is associated with the duration of diabetes"* [184]. Cohort studies with longer follow-up would be beneficial to more reliably estimate the association between diabetes and the onset of frozen shoulder.

All eight studies in this review were judged to be at a high risk of unaccounted confounding (Section 2.3.3, Figure 2.2). Six studies [178, 180–184] appeared to ignore confounders. The reasoning behind why certain covariates were or were not selected was not given in these six studies. Three studies [31, 179, 184] used univariable prefiltering and stepwise selection to select covariates for their multivariable regression models, but such methods are poorly suited for aetiological research since they do not consider the underlying causal structure of the data-generating process being investigated [187–189]. Univariable prefiltering and stepwise selection algorithms cannot distinguish between a variable that is a confounder, which therefore needs to be conditioned on, or a variable with another covariate role such as a mediator which may introduce bias if conditioned on [146, 190]. (Formal definitions of covariate roles and the consequences of conditioning on them will be introduced in Section 3.3.2.)

(Lo et al. 2014) adjusted for stroke in their regression model, but stroke should not be considered a confounder since stroke does not cause diabetes. In fact, diabetes is a risk factor for stroke [191–194], and stroke is a potential risk factor for the onset of frozen shoulder [8, 23, 48], thus meaning stroke may be a mediator and could introduce bias if conditioned on. (Li et al. 2014) adjusted for the history of minor shoulder trauma in their multivariable model. History of minor shoulder trauma is not a cause of diabetes, so again, the variable is not a confounder [190, 195, 196]. However, since the variable is likely only a competing exposure and not a mediator, it is unlikely to bias the diabetes-frozen shoulder association estimate [197].

Additionally, univariable prefiltering and stepwise selection may exclude potential confounders due to them not being statistically significant in the univariable model or during intermediate models in the stepwise selection process. Classic significance testing focuses on minimising type I error rates, but this approach is problematic to testing the inclusion of confounders. Incorrectly excluding a confounder from a model is more harmful than incorrectly including a potential confounder [190, 198].

Further, since the value of a variables coefficient in a multivariable model is dependent upon the other variables included in the model [199], a coefficient may appear to not be associated

with the outcome in some models, but will appear to be associated with the outcome in other models with different covariates [188]. This can also lead to falsely excluding variables that truly are confounders.

Studies in this review that used univariable prefiltering and stepwise selection methods [31, 179, 184] may have missed potentially important confounders[7] that may have distorted the diabetes-frozen shoulder association estimates. The other five studies in the review [178, 180–183] that only adjusted for age, gender, or both age and gender may also have distorted association estimates. These five studies did not explain how covariates were selected so it is difficult to determine why potentially relevant confounders were not adjusted for.

Section 1.3 describes the current evidence and hypotheses for how diabetes may lead to the development of frozen shoulder. Whilst such hypotheses exist, there is still not clear evidence to confirm such hypotheses. This review has also highlighted the lack of reliable epidemiological evidence to suggest that diabetes causes frozen shoulder. Future research with appropriate study design and methods are required. In particular, a transparent covariate selection process that is appropriate for testing a causal hypothesis is needed. It is also worth noting that the only two cohort studies identified by this systematic review were both from the same country, Taiwan. Future cohort studies based in other countries will help to understand whether the findings in the two Taiwanese studies are reproducible in different populations.

To gain a better understanding of the relationship between diabetes and the onset of frozen shoulder, future research could investigate the association between glycaemic control and the risk of developing frozen shoulder. Currently there exists contrasting evidence about this relationship. The results of one study have suggested that poor long-term glycaemic control was associated with an increased incidence of frozen shoulder in people with diabetes [200]. Another study concluded that they found no association between HbA1c and the prevalence of frozen shoulder in their sample of people with diabetes [201]. Cohort studies with a sufficiently

[7] Section 4.3.3 includes a Directed Acyclic Graph (DAG) communicating my own assumptions about potentially confounding variables.

long-term follow-up and repeated measurements of HbA1c (an indicator of glycaemic control) are required to better understand whether poor glycaemic control is associated with the onset of frozen shoulder within people with diabetes. It may also be worth exploring whether specific diabetes drugs are associated with an increased or reduced risk of developing frozen shoulder. For example, researchers could investigate whether the anti-inflammatory action of treatments such as metformin [202, 203] and sodium-glucose cotransporter type 2 (SGLT2) inhibitors [204] reduce the risk of developing frozen shoulder in people with diabetes.

2.5 Conclusion

In summary, this systematic review has found consistent evidence across eight studies that people with diabetes are more likely to develop frozen shoulder than people without diabetes. However, current research is limited due to seven studies being at a high risk of bias and one study being at a moderate risk of bias. Cohort studies with a sufficient duration of follow-up and appropriate adjustment for confounding variables are required to confirm the findings in this review and to better understand why diabetes is associated with the onset of frozen shoulder. Given the results of this review, patients with diabetes and their treating clinicians should be aware of frozen shoulder and other musculoskeletal pain as potential complications of diabetes. Clinicians may consider asking patients with diabetes about any musculoskeletal symptoms during their routine review and supporting them with their assessment and management.

Chapter 4 describes a cohort study conducted in CPRD which aimed to gain a more accurate and comprehensive understanding of the relationship between diabetes and the onset of frozen shoulder. The study explored how causal inference methods can be implemented to provide epidemiological evidence to (potentially) support the hypothesis that type 2 diabetes is a cause of frozen shoulder. Causal mediation analysis was used to investigate one of the hypothesised pathways through which type 2 diabetes may cause frozen shoulder. Chapter 3 will provide a detailed introduction to the causal inference methods and causal mediation analysis methods that will be used in Chapter 4.

Chapter 3

An introduction to causal inference and causal mediation analysis

This chapter will provide an in-depth description of the methods used to address the research objectives in Chapter 4. An introduction to causal inference methods, including causal diagrams and the potential outcomes framework, will be given. Traditional mediation analysis methods will be introduced and the limitations of the methods will be described. This will be followed by the counterfactual approach to defining direct and indirect effects. Then, survival analysis methods will be described and an explanation of how causal mediation analysis can be conducted with a time-to-event outcome will be given.

3.1 Asking causal questions

Anyone who has ever taken a statistics class has probably been taught that correlation is not causation, and rightly so. However, a curious scientist may not be happy with simply knowing that X is "associated" with Y; they may want to understand why. To be able to distinguish between causal and non-causal relationships, it is essential to firstly be clear about what is meant by a causal relationship and why this differs from other types of "associations" that a scientist may encounter.

An example that a teacher may give to demonstrate why correlation is not causation is that the number of shark attacks is correlated with ice cream sales. However, nobody in their right mind would argue that the ice cream salesman should be blamed for shark attacks. In this example, a third variable is connecting the number of shark attacks and the number of ice creams sold. Hotter temperatures will cause more people to be on the beach and therefore, for the number of shark attacks to increase. Also, rising temperatures cause ice cream sales to increase. Hence, the association between the number of shark attacks and the number of ice creams sold is due to a common cause, the weather.

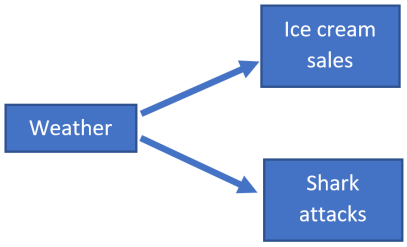

Figure 3.1: Causal diagram for the ice cream sales and shark attack association example

Whilst ice cream sales do not cause shark attacks, they may be a predictor of the number of shark attacks (not the best predictor, but still a predictor). For one variable, X, to be a predictor of another variable, Y, it is not necessary for X to be a cause of Y. It is therefore important that one is clear in defining their research question, and whether they seek to answer a question about prediction, causal inference, or description [205, 206]. As the shark attack example has demonstrated, the answer will change according to which question is being asked, as will the modelling strategies that are required [207, 208].

3.2 An overview of 'traditional statistical modelling' - modelling for prediction

The aim of a prediction model is to optimally predict the value of an outcome, or the risk of an outcome, using a set of covariates. In such studies, the model-selection process is relatively data-driven when compared to causal inference modelling methods [209].

Candidate covariates are initially selected based on prior knowledge of whether they predict or are associated with the outcome, but they are not required to be causes of the outcome. Although, it may be helpful to include predictors that are known to be causally related to the outcome since they often have strong predictive power [210] and will increase the generalisability of the prediction model to new settings or populations [209]. Additionally, avoiding highly correlated variables may help to minimise collinearity. Collinearity can inflate the variance estimate for predictor variables, which may affect whether the variable is correctly selected for inclusion in the model [211, 212].

Once an initial set of candidate predictor variables has been selected, models may be built using, for example:

- **Step-wise selection approaches**, in which an algorithm selects or excludes covariates from a model according to the covariates p-value. (Although it has been proven that this method produces biased parameter estimates, biased p-values and biased standard errors [213, 214].)
- **Best subset selection**, in which all 2^k possible models are fitted using the k candidate covariates, and the 'best' model is selected based on a metric such as adjusted R^2, the Akaike Information Criterion, or Bayesian Information Criterion.
- **Shrinkage techniques (such as Least Angle Selection and Shrinkage Operator (LASSO) regression)**. LASSO uses penalised likelihood estimation in an attempt to prevent overfitting, shrinking some coefficients to zero in the process, thus performing covariate selection.

It is also worth noting that mixed approaches may exist, in which the data-driven approaches above are used, but clinically relevant predictors are forced into the model even if excluded by, say, a step-wise selection algorithm.

After a model has been fitted, it is essential in prediction modelling to test how well the model performs in new data or resampled data [215]. A model may not be generalisable if the initial sample was not representative of the target population, if the model was overfitted, or due to other biases introduced when developing the original model [216]. Ideally, one would test the performance of the model in different data from which it was developed – this testing is called external validation. The test dataset may be from a different centre (but must be sampled from the same target population).

Internal validation is another strategy that can be used, in which the same data for which the model was developed is used to test its performance. The original dataset may have been split and developed using one fraction of the data and tested using the remaining data (although this approach is not recommended, since the splitting of data reduces power and increases overfitting [214]). Alternatively, resampling methods such as bootstrapping and cross-validation can be used for internal model validation [217]. In addition to internal validation, external validation is required to determine how the model performs in data from a different source.

3.3 Modelling for causal inference

3.3.1 Causal effects

Causal inference modelling methods are concerned with explaining whether, or to what extent, the risk of developing an outcome, Y, increases or decreases due to an exposure/treatment/intervention, X.[1] If the risk of developing the outcome does change due to the exposure then we can

[1]To avoid confusion, in this section I will only refer to binary exposures and outcomes, although the ideas are easily generalisable to other outcome types.

say that there exists a causal effect of the exposure on the outcome[2] [218]. The extent to which X causes Y can be quantified in the form of an average causal effect. Studies using causal inference methods may help to guide decision-making about how to manipulate the value of the exposure (through intervention or prevention) to decrease the risk of developing the outcome, or to gain a greater understanding of why an event occurs.

The informal definition of a causal effect given above states that the increase or decrease in the risk of developing the outcome must be *due to* the exposure. The words 'due to' are key to describing whether the question is concerned with prediction or causal inference. If we were to ask whether the risk of developing the outcome increases *given that you have* the exposure, then you are predicting. The words 'due to' explain why there is an association between the exposure and the outcome. Explaining why associations occur requires knowledge of how other variables may be related to the exposure and the outcome. Confounding is a key example of why two variables may be associated, but not directly causally related. The example that is graphically summarised in Figure 3.1 demonstrated how ice cream sales and the number of shark attacks were associated, but only because they were confounded by the weather.

Ideally, a randomised controlled trial (RCT) could be used to make sure that at baseline the two groups are similar in all characteristics aside from the treatment being given or intervention being offered; thus, any difference in the risk of the outcome must be attributable to the treatment. However, RCTs are not suitable for all research questions. For example, one could not randomly allocate individuals to have diabetes or not have diabetes. Additionally, whilst RCTs are often concerned with treating disease or preventing the consequences of a disease, they are often not suitable for investigating the prevention of primary disease [145]. This is because, within the target population of many studies, the risk of incident disease is often relatively low, so an RCT would take too long and be too expensive. In circumstances where an experimental design is not possible, longitudinal observational studies may be used to attempt to answer causal questions, but careful study design and modelling strategies are required to address the

[2]Formal definitions of causal effects are given in Section 3.4.1. The formal definitions will help to clarify how to determine whether an effect was "due to the exposure".

issues arising from a lack of random allocation to study groups. An understanding of observational epidemiological study designs is assumed for this thesis; an overview of study designs can be found in (Rothman et al. 2008) [145].

3.3.2 Causal diagrams

A well-designed study can help to limit potential biases, but additional work is often required to adjust for confounders. When seeking to estimate a causal effect, many variables related to the exposure and outcome of interest may introduce bias, but identifying which variables may bias the effect estimate can be a difficult task. Causal diagrams can be used to summarise knowledge and assumptions about how variables are causally related, and help to make these assumptions clear to readers [190, 195]. Inspection of these diagrams can help to identify potential biases, such as confounding bias and selection bias [219, 220]. Suitable adjustment sets for identifying causal effects can then be identified using the causal diagram [221].

One type of causal diagram used in epidemiological research is the Directed Acyclic Graph (DAG) [222, 223]. DAGs consist of nodes to represent random variables and unidirectional arrows/arcs to represent that a causal relationship could potentially exist between two variables. The requirement for the graph to be acyclic stems from the idea that no variable can cause itself at any one moment in time [198, 224]. Ideally, a DAG will be structured so that arrows flow in one direction (left-to-right or top-to-bottom) to reflect that causal processes occur over time [198]. This temporal structuring will prevent feedback loops/cycles (a unidirectional flow of arrows from a variable back to itself).

The DAG, G, with exposure X and outcome Y in Figure 3.2 will be used to introduce DAG terminology that will be used throughout this thesis. The same terminology used in (M. Glymour, 2013) [196] will be used in this thesis. An arrow from one variable, e.g. C_1, to another variable, e.g. X, denotes a direct causal effect of the former variable C_1 on the latter variable X (direct meaning not mediated through another variable included in the DAG). A sequence of arrows connecting two variables is called a path. A sequence of arrows all flowing in the

same direction between two variables, e.g. $X \rightarrow M \rightarrow Y$, represents an indirect effect and the sequence of directed arrows is called a causal path. Any path that is not a causal path is called a non-causal path, e.g. $X \leftarrow C_1 \rightarrow Y$.

A variable that is directly caused by another variable is said to be a child of that variable, e.g. M and Y are children of X. A variable that directly causes another variable is said to be the parent of that variable, e.g. X is the parent of M and Y. If two variables are connected by a causal path then the variable at the start of the path is said to be the ancestor of the variable at the end of the path, which is called the descendant, e.g. X, C_1 and C_2 are ancestors of M, and X, M and Y are descendants of C_1.

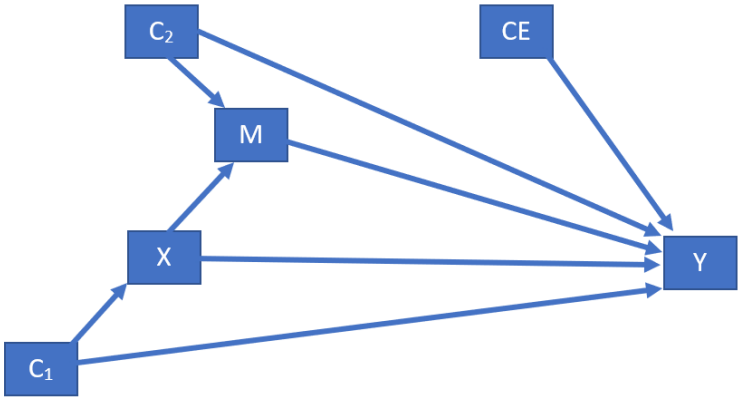

Figure 3.2: A DAG, G, to demonstrate covariate roles

The role that variables play in the data-generating process that a DAG aims to summarise will determine whether that variable may bias the effect estimate of interest. Again, Figure 3.2, will be used to demonstrate how a DAG can be used to identify the role of covariates in a data-generating process. If a variable W lies on a causal path between the exposure and the outcome then W is a mediator, e.g. M is a mediator in the DAG G. If there exists a causal path from a variable W to the exposure and a causal path from W to the outcome (that does not include the exposure), then W is a confounder of the exposure-outcome relationship, e.g C_1 is a confounder of the exposure-outcome relationship in the DAG G. Similarly, if there exists a causal path from a variable W to a mediator and a causal path from W to the outcome (that does not include the mediator), then W is a confounder of the mediator-outcome relationship, e.g C_2

is a confounder of the mediator-outcome relationship in the DAG G. If unstated it is assumed that the term confounder refers to a confounder of the exposure-outcome relationship. If there exists a causal path from a variable W to the outcome, but there is no causal path from W to the exposure or from the exposure to W, then W is a competing exposure, e.g. CE is a competing exposure in the DAG G.

One final type of variable/node that can be identified using a DAG requires special attention. A variable W is a collider if the heads of two arrows pointing in opposite directions meet W, e.g. in the DAG G, M is a collider on the path $X \rightarrow M \leftarrow C_2$. Conditioning on a collider can make two marginally independent variables conditionally dependent[3] [224, 225]. Consider the following example summarised by the DAG in Figure 3.3. Suppose that football teams won trophies either due to superb management or due to the club being rich and being able to afford good players. Also, suppose that whether a club has a lot of money to spend on players is independent of whether the club has good management (even though in reality this may not be the case). Now, suppose that it is known that a club won a trophy; that is, the collider has been conditioned on. Since we know that the club won a trophy, learning that the club had poor management suggests that the club must be rich and can afford good players. Similarly, if the club won a trophy and had bad players then this suggests that the club must have had good management.

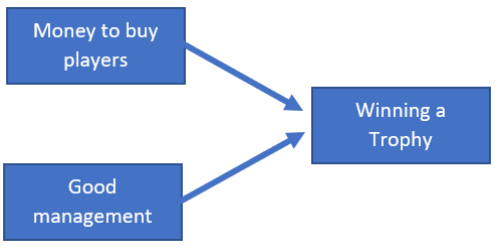

Figure 3.3: A DAG to demonstrate collider bias

By conditioning on a collider, two marginally independent variables become conditionally

[3] A random variable X is said to be marginally independent of another random variable Y if $P(X|Y) = P(X)$.
A random variable X is said to be conditionally independent of another random variable Y, given another random variable Z, if $P(X|Y, Z) = P(X|Z)$.

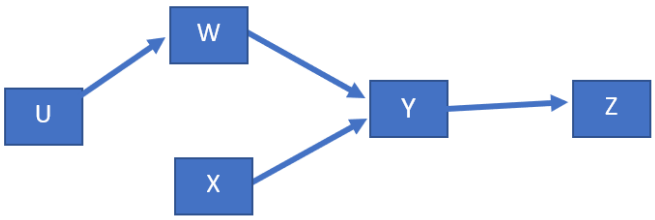

Figure 3.4: A second DAG to demonstrate collider bias

dependent, creating a non-causal association between two variables [190]. In fact, conditioning on any variable that is caused by two independent variables may create a non-causal association between the two common causes [219]. For example, conditioning on either Y or Z in the DAG in Figure 3.4 may transmit a spurious association between U and X or between W and X. The relaying of such non-causal associations brought about by conditioning on a collider, or any descendant of a collider, is called collider bias or selection bias. Collider bias has been able to explain many "paradoxes" in the scientific literature, including some examples where the sign of a regression coefficient completely reverses [226–228].

DAGs are non-parametric diagrams that summarise knowledge or assumptions about how variables may be causally related. However, the motivation for using a DAG is often to guide statistical analysis, and often to decide which variables to condition on in a regression model. The directional-separation criterion, generally referred to as the d-separation criterion, is a set of rules that can be used to read off the statistical independencies implied by a DAG [190, 221].

Two variables W and X are said to be d-separated by a set of variables Z if there is no unblocked path between W and X. A path is said to be blocked if (i) the path contains a non-collider variable V, and V is in Z; or (ii) the path contains a collider variable V, and V nor any of its descendants are in Z [146, 190, 221]. The variable V in each case is said to block the path between X and Y. Open paths transmit correlations, but closed/blocked paths do not. If two variables X and Y are d-separated by a set of variables Z then the DAG implies that X and Y are conditionally independent given Z [196].

The d-separation rule was applied by Pearl in 1993 to create a criterion for identifying a set of variables that are sufficient to condition on in order to estimate a causal effect using observational data, called the backdoor criterion [229]. A set of variables Z satisfies the backdoor criterion relative to a pair of variables X and Y if (i) no variable in Z is a descendant of X, and (ii) Z d-separates every path from X to Y that contains an arrow into X (note: these paths are called backdoor paths) [229]. Rule (i) essentially means that all causal paths from X to Y should be left unblocked. This rule also prevents the creation of non-causal paths between X and Y that may result from conditioning on a collider. Rule (ii) enforces the blocking of backdoor paths that conduct non-causal associations between X and Y. Rule (ii) also ensures that any non-causal pathways between X and Y created by conditioning on colliders are also blocked.

The DAG, call it G, in Figure 3.5 will be used to demonstrate how the backdoor criterion works in practice. G currently contains only one backdoor path from the exposure X to the outcome Y, $X \leftarrow C_1 \rightarrow C_3 \rightarrow Y$. To block this backdoor path we must condition on either C_1 or C_3. Say C_1 is unmeasured (unmeasured variables may be represented by circles in DAGs). Then, the only option to block the backdoor path is by conditioning on C_3. However, by conditioning on C_3, another backdoor path $X \leftarrow C_1 \rightarrow C_3 \leftarrow C_2 \rightarrow Y$ has been unblocked. To satisfy the backdoor criterion, C_2 must also be conditioned on. The set $\{C_2, C_3\}$ satisfies the backdoor criterion relative to X and Y. Note that the backdoor criterion demands that M is not conditioned on since it is a descendant of X. Were M to have been conditioned on, the causal path $X \rightarrow M \rightarrow Y$ would have been blocked. Further, if M had been conditioned on, the non-casual pathway $X \rightarrow M \leftarrow C_4 \rightarrow Y$ would have been unblocked.

The backdoor criterion contains rules that a set of variables Z must meet if a causal effect is to be estimated from observational data through conditioning on Z or 'adjusting' for Z in a regression model. Careful consideration is required when designing studies, building regression models, and presenting/interpreting model coefficients to ensure correct causal inferences can be made. Presenting all covariate-outcome association estimates from a single multivariable regression model is referred to as a 'Table 2 Fallacy' [199]. Presenting model results in

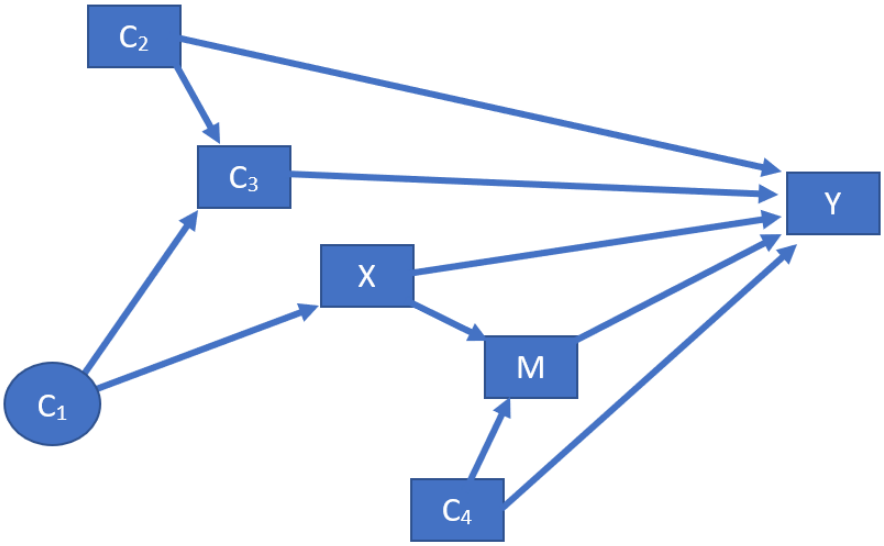

Figure 3.5: A DAG to demonstrate the backdoor criterion

such a way invites the false belief that each coefficient can be interpreted as a total causal effect [199, 230]. A single regression model will likely only include one coefficient that can be correctly interpreted as a total causal effect. Returning to the DAG in Figure 3.5, to estimate the (total) causal effect of X on Y, as stated before, the set of variables $\{C_2, C_3\}$ must be adjusted for in a regression model. Interpreting, say, the C_2 coefficient from the same regression model as the total causal effect of C_2 on Y would be incorrect because $\{C_3, X\}$ does not satisfy the backdoor criterion relative to C_2 and Y.

3.4 Counterfactuals, and the potential outcomes framework

Causal diagrams provide one approach to communicating the assumptions required to make causal inferences. Another approach, using more formal mathematical notation, is to use the potential outcomes framework. The graphical and potential outcome languages can be translated back and forth [190], and modern causal inference textbooks recommend a mixed approach to utilise the best parts of each technique [146, 190, 225, 231].

3.4.1 Potential outcomes

Every individual has once heard a statement equivalent to, "if only I had done x, then y would be different". For example, "if only I had not smoked when I was younger, then I would not have had a heart attack". In this example, the statement compares a factual outcome (the heart attack) that was observed after being exposed to a hypothesised cause (smoking) to a counterfactual outcome that supposedly would not have been observed had the individual not been exposed to the hypothesised cause. Such counterfactual statements are the key idea underlying the potential outcomes framework and are the key to better understanding how causal effects may be estimated from data.

In this section, notation will be introduced to represent individual-level counterfactual/potential outcomes which will allow individual-level causal effects to be defined. Then, in Section 3.4.2, some assumptions will be introduced which will allow population-level (average) causal effects to be estimated.

Consider an experiment in which patients are randomly assigned to either receive a drug or a placebo. Let X be a random variable denoting the treatment that the patient is assigned to and therefore the treatment that the patient actually receives, with $X = 1$ denoting that the patient receives the drug and $X = 0$ denoting that the patient receives the placebo. Let the outcome of interest Y be a binary variable, with $Y = 1$ denoting that the patient experienced the outcome and $Y = 0$ denoting that the patient did not experience the outcome.

Before a patient is assigned to treatment, the patient has two potential outcomes: the outcome that would be observed under the drug treatment, which is denoted $Y^{x=1}$, and the outcome that would be observed under the placebo, which is denoted $Y^{x=0}$. As soon as a patient is either treated with the drug or receives the placebo, it is only possible to observe one of their two potential outcomes. Let the outcome that is observed under the assigned treatment, X, be denoted Y^X. The potential outcome that is not realised, Y^{1-X}, is known as the counterfactual outcome.

Using the potential outcome notation introduced above, the individual-level causal effect can be defined as $Y^{x=1} - Y^{x=0}$ [146, 232]. However, as mentioned above, for any one individual, only one of the potential outcomes $Y^{x=1}$ or $Y^{x=0}$ can be observed. This is known as the "fundamental problem of causal inference" [233]. The individual-level causal effect is well-defined, but the missing data prevents it from being estimatable.

Whilst individual-level causal effects cannot be estimated, their definitions can be used along with some 'identifiability assumptions' to allow average causal effects to be identified. Now, return to the randomised drug/placebo experiment from earlier. An average causal effect on the risk difference scale would be given by $P(Y^{x=1} = 1) - P(Y^{x=0} = 1)$; that is, the difference in the risk of the outcome in the population if everyone had been given the drug versus the risk of the outcome in the population if everyone had been given the placebo [146, 232]. (For other types of causal effects, such as the average treatment effect on the treated or conditional average treatment effect, see (Garrido et al. 2016 [234].)

The identifiability assumptions described below will allow the counterfactual terms $P(Y^{x=1} = 1)$ and $P(Y^{x=0} = 1)$, that are required to identify average causal effects, to be given by measures of association, $P(Y = 1|X = 1)$ and $P(Y = 1|X = 0)$, which can be estimated directly from experimental or observational data [146, 231].

3.4.2 Identifiability conditions

Consistency

The first identifiability assumption, known as consistency, requires that if $X = x$, then $Y^x = Y$ [146, 235]. That is, among people who were observed to have exposure/treatment level $X = x$, the observed outcome equals the outcome that they would have experienced had they been assigned to exposure/treatment level $X = x$ [146, 235].

Consider an observational study investigating the effect of daily fruit consumption on the risk of high blood pressure, in which individuals are considered exposed if they consume fruit

daily and are unexposed if they do not. Consistency will require that there are no variations in the exposure that may affect the outcome. So, for example, the types of fruit that each individual eats must be the same if, say, the nutrients in the fruit or the amount of sugar in the fruit affect the risk of high blood pressure. Variations in the exposure that do not affect the outcome do not violate the consistency assumption. For example, if the time of day that the fruit is eaten does not affect the risk of high blood pressure, then such variations would not violate the consistency assumption.

Exchangeability/conditional exchangeability

The exchangeability assumption is a more formal way of saying that there is no confounding and no collider bias. In the language of counterfactuals, exchangeability is satisfied if $Y^x \perp\!\!\!\perp X, \forall x$ [146, 236]. In other words, if the risk of the outcome Y in the exposed group would be the same as the risk of Y in the unexposed group, had the exposed individuals not been exposed (and vice versa), then the exposed and unexposed groups are exchangeable [231].

Randomisation aims to create exchangeable groups that share the same characteristics on average so that, in the absence of the exposure, both groups would have the same risk of the outcome [237]. In observational studies, methods such as statistical adjustment or matching can be used to obtain conditional exchangeability, which requires that $Y^x \perp\!\!\!\perp X|Z, \forall x$, where Z is a set of variables that guarantees exchangeability within levels of Z [146, 231]. One way to find a set of variables Z that creates conditionally exchangeable groups is by using the backdoor criterion, described in Section 3.3.2. If a set of variables, Z, satisfies the backdoor criterion relative to an exposure X and an outcome Y, then conditional exchangeability, $Y^x \perp\!\!\!\perp X|Z, \forall x$, holds. [146, 231].

Positivity

Once a sufficient set of variables, Z, has been identified that allows the conditional exchangeability assumption to be satisfied, an additional assumption must be met, called positiv-

ity. Positivity requires that $P(X = x|Z = z) > 0, \forall z$ where $P(Z = z) > 0$, where Z is the set of variables that will be adjusted for to satisfy the exchangeability assumption [146]. Less formally, positivity requires that it is possible for each individual in the population of interest to be exposed/unexposed [238, 239].

Consider a study investigating the effects of eating red meat on the risk of myocardial infarction where 'type of diet' needed to be adjusted for to achieve exchangeability. Positivity would not be satisfied if the variable 'type of diet' included the values 'vegan' or 'vegetarian' (since the probability of eating red meat for vegans or vegetarians is zero). The study question may need to be reworded to investigate the effects of red meat on the risk of myocardial infarction, only within meat-eaters.

No spillover effects/interference

To be able to define counterfactuals on the individual level, it is easier to assume no spillover effects; that is, one individual's potential outcome does not depend upon another individual's potential outcome [240, 241]. This may not be the case in scenarios involving, say, infectious diseases where a relative being exposed to a vaccination may reduce your risk of being infected.

Other assumptions in the literature

Another set of assumptions used in causal inference, is the 'Stable-Unit-Treatment-Value-Assumption (SUTVA)' [232]. SUTVA shares many similarities with the identifiability assumptions above. SUTVA requires no interference/spillover effects, and requires no hidden variations in treatment. 'No hidden variations in treatment' is a stricter version of the consistency assumption. Consistency allows for variations in treatment, as long as the risk of the outcome is the same in the different variations of the treatments (coined "treatment variation irrelevance" by VanderWeele) [242]. This alternative potential outcomes approach also considers 'exchangeability' and 'positivity' in the form of 'ignorability' assumptions, which are described in (Rosenbaum and Rubin, 1932) [243] and (Imbens and Rubin, 2015) [232].

It is also worth mentioning why temporality has not been mentioned as an assumption above. Temporality is already implicitly included in the exchangeability assumption since, as mentioned in Section 3.3.2, causal diagrams should be arranged so that the arrows flow in one direction such that only variables in the past can cause variables in the future.

Summary of causal inference assumptions

In Sections 3.3 and 3.4, many assumptions have been introduced that help to clarify what may be required for association estimates to be interpreted causally. The assumptions often require expert knowledge and it may not be possible to check whether they are satisfied, but the assumptions do encourage the researcher to think about the research question they are actually asking and the methods they are using. In some cases, as will be seen in Section 3.7.4, the researcher can investigate the impact that potential violations of assumptions may have on causal effect estimates.

3.5 Mediation analysis

3.5.1 Motivation for using mediation analysis

One may use the causal inference methods described in Section 3.3 to investigate whether an association between two variables may be causal. Suppose that there exists evidence of a cause-effect relationship between two variables, then the next question one may ask is, "why does X cause Y?". A potential causal mechanism through which two variables may be related is mediation. A mediation mechanism between an exposure/cause X and an outcome Y may be described as the exposure of interest, X, causing an intermediate variable, M, which then goes on to affect the outcome, Y. Thus X causes Y through an intermediate variable M, which is called the mediator. In the language of causal models, this may be written as $X \rightarrow M \rightarrow Y$. The path $X \rightarrow M \rightarrow Y$ may be called a mediation path or pathway.

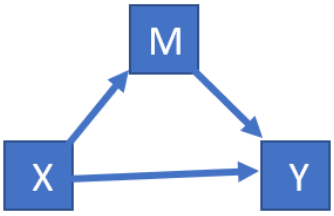

Figure 3.6: A DAG to demonstrate a basic mediation scenario

It may be the case that multiple pathways through mediating variables may exist. Mediation analysis can be used to explain the proportion of an effect that may be explained by the hypothesised mediation pathways and how much of the effect is due to other mechanisms. The total causal effect may be decomposed into "indirect effects" and a "direct effect". The effect of the exposure on the outcome that is mediated through the mediator of interest is called an indirect effect. The remaining effect of the cause on the outcome that is not mediated by the mediator(s) of interest is called the direct effect.

3.5.2 Traditional mediation analysis methods

Traditionally, two methods have commonly been used to decompose a total effect into a direct effect and an indirect effect - the difference method and the product method. Both methods use a structural equation modelling approach. The difference method has been more popular in epidemiological research and the product method has been more popular in the social sciences [244].

The difference method [244, 245]

Let X, M, and Y, respectively, denote the exposure, mediator, and outcome of interest. Let C denote a set of covariates. Then, regress the outcome Y on the exposure X and the set of covariates C; that is,

$$\mathbb{E}(Y|X = x, C = c) = \phi_0 + \phi_1 x + \phi_2 c. \tag{3.1}$$

Now, regress the outcome Y on the exposure X, the mediator M, and the set of covariates C;

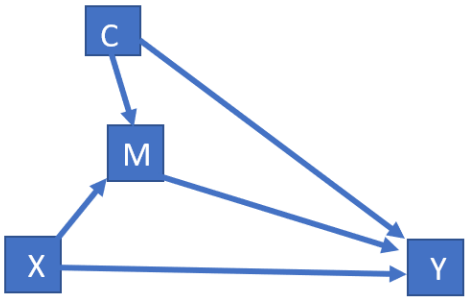

Figure 3.7: A DAG to demonstrate mediator-outcome confounding

that is,

$$\mathbb{E}(Y|X = x, M = m, C = c) = \theta_0 + \theta_1 x + \theta_2 m + \theta_3 c. \tag{3.2}$$

The coefficient ϕ_1 in Equation 3.1 denotes the total effect of the exposure, X, on the outcome, Y (where the total effect = indirect effect + direct effect). The coefficient θ_1 in Equation 3.2 denotes the direct effect of the exposure, X, on the outcome, Y. Thus, the indirect effect of X on Y is given by $\phi_1 - \theta_1$. Therefore, if $\phi_1 \neq \theta_1$ then this may suggest that some of the $X - Y$ effect is mediated by M.

The product method [244–247]

The regression model in Equation 3.2 is also used in the product method, alongside another regression model. Regress the mediator M on exposure X and the set of covariates C:

$$\mathbb{E}(M|X = x, C = c) = \beta_0 + \beta_1 x + \beta_2 c. \tag{3.3}$$

Recall, in Equation 3.2, that the direct effect is given by θ_1. Also in Equation 3.2, θ_2 gives the effect of the mediator M on the outcome Y. Now, in Equation 3.3, β_1 gives the effect of the exposure X on the mediator M. Thus, $\beta_1\theta_2$ gives the indirect effect.

Limitations of traditional mediation analysis methods

The first limitation that arises when using traditional mediation analysis methods is their

limitations in addressing confounding. In fact, the original articles that motivated the product method and difference method do not mention accounting for confounders [245, 248]. However, confounders may be incorporated into the product and difference methods (as they have in equations 3.1 – 3.3) and allow for valid causal inferences, but only if some key assumptions hold[4] [248].

The first assumption is that there is no unmeasured exposure-outcome confounding. The second assumption is that there is no unmeasured mediator-outcome confounding.[5] Mediator-outcome confounding can bias the direct and indirect effect estimates for both the product and difference methods. In both methods, the direct effect is given by θ_1 from Equation 3.2. However, if mediator-outcome confounding is present, then the estimate for the direct effect is collider biased. This is due to the fact that the mediator M is conditioned on in Equation 3.2, which would open a path from the exposure to the outcome through the mediator-outcome confounder (the path $X \rightarrow M \leftarrow C \rightarrow Y$ in Figure 3.7). The indirect effect is also biased in the product and difference methods when there is unmeasured mediator-outcome confounding. In the difference method, the indirect effect is calculated using the direct effect, θ_1, which is confounded and therefore biases the indirect effect. In the product method, the indirect effect is given by the product of the effect of the exposure on the mediator and the effect of the mediator on the outcome. Thus, unmeasured mediator-outcome confounding will bias the indirect effect in the product method.

The third assumption required to allow causal inferences from the product and difference methods is that there is no unmeasured exposure-mediator confounding. To identify the indirect effect using the difference method, the total effect (given by ϕ_1 in Equation 3.1) must first be identified. However, an exposure-mediator confounder is also an exposure-outcome confounder because it causes both the exposure and the outcome (Figure 3.8). Therefore, all exposure-mediator confounders need to be accounted for in the difference method. The product method

[4] VanderWeele (2015) provides a more detailed discussion about the assumptions required for direct and indirect effects to be interpreted causally.

[5] Note that this assumption is also of importance when doing mediation analysis with randomised experiments since the value of the mediator is not randomly allocated.

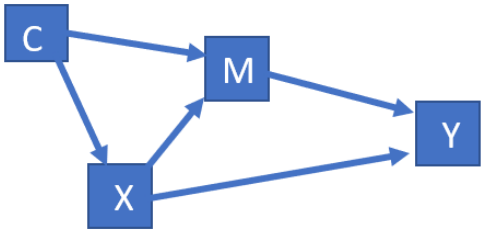

Figure 3.8: A DAG to demonstrate exposure-mediator confounding

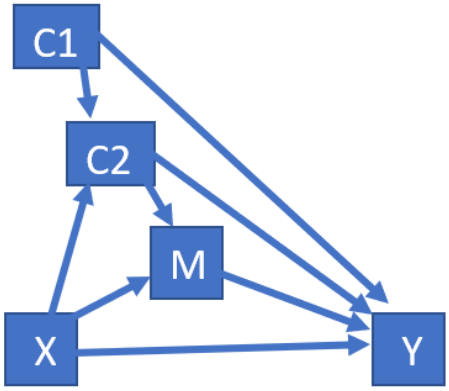

Figure 3.9: A DAG to demonstrate collider bias introduced by conditioning on a mediator-outcome confounder, $C2$, that is caused by the exposure

also requires that all exposure-mediator confounders are accounted for, since the effect of the exposure on the mediator (β_1 in Equation 3.3) is used in the estimation of the indirect effect.

The fourth assumption required is that no mediator-outcome confounder is caused by the exposure because otherwise the mediator-outcome confounder is also a mediator and may collider bias effect estimates that are required in the product and difference methods. For example, the path $X \rightarrow C_2 \leftarrow C_1 \rightarrow Y$ in Figure 3.9 is unblocked by conditioning on the mediator-outcome confounder C_2, which will collider bias the direct effect in the product and difference methods.

The fifth assumption is that no exposure-mediator interactions are present. To understand the reasoning for this, suppose that Equation 3.2 included an exposure-mediator interaction term such that

$$\mathbb{E}(Y|A = a, M = m, C = c) = \theta_0^* + \theta_1^* x + \theta_2^* m + \theta_3^* am + \theta_4^* c.$$

Then, the direct effect in the product and difference methods, originally given by θ_1 in Equation 3.2 would in theory now be given by $\theta_1^* + \theta_3^* m$. Therefore, the direct effect will change according to the value of the mediator, but this goes against the definition of, or intuition behind, a direct effect. If the direct effect includes the value of the mediator, then by intervening on the mediator some of the effect can be removed. Therefore, the proposed direct effect, $\theta_1^* + \theta_3^* m$, is not actually a true direct effect because it includes part of the indirect effect [249, 250]. Hence, for the product and difference methods to be used, it must be assumed that no exposure-mediator interactions are present [248].

Assuming that the model is correctly specified and the assumptions above hold, then the indirect effect in the product method will be algebraically equivalent to the indirect effect in the difference method for ordinary least squares linear regression [251, 252]. However, for logistic regression, indirect effects with a causal interpretation cannot be identified using the product and difference methods[6] [253].

To summarise, the product method and the difference method can be used to produce estimates with a causal interpretation conditional on confounding assumptions (assumptions 1–4) holding, and if no exposure-mediator interactions are present (assumption 5), and generally only for ordinary least squares linear regression [248].

3.5.3 Causal mediation analysis methods

To overcome the issues described above, counterfactual definitions can be used to make mediation analysis more formal and provide clarity over the exact meaning of direct and indirect effects.

Consider an outcome Y, mediator M, and exposure X. Let Y^x denote the subjects outcome whilst fixing $X = x$. Let M^x denote the subjects value of the mediator whilst fixing $X = x$. Let

[6](VanderWeele, et al. 2015) explains that in some specific scenarios the product and different methods can be used with logistic regression to approximate causal indirect effects.

Y^{xm} denote the value the outcome would take whilst fixing $X = x$ and $M = m$. Then, consider the following counterfactual definitions of direct and indirect effects given in (VanderWeele 2016) [244]. So, define the controlled direct effect, CDE^m, the natural direct effect, NDE, the natural indirect effect, NIE, as:

$$\begin{aligned} \text{CDE}^m &= Y^{xm} - Y^{\dot{x}m}, \\ \text{NDE} &= Y^{xM^{\dot{x}}} - Y^{\dot{x}M^{\dot{x}}}, \\ \text{NIE} &= Y^{xM^{x}} - Y^{xM^{\dot{x}}}. \end{aligned} \tag{3.4}$$

The controlled direct effect, CDE^m, is the effect of the exposure on the outcome (setting $X = x$ versus setting $X = \dot{x}$) whilst fixing $M = m$. Thus, the controlled direct effect is the effect of the exposure on the outcome that is not mediated through the mediator. The natural direct effect, NDE, is the effect of the exposure on the outcome (setting $X = x$ versus setting $X = \dot{x}$) whilst fixing $M = M^{\dot{x}}$; that is, the effect of the exposure on the outcome whilst fixing the mediator to the value it would have been had $X = \dot{x}$. The natural indirect effect, NIE, is the effect of setting $M = M^{x}$ versus $M = M^{\dot{x}}$, whilst fixing $X = x$. The NIE disables the direct effect by fixing $X = x$ and compares the effect of the mediator on the outcome if the mediator had taken the value it would when $X = x$ versus the effect of the mediator on the outcome if the mediator had taken the value it would when $X = \dot{x}$.

For the effects above to be identifiable, some assumptions must hold - namely, the confounding assumptions 1–4 from Section 3.5.2 [244, 248]. To recap, those assumptions were: (i) no unmeasured exposure-outcome confounding ($Y^{xm} \perp\!\!\!\perp X|C$), (ii) no unmeasured mediator-outcome confounding ($Y^{xm} \perp\!\!\!\perp M|X, C$), (iii) no unmeasured exposure-mediator confounding ($M^{x} \perp\!\!\!\perp X|C$), (iv) no mediator-outcome confounder is caused by the exposure ($Y^{xm} \perp\!\!\!\perp M^{\dot{x}}|C$).

If those assumptions hold then the controlled direct effect, natural direct effect, and natural indirect effect are identifiable. Further, the total causal effect decomposes into the natural direct

effect and natural indirect effect:

$$Y^{xM^x} - Y^{\dot{x}M^{\dot{x}}} = [Y^{xM^x} - Y^{xM^{\dot{x}}}] + [Y^{xM^{\dot{x}}} - Y^{\dot{x}M^{\dot{x}}}].$$

The counterfactual-based approach to mediation analysis does not require that no exposure-mediator interactions are present, and the definitions and properties are not limited to any specific model. In Section 3.7, this counterfactual approach will be used to describe how mediation analysis may be conducted with a survival outcome. First, Section 3.6 will provide a brief overview of the survival analysis methods that will be used throughout this thesis.

3.6 Survival analysis

Survival analysis is used for the investigation of time-to-event data. If the study outcome is an event that either occurs or doesn't occur within a set period of time, then a logistic regression model could be used. However, simply stating whether an event takes place or not is hiding some potentially important information. How soon after the start of the study did the event occur? If we are comparing the survival of patients after an operation and follow them up for 6 years then it is important to differentiate between those that died after 5 weeks and those that died after 5 years.

The aims of survival analyses may be to describe the distribution of survival times and compare these between study groups, or to investigate how an exposure/predictor/cause is associated with survival time [214].

3.6.1 Censoring

Many epidemiological studies will encounter the problem in which participants are lost to follow-up, follow-up ends, or the individual is removed from follow-up after suffering a competing risk. The loss or removal of individuals from the observed group of participants for whom the outcome is still possible is called censoring. Censoring can introduce problems for

survival analysis because the outcome of interest, the time-to-event, is unknown for censored individuals. Excluding these individuals from the analysis would result in loss of power, and may introduce bias [254].

There are three different types of censoring. The most common type is right censoring, in which the patient survives until a particular time point without experiencing the outcome of interest, but the actual survival time is unknown [214]. The survival time may not be known because the patient is lost to follow-up, the study follow-up period may end, or a competing outcome may remove them from the pool of susceptible individuals. Consider a study in which the outcome of interest is incident frozen shoulder. If a patient dies in a car crash at 6 months without ever developing frozen shoulder, then it is known that the survival time exceeds 6 months, but we do not know if or when the patient would have developed frozen shoulder if they did not die.

Interval-censoring refers to a scenario in which it is known that a unit has failed during a particular interval of time but the specific failure time is unknown [214]. For an example of interval censoring, consider a study in which the outcome of interest is death and participants are followed up at 6 month intervals and a patient dies between 12 and 18 months, but the exact date of death is unknown; this patient would be considered interval censored.

Left-censoring means that the survival time is known to be less than a certain time [214]. Consider a study in which incident type 2 diabetes is the outcome. Some patients may already have undiagnosed type 2 diabetes at the onset of study. These individuals would have 'failed' before the start of the study and would be ineligible for inclusion, meaning that they are left-censored.

Often in survival analyses it is assumed that censoring is non-informative, meaning that any censoring is independent of the survival time, had the survival time been observed. Assessing the believability of this assumption will require subject-knowledge. If censoring occurs due to the study follow-up period ending, then censoring will be non-informative. If a patient is miss-

ing monthly follow-up appointments because they are too ill to attend due to the side effects of the study exposure then this may be informative censoring. For the rest of Section 3.6, it will be assumed that censoring is non-informative.

3.6.2 Survival analysis notation

Let T be the time to the event of interest, also known as the survival time. Then, define the survival function, $S(t)$, as

$$S(t) = P(T > t), \tag{3.5}$$

thus denoting the probability of surviving (at least) until time T.

The probability density function, $f(t)$, is given by

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t},$$

which can be thought of as the risk of the event occurring during the infinitely small time interval (t, $t + \Delta t$), where Δt tends to zero.

The cumulative density function is therefore given by

$$\begin{aligned} F(t) &= \int_0^t f(t).dt \\ &= P(T \leq t). \end{aligned}$$

Returning back to Equation 3.5, the survival function may now be written in terms of the cumulative density function as

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t). \end{aligned}$$

The hazard function is the instantaneous risk of the event, conditional on the event having

not occurred previously. So, formally, the hazard function, $\lambda(t)$, is given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t | T > t)}{\Delta t}. \tag{3.6}$$

Using the law of total probability, Equation 3.6 can be rewritten as

$$\begin{aligned}
\lambda(t) &= \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t)}{P(T > t)\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \frac{1}{P(T > t)} \\
&= F'(t) \frac{1}{S(t)} \\
&= \frac{f(t)}{S(t)}.
\end{aligned}$$

3.6.3 Cox proportional-hazards regression

Cox regression [255] will be used throughout this thesis. It allows for the investigation of associations between exposures, predictors or causes and survival time. The Cox model for the hazard function, $\lambda(t|\mathbf{x}_j)$, may be defined as:

$$\lambda(t|\mathbf{x}_j) = \lambda_0(t)\text{exp}(\boldsymbol{\beta}^T \mathbf{x}_j), \tag{3.7}$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients, $\mathbf{x}_j$ is the vector of observed covariates for individual j, $j = 1, \ldots, n$, and $\lambda_0(t)$ is the baseline hazard function.

The baseline hazard function, $\lambda_0(t)$, is the value $\lambda(t)$ would take if $\mathbf{x}_j = \mathbf{0}$. The model requires that there is a multiplicative relationship between the covariates and the hazard function, and that the relationship is constant over time (i.e the hazards are proportional) [214, 255]. It is also necessary to assume that each individual's survival time is independent [214, 255].

The Cox model is said to be semi-parametric because the standard regression portion of the model, $\boldsymbol{\beta}^T \mathbf{x}_j$, is parametric, but the nature of the baseline hazard function, $\lambda_0(t)$, is left unspecified. If the nature of the underlying hazard function is unknown or difficult to model, then the

Cox model will be advantageous since the form of $\lambda_0(t)$ can be ignored [214]. If the main goal of the survival analysis is to understand the association of a variable with survival time then the nature of the hazard function is likely not important. A coefficient β_k may be interpreted as the multiplicative change in log-hazard expected with a one unit change in $x_{j,k}$, for an arbitrary individual j, whilst holding all other $x_{j,l}$, $k \neq l$, constant. Thus, $\exp(\beta_k)$ gives the corresponding hazard ratio.

The assumption of proportional hazards, that is required for the Cox model, can be assessed using Schoenfeld residuals for each model covariate. The Schoenfeld residual for a given variable for a given patient is given by the observed covariate value minus its expected value at the patient's event time. If the proportional hazards assumption is met, then the Schoenfeld residuals should be independent of time. Thus, a plot of the residuals for a given covariate against time should have zero slope and no pattern.

3.7 Mediation analysis with survival outcomes

3.7.1 Definitions for mediation analysis with a survival outcome

Various methods exist to approach mediation analysis with survival outcomes [248, 256], each of which require various assumptions. This thesis will focus on mediation analysis methods that do not require a rare-outcome assumption (as other approaches do [248]). To avoid being restricted to rare outcomes, the best approach is a weighting method. Before describing the details of the weighting approach, the definitions in Equation 3.4 may be converted to the log-hazard scale as [248]:

$$\begin{aligned}
\text{CDE}^m &= \log[\lambda^{xm}(t)] - \log[\lambda^{\dot{x}m}(t)], \\
\text{NDE} &= \log[\lambda^{xM^{\dot{x}}}(t)] - \log[\lambda^{\dot{x}M^{\dot{x}}}(t)], \\
\text{NIE} &= \log[\lambda^{xM^{x}}(t)] - \log[\lambda^{xM^{\dot{x}}}(t)],
\end{aligned}$$

again, assuming assumptions 1–4 from Section 3.5.2 hold. Here, it is also required that the

mediator is measured prior to the event of interest occurring [248]. For example, if the mediator is measured at 6 months and the outcome is death, but people die before the 6 month follow-up appointment, then the assumption is violated.

The total effect on the log-hazard scale can be decomposed into the NIE and NDE as:

$$\log[\lambda^{x}(t)] - \log[\lambda^{\dot{x}}(t)] = \left[\log[\lambda^{xM^{x}}(t)] - \log[\lambda^{xM^{\dot{x}}}(t)]\right] + \left[\log[\lambda^{xM^{\dot{x}}}(t)] - \log[\lambda^{\dot{x}M^{\dot{x}}}(t)]\right],$$

or exponentiated in terms of hazard ratios:

$$\frac{\lambda^{x}(t)}{\lambda^{\dot{x}}(t)} = \frac{\lambda^{xM^{x}}(t)}{\lambda^{xM^{\dot{x}}}(t)} \times \frac{\lambda^{xM^{\dot{x}}}(t)}{\lambda^{\dot{x}M^{\dot{x}}}(t)}.$$

3.7.2 The weighting approach to mediation analysis with a survival outcome

This section will introduce an inverse probability weighting approach to causal mediation analysis using a Cox model. Other approaches to mediation analysis with a survival outcome exist, but the weighting approach allows for a non-rare outcome which other methods do not [248]. This weighting approach has been described in (Hong, 2010) [257], (Hong et al., 2015) [258] and (Lange et al., 2012) [259]. Below, it will be assumed that the exposure is binary, but the method can be adapted to continuous or categorical exposures [259].

The first step in this method is to create a new dataset which includes two copies of each individual. Then, create a new variable, X^*, which may be defined as:

$$X_i^* = \begin{cases} X_i & \text{for the first copy of individual } i, \\ 1 - X_i & \text{for the second copy of individual } i. \end{cases}$$

Then two weights for each copy of each individual need to be created. First, the weight for

the exposure can be computed as

$$W_i^X = \frac{P(X = x_i)}{P(X = x_i|C = c_i)}, \tag{3.8}$$

which is a stabilised[7] inverse probability (IP) weight to adjust for confounding. If the exposure is binary then $P(X = x_i|C = c_i)$ can be estimated using logistic regression, adjusting for the covariate set C.

The weight for the mediator may be computed as

$$W_i^M = \frac{P(M = m_i|X = x_i^*, C = c_i)}{P(M = m_i|X = x_i, C = c_i)}. \tag{3.9}$$

Then, obtain the overall weight, W_i, for each copy of each individual in the new dataset by multiplying W_i^X and W_i^M together; that is,

$$W_i = W_i^X \times W_i^M. \tag{3.10}$$

Next, apply the weight W_i to each individual, i, in the new dataset. Then, fit the proportional hazards model

$$\lambda(t|x, x^*) = \lambda_0(t)\text{exp}(\psi_1 x + \psi_2 x^*). \tag{3.11}$$

The natural direct effect is given by ψ_1 and the natural indirect effect is given by ψ_2 (both on the log-hazard ratio scale). Confidence intervals for the direct, indirect and total effect must be obtained through bootstrapping [248]. (The bootstrapping procedure is explained in Section B.4.)

[7]The weight is stabilised because the term $P(X = x_i)$ replaces the number 1 in the numerator. The unstabilised version of the weight, $1/P(X = x_i|C = c_i)$ can lead to some individuals being assigned very large weights and dominating the analysis [260].

3.7.3 Intuition behind weighted mediation analysis

Section 3.7.2 described the procedure which allows for effect decomposition using Cox models. The method relies on inverse probability weighting for both the exposure and the mediator. First, consider the exposure weight, W_i^X, given in Equation 3.8. In an observational study, there will be covariate imbalances in the exposed and unexposed groups that prevent a measure of association directly being interpreted as a causal effect. In other words, the exposed and unexposed groups are not exchangeable. However, given the confounding assumptions 1–4 from Section 3.5.2 hold for a covariate set C, then there is conditional exchangeability conditional on C. The covariate set C may be accounted for using inverse probability weighting to account for any confounding of the causal effects of the exposure on the mediator or the outcome.

In the original population C and X are statistically dependent, meaning individuals within different levels of C will be more/less likely than others to be exposed. By weighting each individual in the population by the weight W_i^X, each individual in the weighted population will have an equal chance of being exposed. In the re-weighted population, the exposure will no longer be associated with C, deleting all backdoor paths from X to the mediator and from X to the outcome [146, 190]. This procedure attempts to create exchangeable exposed and unexposed groups, sharing the same pre-treatment/pre-exposure risk of outcome, as is achieved by randomisation.[8]

Now, for the mediation analysis, there needs to be a way to decompose the total causal effect into the natural direct effect and the natural indirect effect whilst accounting for mediator-outcome confounding. This is achieved by applying the mediator weight W_i^M, given in Equation 3.9, in a method called Ratio-of-Mediator-Probability Weighting (RMPW) [258].

First, recall that definitions of the natural direct effect and natural indirect effect for a binary exposure are given by $\lambda^{1M^0}(t)/\lambda^{0M^0}(t)$ and $\lambda^{1M^1}(t)/\lambda^{1M^0}(t)$, respectively (when written as hazard ratios). So, we need contrasts of the three terms $\lambda^{0M^0}(t)$, $\lambda^{1M^1}(t)$, and $\lambda^{1M^0}(t)$. The

[8] Although, IP weighting requires that C blocks all backdoor paths, which is a strong assumption, and a limitation that does not exist in randomised experiments.

first two terms, $\lambda^{0M^0}(t)$ and $\lambda^{1M^1}(t)$, can be obtained from the unexposed group and exposed group, respectively, within the weighted-population in which individuals were weighted by W_i^X [257]. This is because $\lambda^{0M^0}(t) = \lambda^0(t)$ and $\lambda^{1M^1}(t) = \lambda^1(t)$.

Contrarily, $\lambda^{1M^0}(t)$ needs an additional weight to account for the mediator value being set counterfactually to M^0 instead of the observed value M^1. If assumptions 1–4 from Section 3.5.2 hold then, within levels of C,[9] the only factor determining whether the mediator takes value $M = m$ is the value of the exposure. So, in the weighted population $P(M^x = m|C = c) = P(M = m|X = x, C = c)$. Therefore, to obtain $\lambda^{1M^0}(t)$, an additional weight, $P(M = m|X = 0, C = c)/P(M = m|X = 1, C = c)$, can be applied to the exposed group so that the probability of having mediator value $M = m$ is equal to what it would have been had they not been exposed. Then, $\lambda^{1M^0}(t)$ can be estimated in the new population of exposed individuals that have been weighted by both W_i^X and $P(M = m|X = 0, C = c)/P(M = m|X = 1, C = c)$.

Now, to estimate the direct and indirect effects, we need contrasts of the term $\lambda^{1M^0}(t)$, which requires individuals to be weighted by W_i^X and $P(M = m|X = 0, C = c)/P(M = m|X = 1, C = c)$, with $\lambda^{0M^0}(t)$ and $\lambda^{1M^1}(t)$, which require individuals to be weighted only by W_i^X. This can be achieved using the model in Equation 3.11 which includes the dummy indicator X^* to allow for the estimation of the term $\lambda^{1M^0}(t)$ from the same model as $\lambda^{1M^1}(t)$ and $\lambda^{0M^0}(t)$.

Consider the weight W_i^M, which includes the dummy indicator X_i^*. When $X_i^* = X_i$, $W_i^M = 1$. When $X_i = 1$ and $X_i^* = 1 - X_i = 0$, then $W_i^M = P(M = m|X = 0, C = c)/P(M = m|X = 1, C = c)$. So, using Equation 3.11, $\lambda^{0M^0}(t)$ is given by $\lambda_0(t)$; $\lambda^{1M^0}(t)$ is given by $\lambda_0(t)\exp(\psi_1)$; $\lambda^{1M^1}(t)$ is given by $\lambda_0(t)\exp(\psi_1 + \psi_2)$. Thus, on the hazard ratio scale, NIE $= \lambda^{1M^1}(t)/\lambda^{1M^0}(t) = \exp(\psi_2)$ and NDE $= \lambda^{1M^0}(t)/\lambda^{0M^0}(t) = \exp(\psi_1)$.

For a formal proof of the RMPW method, see (Lange et al. 2012) [259] and (Hong et al.

[9] Whilst X is independent of C in the population that has been weighted by W_i^X, there still may be mediator-outcome confounders that need to be accounted for.

2015) [258].

3.7.4 Sensitivity analysis

The assumptions of no unmeasured confounding are unlikely to hold completely. Through using a DAG and being transparent in the assumptions that are made, other researchers can scrutinise the assumptions and make judgements about their believability. In observational research, whilst it may be unlikely to eliminate confounding completely, the use of appropriate causal inference methods can aid the researcher in reducing confounding. Sensitivity analysis can then be used to assess the extent to which the total effect estimate may differ if there were some unmeasured confounder(s). One way to conduct sensitivity analysis for unmeasured confounding of the total causal effect is to calculate an E-Value. The E-Value can be defined as the minimum strength of the effect of an unmeasured confounder on the exposure and on the outcome (conditional on measured covariates) required in order to completely explain away the association estimate[10] [244, 261]. The E-Value for the association estimate is given by [261]

$$\text{E-Value} = \begin{cases} \omega + \sqrt{\omega(\omega-1)}, & \text{if HR}>1, \\ 1/\omega + \sqrt{1/\omega(1/\omega-1)}, & \text{if HR}<1, \end{cases}$$

where

$$\omega = \frac{1-0.5^{\sqrt{HR}}}{1-0.5^{\sqrt{1/HR}}}.$$

An E-value can also be obtained for the confidence interval limit closest to the null, to explain how strong the effect of the confounder on both the exposure and outcome must be in order for the confidence interval to include the null. The E-value for the confidence interval limit closest to the null is given by [261]

$$\text{E-Value} = \begin{cases} \omega_L + \sqrt{\omega_L(\omega_L-1)}, & \text{if HR}>1, \\ 1/\omega_U + \sqrt{1/\omega_U(1/\omega_U-1)}, & \text{if HR}<1, \end{cases}$$

[10]Note that, throughout this thesis, the E-Value will be relative to a binary exposure and a survival outcome. Thus, the effects of the confounder on the exposure and the confounder on the outcome will be estimated on the relative risk and hazard ratio scales, respectively.

where

$$\omega_L = \frac{1 - 0.5^{\sqrt{LL}}}{1 - 0.5^{\sqrt{1/LL}}}$$

and

$$\omega_U = \frac{1 - 0.5^{\sqrt{UL}}}{1 - 0.5^{\sqrt{1/UL}}}.$$

LL and UL, respectively, denote the lower and upper confidence interval limits. Where the confidence interval already includes the null, the E-value is defined to be equal to 1.

Chapter 4

Type 2 diabetes and the risk of developing frozen shoulder: a cohort study

The Independent Scientific Advisory Committee (ISAC) protocol for this study (19_219R) was accepted on 16th December 2020.

This study has been reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [262]. A completed STROBE checklist can be found in Appendix Section B.1.

4.1 Introduction

The systematic review in Chapter 2 identified limitations in current evidence for the relationship between diabetes and the development of frozen shoulder. This chapter describes a cohort study conducted in UK electronic health records that addresses these limitations. The main limitation of studies in the systematic review was that they either did not consider confounders or used covariate-selection methods that were unsuitable for an aetiologic model. This study uses causal inference methods with the aim of estimating the causal effect of type 2 diabetes on

the development of frozen shoulder[1] with a longer follow-up than existing studies. This study focuses specifically on type 2 diabetes, rather than all types of diabetes (as was done in Chapter 2), to avoid violating the causal identifiability assumptions discussed in Section 3.4.2. An explanation of how the assumptions may be violated is given in Section 4.3.3.

Additionally, the association between other metabolic factors and the development of frozen shoulder has previously only been investigated in cross-sectional research [28]. Authors have highlighted that the role of metabolic syndrome in the development of musculoskeletal conditions is often overlooked [102, 105]. It has been hypothesised that inflammation may play a key role in the pathogenesis of frozen shoulder (Sections 1.1.3 and 1.3). Since metabolic syndrome is associated with chronic inflammation [263–265], it has been hypothesised that metabolic syndrome could be part of the reason why people with type 2 diabetes are more likely to develop frozen shoulder [24, 106]. Thus, in this study, a mediation analysis was conducted to understand whether people with type 2 diabetes who develop additional metabolic factors are significantly more at risk of developing frozen shoulder.

4.2 Objectives

Objective i

To estimate the causal effect of type 2 diabetes on the development of frozen shoulder.

Objective ii

To estimate the proportion of the effect of type 2 diabetes on the risk of developing frozen shoulder that is mediated through other metabolic factors.

[1] Although the limitations of using observational data to estimate causal effects have been acknowledged and the limitations are discussed throughout this chapter.

4.3 Methods

4.3.1 Study sample

This matched cohort study was conducted in CPRD GOLD with linkage to IMD and HES data. Patients were eligible if they were aged 18 years or over and presented with their first ever Read code for type 2 diabetes between 1st May 2004 and 31st December 2017. Index date was defined as the date of the patient's first ever type 2 diabetes Read code. It was required that patients did not have any shoulder-related Read codes before index date, and patients were required to have at least two years of up-to-standard data at the index date. Exposure, outcome and covariate Read code lists, which were either constructed by general practitioners or obtained from previous studies within the Primary Care Centre Versus Arthritis, can be found in Appendix B.2. Confounders had to be reported in CPRD before the index date and mediators had to be recorded in CPRD after the index date, but before the end of the patient's follow-up.

Each patient was matched to one individual with the exact same year of age, same gender, and from the same practice, but without a diabetes diagnosis prior to the index date of their matched pair. The matched individuals without diabetes were also required to not have any shoulder-related Read codes before index date, and were also required to have at least two years of up-to-standard data at the index date. The matched individuals without diabetes were also required to be alive and at a CPRD practice on the index date.

Patients were followed from their index date until the earliest of: end of follow up (17th February 2020), date of frozen shoulder diagnosis, date of death (derived from CPRD data), date of transfer to a non-CPRD practice, or date of last CPRD data collection.

4.3.2 Analysis plan

The causal mediation analysis methods described in Section 3.7.2 were used to estimate the causal effect of type 2 diabetes on the development of frozen shoulder, and to analyse how much of the effect was mediated by metabolic health. The number of metabolic factors (hypertension, hyperlipidaemia, obesity) developed during follow-up (post-index date) was used as an indicator of the patient's metabolic health. Due to only a small proportion of patients developing all three metabolic factors during follow-up (Table 4.1), the number of metabolic factors were classed as 0, 1, $\geq$2. Weights for the mediator were obtained through ordinal logistic regression and weights for the exposure were obtained using logistic regression. In both the exposure weight model and mediator weight model, generalised estimating equations (GEEs) [266] were used to estimate parameters whilst accounting for the correlation in outcomes for matched individuals. Standard errors for the causal mediation analysis effect estimates were obtained through bootstrapping (an explanation of how bootstrap confidence intervals were obtained is given in Appendix Section B.4). An interaction term between 'type 2 diabetes' and 'the number of metabolic factors developed during follow-up' was initially included in the mediation analysis but was removed since there was no evidence of interaction. The proportional hazards assumption was checked using graphical diagnostics using Schoenfeld residuals. E-values were computed to assess the impact that unmeasured confounding may have had on the total causal effect estimate. The Cox model was re-run using truncated weights [267, 268], as a sensitivity analysis, to assess the impact that extreme weights may have on the results. Exposure weights (W_i^X from equation 3.8) exceeding the 95$^{\text{th}}$ percentile were truncated at the value of the 95$^{\text{th}}$ percentile [267, 268].

The covariate adjustment set for the model was identified using the DAG described in Section 4.3.3. As well as being matched on, age and gender were also adjusted for in the analysis to avoid bias, as recommended in (Sjölander et al. 2013 [269]). Due to some factor levels of categorical variable levels having small cell counts, some categories were collapsed. A breakdown of the collapsing of categories is given in Appendix Section B.3.

The missing data indicator method was used to handle missing data for smoking, alcohol, and obesity. The approach requires that an extra category is created to indicate whether data on a variable are missing or not. This approach allows for participants to be included in the analysis despite having incomplete data; thus, the approach reduces the loss of statistical power. Otherwise, missing Read codes were assumed to indicate that the patient did not have the corresponding disease e.g., if a patient has no record of a frozen shoulder Read code then it was assumed that they never had a frozen shoulder. As a sensitivity analysis, a complete case analysis was conducted to assess the extent to which missing data may have affected the results.

Participants were censored upon their death, transfer to a non-CPRD practice, or at the end of follow-up (17th February 2020). Patients in the control group were censored if they were diagnosed with type 2 diabetes after the index date.

Data were prepared for analysis in Stata version 14.0 [175], and analysis was conducted in RStudio version 1.2.5033 [270]. R code can be found in Appendix Section B.4.

4.3.3 Specifying the knowledge and assumptions about the process which leads to type 2 diabetes (potentially) causing frozen shoulder

Figure 4.1 contains a DAG illustrating any knowledge and assumptions about the causal relationships between covariates that could potentially affect (I) the estimated total effect of type 2 diabetes on the risk of developing frozen shoulder, or (II) the indirect effect of type 2 diabetes on the risk of developing frozen shoulder, mediated by other metabolic factors. The DAG in Figure 4.1 has been constructed consistent with the recommendations given in (Tennant et al. 2019 [198]). Some key recommendations from this paper that will provide an insight into how the DAG in Figure 4.1 was constructed are:

- Variables are arranged spatially to reflect the passage of time.
- Arcs should only flow in one direction since causal processes can only occur with time and not against it.

- Assuming that there is zero causal effect of one variable on another variable is a much stronger assumption than assuming a very small effect may exist between two variables. Thus, it should generally be assumed that arcs exist between two variables and excluding an arc should be justified by theory and/or evidence.

The only arcs that were excluded from the DAG were those from gender to deprivation and gender to practice since there is no reason to believe that any geographical areas or practices significantly differ from a 50/50 gender split. It is also worth noting, thyroid dysfunction is included because it is known to be associated with an elevated risk of developing type 2 diabetes [271–273] (and other metabolic factors [274–277]) and frozen shoulder [29, 30, 181, 278]. Age and gender are included since, as mentioned in Chapter 1, they are known to be associated with both type 2 diabetes and frozen shoulder. Ethnicity, deprivation, practice, smoking, and alcohol were included because there was a lack of evidence/theory to argue that they have zero effect on type 2 diabetes, metabolic factors, and frozen shoulder; further, it is plausible that these variables could affect the risk of developing type 2 diabetes, metabolic factors, and frozen shoulder.

To meet the causal identifiability assumptions described in Section 3.4.2 it was required that the study focused on one specific type of diabetes as the exposure of interest. The different types of diabetes are associated with different covariates and thus will lead to a different DAG and different adjustment set. So, exchangeability could not be met if 'diabetes' was the exposure of interest rather than 'type 2 diabetes'. Furthermore, the consistency assumption may be violated if the different types of diabetes are associated with different risks of developing frozen shoulder. Considering that the different types of diabetes are associated with different levels of hyperglycaemia and inflammation (the two most popular hypotheses for why diabetes may cause frozen shoulder), it is likely that the consistency assumption would not be met. Restricting the exposure to 'type 2 diabetes' rather than all types of diabetes will provide a better chance of the consistency assumption being satisfied. (The consistency assumption is further discussed in the study limitations in Section 4.5.)

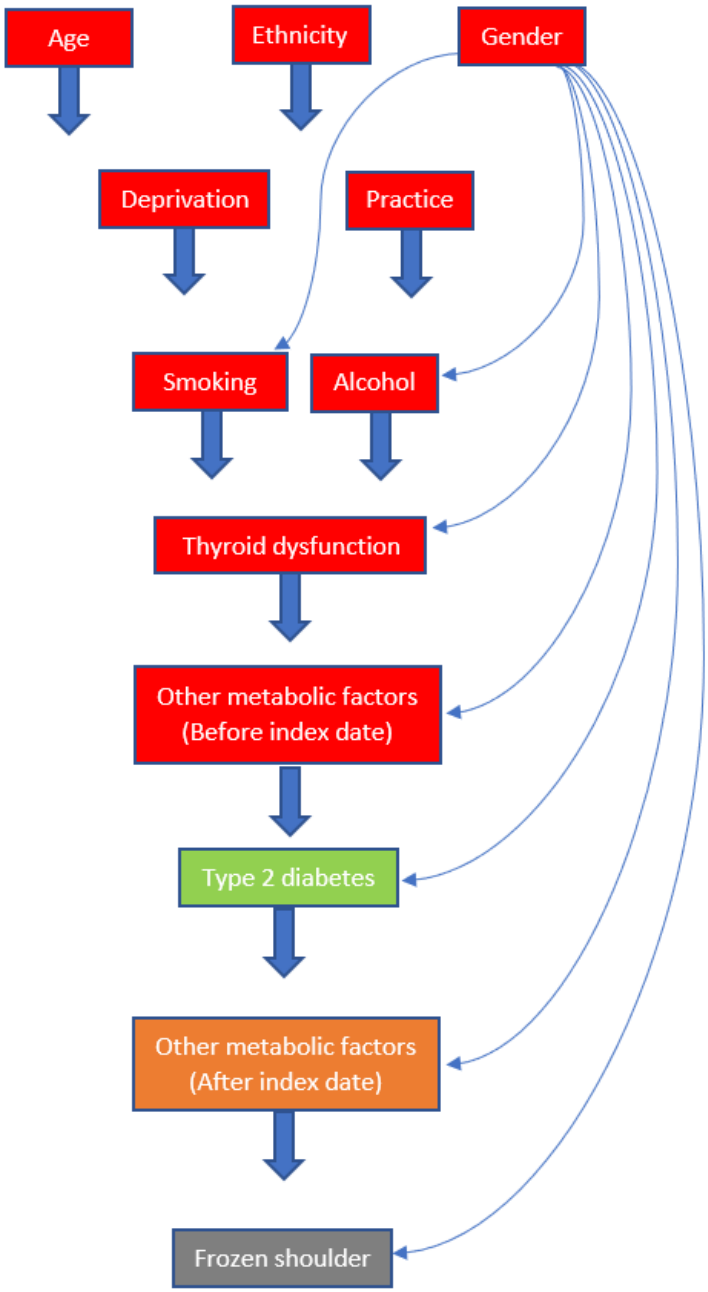

Figure 4.1: A DAG to illustrate known or potential causal relationships that could affect the process from which type 2 diabetes leads to the development of frozen shoulder.

To avoid overcrowding the diagram, thick block arrows have been used to indicate arcs to all nodes below. Thin arrows are used to indicate potential direct causal relationships between variables. The exposure, type 2 diabetes, is coloured green; the outcome, frozen shoulder, is coloured grey; the mediator, other metabolic factors (post-index date), is coloured orange; exposure-outcome confounders are coloured red. All exposure-outcome confounders were also mediator-outcome confounders (and vice versa). Note: the 'other metabolic factors' are: obesity, hypertension, and hyperlipidaemia

The DAG-implied adjustment set for the estimation of the total effect, direct effect, and indirect effect is: {age; ethnicity; gender; deprivation (IMD score); practice; smoking; alcohol; thyroid dysfunction; obesity, hypertension, hyperlipidaemia pre-index date}. Note that all confounders in this adjustment set were measured before the index date. Also, note that it was not necessary to create a composite variable of 'number of metabolic factors pre-index' in order to satisfy the no-confounding assumptions required for the mediation analysis; thus, obesity pre-index date, hypertension pre-index date, hyperlipidaemia pre-index date are left as three separate variables. The variables are represented as one node in the DAG to avoid overcrowding the DAG.

4.4 Results

4.4.1 Sample characteristics

The study sample consisted of 43,977 patients with incident type 2 diabetes diagnosed between 1st May 2004 and 31st December 2017, matched to 43,977 age-, gender-, and practice-matched adults that did not have a diabetes Read code up to or on their index date (Table 4.1). Fifty-seven percent of the patients with type 2 diabetes were male. The mean age at which the patients with type 2 diabetes were diagnosed with type 2 diabetes was 59.43 years (SD=14.01). The median follow-up duration was 8.24 years (IQR: 4.90–11.65) in people with type 2 diabetes and 9.06 (IQR: 5.95–12.19) in people without diabetes. People with diabetes were more likely to have alcohol and smoking status recorded in CPRD.[2]

	Type 2 diabetes n=43,977 (50%)	**No diabetes n=43,977 (50%)**	**All n=87,954**
Median follow-up[3] in years (IQR)	8.24 (4.90–11.65)	9.06 (5.95–12.19)	8.69 (5.42–11.93)
Mean age (years)	59.43 (SD=14.01)	59.43 (SD=14.01)	59.43 (SD=14.01)
Gender			
Male	25,236 (57.38%)	25,236 (57.38%)	50,472 (57.38%)
Female	18,741 (42.62%)	18,741 (42.62%)	37,482 (42.62%)
Hypertension (pre-index date)			
Diagnosed	21,538 (48.98%)	9,803 (22.29%)	31,341 (35.63%)
Not diagnosed	22,439 (51.02%)	34,174 (77.71%)	56,613 (64.37%)
Hyperlipidaemia (pre-index date)			
Diagnosed	8,649 (19.67%)	3,782 (8.60%)	12,431 (14.13%)
Not diagnosed	35,328 (80.33%)	40,195 (91.40%)	75,523 (85.87%)
Obesity (pre-index date)			

Continued on next page

[2]This is consistent with previous research suggesting that people with comordbities (especially those which require routine health monitoring, such as diabetes) are more likely to have more complete recording of alcohol and smoking status [279].

[3]Defined as time from index data to the earliest of: end of study (17th February 2020), date of death, date of transfer to a non-CPRD practice, or date of last CPRD data collection.

	Type 2 diabetes n=43,977 (50%)	**No diabetes n=43,977 (50%)**	**All n=87,954**
Obese	23,603 (53.77%)	6,687 (15.23%)	30,290 (34.50%)
Not obese	17,599 (40.09%)	25,847 (58.88%)	43,446 (49.49%)
Missing	2,693 (6.14%)	11,361 (25.88%)	14,054 (16.01%)
Number of metabolic factors[4] developed during follow-up			
0	30,542 (69.45%)	38,609 (87.79%)	69,151 (78.62%)
1	11,052 (25.13%)	4,466 (10.16%)	15,518 (17.61%)
2	2,173 (4.94%)	828 (1.88%)	3,001 (3.41%)
3	210 (0.48%)	74 (0.17%)	284 (0.32%)
Thyroid dysfunction			
Diagnosed	4,014 (9.13%)	2,396 (5.45%)	6,410 (7.29%)
Not diagnosed	39,963 (90.87%)	41,581 (94.55%)	81,544 (92.71%)
Type of thyroid dysfunction Read code			
Congenital	13 (0.03%)	7 (0.02%)	20 (0.02%)
Hyperthyroidism	625 (1.42%)	418 (0.95%)	1,043 (1.19%)
Hypo/Hyperthyroidism[5]	588 (1.34%)	319 (0.73%)	907 (1.03%)
Hypothyroidism	3,043 (6.92%)	1,622 (3.69%)	4,665 (5.30%)
Malignant	17 (0.04%)	19 (0.04%)	36 (0.04%)
Surgery	398 (0.91%)	287 (0.65%)	685 (0.78%)
Other	1,010 (2.30%)	703 (1.60%)	1,713 (1.95%)
Smoking			
Yes	8,149 (18.53%)	8,498 (19.32%)	16,647 (18.93%)
No	20,686 (47.04%)	20,292 (46.14%)	40,978 (46.59%)
Ex	14,398 (32.74%)	8,599 (19.55%)	22,997 (26.15%)
Missing	744 (1.69%)	6,588 (14.98%)	7,332 (8.34%)
Alcohol			
Yes	29,468 (67.01%)	27,074 (61.56%)	56,542 (64.29%)
No	8,090 (18.40%)	4,814 (10.95%)	12,904 (14.67%)
Ex	1,487 (3.38%)	584 (1.33%)	2,071 (2.35%)
Missing	4,932 (11.21%)	11,505 (26.16%)	16,437 (18.69%)
Ethnicity			

Continued on next page

[4]Hyperlipidaemia, hypertension, or obesity.
[5]Referring to non-specific codes that could relate to hypothyroidism or hyperthyroidism.

	Type 2 diabetes n=43,977 (50%)	**No diabetes n=43,977 (50%)**	**All n=87,954**
Bangladeshi	124 (0.28%)	47 (0.11%)	171 (0.19%)
Black African	184 (0.42%)	102 (0.23%)	286 (0.33%)
Black Caribbean	223 (0.51%)	141 (0.32%)	364 (0.41%)
Black – other	78 (0.18%)	59 (0.13%)	137 (0.16%)
Chinese	79 (0.18%)	56 (0.13%)	135 (0.15%)
Indian	590 (1.34%)	184 (0.42%)	774 (0.88%)
Mixed	168 (0.38%)	93 (0.21%)	261 (0.30%)
Other Asian	316 (0.72%)	98 (0.22%)	414 (0.47%)
Other	356 (0.81%)	286 (0.65%)	642 (0.73%)
Pakistani	320 (0.73%)	141 (0.32%)	461 (0.52%)
Missing	11,020 (25.06%)	12,331 (28.04%)	23,351 (26.55%)
White	30,519 (69.40%)	30,439 (69.22%)	60,958 (69.31%)
IMD Quintile			
Least deprived quintile	8,535 (19.41%)	9,899 (22.51%)	18,434 (20.96%)
2nd least deprived quintile	9,178 (20.87%)	9,675 (22.00%)	18,853 (21.44%)
3rd least deprived quintile	9,376 (21.32%)	9,320 (21.19%)	18,696 (21.26%)
4th least deprived quintile	8,811 (20.04%)	8,140 (18.51%)	16,951 (19.27%)
Most deprived quintile	8,052 (18.31%)	6,885 (15.66%)	14,937 (16.98%)
Missing	25 (0.06%)	58 (0.13%)	83 (0.09%)

Table 4.1: Table summarising baseline characteristics for study participants

4.4.2 Results for the effect of type 2 diabetes on the development of frozen shoulder

During the study, 1076 (1.22%) patients developed frozen shoulder. Within the type 2 diabetes group, 797 out of 43,977 (1.81%) patients developed frozen shoulder, and 279 out of 43,977 (0.63%) patients without diabetes developed frozen shoulder. The Kaplan-Meier plot[6] in Figure 4.2 shows the difference in survival probabilities between the type 2 diabetes group and the control group. The 1–, 3–, 5–, 10–, 15.8–year[7] Kaplan-Meier estimates for the patients with

[6]Note that the y-axis does not start from zero. This choice was made so that the reader can see the shape of the curves more clearly.

[7]Note that 15.8 years represents the full duration of the follow-up period.

type 2 diabetes and the patients without diabetes were 99.7%, 99.0%, 98.6%, 97.8%, 97.7% and >99.9%, 99.7%, 99.5%, 99.3%, 99.3% respectively.

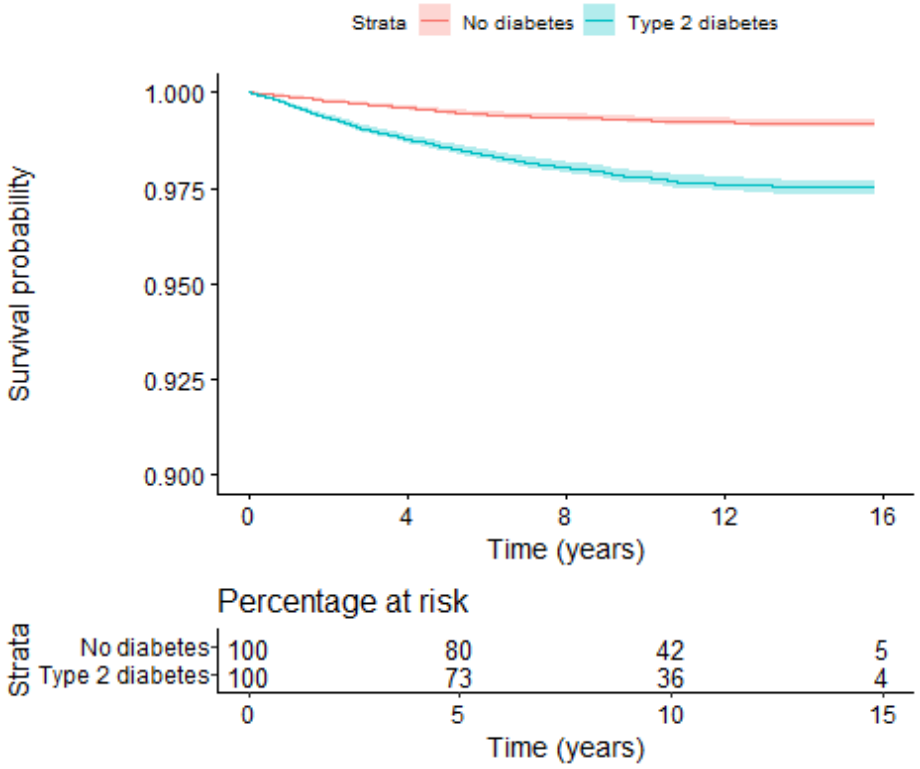

Figure 4.2: Plot of Kaplan-Meier estimates (with 95% confidence intervals) for surviving a frozen shoulder diagnosis among patients with type 2 diabetes compared to patients without diabetes, with a percentage at risk (of developing frozen shoulder) table

The total effect of type 2 diabetes on developing frozen shoulder was estimated to be HR = 4.38 (95% CI: 3.70 – 5.21). The corresponding E-value (the strength of effect that an unmeasured confounder would need to have on the exposure, conditional on measured covariates, and the outcome to completely explain away the association estimate) was estimated to be 4.87 for the point estimate and 4.30 for the lower bound of the 95% CI.

The total effect of type 2 diabetes on the development of frozen shoulder was decomposed into the natural direct effect, estimated to be HR = 4.46 (95% CI: 3.68 – 5.41), and the natural indirect effect, estimated to be HR = 0.98 (95% CI: 0.93 – 1.03). Therefore, there was no evidence of mediation.

There was little evidence of any violation of the proportional hazards assumptions for either the direct effect or the indirect effect. The smoothing splines in Figures 4.3 and 4.4 form horizontal lines with no pattern; this suggests that the residuals are independent of time. Further, the Kaplan-Meier plot in Figure 4.2 showed no crossover or divergence of Kaplan-Meier curves.

Truncated weights were used to assess the impact that extreme weights (exceeding the 95$^{\text{th}}$ percentile) may have had on the effect estimates. The results were similar when using truncated weights; the total effect was estimated to be HR = 3.10 (95% CI: 2.78 – 4.13), the direct effect was estimated to be HR = 3.32 (95% CI: 2.90 – 4.15), and the indirect effect was estimated to be HR = 0.93 (95% CI: 0.90 – 1.07). The distribution of the exposure weights (W_i^X in equation 3.8), mediator weights (W_i^M in equation 3.9), final weights (W_i in equation 3.10) and truncated final weights can be seen in Figures B.1–B.4.

When repeating the analysis only on patients with complete data (80.5% of patients had complete data), the results were similar, aside from the direct effect (and therefore also the total effect) being smaller in magnitude. The total effect was estimated to be HR = 2.59 (95% CI: 2.20 – 3.05), the direct effect was estimated to be HR = 2.62 (95% CI: 2.21 – 3.09), and the indirect effect was estimated to be HR = 0.99 (95% CI: 0.94 – 1.04).

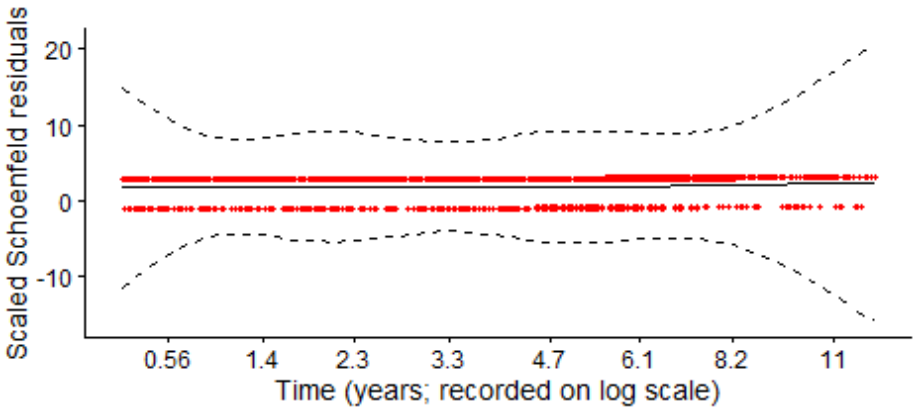

Figure 4.3: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the direct effect coefficient in the Cox model

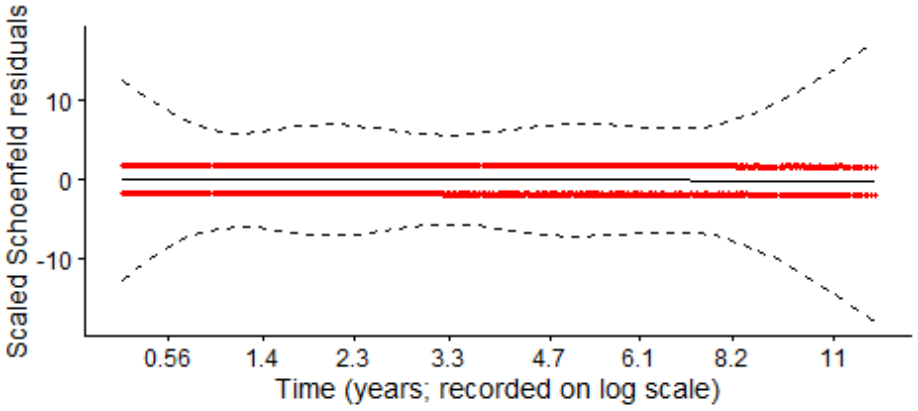

Figure 4.4: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the indirect effect coefficient in the Cox model

4.5 Discussion

Previous research has suggested that patients with diabetes are more likely to develop frozen shoulder than patients without diabetes (Chapter 2). Further, it had also been hypothesised that type 2 diabetes may be a cause of frozen shoulder. Researchers have suggested that the relationship could be due to glycation processes causing changes to capsule tissues, or that diabetes-related inflammation may lead to fibrotic changes in tissues. This cohort study used causal inference methods to provide evidence to support the hypothesis that type 2 diabetes could be a cause of frozen shoulder (HR = 4.38; 95% CI: 3.70 – 5.21). A causal mediation analysis was conducted to investigate whether the development of other metabolic factors could mediate the relationship between type 2 diabetes and frozen shoulder. This research was conducted to explore the mechanisms through which type 2 diabetes may lead to the development of frozen shoulder. However, the evidence did not suggest that the number of metabolic factors a patient developed mediated the effect of type 2 diabetes on frozen shoulder.

The results of this study are consistent with those in the systematic review in Chapter 2. To recap, six case-control studies and two cohort studies all found that people with diabetes were more likely to develop frozen shoulder. When compared to the two cohort studies from the systematic review, this study estimated a larger hazard ratio. One could speculate that this could be due to the differences in population; this study only included people with newly diagnosed type 2 diabetes, whereas the other two cohort studies did not restrict the type of newly diagnosed diabetes. It may be expected that the two studies including people with newly diagnosed type 1 diabetes would have had a smaller hazard ratio because many of the patients with type 1 diabetes would not be at the optimal age to develop frozen shoulder by the end of follow-up.[8] Further, this study is the first cohort study on this topic to be conducted outside of Taiwan. Different countries will have different population and environment/lifestyle characteristics, but also different health care strategies (for example, screening strategies, routine check-ups, and performance management strategies such as the QOF) to diagnose and manage patients with

[8]The follow-up duration in (Huang et al. 2013) was 3 years [184] and in (Lo et al. 2014) was 8 years [31].

diabetes.

This study is the first to investigate the pathways through which type 2 diabetes may lead to the onset of frozen shoulder. Further, this study is the first to apply causal inference methods to estimate the effect of type 2 diabetes on the development of frozen shoulder. DAGs and the backdoor criterion were used to identify confounders and avoid blocking causal pathways. A sensitivity analysis was conducted and demonstrated that it would require strong unmeasured confounding to explain away the estimated effect of type 2 diabetes on frozen shoulder. Whilst it is possible to calculate E-values for direct and indirect effects, they were not required for this study since there was no evidence of mediation.

Within this primary care cohort, the proportion of participants that had a Read code for frozen shoulder during follow-up (1.22%) was smaller than expected. This is likely a result of general practitioners being hesitant to record a specific diagnosis for shoulder pain and instead opt to use non-specific shoulder pain codes [280, 281]. The limitation of using electronic health records to identify patients with frozen shoulder will apply throughout this thesis so is discussed in the general thesis discussion in Section 8.4.

Whilst CPRD provided a large sample of patients with type 2 diabetes that should have been broadly representative of the UK population of people with type 2 diabetes, a limitation of using CPRD is the amount of missing data. The missing indicator method was used to avoid losing a large proportion of the sample and potentially introducing selection bias. However, the missing indicator method can produce biased estimates. A complete case analysis was conducted to assess the sensitivity of the results to the impact of missing data. However, the conclusions drawn from the analysis would not have been different if conducting a complete case analysis, compared to the missing indicator method.

To assess the extent to which metabolic health may mediate the effect of type 2 diabetes on the development of frozen shoulder, the number of metabolic factors identified and recorded in the primary care records during follow-up was used as an indicator of metabolic health. Whilst

the number of metabolic factors that are recorded for an individual is likely to be a good indicator of the patient's metabolic health, there are variations within levels of this variable i.e., there is a lack of consistency (see Section 3.4.2). For example, patients may have developed the same number of metabolic factors, but developed different metabolic factors, so they would be classed within the same level of the mediator variable. Further, this definition of metabolic health requires that the metabolic factor measurements be categorised (for example, blood pressure is categorised into hypertensive/not hypertensive) which also leads to a loss of information. The mediator was defined as 'the number of metabolic factors that the patient developed' rather than a binary variable indicating whether the patient developed metabolic syndrome since this provided extra information about the patient's metabolic health. (For example, a patient that develops two metabolic factors but does not have metabolic syndrome could be said to have worse metabolic health than someone that does not develop a single metabolic factor.) It is also worth noting that it was not possible to include each metabolic factor as a separate mediator since the metabolic factors are interrelated and cannot be separated in a way that would satisfy the no unmeasured confounding assumptions required for causal mediation analysis.

A further limitation of the mediation method used is that it assumes that once a patient has a metabolic factor then they have it for the rest of follow-up. This is a simplifying assumption that is required for the mediation analysis, although the assumption will not hold for all participants. For example, a participant may become obese and then return to a healthy weight. The period of being obese will have a negative impact on the patient's metabolic health, although not as much as someone who remains obese throughout follow-up. The exact impact of temporary obesity on a patient's metabolic health is difficult to determine without the inclusion of time-dependent mediation methods [248]. Such methods would greatly complicate the interpretability of results and thus I believe that making the simplifying assumption that a patient keeps the metabolic factor throughout follow-up is the best approach to answer the study's research question.

The limitations in the way that metabolic health is measured in this study mean that readers should have some hesitation in completely ruling out the hypothesis that metabolic health may mediate the effect of type 2 diabetes on the development of frozen shoulder. Future research

may be conducted with a different approach to measuring metabolic health and may find some evidence of mediation, however this study did not find any evidence to support this hypothesis.

In this study, the exposure of interest, type 2 diabetes, was identified using Read codes which indicate a clinical diagnosis of type 2 diabetes. As mentioned in Chapter 1, it has been hypothesised that glycation may lead to fibrotic changes in the shoulder capsule of patients with frozen shoulder. If this hypothesis is correct then one would expect that patients with worse glycaemic control would be more likely to develop frozen shoulder. Previous attempts to investigate the association between HbA1c and frozen shoulder have been limited by their cross-sectional design [200, 201]. One study created a marker of historical glycaemic control in their analysis and did find that their marker was positively associated with frozen shoulder [200]. Future research could focus on investigating whether prospective longitudinal measurements of HbA1c predict the occurrence of frozen shoulder; this could be achieved using joint modelling strategies.

Lastly, one could investigate whether certain coexisting metabolic factors moderate the effect of type 2 diabetes on the development of frozen shoulder. It should be noted that the results of such a study should not be used to draw conclusions about whether intervening on the metabolic factors would help to reduce the risk of developing frozen shoulder. (See VanderWeele 2009 [282] or Rothman et al. 2008 [145] for more on the distinction between interaction and effect modification.) Rather, the results would just help to gain an understanding of whether there are certain groups of people (i.e. those with certain metabolic factors) for which type 2 diabetes may have an extra effect on the development of frozen shoulder.

4.6 Conclusion

This study strengthens the evidence suggesting that people with diabetes are at a greater risk of developing frozen shoulder and supports the hypothesis that type 2 diabetes may potentially be a cause of frozen shoulder. Future work should focus on understanding why people with type 2 diabetes are more likely to develop frozen shoulder. Although previous research investigating

the association between glycaemic control and the development of frozen shoulder has produced mixed results, this should remain a focus.

The next chapter describes a cohort study conducted in CPRD which aimed to describe the association between newly diagnosed frozen shoulder and a subsequent diagnosis of type 2 diabetes in primary care. Following some small studies suggesting that people with frozen shoulder are more at risk of having undiagnosed type 2 diabetes than people without frozen shoulder [130, 283, 284], there has been debate about whether people with frozen shoulder should be screened for type 2 diabetes [130, 131]. The following study will add new evidence to this debate using a large, representative sample of patients presenting with frozen shoulder in primary care.

Chapter 5

Are patients with newly diagnosed frozen shoulder more likely to be diagnosed with type 2 diabetes? A cohort study

The ISAC protocol for this study (19_219R) was accepted on 16th December 2020.

This study has been reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [262]. A completed STROBE checklist can be found in Appendix Section C.1.

5.1 Introduction

It has been estimated that around 1 million people in the UK have undiagnosed type 2 diabetes [285], and the prevalence of diabetes in people with frozen shoulder has been estimated to be 30% (95% CI: 24 – 37%) [25]. Until now there has been limited high-quality evidence investigating whether patients with newly diagnosed frozen shoulder are more likely to be diagnosed with type 2 diabetes than patients without frozen shoulder.

Patients with type 2 diabetes may not experience any symptoms, meaning that (in the absence of routine health monitoring) they may go undiagnosed for many years. It has been estimated that the onset of type 2 diabetes may occur, on average, more than 7 years before its diagnosis [286, 287]. During the time that hyperglycaemia is uncontrolled, patients may develop serious health complications such as nephropathy [288, 289], retinopathy [290], neuropathy [291–293] and cardiovascular complications [294, 295]. Thus, it is important to detect type 2 diabetes as early as possible so patients can be monitored and receive treatment to reduce the likelihood of any complications occurring.

Previous studies have suggested that people with frozen shoulder should be screened for type 2 diabetes [130, 131]. This cohort study conducted in CPRD will enhance the understanding of whether, and the extent to which, people with frozen shoulder may be more likely to be diagnosed with type 2 diabetes. This may motivate future research to investigate the efficacy and cost-effectiveness of testing patients with newly diagnosed frozen shoulder for type 2 diabetes.

5.2 Objectives

To determine the association of newly diagnosed frozen shoulder with a subsequent diagnosis of type 2 diabetes in primary care.

5.3 Methods

5.3.1 Study sample

This matched cohort study was conducted in CPRD GOLD with linkage to IMD and HES data. To be included in the frozen shoulder group for this study, patients needed to be at least 18 years old and have had their first ever Read code of frozen shoulder between 1st May 2004 and 31st December 2017 (the date of frozen shoulder diagnosis was defined as the index date). It was required that patients did not have any diabetes Read code prior to the index date. Patients

needed to have at least two years of up-to-standard data at the index date. Variable Read code lists can be found in Appendix B.2.

Each patient with frozen shoulder on their index date was age-, gender- and practice-matched to an individual without a frozen shoulder diagnosis prior to the index date of the matched pair. Each matched individual without frozen shoulder was required to not have any diabetes Read code prior to the index date and was also required to have at least two years of up-to-standard data at the index date. The matched individuals without frozen shoulder needed to be alive and at a CPRD practice on the index date.

Patients were followed from their index date until the earliest of: end of study follow-up (17^{th} February 2020), date of type 2 diabetes diagnosis, date of death (derived from CPRD data), date of transfer to a non-CPRD practice, or date of last CPRD data collection.

5.3.2 Analysis

A Cox proportional hazards model was used to obtain hazard ratios to estimate the association between incident frozen shoulder and the likelihood of a following type 2 diabetes diagnosis. For the primary analysis, only age and gender were adjusted for in the Cox model (age, gender and practice were also matched on). A shared frailty term with a gamma distribution was used to account for the matching of individuals with frozen shoulder to people without frozen shoulder (an introduction to shared frailty models is given in Section 5.3.3).

A second Cox model with a shared frailty term was used to investigate whether the association between incident frozen shoulder and a type 2 diabetes diagnosis could be explained by other factors that are known or hypothesised to be associated with frozen shoulder and the diagnosis of type 2 diabetes (mean number of consultations per year, hyperlipidaemia, hypertension, obesity, thyroid dysfunction, ethnicity, deprivation, age, and gender). Missing data categories were created for any missing data. To avoid excessively small cell counts, the ethnicity variable levels were collapsed into “white”, “other” and “missing” categories. A breakdown of the cate-

gories of categorical variables is given in Appendix Section B.3. A complete case analysis was conducted to assess the sensitivity of the results to missing data.

For each Cox model, participants were censored upon their death, transfer to a non-CPRD practice, or at the end of follow-up (17th February 2020). Patients without frozen shoulder were censored if they were diagnosed with frozen shoulder after the index date. The proportional hazards assumption in each Cox model was checked through inspection of scaled Schoenfeld residual plots [296, 297].

Kaplan-Meier plots for the frozen shoulder group and the group without frozen shoulder are presented, along with 1–, 3–, 5–,10–, 15.8–year Kaplan-Meier estimates.

Data were prepared for analysis in Stata version 14.0 [175] and analysis was conducted in RStudio version 1.2.5033 [270]. R code can be found in Section C.2.

5.3.3 Shared frailty models

In this study, a shared frailty model was used to account for the lack of independence between matched individuals within the Cox regression. This section details how the standard Cox model from Section 3.6.3 can be extended using random effects to account for the lack of independence between patients in clustered data.

In the generic Cox model defined in Equation 3.7, the vector of covariates $\mathbf{x}_j$ is used to model the heterogeneity in hazard between individuals. One may believe that there are additional unmeasured factors that influence a patients survival. Random effects, called frailties, can be introduced into the model in an attempt to mop up any heterogeneity in hazards that has not been accounted for by the model's covariates.

The Cox model in Equation 3.7 can be extended to a frailty model by including the unobserved random variable Z_j which acts multiplicatively on the hazard to explain the unobserved

heterogeneity for each individual j, $j = 1, \ldots, n$; that is,

$$\lambda(t|\mathbf{x}_j, Z_j) = Z_j \lambda_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x}_j).$$

It is most commonly assumed that Z follows a gamma or log-normal distribution, although the choice of distribution is generally based on computational ease rather than its real-world application [298].

Frailty terms can also be included in Cox models to account for the lack of independence between individuals in clustered data. The resulting model is called a shared frailty model and can be defined as

$$\lambda(t|\mathbf{x}_{ij}, Z_i) = Z_i \lambda_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x}_{ij}),$$

where the shared frailty term Z_i, is the same for each individual j, $j = 1, \ldots, n_i$, that shares the same cluster, and where $i = 1, \ldots, n$, such that there are n clusters. The vector $\mathbf{x}_{ij}$ is the vector of observed covariates for individual j from cluster i, and $\boldsymbol{\beta}$ is the corresponding vector of regression coefficients. Again, gamma and log-normal distributions are commonly assumed for the frailties.

5.4 Results

5.4.1 Sample characteristics

The total sample analysed in this study consisted of 31,226 adults with a frozen shoulder Read code, matched to 31,226 adults of the same age, gender and practice that did not have a frozen shoulder Read code on or before the index date (Table 5.1). Sixty-two percent of the frozen shoulder group were female, and the mean age of diagnosis was 59.78 years (SD=13.24). The mean BMI of the frozen shoulder group, 27.59 (SD=6.00), was greater than that of the group without frozen shoulder, 26.97 (SD=6.14). The mean duration of follow-up was 8.58 years (SD=4.27), being approximately equal in the frozen shoulder group and the group without frozen shoulder.

	Frozen shoulder n=31,226 (50%)	**No frozen shoulder n=31,226 (50%)**	**All n=62,452**
Mean follow-up duration[1] (years)	8.58 (SD=4.29)	8.59 (SD=4.24)	8.58 (SD=4.27)
Mean age (years)	59.78 (SD=13.24)	59.78 (SD=13.24)	59.78 (SD=13.24)
Gender			
Male	11,825 (37.87%)	11,825 (37.87%)	23,650 (37.87%)
Female	19,401 (62.13%)	19,401 (62.13%)	38,802 (62.13%)
Obesity			
Obese	7,678 (24.59%)	6,182 (19.80%)	13,860 (22.19%)
Not obese	21,094 (67.55%)	20,837 (66.73%)	41,931 (67.14%)
Missing	2,454 (7.86%)	4,207 (13.47%)	6,661 (10.67%)
Hypertension			
Diagnosed	9,159 (29.33%)	8,632 (27.64%)	17,791 (28.49%)
Not diagnosed	22,067 (70.67%)	22,594 (72.36%)	44,661 (71.51%)
Hyperlipidaemia			
Diagnosed	4,580 (14.67%)	3,694 (11.83%)	8,274 (13.25%)
Not diagnosed	26,646 (85.33%)	27,532 (88.17%)	54,178 (86.75%)
Thyroid dysfunction			
Diagnosed	2,986 (9.56%)	2,804 (8.98%)	5,790 (9.27%)
Not diagnosed	28,240 (90.44%)	28,422 (91.02%)	56,662 (90.73%)
Ethnicity			
Bangladeshi	28 (0.09%)	23 (0.07%)	51 (0.08%)
Black African	87 (0.28%)	78 (0.25%)	165 (0.26%)
Black Caribbean	129 (0.41%)	111 (0.36%)	240 (0.38%)
Black – other	38 (0.12%)	38 (0.12%)	76 (0.12%)
Chinese	61 (0.20%)	26 (0.08%)	87 (0.14%)
Indian	247 (0.79%)	138 (0.44%)	385 (0.62%)
Mixed	71 (0.23%)	78 (0.25%)	149 (0.24%)
Other Asian	127 (0.41%)	61 (0.20%)	188 (0.30%)
Other	220 (0.70%)	222 (0.71%)	442 (0.71%)
Pakistani	103 (0.33%)	87 (0.28%)	190 (0.30%)
Missing	7,227 (23.14%)	7,407 (23.72%)	14,634 (23.43%)
White	22,888 (73.30%)	22,957 (73.52%)	45,845 (73.41%)
IMD Quintile			

Continued on next page

[1]Defined as the time from the index date to the earliest of: end of study (17th February 2020), date of death, date of transfer to a non-CPRD practice, or date of last CPRD data collection.

	Frozen shoulder n=31,226 (50%)	**No frozen shoulder n=31,226 (50%)**	**All n=62,452**
Least deprived quintile	7,634 (24.45%)	7,875 (25.22%)	15,509 (24.83)%)
2nd least deprived quintile	6,923 (22.17%)	7,029 (22.51%)	13,952 (22.34%)
3rd least deprived quintile	6,771 (21.68%)	6,592 (21.11%)	13,363 (21.40%)
4th least deprived quintile	5,496 (17.60%)	5,341 (17.10%)	10,837 (17.35%)
Most deprived quintile	4,378 (14.02%)	4,005 (12.83%)	8,383 (13.42%)
Missing	24 (0.08%)	384 (1.23%)	408 (0.65%)

Table 5.1: Table summarising characteristics of study participants

5.4.2 Results for the association between newly diagnosed frozen shoulder and a subsequent diagnosis of type 2 diabetes

In total, 1647 study participants had a Read code of type 2 diabetes during follow-up. In the frozen shoulder group, 1559 out of 31,226 patients (5%) were diagnosed with type 2 diabetes post-index. In the group without frozen shoulder, 88 out of 31,226 patients (0.28%) were diagnosed with type 2 diabetes post-index. The difference in survival probability between the frozen shoulder group and the group without frozen shoulder can be seen in the Kaplan-Meier plot[2] (Figure 5.1). The 1–, 3–, 5–, 10–, 15.8–year Kaplan-Meier estimates for the frozen shoulder group and the group without frozen shoulder were 99.1%, 97.6%, 96.2%, 93.9%, 93.1% and $>$99.9%, $>$99.9%, $>$99.9%, 99.7%, 99.2% respectively.

The hazard ratio for a diagnosis of type 2 diabetes, comparing the frozen shoulder group to the group without frozen shoulder, was estimated to be 19.37 (95% CI: 15.62 – 24.01). When accounting for other factors (mean number of consultations per year, hyperlipidaemia, hypertension, obesity, ethnicity, thyroid dysfunction, and deprivation), the association between incident frozen shoulder and the hazard of a following type 2 diabetes diagnosis remained, with a

[2]Note that the y-axis does not start from zero. This choice was made so that the reader can see the shape of the curves more clearly.

hazard ratio of 19.98 (95% CI: 15.99 – 24.97). When re-running the mean number of consultations per year-, hyperlipidaemia-, hypertension-, obesity-, ethnicity-, thyroid dysfunction-, and deprivation-adjusted Cox model in a complete case analysis (76.1% of patients had complete data), the results of the Cox model were similar (HR = 21.34; 95% CI: 16.51 – 27.57).

There was little evidence to suggest a violation of the proportional hazards assumption in either Cox model. The two plots of scaled Schoenfeld residuals against time with smoothing splines showed little evidence of systematic departures from a horizontal line (Figures 5.2 and 5.3). The only slight deviations were at the extreme right of the two graphs where there were fewer events occurring.

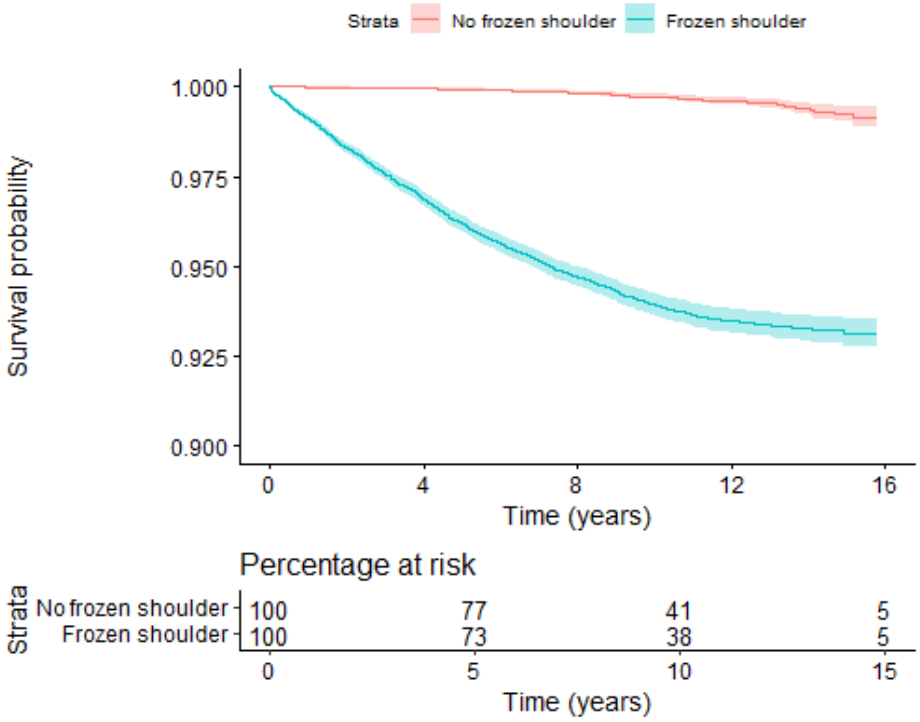

Figure 5.1: Plot of Kaplan-Meier estimates (with 95% confidence intervals) for surviving a type 2 diabetes diagnosis among patients with frozen shoulder compared to patients without frozen shoulder, with a percentage at risk (of developing type 2 diabetes) table

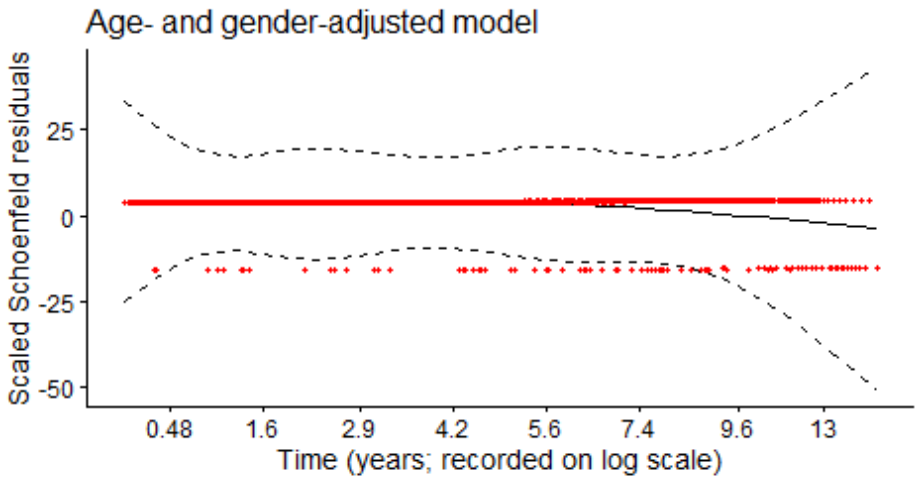

Figure 5.2: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the frozen shoulder coefficient in the age- and gender-adjusted Cox model

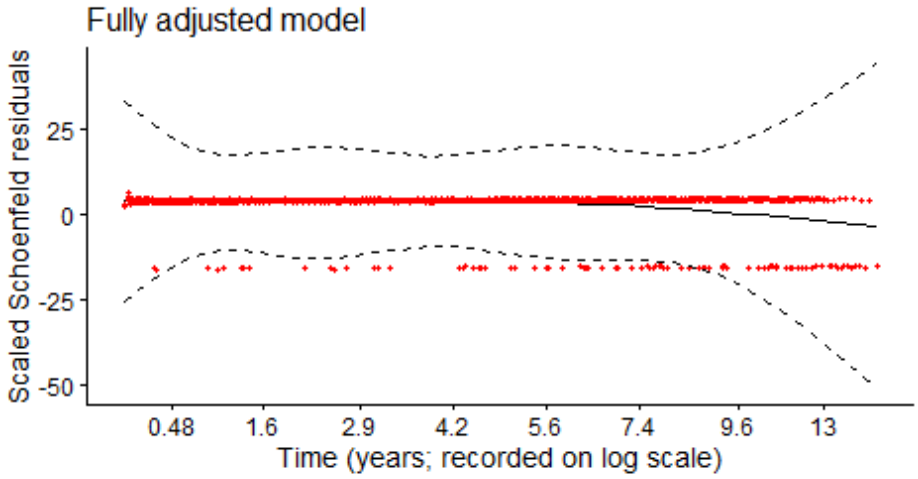

Figure 5.3: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the frozen shoulder coefficient in the age-, gender-, mean number of consultations per year-, ethnicity-, hyperlipidaemia-, hypertension-, obesity-, thyroid dysfunction-, and deprivation-adjusted Cox model

5.5 Discussion

Patients who have been newly diagnosed with frozen shoulder are considerably more likely to be diagnosed with type 2 diabetes than people without frozen shoulder (Kaplan-Meier estimate: 93.1% vs. 99.2%). Further, when accounting for other factors that could explain the association, incident frozen shoulder remained a predictor of a subsequent diagnosis of type 2 diabetes; sensitivity analysis suggested that the model's results were robust to missing data.

Previous studies estimating association between undiagnosed diabetes and frozen shoulder have often been limited by small sample sizes and have shown somewhat contradictory results. Tighe et al. conducted blood tests on 88 patients presenting with frozen shoulder that had no history of diabetes [130]. They found that 39% of the patients had diabetes and 33% of the patients had pre-diabetes [130]. Another study contained 77 patients with frozen shoulder that had not previously been diagnosed with diabetes [283]. After testing the participants, they discovered that 17% had undiagnosed diabetes [283]. Further, a small study of 18 patients presenting with frozen shoulder, but no previous diagnosis of diabetes, were tested for type 2 diabetes; 17% of the patients tested positive [284]. In another study, when tested for type 2 diabetes, zero out of 122 patients presenting frozen shoulder, but with no history of diabetes, tested positive for type 2 diabetes [299].

Previous research had not investigated whether any elevated risk of having type 2 diabetes in people with frozen shoulder could be explained by other factors, such as a patients BMI and other cardiovascular risk factors. The results of this analysis suggest that there is an association between incident frozen shoulder and a future type 2 diabetes diagnosis after adjusting for other factors. It should be noted that the results do not suggest that frozen shoulder causes type 2 diabetes, but rather that it is likely that patients presenting with frozen shoulder are more likely to have undiagnosed type 2 diabetes than people without frozen shoulder.

The use of electronic health records for this study allowed me to obtain a large sample of patients that were representative of the frozen shoulder population in the UK. A main limitation

of using electronic health records is that not every participant in the study was tested for type 2 diabetes. The reliance on Read codes to detect which patients had type 2 diabetes may have meant that the true number of patients that developed type 2 diabetes post index may be underestimated. Additionally, it is difficult to know the extent to which people with frozen shoulder are already being tested for type 2 diabetes due to pre-existing knowledge of the strong association between type 2 diabetes and frozen shoulder. A survey completed by 714 UK GPs suggested that 60% of GPs would run blood tests on patients presenting with frozen shoulder; however, the response rate to the survey was low (14.7%) so results may contain some bias [300]. Future studies could investigate the clinical and cost-effectiveness of testing patients with frozen shoulder for type 2 diabetes upon their frozen shoulder diagnosis.

5.6 Conclusion

Patients that have been newly diagnosed with frozen shoulder are considerably more likely to be diagnosed with type 2 diabetes in the years following their frozen shoulder diagnosis. Further, the association between incident frozen shoulder and a subsequent type 2 diabetes diagnosis could not be explained by other factors. This study should motivate further research to understand whether testing patients with frozen shoulder upon their diagnosis is an effective approach to detecting type 2 diabetes early and reducing the likelihood of complications.

The following chapter describes a systematic review which summarises evidence from longitudinal observational studies to investigate whether diabetes is associated with the course of symptoms in people with frozen shoulder. The knowledge gained from the systematic review may help to inform clinicians and patients of how to best manage frozen shoulder in people with diabetes.

Chapter 6

Diabetes as a prognostic factor in frozen shoulder: a systematic review

B. P. Dyer, C. Burton, T. Rathod-Mistry, M. Blagojevic-Bucknall, D. A. van der Windt. Diabetes as a Prognostic Factor in Frozen Shoulder: A Systematic Review. *Archives of Rehabilitation Research and Clinical Translation*, 3(3):100141, 2021. https://doi.org/10.1016/j.arrct.2021.100141

The protocol for this systematic review was registered on the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42019122963; available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019122963).

The review was conducted and reported using the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines [132]. A completed PRISMA checklist can be found in Appendix Section D.1.

6.1 Introduction

Prognosis research concerns itself with studying the likelihood of future outcomes (referred to as endpoints) in people with a given disease (referred to as the startpoint) [301]. Prognostic factor research is concerned with investigating whether some baseline factor, measured at the startpoint, is a predictor of a future outcome/endpoint [302]. Thus, prognostic factors identify strata of people with a disease that can expect to experience different outcomes from their disease. Such knowledge may help to guide strata-specific clinical decision making, identify targets for intervention that may improve the outcomes of disease, or prognostic factors may predict treatment response [302].

Given that patients with frozen shoulder experience varying degrees of improvement in pain and function, it is important to understand whether there are predictors of the outcomes of frozen shoulder. Furthermore, since 30% of people with frozen shoulder have diabetes [25], it is crucial to investigate whether people with diabetes experience worse outcomes from frozen shoulder than people without diabetes. This systematic review aims to understand whether diabetes is a predictor of the course of symptoms in people with frozen shoulder.

6.1.1 Systematic review objective

To summarise evidence from longitudinal observational studies to investigate whether diabetes is a prognostic factor for the outcomes of frozen shoulder.

6.2 Methods

6.2.1 Defining the eligibility criteria

Studies satisfying the inclusion criteria in Table 6.1 and the exclusion criteria in Table 6.2 were eligible for this review.

	Inclusion criteria
Population/startpoint	Adults with frozen shoulder/adhesive capsulitis.
Prognostic factor	Diabetes. (All types of diabetes were considered, and Diabetes could be identified via clinical diagnosis, blood testing, or self-reporting.)
Outcome/endpoint of interest	All outcomes of frozen shoulder, measured at follow-up (>2 weeks), were considered. Potential outcomes/endpoints of interest include, but are not restricted to, ROM, pain and disability.
Setting	No restrictions to study setting or treatment received; population based as well as clinical cohorts were eligible.
Study design	Longitudinal observational studies (cohort or case-control).

Table 6.1: Table summarising inclusion criteria

6.2.2 Identification of suitable literature

6.2.2.1 Search strategy

The same search described in Section 2.2.2.3 was used to identify studies for this review. The updated search for this systematic review was conducted on June 2021. The same 11 bibliographic databases were searched, and additional studies were identified using reference screening and through emailing a professional contact of DvdW.

6.2.2.2 Study selection

Mendeley [144] was used to download citations, and Excel was used to check for duplicate citations and allowed reviewers to provide their reasoning for excluding studies. One reviewer (BPD) screened all titles and abstracts; two reviewers (MB-B and CB) screened a 20% random sample of the titles and abstracts. Full-texts were screened by reviewer BPD and independently screened by another reviewer (MB-B, CB or TR-M). All stages of screening were conducted using the pre-defined inclusion and exclusion criteria (Tables 6.1 and 6.2). Discrepancies during the screening process were resolved through discussion with a third reviewer DvdW.

Exclusion criteria
• If the full text was not available then the study was excluded. (Authors were contacted in an attempt to access full-text documents.) • Non–English language papers were eligible dependent upon finding a translator within the research institute. • Cross–sectional studies, trials and case series were excluded. • If the paper did not present an association estimate (odds ratio, risk ratio, hazard ratio) or present sufficient data to estimate an association estimate then the study was excluded. • Outcomes measured at less than or equal to two weeks follow-up were not eligible. If a study included results at multiple follow-up points then only the measurements taken at more than two weeks follow-up were included.

Table 6.2: Table summarising exclusion criteria

6.2.3 Data extraction

Data were extracted by one reviewer (BPD) and another reviewer (MB or TR-M) checked that the data were extracted correctly. The data-extraction sheet was piloted using the first three studies to ensure all important data were included in the extraction sheet. Extracted data included details of study design, setting, treatment type, sample characteristics, sample size, variable measurement, attrition, statistical analysis, association estimates and confidence intervals.

6.2.4 Risk of bias assessment

Risk of bias was judged using the QUIPS tool [147]. The tool was described in Section 2.2.4. The only change to the tool for this review compared to the review in Chapter 2 is that diabetes is being investigated as a prognostic factor in people with frozen shoulder rather than a risk factor for the onset of frozen shoulder.

Risk of bias in each study was judged by BPD and also independently judged by either MB-B or TR-M. Disagreements were resolved through discussion between the two reviewers and DvdW. As in Chapter 2, the overall risk of bias score was based on reviewer judgement to avoid the use of a tallied score.[1]

[1]The reason a tallied score was avoided is because only one major flaw on one of the bias domains can mean a study is at a high risk of bias.

6.2.5 Systematic review analysis

Due to variation in study outcome measures and follow-up duration, pooled association estimates were not calculated. Forest plots of mean differences in outcomes scores (such as Visual Analogue Scores (VAS) [303]) were plotted to help visualise results. If studies provided sufficient data then confidence intervals were calculated and included in the plot; if studies did not provide the required data to estimate a confidence interval then only the point estimate was included in the forest plot. Forest plots were plotted using R version 4.0.2 [304].

Evidence synthesis and assessment of the quality of evidence were conducted using a version of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework that has been adapted for prognostic factor research [305]. The adapted GRADE approach provides a systematic and transparent framework to summarise key factors that determine the quality of prognostic evidence. The adapted GRADE framework proposes that six factors may lower the quality of prognostic evidence: phase of investigation, study limitations, inconsistency, indirectness, imprecision, publication bias; and two factors may increase the quality of prognostic evidence: effect size, exposure-response gradient.

To grade the GRADE factors, the following evidence was considered:

Phase of investigation: Determined from the research article[2].

Study limitations: Determined by QUIPS scoring.

Inconsistency: Graded after inspecting forest plots, tallying the direction of association in each study, and examining raw extracted data.[3]

Indirectness: Determined from the research article and through QUIPS scoring to see if the

[2]The prognostic factor GRADE framework classes studies as phase 3, 2, or 1 according to the following criteria:
Phase 3: *"Explanatory research aimed to understand prognostic pathways"*.
Phase 2: *"Explanatory research aimed to confirm independent associations between potential prognostic factor and the outcome"*.
Phase 1: *"Outcome prediction research or explanatory research aimed to identify associations between potential prognostic factors and the outcome"*[305].

[3]Raw data were analysed when the outcome scores were not able to be included in the forest plot, for example, when the outcome was categorised.

sample, prognostic factor, and outcome accurately reflected the review question.

Imprecision: Graded after inspecting forest plots and examining raw extracted data (including the sample size).

Publication bias: Determined by QUIPS scoring.

Effect size: Graded after inspecting forest plots and examining reported association estimates.

Exposure-response gradient: This GRADE factor was not appropriate for the type of evidence collected in this systematic review, thus certainty in evidence was not upgraded due to there being evidence of an exposure-response gradient.

Some studies within this review have reported multiple measurements for the same outcome at different follow-up points. Each study was only included once into the GRADE assessment and only contributed once to the tally of association direction. Where multiple follow-up measurements on the same outcome domain were reported, the most common association direction across follow-up points was used for the 'direction of association' for that study. Similarly, the 'direction of association' for studies investigating ROM measurement as an outcome was the most common direction of association observed across all ROM movements. Some studies used multiple instruments to measure the same domain, for example, using both American Shoulder and Elbow Surgeons (ASES) scores [306] and Oxford Shoulder Scores (OSS) [307]. In this scenario, the results for the instrument that was most commonly used across all studies was included in the evidence synthesis and GRADE assessment.

6.3 Results

6.3.1 Summary of search results

The search of electronic bibliographic databases identified 1784 unique citations. Eight additional citations were identified by the professional contact. Forty-six full-texts were assessed for eligibility and 28 met the criteria to be included in the review. A PRISMA flow chart summarising the citation identification, screening and selection process can be found in Figure 6.1.

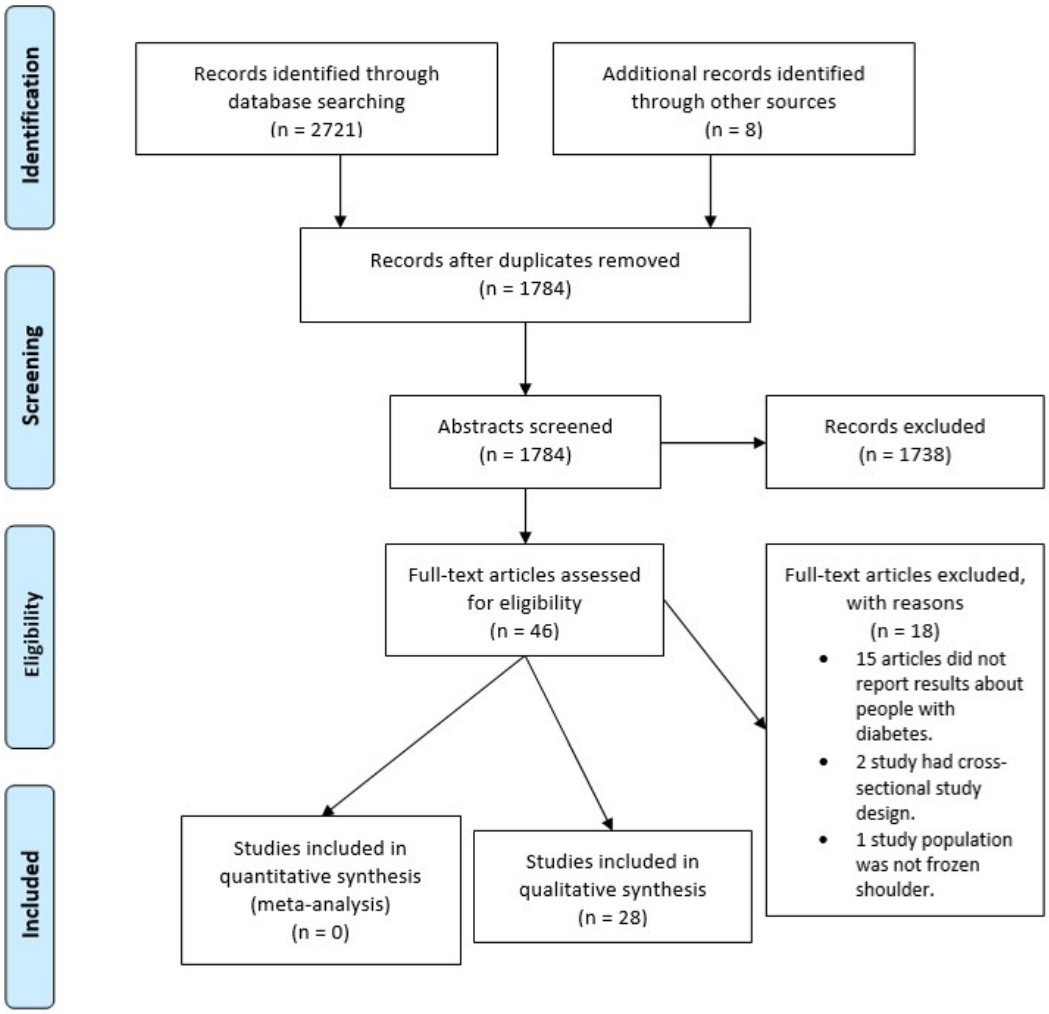

Figure 6.1: PRISMA flow diagram summarising the process of citation identification and study selection

6.3.2 Study characteristics

Table 6.3 summarises the characteristics of all 28 studies that were included in the review. All studies in the review were cohort studies. Thirteen studies reported ROM results, 10 reported pain results, 18 reported multi-dimensional scores, four reported function and disability scores, in four studies the requirement for surgery was the outcome of interest, in two studies the need for a second MUA was the outcome of interest, in one study the development of contralateral frozen shoulder and requiring a second glenohumeral joint injection were both outcomes of interest, and one defined the failure to recover from frozen shoulder as the outcome.

Patients received arthroscopic capsular release in seven studies, MUA in six, physiotherapy in five, hydrodilatation in three, manipulation and arthroscopic capsular release in one, manipulation under ultrasound-guided brachial plexus block in one, a mixture of physiotherapy and arthroscopic capsular release in two, a mixture of MUA and conservative treatment in one, ultrasound-guided intra-articular corticosteroid injection in one and treatment was not reported in one study.

Eleven studies were based in Europe, ten were from Asia, six were from North America and one was from Oceania. Fifteen studies were based in hospitals, three in medical centres, one in a physiotherapy clinic, two in sports medicine clinics, one was based in electronic health records, and six did not specify setting.

The median sample size was 56 people, with range 15–2190. The mean percentage of people with diabetes in each study was 26% (SD=0.12%) and ranged from 12% to 57%.

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
G. P. Nicholson, 2003 (USA) [308]	High	Cohort study. Hospital-based	Arthroscopic Capsular Release	ROM Pain (VAS), Multi-dimensional score (ASES), Function and disability (SST)	Mean 3 years post-capsular release (range 2-8 years)	Diabetes: 8 shoulders; Non-diabetes: 17 shoulders
G. L. Cvetanovich et al., 2018 (USA) [309]	High	Cohort study. Medical centre	Arthroscopic Capsular Release	ROM	Mean 3.7 years post-capsular release (range 2-6 years)	Diabetes: 8 people; Non-diabetes: 19 shoulders
R. G. E. Clement et al., 2013 (UK) [310]	High	Cohort study. Hospital-based	Hydrodilatation	ROM, Pain (VAS), Multi-dimensional score (OSS)	1 month post-hydrodilatation	Diabetes: 12 people; Non-diabetes: 39 people
S. Bell et al., 2003 (Australia) [311]	High	Cohort study. Setting unclear	Hydrodilatation	ROM, Pain (VAS) scored as nil, mild, moderate or severe	2 months post-hydrodilatation	Diabetes: 15 people; Non-diabetes: 94 people

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
H. Vastamäki et al., 2013 (Finland) [312]	High	Cohort study. Hospital-based	MUA	ROM, Pain (VAS)	Mean 23.1 years post-MUA (range 19-30 years)	Diabetes: 4 people; Non-diabetes: 11 people
C-H Cho et al., 2016 (Republic of Korea) [313]	Moderate	Cohort study. Setting unclear	Arthroscopic capsular release	ROM, Pain (VAS), Multi-dimensional score (ASES)	3 months, 6 months, 12 months post-capsular release and at a final follow-up of mean 48.4 months (SD=15.8 months)	Diabetes: 17 shoulders pre-capsular release and final follow-up, 15 at 3 months, 9 at 6 months, 13 at 12 months; Non-diabetes: 20 shoulders pre-capsular release, at 3 months and final follow-up, 17 at 6 months, 15 at 12 months
A. Ando et al., 2018 (Japan) [314]	High	Cohort study. Setting unclear	Manipulation under ultrasound-guided brachial plexus block	ROM, Pain (VAS), Multi-dimensional score (Constant score)	Mean 4.8 years (SD=3.5 years) for the diabetes group, and mean 5.1 years (SD=2.4 years) for the non-diabetes group	Diabetes: 10 shoulders; Non-diabetes: 42 shoulders

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
İ. Düzgün et al., 2012 (Turkey) [315]	Moderate	Cohort study. Physiotherapy centre	Physiotherapy	ROM, Multi-dimensional score (Constant score)	Following the treatment protocol averaging 8 weeks duration.	Diabetes: 12 people; Non-diabetes: 38 people
H. Vastamäki et al., 2016 (Finland) [316]	High	Cohort study. Hospital-based	Diabetes group: 69% had MUA and 31% had conservative treatment; Non-diabetes group: 53.3% had MUA and 37.3% had conservative treatment	ROM, Pain (VAS), Multi-dimensional score (Constant score)	Mean 10 years (SD=8 years) for the diabetes groups, and mean 9.7 years (SD=7 years) for the non-diabetes group.	Diabetes: 29 shoulders; Non-diabetes: 169 shoulders
C-H. Cho et al., 2020 (Republic of Korea) [317]	High	Cohort study. Hospital-based	Ultrasound-guided intra-articular corticosteroid injection	ROM, Pain (VAS), Multi-dimensional score (ASES)	3 weeks, 6 weeks, 12 weeks post treatment	Diabetes group: 32 shoulders; Non-diabetes group: 110 shoulders

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
Y. W. Ko et al., 2021 (Republic of Korea) [318]	Moderate	Cohort study. Hospital-based	MUA	ROM, Pain (VAS), Multi-dimensional score (Constant score)	6 weeks, 3 months post treatment	Diabetes group: 32 shoulders; Non-diabetes group: 203 shoulders
G. L. Yanlei et al., 2019 (Singapore) [319]	High	Cohort study. Hospital-based	Arthroscopic capsular release	ROM, Pain (VAS), Multi-dimensional scores (Constant score)	12 months post treatment	Diabetes group: 32 shoulders; Non-diabetes group: 24 shoulders
F. Barbosa et al., 2019 (UK) [320]	High	Cohort study. Hospital-based	Mixture of conservative or surgical	ROM, Multi-dimensional score (OSS)	3, 6, 12 months follow-up	Diabetes group: 46 shoulders; Non-diabetes group: 164 shoulders
S. S. Mehta et al., 2014 (UK) [321]	High	Cohort study. Hospital-based	Arthroscopic capsular release.	Multi-dimensional score (Constant score)	6 weeks, 6 months and 2 years post-capsular release	Diabetes: 21 people; Non-diabetes: 21 people
M. Çinar et al., 2010 (Turkey) [322]	High	Cohort study. Setting unclear	Arthroscopic capsular release	Multi-dimensional score (Constant score)	Mean 48.5 months for the diabetes group and mean 60.2 months for the non-diabetes group	Diabetes: 15 shoulders; Non-diabetes: 13 shoulders

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
J-P. Wang et al., 2010 (Taiwan) [323]	High	Cohort study. Medical centre	MUA	Multi-dimensional score (Adjusted constant score, excluding the 25 points for assessment of muscle strength)	3 weeks post-MUA and an average of 95 months (range 18-189 months) post-MUA	Diabetes: 21 shoulders; Non-diabetes: 42 shoulders
H. Celik et al., 2017 (Turkey) [324]	High	Cohort study. Setting unclear	Manipulation and arthroscopic capsular release	Multi-dimensional score (Constant score)	Mean 49.5 months (range: 24–90 months)	Diabetes: 12 shoulders; Non-diabetes: 20 shoulders
R. Sinha et al., 2017 (UK) [325]	Moderate	Cohort study. Hospital-based	Hydrodilatation	Multi-dimensional score (OSS)	Improvement in OSS between pre-procedure and 4 weeks post-procedure	Diabetes: 26 people; Non-diabetes: 90 people
J. M. Lyhne et al., 2018 (Denmark) [326]	High	Cohort study. Hospital-based	Arthroscopic capsular release	Multi-dimensional score (OSS), Function and disability (Visual Quality Scale (VQS))	Improvement between pre-procedure and 6-month post-op OSSs and VQS scores	Diabetes: 18 people; Non-diabetes: 75 people

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
A. A. Theodorides et al., 2014 (UK) [327]	Moderate	Cohort study. Hospital-based	MUA	Multi-dimensional score (OSS)	Mean follow-up 28 days post-MUA and at mean follow-up 3.6 years post-MUA (IQR 1.7 – 5.0 years)	Diabetes: 39 people; Non-diabetes: 256 people
J. D. Lamplot, et al., 2018 (USA) [328]	High	Cohort study. Sports medicine clinic	Conservative treatment	Multi-dimensional score (ASES), function and disability (shoulder activity scale), being diagnosed with frozen shoulder in the contralateral shoulder, requiring a second glenohumeral joint injection	Minimum 2-year follow-up (mean, 3.4 years)	Diabetes: 9 people; Non-diabetes: 51 people

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
B. K. Rill et al., 2011 (USA) [329]	High	Cohort study. Setting unclear	Physiotherapy and home exercise for all patients. Arthroscopic capsular release for the surgery group.	Function and disability (SST)	Minimum 2 years, mean 40 months, range 24-68 months	Diabetes nonoperative group: 19 patients; Non-diabetes nonoperative group: 49 shoulders. Diabetes surgery group: 9 shoulders; Non-diabetes surgery group: 15 shoulders.
W. N. Levine, et al., 2007 (USA) [330]	High	Cohort study. Medical centre.	Nonoperative treatment.	Whether patient required surgery after treatment programme	Treatment programme averaged 4.7 months (range 0.2-43.9 months)	Diabetes group: 19 shoulders, Non-diabetes group: 86 shoulders.
K. Kingston et al., 2018 (USA) [183]	High	Cohort study. Electronic health records	Not reported	Requiring surgery	Follow-up duration not reported	Diabetes group: 572 patients, Non-diabetes group: 1618 patients.

Continued on next page

Author, Year (Country of study)	Overall QUIPS risk of bias	Study design and setting	Treatment type	Outcomes measured (tools used)	Follow-up measurements taken	Sample size
P. H. Gundtoft et al., 2020 (Denmark) [331]	Moderate	Cohort study. Hospital-based	Conservative treatment methods	Requiring surgery	Follow-up duration not reported	Diabetes group: 34 shoulders; Non-diabetes group: 201 shoulders
D. A. Woods et al., 2017 (UK) [332]	High	Cohort study. Hospital-based.	MUA	Requiring a second MUA	Follow-up duration unclear.	Diabetes group: 96 shoulders (56 type 1, 40 type 2), Non-diabetes group:696 shoulders.
E. F. Jenkins et al., 2012 (UK) [333]	High	Cohort study. Hospital-based	MUA	Requiring a second MUA	Follow-up duration unclear	Diabetes group: 39 shoulders, Non-diabetes group: 274 shoulders.
A. Ando et al., 2013 (Japan) [334]	Moderate	Cohort study. Hospital-based.	Nonoperative treatment	Failure to recover from frozen shoulder	30 months follow-up duration	Diabetes group: 61 shoulders, Non-diabetes group: 356 shoulders.

Table 6.3: Summary of study characteristics for studies reporting results for the association between diabetes and outcome in patients with frozen shoulder. Abbreviations: ROM – range of motion, VAS – Visual Analogue Score, ASES – American Shoulder and Elbow Surgeons score, SST – simple shoulder test, OSS – Oxford shoulder score, MUA – manipulation under anaesthesia, VQS – Visual Quality Scale

6.3.3 Risk of bias

Complete QUIPS risk of bias scores can be found in Table 6.4. A bar graph of QUIPS domain scores for studies reporting common (>5 studies) outcome types can be found in Figure 6.2. Reviewers agreed on 82% of QUIPS domain scores and agreed on 26 of the 28 overall risk of bias scores. After discussion 100% agreement was achieved for all domain and overall risk of bias scores.

Ten of the 13 studies reporting ROM outcomes were rated as being at a high risk of bias and three studies were rated as being at a moderate risk of bias. Seven of the 10 studies reporting pain outcomes were rated as being at a high risk of bias and three studies as being at a moderate risk of bias. Eleven of the 18 studies reporting multi-dimensional clinical scores were rated as being at a high risk of bias and seven studies as being at a moderate risk of bias. All four studies reporting function and disability outcomes were rated as being at a high risk of bias. Five of the six studies that reported less-common outcomes were rated as being at a high risk of bias, and one was rated as being at a moderate risk of bias. Common[4] reasons given by reviewers for potential bias in each QUIPS domain are given below.

Participation

Three studies were deemed to be at a high risk of bias, 17 at a moderate risk of bias, and eight at a low risk of bias. The source of the target population was unclear in 16 studies. Fourteen studies provided insufficient baseline statistics to allow for the diabetes and non-diabetes group characteristics to be compared. The place of recruitment was unclear in 11 studies, and the recruitment procedure was unclear in 10.

[4]Here, 'common' refers to the reason being given for more than two studies.

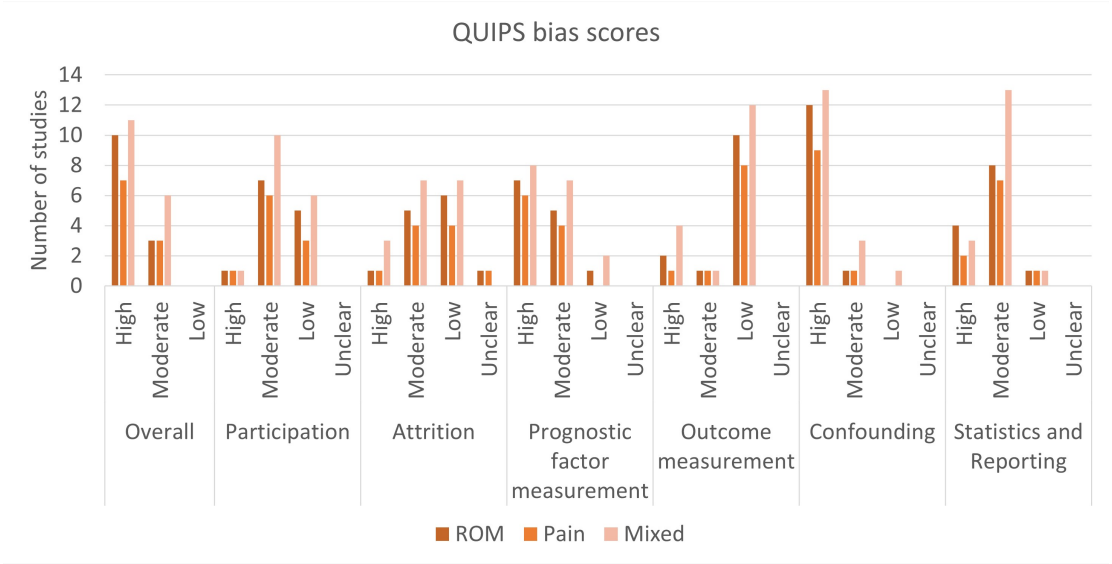

Figure 6.2: Bar graph of QUIPS risk of bias scores for overall risk of bias and for the six QUIPS risk of bias domains: study participation, prognostic factor measurement, outcome measurement, study confounding, statistical analysis and reporting for studies reporting ROM, VAS, multi-dimensional (mixed) scores

Attrition

Five studies were scored as being at a high risk of bias, 17 at a moderate risk of bias, five at a low risk of bias, and rated as unclear in one study. Fourteen studies provided inadequate detail on the numbers lost to dropout, and seven studies did not provide reasons for dropout or explain the potential impact of losing participants to follow-up.

Prognostic factor/diabetes measurement

Eleven studies were scored at a high risk of bias, 14 at a moderate risk of bias, and three at a low risk of bias. Twenty studies did not describe the method used to identify diabetes. In 17 studies the definition of diabetes was vague, including the type of diabetes that participants had.

Outcome measurement

Seven studies were scored as being at a high risk of bias, three at a moderate risk of bias, 17 at a low risk of bias, and one study was rated as being at a moderate risk of bias for measurement of pain, but low for ROM measurement. Six studies had very unequal mean follow-up durations between the diabetes group and the non-diabetes group.

Confounding

Twenty-three studies were deemed to be at a high risk of unaccounted confounding since they did not account for any confounders, four were deemed to be at a moderate risk of unaccounted confounding, and one at a low risk of unaccounted confounding.

Statistics and reporting

Seven studies were scored as being at a high risk of bias, 18 at a moderate risk of bias, and three at a low risk of bias. Twenty studies used basic analysis that did not account for any covariates. Additionally, it was not always clear whether it was the intention at the onset of the study to compare outcomes between people with diabetes and people without diabetes or if an association was spotted which lead to the results being reported; thus, there may have been potential publication bias present.

Author, Year	Summary participation	Study attrition	Prognostic factor measurement	Outcome measurement	Confounding	Statistical analysis and presentation	Overall risk of bias
G. P. Nicholson, 2003	Moderate	Moderate	High	High	High	Moderate	High
G. L. Cvetanovich et al., 2018	Moderate	Low	High	High	High	High	High
R. G. E. Clement et al., 2013	Moderate	Low	High	Low	High	Moderate	High
S. Bell et al., 2003	High	Unclear	High	Moderate for pain, low for ROM	High	High	High
H. Vastamäki et al., 2013	Moderate	Moderate	Moderate	Low	High	Moderate	High
C-H Cho et al., 2016	Moderate	Moderate	Moderate	Low	High	Moderate	Moderate
A. Ando et al., 2018	Moderate	Low	High	Low	High	Moderate	High
İ. Düzgün et al., 2012	Moderate	Low	Low	Low	High	Moderate	Moderate
H. Vastamäki et al., 2016	Moderate	Moderate	High	Low	High	Moderate	High
C-H. Cho et al., 2020	Low	High	High	Low	High	Moderate	High

Continued on next page

Author, Year	Summary participation	Study attrition	Prognostic factor measurement	Outcome measurement	Confounding	Statistical analysis and presentation	Overall risk of bias
Y. W. Ko, et al., 2021	Low	Low	Moderate	Low	Moderate	Low	Moderate
G. L. Yanlei, et al., 2019	Low	Low	Moderate	Low	High	High	High
F. Barbosa, et al., 2019	Low	Low	High	Low	High	High	High
S. S. Mehta et al., 2014	Moderate	Moderate	High	Low	High	High	High
M. Çinar et al., 2010	High	Low	Moderate	High	High	Moderate	High
J-P. Wang et al., 2010	Low	Moderate	Low	Low	Moderate	Moderate	High
H. Celik et al., 2017	Moderate	Moderate	Moderate	High	High	Moderate	High
R. Sinha et al., 2017	Moderate	Moderate	Moderate	Moderate	High	Moderate	Moderate
J. M. Lyhne et al., 2018	Moderate	High	High	Low	Low	Moderate	High
A. A. Theodorides et al., 2014	Moderate	High	Moderate	Low	High	Low	Moderate

Continued on next page

Author, Year	Summary participation	Study attrition	Prognostic factor measurement	Outcome measurement	Confounding	Statistical analysis and presentation	Overall risk of bias
J. D. Lamplot, et al., 2018	Low	Moderate	Moderate	High	Moderate	Moderate	High
B. K. Rill et al., 2011	High	Moderate	Moderate	High	High	High	High
W. N. Levine, et al., 2007	Moderate	Moderate	Moderate	High	High	Moderate	High
K. Kingston et al., 2018	Low	Moderate	Moderate	Low	High	Moderate	High
P. H. Gundtoft, et al., 2020	Low	High	Moderate	Low	High	Moderate	Moderate
D. A. Woods et al., 2017	Moderate	Moderate	Low	Moderate	High	High	High
E. F. Jenkins et al., 2012	Moderate	High	Moderate	Low	High	Moderate	High
A. Ando et al., 2013	Moderate	Moderate	High	Moderate	Moderate	Low	Moderate

Table 6.4: QUIPS domain risk of bias scores

6.3.4 Results for diabetes as a prognostic factor in frozen shoulder

6.3.4.1 Summary of results

Tables D.1–D.5 in Appendix Section D.2 include summaries of results from studies investigating the association between diabetes and outcomes of follow-up. Tables D.1–D.3 include result summaries from studies reporting ROM results, pain scores and multi-dimensional clinical scores. Results from studies reporting function and disability scores and other less common outcomes ($\leq$4 studies) are reported in Tables D.4 and D.5, respectively.

Studies reporting results for ROM, pain and multi-dimensional clinical scores provided very little evidence to suggest that people with diabetes had worse baseline/pre-treatment scores than people without diabetes.

The forest plot in Figure 6.3 contains mean differences in measurements of: abduction for 673 patients from eight studies, external rotation for 1581 patients from 13 studies, and flexion for 997 patients from 12 studies. People with diabetes generally had worse ROM at follow-up than people without diabetes (Figure 6.3, Table D.1), although there was an inconsistency in association sizes for each movement (abduction, external rotation and flexion). Further, there are some studies that suggest diabetes is associated with better ROM, so there is some inconsistency in results. Studies with follow-up duration less than three months showed no evidence of an association between diabetes and ROM at follow-up (Figure 6.3, Table D.1).

The forest plot of mean differences in VAS scores (0-10 scale with 10 representing the most pain) [303] in Figure 6.4 contains 784 patients from eight studies. People with diabetes consistently had worse pain at follow-up (Figure 6.4, Table D.2). The magnitude of the association size was often small and confidence intervals were often wide, which could have been due to small sample sizes. Imprecise estimates and small association sizes could explain why differences in pain often did not meet the statistical significance thresholds in the primary studies (Table D.2).

The forest plot of mean differences for multi-dimensional clinical scores in Figure 6.5 contains results for 1170 patients. Within the forest plot, nine studies reported results using the Constant score [335] for 758 patients, two studies containing 148 patients reported Oxford Shoulder Scores (OSS) [307], and four studies reported ASES scores [306] for 264 patients. People with diabetes consistently had worse multi-dimensional clinical scores, although the mean difference did not always meet statistical significance in some smaller studies (Figure 6.5, Table D.3).

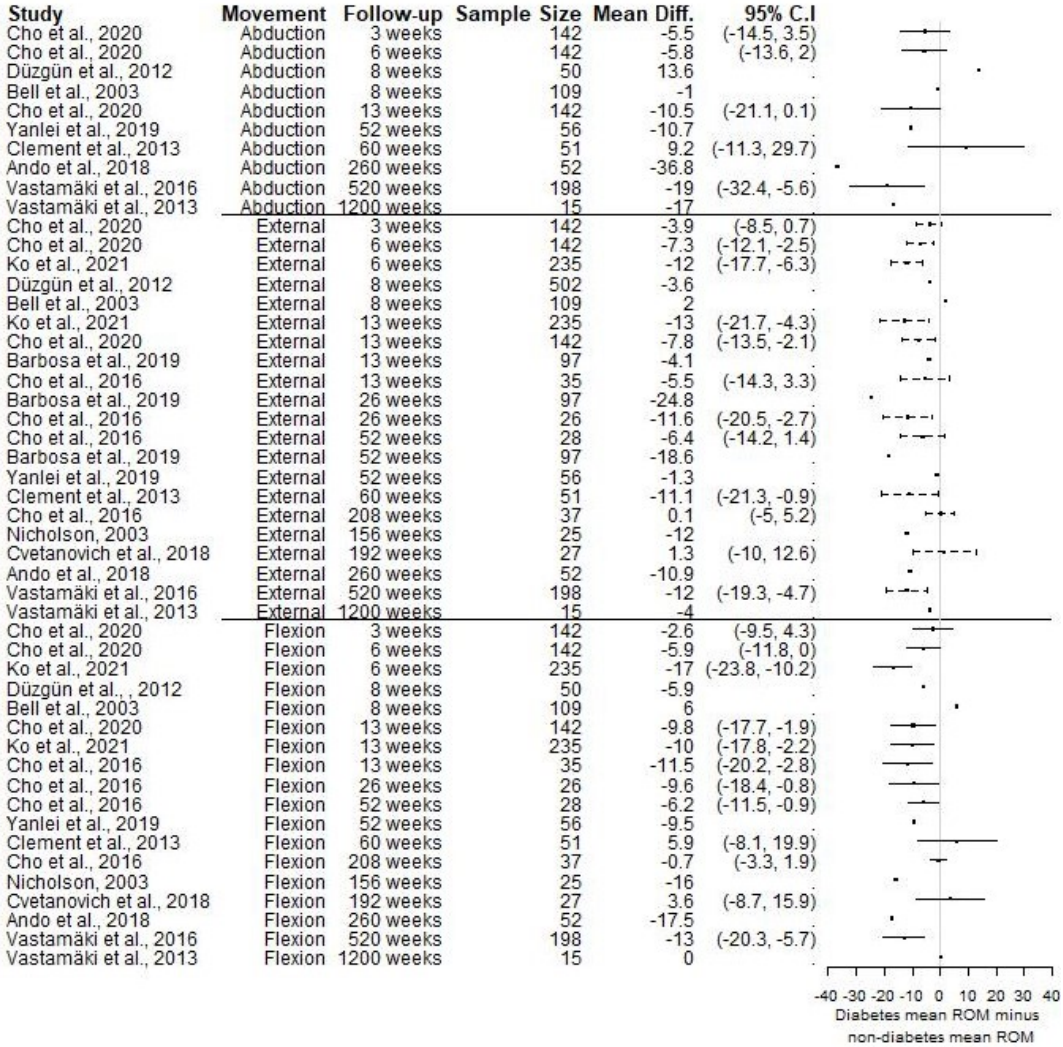

Figure 6.3: Forest plot of mean differences in ROM measurements (abduction, external rotation, flexion) between people with diabetes compared to people without diabetes

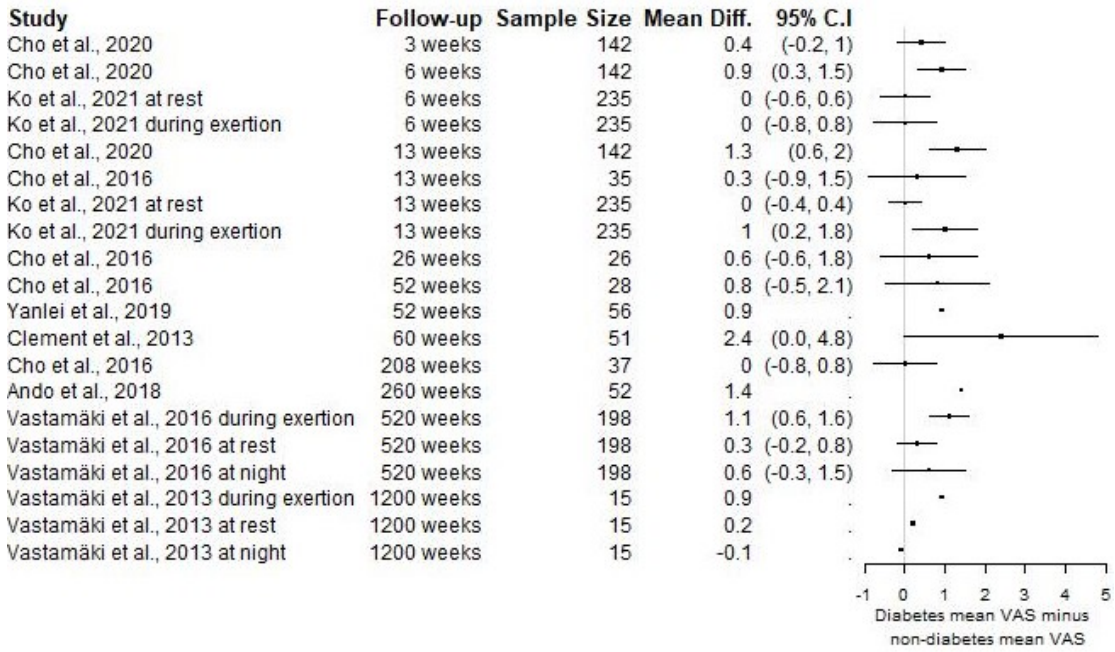

Figure 6.4: Forest plot of mean differences in VAS scores (measured on a 0-10 scale with 10 representing the worst pain) between people with diabetes compared to people without diabetes

6.3.4.2 GRADE factor scoring

The phase of investigation for the studies in this review for each outcome domain was graded as being 'phase 2', meaning that studies were generally cohort studies seeking to confirm independent associations between the prognostic factor (diabetes) and the study outcome(s) [305]. The GRADE factor 'study limitations' was downgraded for ROM, pain and multi-dimensional scores because studies were generally deemed to be at a high risk of bias. Evidence for ROM outcomes was downgraded for inconsistency since, whilst the majority of studies did suggest that people with diabetes had better ROM at follow-up, some studies suggested that people with diabetes had worse ROM at follow-up (Table 6.5, Figure 6.3). The 'imprecision' GRADE factor was downgraded for all three domains since authors did not describe how sample size was calculated and some studies produced underpowered estimates with wide confidence intervals, making it difficult to determine if a true association likely existed. The certainty in evidence for all three domains was also downgraded for the 'publication bias' GRADE factor since it was difficult to determine if the decision to report results for the association between diabetes and outcomes was made based on a hypothesis made at the onset of the study or if it was post-hoc

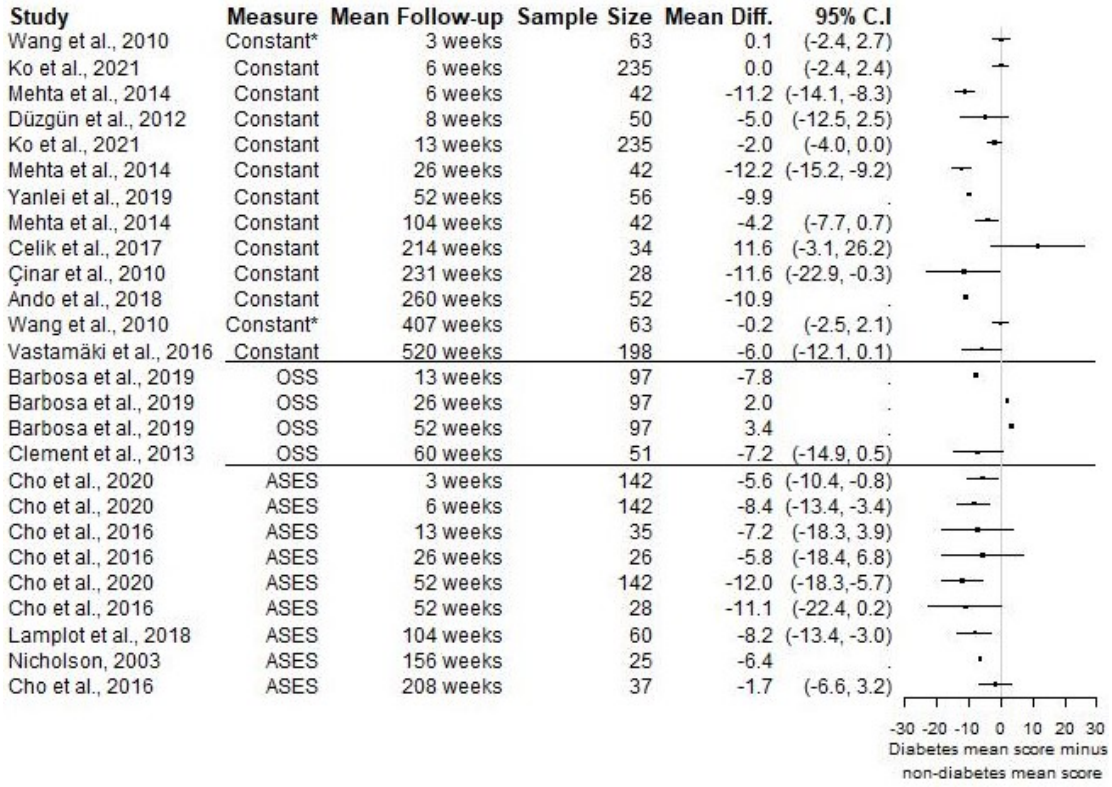

Figure 6.5: Forest plot of mean differences in multi-dimensional clinical scores between people with diabetes compared to people without diabetes. Constant scores and ASES scores are measured on a 0-100 scale and OSS are measured from 0-48. For each instrument a larger score represents a better outcome from frozen shoulder. *Wang et al. 2010 used an adjusted version of the Constant score, excluding the 25 points for the assessment of muscle strength

analysis. Certainty in evidence for ROM and pain outcomes were not upgraded for the 'effect size' GRADE factor since differences in outcomes between people with diabetes and without diabetes were generally small in magnitude.

The overall GRADE certainty in evidence for the association between diabetes and the outcomes of frozen shoulder was graded as very low for ROM outcomes, low for pain outcomes, and moderate for multi-dimensional clinical scores. For each domain, people with diabetes experienced worse outcomes than people without diabetes.

6.3.4.3 Summary of results reporting less-common outcome measures

Four studies reported results comparing function and disability in people with and without diabetes. Results from these studies are summarised in Table D.4. Two studies used the Simple Shoulder Test (SST), one study used the Visual Quality Scale and one study used the shoulder activity scale to measure function and disability. Nicholson et al. reported that people with diabetes had slightly worse SST scores at follow-up than people without diabetes [308]. Rill et al. reported equal mean SST scores at follow-up [329]. Lyhne et al. reported similar mean improvement in Visual Quality Scale scores, but did not report baseline scores so it is difficult to draw sensible conclusions from the results [326]. Lamplot et al. did not find evidence to suggest that people with diabetes and non-diabetes had different shoulder activity scale scores [328].

Results from the remaining eight studies that reported less-common outcomes are summarised in Table D.5. Two studies reported that people with diabetes were not more likely to require surgery than people without diabetes (10.5% vs. 10.5% and 14.0% vs. 17.4%) [183, 330]. Two studies found that people with diabetes were more likely to have surgery (70% vs. 44% and 14.7% vs. 5.5%) [336, 337]. Lamplot et al. provided evidence that people with diabetes were more likely to develop frozen shoulder in their contralateral shoulder than people without diabetes (77.8% vs. 29.4%, p=0.009) [328]. Lamplot et al. also reported that their patients with diabetes were more likely to require an additional glenohumeral joint injection than their patients that did not have diabetes (55.6% vs. 29.4%, p=0.15) [328]. Woods et al. and Jenkins et al. both reported that their patients with diabetes were more likely to require a second MUA than patients without diabetes [332, 333]. Ando et al. 2013 used a Cox regression model to analyse the association between diabetes and the recovery rate from frozen shoulder. They found evidence to suggest that diabetes is associated with worse recovery from frozen shoulder (HR=0.54; 95% CI: 0.36 – 0.96) [334].

			Direction of Association		
Outcome Domain	**Number of Participants**	**Number of Studies**	**Diabetes Group Generally Had Better Outcomes**	**Tie in Direction of Association**	**Diabetes Group Generally Had Worse Outcomes**
ROM	2107	13	3	0	10
Pain	920	10	0	0	10
Multi-dimensional scores	1785	18	2	1	15

Table 6.5: Summary of direction of the association in studies investigating the association between diabetes and frozen shoulder outcomes

Outcome Domain	**Phase of Investigation**	**Study Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication Bias**	**Effect Size**	**Exposure-response gradient**	**Overall Certainty in Evidence**
ROM	Phase 2	×	×	✓	×	×	×	×	Very low
Pain	Phase 2	×	✓	✓	×	×	×	×	Low
Multi-dimensional scores	Phase 2	×	✓	✓	×	×	✓	×	Moderate
GRADE scoring: ✓ means that no serious limitations were present, or ✓ means 'present' for the moderate/large association size and exposure-response gradient. × means that serious limitations were present, or × means 'absent' for the moderate/large association size and exposure-response gradient. A study is classed as a 'phase 2 study' if it is *"a cohort study that seeks to confirm independent associations between the prognostic factor and the outcome"*.									

Table 6.6: Summary of GRADE results

6.4 Discussion

This systematic review has summarised evidence from longitudinal observational studies to investigate whether patients with diabetes experience worse outcomes from frozen shoulder than patients without diabetes. Common outcome types that were reported in studies were ROM, pain and multi-dimensional clinical scores. The quality of evidence to suggest that people with diabetes experienced worse ROM at follow-up than people without diabetes was very low. GRADE certainty in evidence was graded as low for the association between diabetes and pain at follow-up. Certainty in evidence for diabetes as a prognostic factor for frozen shoulder outcomes measured using multi-dimensional clinical scores was graded as moderate. Additionally, 12 studies reported less-common outcome types ($\leq$4 studies). These studies also provided results to suggest that people with diabetes may experience less favourable outcomes from frozen shoulder.

The high variation in the length of follow-up and in the outcome measures that were used meant that the pooling of results would have been inappropriate and thus a narrative synthesis was used. The review used a transparent GRADE approach to evidence synthesis and grading certainty in evidence that was adapted for prognosis research. Forest plots were used to help with the visualisation of study results. Forest plots helped to complement the interpretation of raw data and tallies of association direction to guide the scoring of GRADE factors. Not all studies could be included in the forest plots as they reported results as categorical data or reported mean improvement in outcome scores from baseline to follow-up. Tallying association directions was a transparent method that allowed all results to be summarised collectively regardless of the scale that outcomes were reported on. Collectively, forest plots, tallies of association, interpretation of raw data, and the use of QUIPS risk of bias assessments allowed evidence to be synthesised in a way that all eight GRADE factors could be scored and conclusions about the certainty of evidence for each outcome domain could be made. A limitation of this approach is that the GRADE factors must be scored using the reviewers judgement so the approach is less transparent than quantitative pooling methods. The evidence synthesis and GRADE scoring was made as transparent as possible by explaining the reasons why GRADE factors were upgraded or downgraded and providing all the results and plots that were considered when scoring each

GRADE factor.

Whilst current evidence suggests that people with diabetes experience worse outcomes from frozen shoulder, the current evidence is generally at a high risk of bias. The QUIPS scoring for risk of unaccounted confounding was scored as high in 23 of the 28 studies, making it difficult to determine the added prognostic value of diabetes over and above existing prognostic factors (such as age, gender, ethnicity, and deprivation [183]). Another limitation of current evidence is that it was not clear in all studies whether the researcher's analysis was based on an a priori hypothesis or an a posteriori hypothesis that people with diabetes may experience worse outcomes from frozen shoulder, meaning that publication bias could have been introduced. It is recommended that study protocols and analysis plans should be published to improve the transparency of prognosis research [338]. Researchers that are planning to partake in research investigating the association between diabetes and outcomes in frozen shoulder should attempt to avoid these limitations in their research.

The studies identified in this review were generally conducted using clinical cohorts of patients that were receiving operative treatment. Such cohorts are generally towards the worse end of the frozen shoulder disease spectrum, either due to having a more severe form of frozen shoulder or due to being in a later phase of frozen shoulder. Since the majority of patients with frozen shoulder are based in primary care settings, researchers should consider investigating whether patients with frozen shoulder and coexisting diabetes also experience worse outcomes in the primary care setting. Additionally, it would be worthwhile investigating whether diabetes affects the effectiveness of specific frozen shoulder treatments; such research would need to be conducted using a randomised controlled trial or an individual participant data systematic review.

6.5 Conclusion

The evidence in this review suggests that people with diabetes may experience less favourable outcomes from frozen shoulder than people without diabetes. However, certainty in evidence for diabetes as a predictor of outcomes in frozen shoulder is currently very low for ROM outcomes, low for pain outcomes and moderate for multi-dimensional clinical outcome scores.

Current evidence is largely at a high risk of bias. To improve on current evidence, future research should account for covariates (such as age and gender) that could distort association estimates. Studies that compare outcome scores in people with and without diabetes should also clarify whether this comparison was intended at the onset of research or if an association was observed after data collection; this will help readers to judge whether publication bias could be present. Additionally, careful consideration should be taken to ensure studies include an adequately sized sample to allow for the comparison of outcomes between people with and without diabetes.

Whilst the certainty in evidence is very low, low, and moderate, the evidence does suggest that people with diabetes may have worse outcomes from frozen shoulder than people without diabetes. Thus, clinicians should consider monitoring patients with frozen shoulder that have diabetes and offer them further treatment if either pain or shoulder function do not improve.

Following this systematic review identifying the need for more prognosis research conducted in the primary care setting, the cohort study presented in the next chapter was conducted. The study used data from CPRD to determine whether diabetes is a predictor of surgery in people presenting with frozen shoulder in primary care.

Chapter 7

Is diabetes a predictor of surgery in people with frozen shoulder? A cohort study

The ISAC protocol for this study (19_219R) was accepted on 16th December 2020.

This study has been reported using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [262]. A completed STROBE checklist can be found in Appendix Section E.1.

7.1 Introduction

The systematic review in Chapter 6 found evidence to suggest that diabetes is a predictor of poorer outcomes in people with frozen shoulder. However, the majority of existing research is conducted in clinical cohorts of patients receiving surgery. People with frozen shoulder that require surgical intervention will be at a later stage of the disease and will likely have a more severe form of frozen shoulder, so results from studies including only people receiving surgery may not be generalisable to the entire frozen shoulder population. Since the majority of people with frozen shoulder will only ever be treated in primary care, it is important to understand

whether diabetes is also a prognostic factor in people presenting with frozen shoulder in primary care. This study, conducted using CPRD's primary care electronic health records, will investigate whether diabetes (any type) is a predictor of surgical intervention (an indicator of worse outcome from primary care management) in frozen shoulder.

7.2 Objective

To investigate the association between diabetes (the candidate prognostic factor) and receiving surgical intervention (a proxy for poor outcome) in people presenting with frozen shoulder in primary care.

7.3 Methods

7.3.1 Study sample and startpoint

To be eligible for this cohort study conducted in CPRD GOLD, patients needed to be aged 18 years or more and have presented with an incident episode of frozen shoulder between 1st May 2004 and 31st December 2017 (the date of frozen shoulder diagnosis was defined as the startpoint/index date). Patients that had any shoulder-related Read code before the startpoint were excluded, and patients needed to have a minimum of two years of up-to-standard data at the startpoint. It was necessary to have access to linked hospital episode statistic (HES) data in order to ensure optimal data on the endpoint (surgery). Participants also needed to have IMD data as deprivation was one of the covariates.

7.3.2 Establishing the presence of the candidate prognostic factor

At the startpoint, a patient was said to have the candidate prognostic factor (diabetes) if they had any diabetes Read code prior to, or on, the patient's startpoint.

7.3.3 Endpoint

The endpoint of interest in this study was the time to first ever frozen shoulder surgery (arthroscopic capsular release, MUA, hydrodilatation). The shoulder surgery code list can be found in Appendix Section E.2. Patients were followed from the startpoint until the earliest of: date of first frozen shoulder surgery (the endpoint), end of follow-up (17th February 2020), date of death (derived from CPRD data), date of transfer to a non-CPRD practice, or date of last CPRD data collection.

7.3.4 Covariates

To understand the prognostic value of diabetes it is important to adjust for other variables which have already been shown to be prognostic factors in people with frozen shoulder [302]. Covariates that were selected a priori as potential prognostic factors in people with frozen shoulder were: age, gender, ethnicity, and deprivation (IMD score) (for which previous research has shown the covariates to be potential prognostic factors in people with frozen shoulder [183]). It should be noted that other variables may be predictors of surgery; however, since the aim of this analysis was to understand whether diabetes was a predictor of poor outcomes in frozen shoulder (recall that surgery is used as a proxy for a poor outcome from frozen shoulder) it was decided to only adjust for variables that have been shown to be predictors of poor outcomes in people with frozen shoulder. Discussion of the limitations of using surgery as a proxy for a poor outcome from frozen shoulder is provided in Section 7.5.

Factor levels for the ethnicity variable were collapsed into white/not white/missing categories since only a small proportion (<3%) of patients fell into the non-white ethnicity categories.

Note: diabetes and covariate Read code lists can be found in Appendix B.2.

7.3.5 Analysis

A Kaplan-Meier plot was used to graph Kaplan-Meier estimates between people with diabetes and people without diabetes. A Cox model was used to determine whether diabetes was a predictor of time to surgery over and above other potential prognostic factors in people with frozen shoulder. The proportional hazards assumption was checked by inspecting Schoenfeld residual plots and Kaplan-Meier curves. Participants were censored upon their death, transfer to a non-CPRD practice, or at the end of follow-up (17th February 2020). Patients without diabetes were censored if they were diagnosed with diabetes after the index date. Missing data were handled using the missing data indicator method. A complete case analysis was conducted to assess the sensitivity of the results to missing data.

Within CPRD it was difficult to determine whether the surgery being conducted was for the first frozen shoulder which was diagnosed on the index date or if the surgery could have been for a second frozen shoulder on the same or contralateral side. The closer the surgery occurs to the index date, the more likely it is that the surgery was for the first frozen shoulder. So, as a sensitivity analysis, an additional two Cox models with a maximum of 3 and 5 years of follow-up were used to understand whether patients having a second frozen shoulder could affect the interpretation of the results.

Data were prepared for analysis in Stata version 14.0 [175] and analysis was conducted in RStudio version 1.2.5033 [270]. R code can be found in Appendix Section E.3.

7.4 Results

7.4.1 Sample characteristics

The study sample comprised of 40,644 patients with incident frozen shoulder diagnosed between 1st May 2004 and 31st December 2017. Within this sample, 6,319 patients (15.55%) had diabetes (Table 7.1). Sixty percent of the patients with frozen shoulder were female. The mean age at which the incident frozen shoulders were diagnosed was 61 years; the mean age of di-

agnosis was older in the patients with diabetes than in the patients without diabetes (64 years vs. 61 years). The median follow-up duration was 7.99 years (IQR: 4.73–11.10) in people with diabetes and 8.86 years (IQR: 5.40–12.04) in people without diabetes.

	Diabetes n=6,319 (15.55%)	**No diabetes n=34,325 (84.45%)**	**All n=40,644**
Median follow-up duration[1] (years)	7.99 (IQR: 4.73–11.10)	8.86 (IQR:5.40–12.04)	8.71 (IQR: 5.28–11.90)
Mean age	63.65 (SD=12.08)	61.08 (SD=12.78)	61.48 (SD=12.71)
Gender			
Male	3,148 (49.82%)	13,011 (37.91%)	16,159 (39.76%)
Female	3,171 (50.18%)	21,314 (62.09%)	24,485 (60.24%)
Ethnicity			
Bangladeshi	34 (0.54%)	33 (0.10%)	67 (0.16%)
Black African	47 (0.74%)	150 (0.44%)	197 (0.48%)
Black Caribbean	75 (1.19%)	217 (0.63%)	292 (0.72%)
Black – other	22 (0.35%)	60 (0.17%)	82 (0.20%)
Chinese	14 (0.22%)	87 (0.25%)	101 (0.25%)
Indian	207 (3.28%)	352 (1.03%)	559 (1.38%)
Mixed	40 (0.63%)	106 (0.31%)	146 (0.36%)
Other Asian	74 (1.17%)	190 (0.55%)	264 (0.65%)
Other	77 (1.22%)	310 (0.90%)	387 (0.95%)
Pakistani	98 (1.55%)	135 (0.39%)	233 (0.57%)
Missing	402 (6.36%)	4,537 (13.22%)	4,939 (12.15%)
White	5,229 (82.75%)	28,148 (82.00%)	33,377 (82.12%)
IMD Quintile			
Least deprived quintile	1,132 (17.91%)	8,461 (24.65%)	9,593 (23.60%)
2nd least deprived quintile	1,284 (20.32%)	7,580 (22.08%)	8,864 (21.81%)
3rd least deprived quintile	1,343 (21.25%)	7,405 (21.57%)	8,748 (21.52%)
4th least deprived quintile	1,309 (20.72%)	5,958 (17.36%)	7,267 (17.88%)
Most deprived quintile	1,247 (19.73%)	4,897 (14.27%)	6,144 (15.12%)
Missing	4 (0.06%)	24 (0.07%)	28 (0.07%)

Table 7.1: Table summarising baseline characteristics for study participants

[1]Defined as the time from the index date to the earliest of: end of study (17th February 2020), date of death, date of transfer to a non-CPRD practice, or date of last CPRD data collection.

7.4.2 Are people with diabetes more likely to have frozen shoulder surgery?

During the study, 3269 (8.04%) patients with frozen shoulder received surgery. Within the people with diabetes, 628 out of 6,319 patients (9.93%) had surgery, and within the people without diabetes, 2641 out of 34,325 patients (7.69%) received surgery. The Kaplan-Meier plot[2] in Figure 7.1 shows the difference in survival probabilities for people with diabetes compared to people without diabetes. The 1–, 3–, 5–, 10–, 15.8–year Kaplan-Meier estimates for the patients with diabetes and the patients without diabetes were 94.5%, 92.2%, 91.0%, 88.8%, 87.8% and 96.3%, 94.7%, 93.6%, 91.3%, 89.8% respectively.

The hazard ratio for surgery in patients with diabetes compared to patients without diabetes was estimated to be 1.38 (95% CI: 1.26 – 1.51), adjusted for covariates. There was little evidence of any violation of the proportional hazards assumption. The smoothing spline in Figure 7.2 forms a horizontal line with no pattern, indicating that the residuals are independent of time. The Kaplan-Meier plot in Figure 7.1 shows no crossover or divergence of Kaplan-Meier curves. When the complete case analysis was conducted (87.8% of patients had complete data), the results of the Cox model were similar (HR = 1.34; 95% CI: 1.24 – 1.49).

When restricting follow-up to 3 years or 5 years, the results from the Cox model remained similar. In the Cox model with a restricted 3-year follow-up, the hazard ratio for surgery in patients with diabetes compared to patients without diabetes was estimated to be 1.56 (95% CI: 1.41 – 1.73). In the Cox model with a restricted 5-year follow-up, the hazard ratio was estimated to be 1.51 (95% CI: 1.37 – 1.66). There was no evidence of a violation of the proportional hazards assumption in either model (Schoenfeld residual plots for the 3-year and 5-year follow-up Cox models can be found in Appendix E.4).

[2]Note that the y-axis does not start from zero. This choice was made so that the reader can see the shape of the curves more clearly.

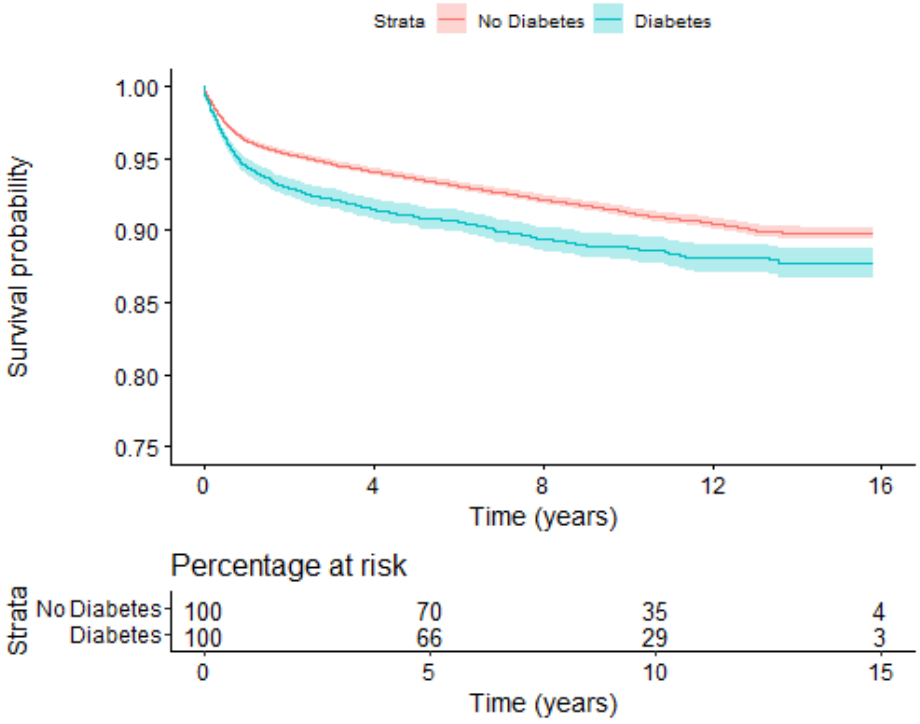

Figure 7.1: Plot of Kaplan-Meier estimates (with 95% confidence intervals) for surviving referral for frozen shoulder surgery among patients with diabetes compared to patients without diabetes, with a percentage at risk (of being referred for frozen shoulder surgery) table

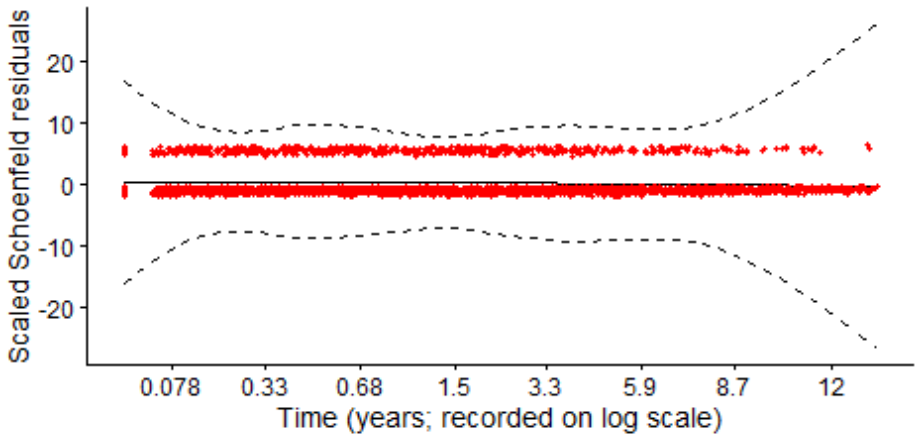

Figure 7.2: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the diabetes coefficient in the Cox model

7.5 Discussion

This study found evidence to suggest that people with diabetes are more likely to have surgery for their frozen shoulder than people without diabetes (9.93% vs. 7.69%). These results are consistent with the results of the systematic review in Chapter 6, suggesting that people with diabetes may experience worse outcomes from their frozen shoulder. However, it should be noted that the majority of patients with frozen shoulder (92%) do not need, or are not referred for surgery.

The systematic review in Chapter 6 identified four studies that had investigated whether diabetes was a predictor of the need for surgery in people presenting with frozen shoulder. All four studies were conducted within the secondary care setting, so the proportion of patients receiving surgery in each study was much greater than in this study, which was conducted in the primary care setting. In a prospective study with a 12-month follow-up in the UK, Barbosa et al. [336] found that 32 of the 46 (70%) patients presenting with frozen shoulder and coexisting diabetes received surgery, compared to 72 of the 164 (44%) in patients presenting with frozen shoulder but without diabetes. In another prospective study with a 2-year follow-up in Denmark, Gundtoft et al. [337] had five of 34 (14.7%) patients with frozen shoulder and diabetes that received surgery, compared to 11 out of 201 (5.5%) patients with frozen shoulder that did not have diabetes. Levine et al. [330] retrospectively studied 105 frozen shoulders in 98 patients from the USA and concluded that there was no difference in the likelihood of requiring surgery between patients with and without diabetes. Two of the 19 (10.5%) shoulders from patients with diabetes received surgery, and nine of the 86 (10.5%) shoulders from patients without diabetes received surgery. Kingston et al. [183] retrospectively studied 2190 patients with frozen shoulder in the USA. They found that 80 of the 572 (14%) patients with diabetes received surgery and 281 of the 1618 (17%) patients without diabetes had surgery.

In addition to people with diabetes potentially being more likely to require surgery, three studies from the systematic review in Chapter 6 found evidence to suggest that people with diabetes are more likely to need a second surgery [332, 333, 336].

As identified in the systematic review in Chapter 6, previous research had focused on secondary care cohorts of people with frozen shoulder receiving surgery. A strength of this study is that, through the use of primary care electronic health records, the sample analysed in this study should be representative of the general UK frozen shoulder population and not limited strictly to the more severe cases of frozen shoulder that require surgical intervention.

A limitation of using CPRD is that it is difficult to determine whether the surgery being conducted is for the first frozen shoulder being diagnosed on the index date or for a reoccurring or contralateral frozen shoulder. Lamplot et al. [328] found that people with diabetes were more likely to develop a second frozen shoulder on the contralateral side, however the sample size was small with nine patients with diabetes and 51 without diabetes. However, restricting follow-up in the analysis had little impact on the Cox model results. This would suggest that reoccurring or contralateral frozen shoulders do not substantially affect the ability to infer that patients with diabetes are more likely to require surgery for their frozen shoulder.

This study aimed to determine whether diabetes was a predictor of poor outcomes from frozen shoulder in patients presenting to primary care. However, direct measures of the clinical course of frozen shoulder, such as ROM or pain intensity, are not routinely collected in CPRD. Thus, to achieve the aim, surgical intervention was used as an indicator of a poor outcome for patients presenting to primary care with frozen shoulder. Whilst surgical intervention is likely to be a good indicator of a patient having worse outcomes from their frozen shoulder, this also will depend on local commissioning of services, practitioner opinion and patient choice. To gain a better understanding of the prognostic value of diabetes in people with frozen shoulder in primary care, a more bespoke cohort with patient reported outcomes would be required.

A Patient and Public Involvement and Engagement (PPIE) meeting would have been helpful to improve the understanding of patient's perceptions of surgery for frozen shoulder and discuss the findings of the analysis with people with lived experience of the condition. This was originally planned at the onset of the PhD but we were not able to do this due to the COVID-19

pandemic. Such a study would aid the interpretation of the findings of this chapter. [3]

Future research could also focus on investigating the reasons why patients with diabetes may have worse outcomes from frozen shoulder. Researchers could investigate whether glycaemic control is associated with worse outcomes from frozen shoulder. Such findings could help guide interventions to improve frozen shoulder outcomes in people with diabetes.

[3] Other than not being able to work with a patient group to understand their treatment preferences, all other research that was planned before the COVID-19 pandemic went ahead.

7.6 Conclusion

This cohort study conducted in CPRD has demonstrated that diabetes is a predictor of surgical intervention in people with frozen shoulder, further suggesting that patients with diabetes may have worse outcomes from frozen shoulder. However, the majority of patients in primary care are not referred to have surgery.

The next chapter will summarise the key findings from this thesis and discuss the main strengths and limitations of the studies presented. Some recommendations for future research will be made, and a description of the importance of the research presented in this thesis will be given. The chapter will then close with the thesis conclusion.

Chapter 8

Discussion

8.1 Summary of key findings

8.1.1 Diabetes as a risk factor for the onset of frozen shoulder: a systematic review and meta-analysis

The systematic review in Chapter 2 identified six case-control studies and two cohort studies that estimated the association between diabetes and the development of frozen shoulder. The six case-control studies, which included a total of 5388 patients, were pooled in a meta-analysis and the odds of developing frozen shoulder for people with diabetes was estimated to be 3.69 (95% CI: 2.99, 4.56) times larger than the odds for people without diabetes. The influence analysis demonstrated that the pooled odds ratio estimate was robust to the exclusion of any single study. A funnel plot was used to assess if any small-study bias may have been present, but the evidence was inconclusive due to only having a small number of studies included in the plot.

Similar to the case-control studies, the two cohort studies also estimated that people with diabetes were more likely to develop frozen shoulder than people without diabetes. The estimated hazard ratios in the cohort studies were 1.32 (95% CI: 1.22, 1.42) and 1.67 (95% CI: 1.46, 1.91).

The systematic review demonstrated that people with diabetes are more likely to develop frozen shoulder, although risk of bias was high in seven studies and moderate in one study. The

main reason for studies being at a high risk of bias was that all studies were at a high risk of unaccounted confounding. Thus, it was concluded that more research was required to determine whether the association between diabetes and frozen shoulder remained after accounting for common causes of the two conditions. It has been hypothesised that diabetes may be a cause of frozen shoulder due to glycation processes and/or inflammatory processes causing changes in the tissues of the glenohumeral joint capsule. The systematic review helped to identify a gap in the literature – that high quality epidemiological evidence with appropriate causal inference methods was needed to support or potentially oppose the claim that a causal relationship between diabetes and frozen shoulder could exist.

8.1.2 Type 2 diabetes and the risk of developing frozen shoulder: a cohort study

This cohort study aimed to estimate the total causal effect of type 2 diabetes on the development of frozen shoulder and decompose the total effect into an indirect effect (the effect mediated by the development of additional metabolic factors) and the remaining direct effect. The total effect of type 2 diabetes on frozen shoulder was large (HR=4.38; 95% CI: 3.70 – 5.21). The sensitivity analysis estimated that it would require very strong unmeasured confounding to completely explain away the total effect, with the E-value for the point estimate calculated as 4.87, and the E-value for the lower bound of the 95% CI calculated as 4.30. It would be unlikely that a confounder with such a strong association with both type 2 diabetes and frozen shoulder exists and would not have been identified during the examination of the literature that was conducted when constructing the DAG. If such a confounder does exist then it would likely be an unmeasurable (or difficult to measure) confounder such as genetics, diet, or an environmental factor.

When the total effect of type 2 diabetes on frozen shoulder was decomposed into the natural direct effect and natural indirect effect, there was no evidence that the number of metabolic factors developed during follow-up mediated the total effect (the indirect effect HR was 0.98; 95% CI: 0.93 – 1.03). Thus, the evidence in this study does not support the theory that inflammatory processes associated with metabolic syndrome lead to the fibrotic changes in the glenohumeral joint capsule of people with frozen shoulder.

8.1.3 Are patients with newly diagnosed frozen shoulder more likely to be diagnosed with type 2 diabetes? A cohort study

The cohort study in Chapter 5 investigated whether people with newly diagnosed frozen shoulder are more likely to have a subsequent diagnosis of type 2 diabetes than matched controls without frozen shoulder. The Kaplan-Meier survival probabilities at 1, 3, 5, 10, 15.8 years for the people with newly diagnosed frozen shoulder were 99.1%, 97.6%, 96.2%, 93.9%, 93.1% respectively. These were lower than the corresponding Kaplan-Meier survival probabilities, >99.9%, >99.9%, >99.9%, 99.7%, 99.2%, for the people without frozen shoulder.

In the age-, and gender-adjusted Cox regression model, newly diagnosed frozen shoulder was strongly associated with a following type 2 diabetes diagnosis (HR=19.37; 95% CI: 15.62 – 24.01). Further, when adjusting for other variables that could explain the association, the strong association between newly diagnosed frozen shoulder and a subsequent type 2 diabetes diagnosis was still present (HR=19.98; 95% CI: 15.99 – 24.97).

It should be noted that this evidence does not suggest that frozen shoulder causes type 2 diabetes. However, this evidence could suggest that people with frozen shoulder may be more likely to have undiagnosed type 2 diabetes than matched controls, but more research is needed to confirm this (see Section 8.5).

8.1.4 Diabetes as a prognostic factor in frozen shoulder: a systematic review

Twenty-eight longitudinal observation studies comparing frozen shoulder outcomes between patients with and without diabetes were identified in the systematic review in Chapter 6. Thirteen studies compared ROM measurements at follow-up; the total sample size across all 13 studies was 2107 patients. The evidence suggested that patients with frozen shoulder and coexisting diabetes had worse ROM at follow-up than people with frozen shoulder that did not have diabetes, although GRADE certainty in evidence was very low.

Ten studies reported pain (VAS) outcomes, and the total number of participants with frozen shoulder was 920. The results suggested that patients with diabetes had worse VAS scores at follow-up than patients without diabetes, but the GRADE certainty in evidence was low.

Multi-dimensional clinical scores were compared in 18 studies with a total of 1785 patients. The people with frozen shoulder and coexisting diabetes consistently had worse scores at follow-up than patients with frozen shoulder and without diabetes; the GRADE certainty in evidence was moderate.

In the studies reporting less-common outcomes, eleven results suggested people with diabetes had worse outcomes, one result suggested people with diabetes had better outcomes, and three studies suggested that there were no differences in frozen shoulder outcomes between patients with diabetes and patients without diabetes.

Risk of bias was judged to be high in 21 studies and moderate in seven. Other prognostic factors were often not accounted for and a comparison of baseline characteristics for the people with diabetes versus the people without diabetes were often not presented. Additionally, there was potentially some reporting bias present in some studies as it was not always clear whether the decision to compare outcomes between people with and without diabetes was made at the onset of the study or if an association was spotted once the data were analysed.

Lastly, it was noticed that the majority of existing research was conducted in cohorts of patients receiving surgical treatments and that there was limited research conducted in the primary care setting. The lack of understanding of how patient outcomes may differ for primary care patients was a concern because, for the majority of patients with frozen shoulder, primary care will be the only place where they are treated. Furthermore, cases of frozen shoulder that require surgery may be more severe, thus the results observed in secondary care populations may not be generalisable to the wider frozen shoulder population.

8.1.5 Is diabetes a predictor of surgery in people with frozen shoulder? A cohort study

Following the identification of a gap in the literature in the aforementioned systematic review, a cohort study was conducted in CPRD to investigate whether diabetes is an indicator of an unfavourable outcome of frozen shoulder in the primary care setting. Outcome measures, such as pain intensity and ROM are not routinely recorded in CPRD, so surgical intervention was used as an indicator of the outcome of a patient's frozen shoulder. The study demonstrated that diabetes was a predictor of a patient having surgery for their frozen shoulder (HR=1.38; 95% CI: 1.26 – 1.51).

The 1–, 3–, 5–, 10–, 15.8–year Kaplan-Meier estimates for the patients with diabetes and the patients without diabetes were 94.5%, 92.2%, 91.0%, 88.8%, 87.8% and 96.3%, 94.7%, 93.6%, 91.3%, 89.8% respectively. So, whilst diabetes was a predictor of surgery, a large proportion of the patients with frozen shoulder did not have surgery.

8.2 Methodological contribution

To my knowledge, no study has applied the weighted mediation analysis approach from Section 3.7.2 to data which have entirely come from electronic health records. Attempting to answer causal questions using observational data is always a challenge, but using electronic health records can add an extra layer of complexity. The following paragraph will describe one important learning point that relates specifically to using electronic health record data for causal inference.

The reliance upon Read codes to identify diseases may lead to diagnoses being more delayed than in a more bespoke cohort where participants are regularly being checked for the diseases of interest. The impact that this may have on the ability to draw causal conclusions will depend upon the specific causal mechanism being investigated. Diseases which have a long latent period (time from disease occurrence to disease detection) can cause differences in how well the tempo-

ral ordering described by the data relate to the temporal ordering of the real life causal processes (which are described in the DAG). A longer induction period (time from causal action to disease occurrence) will increase confidence that the temporal ordering in the data matches the temporal ordering of the causal process. For example, if the exposure generally occurs in children aged $<$18 and the outcome occurs around 50-60 years of age then, even if the outcome is diagnosed 10 years later than its onset, the temporal ordering of the diagnoses and the temporal ordering of the causal process will still coincide. Within a causal mediation analysis, the temporal ordering requires even more attention since not only is it required that confounders proceed the exposure, which proceeds the outcome, it is also required that the mediator occurs between the exposure and outcome. Understanding how the misclassification and/or delayed diagnosis brought about by reliance on Read codes to identify diseases requires careful consideration. This should be considered once the DAG has been constructed and prior to data collection.

Something else that researchers may find useful when designing a study that aims to answer a causal question is to think about what a hypothetical randomised trial may look like if it were to be conducted. This "target trial" framework is described in detail in Hernán et al. 2016 [339] and Hernán et al. 2020 [146]. A tool that utilises the target trial framework to assess risk of bias in non-randomised studies of interventions can be found in Sterne et al. 2016 [340].

Readers that wish to conduct a causal mediation analysis using the RMPW approach with a survival outcome described in Section 3.7.2 may find the code in Appendix Section B.4 helpful. The code was adapted from Rochon et al. [341].

8.3 Strengths

After identifying the confounding limitations of previous research in Chapter 2's systematic review, causal inference methods were used in the cohort study to investigate the association between type 2 diabetes and the development of frozen shoulder. The study was the first to investigate one of the hypothesised pathways through which type 2 diabetes may lead to the onset

of frozen shoulder. CPRD data allowed for the analysis of a large sample of patients with an incident diagnosis of type 2 diabetes over a period of 15.8 years. CPRD data provided information on all the variables that were identified a priori as confounders in the DAG, and allowed me to identify the dates at which data were entered/diagnoses were given.

The study in Chapter 5 built on previous small studies investigating the association between newly developed frozen shoulder and a later diagnosis of type 2 diabetes. My study utilised CPRD data to analyse a large sample of patients with frozen shoulder which are likely to be broadly representative of the general UK frozen shoulder population. The study supported the hypothesis that people with newly diagnosed frozen shoulder are more likely to have a subsequent diagnosis of type 2 diabetes than matched controls.

Following the systematic review in Chapter 6 suggesting that there was a need for more research conducted in the primary care setting to understand whether diabetes is a prognostic factor in frozen shoulder, the cohort study in Chapter 7 was conducted. The study utilised primary care records to analyse a sample that should be broadly representative of the entire population of people with frozen shoulder in the UK. The research allowed me to determine the prognostic value of diabetes within patients presenting with frozen shoulder in UK primary care without having to restrict the setting to only include people that had surgery, as the majority of pre-existing studies had done.

An additional strength of the studies presented in this thesis is that the robustness of the results was tested rigorously. Sensitivity analyses were conducted to test the robustness of results to missing data (Chapters 4, 5, 7), unobserved confounding (Chapter 4), extreme weights (Chapter 4), duration of follow-up (Chapter 7), outliers and/or influential estimates (Chapters 2). In each of the aforementioned sensitivity analyses, the results were shown to be robust.

8.4 Limitations

The use of CPRD for the studies included in this thesis allowed for the analysis of large samples of patients with diabetes and/or frozen shoulder that should be broadly representative of people in the UK with those respective diseases. However, the use of CPRD data does have limitations. Firstly, the use of Read codes to identify diseases may lead to some misclassification. In Chapter 4, the Kaplan-Meier estimates for being diagnosed with frozen shoulder at final follow-up in the people with type 2 diabetes and in the people without diabetes were 2.3% and 0.7%, respectively. This proportion is much lower than expected at the onset of the study, given that the prevalence of frozen shoulder in people with diabetes has been estimated to be 13.4% [25]. Previous research has suggested that UK general practitioners [280] and Dutch general practitioners [281] often use non-specific shoulder pain codes to record shoulder pain diagnoses and avoid recording a specific diagnosis. The potential issue of the underdiagnosis, or non-specific coding, of frozen shoulder is most likely to affect the study in Chapter 4 and has likely caused the lower than expected incidence of frozen shoulder. However, it is difficult to know whether the probability of outcome misclassification would have been differential with respect to the exposure and/or mediator; thus, it is difficult to determine how the misclassification may have effected the association estimates. It is also worth noting that people who do not consult their GP for shoulder pain are missed in CPRD. However, given the prolonged pain associated with frozen shoulder, it is unlikely that a large proportion of people with frozen shoulder do not consult their GP.

Misclassification may also be a concern for the diagnosis of type 2 diabetes. It has been estimated that one million people in the UK have undiagnosed type 2 diabetes, so reliance on Read codes rather than testing participants will have meant that some misclassification will have occurred. In Chapters 4 and 7, some participants would have had undiagnosed diabetes which likely would have caused the association estimates to be closer to the null than they would have been if the participants were classed as exposed within the studies (since the people with diabetes were more likely to develop frozen shoulder/have surgery in the respective studies).

In Chapter 4, the number of metabolic factors developed during follow-up was used as an indicator of metabolic health. Similar to type 2 diabetes, the other metabolic factors (hypertension, hyperlipidaemia, obesity) may go undiagnosed which could have affected the estimates of the direct and indirect effects. Further, when patients are diagnosed with type 2 diabetes they are more likely to be monitored and diagnosed for the presence of the other metabolic factors. This may lead to differential mediator misclassification since the probability of metabolic factors being undiagnosed will likely be reduced following a diagnosis of type 2 diabetes.

Additionally, covariates could have been misclassified. Throughout the thesis, diagnoses were identified using only Read codes, and not prescription codes. This may have led to reduced sensitivity for the identification of diseases such as hyperlipidaemia and hypertension, although will likely have had negligible impact on the identification of people with diabetes [342].

A further limitation of using electronic health records is that missing data are common. To avoid losing patients and therefore losing statistical power, the missing data indicator method was used to account for missing data. In some circumstances, the missing data indicator may lead to biased estimates. However, the approach is often used in the analysis of electronic health records to avoid loss of power and due to many multiple imputation methods being inappropriate since they assume that data are missing at random, which is unlikely for variables such as BMI, smoking, and alcohol. In Chapters 4, 5, and 7, complete case analyses were conducted to assess how missing data may have affected the results. Each sensitivity analysis suggested that the results were robust to missing data.

8.5 Recommendations for future research

The evidence in this thesis did not support the hypothesis that the reason type 2 diabetes is potentially a cause of frozen shoulder is due to the inflammation associated with type 2 diabetes and poor metabolic health [24, 45, 106]. Future work could focus on the other main hypothesis surrounding the association between type 2 diabetes and the onset of frozen shoulder – that

glycation processes may lead to the capsular fibrosis seen in frozen shoulder [113, 114]. If the hypothesis is true, then it should be expected that patients with worse glycaemic control are more likely to develop frozen shoulder. The association between longitudinal glycaemic control (measured using HbA1c tests) and the onset of frozen shoulder could be investigated using joint modelling strategies.

Whilst the evidence in this thesis suggests that patients with newly diagnosed frozen shoulder are more likely to have a subsequent diagnosis of type 2 diabetes, the research is limited by the inability to test all study participants. Future work could investigate whether the routine testing of patients with newly diagnosed frozen shoulder is an effective strategy to detect undiagnosed type 2 diabetes or pre-diabetes.

To further understand the prognostic value of diabetes in patients presenting with frozen shoulder in primary care, a bespoke cohort with direct measures of the clinical course of frozen shoulder (e.g. pain, function, ROM) is required. Future work could also focus on understanding the reasons why patients with diabetes may have worse outcomes from frozen shoulder. Researchers could investigate whether behavioural factors, such as adherence to treatment, could explain the difference in frozen shoulder outcomes between people with diabetes and people without diabetes. Other research has found that patients with diabetes are less-likely to adhere to rehabilitation programmes [343–347], although the reason behind the association is unclear. It could also be worthwhile investigating whether glycaemic control is associated with worse frozen shoulder outcomes. Diabetic control is monitored in patients with diabetes to reduce the likelihood of complications and has also been shown to be associated with the prognosis of diseases other than diabetes [348–350].

If a cohort study with direct measures of the clinical course of frozen shoulder were to be conducted, the study could include measurements of glycaemic control, metabolic factors, inflammatory and immunological markers, alongside treatment type and adherence to treatment to understand the reasons for people with diabetes potentially having poorer outcomes from frozen shoulder. If future research can confirm that diabetes is a prognostic factor in frozen shoulder

and determine the reasons for patients with diabetes having worse outcomes, then interventions can be introduced to improve patient recovery.

8.6 Diabetes and its complications

Frozen shoulder is one of many potential complications of diabetes. The results of this thesis should not be generalised to the other complications of diabetes, although the results of this thesis could help to identify gaps in the literature. Many musculoskeletal conditions have been shown to be associated with diabetes, and it has been hypothesised that the relationship in some cases may be causal. Glycation processes may be involved in the pathogenesis of limited joint mobility of the hand [351, 352], Dupuytren's contracture [352, 353], and trigger finger [352, 354]. Thus, if one were to use joint modelling strategies to investigate the association between HbA1c and the development of frozen shoulder then it would be valuable to also research the relationship between HbA1c and limited joint mobility of the hand, Dupuytren's contracture, and trigger finger too. Similar research has shown higher HbA1c levels to be associated with an increased risk of developing cardiovascular disease amongst people with type 1 diabetes [355]. Understanding the association between glycaemic control and the complications of diabetes would help to improve the understanding of the pathogenesis of diseases and emphasise the importance of maintaining a healthy HbA1c.

8.7 Importance of findings

This thesis demonstrates that people with diabetes are at a greater risk of developing frozen shoulder than people without diabetes and that type 2 diabetes may indeed be a cause of frozen shoulder. Given that musculoskeletal conditions are common amongst people with diabetes, clinicians may wish to consider asking patients whether they are experiencing any symptoms at their diabetes reviews.

Additionally, patients with newly diagnosed frozen shoulder are much more likely to be diagnosed with type 2 diabetes than people without frozen shoulder. In the study, the difference in 1–, 3–, 5–, 10–, 15.8–year Kaplan-Meier estimates between the people with frozen shoulder and the matched controls were 0.9%, 2.3%, 3.7%, 5.8% and 6.1%. Therefore, testing patients with frozen shoulder upon their diagnosis could have the potential to be an effective way of detecting type 2 diabetes earlier in its course, thus reducing the likelihood of any complications. However, more research, in which all participants are tested for type 2 diabetes, is needed to determine whether routinely testing patients with newly diagnosed frozen shoulder is an effective approach to detect type 2 diabetes or pre-diabetes. Until more research is conducted, clinicians should be aware of the association between frozen shoulder and type 2 diabetes and be alert to the possibility of undiagnosed type 2 diabetes in patients with frozen shoulder.

Lastly, current evidence suggests that diabetes may be a predictor of poor outcomes from frozen shoulder. However, the certainty in existing evidence was graded as being moderate to very low and my cohort study was limited by the lack of patient reported outcomes recorded in CPRD. If high quality studies can confirm that diabetes is a predictor of poor outcomes in frozen shoulder then clinicians should consider monitoring patients with diabetes and check whether further treatment may be required.

8.8 Thesis conclusion

Frozen shoulder is a painful condition that can severely restrict shoulder function. People with diabetes are more at risk of developing frozen shoulder. Further, the evidence in this thesis supports the hypothesis that type 2 diabetes could potentially be a cause of frozen shoulder, although the thesis has not been able to provide evidence of the mechanism underlying the association. Additionally, people with newly diagnosed frozen shoulder are more likely to have a subsequent diagnosis of type 2 diabetes than people without frozen shoulder. Lastly, current evidence suggests that people with diabetes may experience worse outcomes from frozen shoulder than people without diabetes.

Appendices

Appendix A

Diabetes as a risk factor for the onset of frozen shoulder: a systematic review and meta-analysis

A.1 PRISMA checklist

The following pages contain a completed PRISMA checklist corresponding to the work presented in Chapter 2.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Pg 15
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	n/a
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Sec 2.1
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Sec 2.1.1
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Tab 2.1, 2.2
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Sec 2.2.2.1, Sec 2.2.2.2, Sec A.2
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Sec A.2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Sec 2.2.2.4
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Sec 2.2.3
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Tab 2.1, Sec 2.2.3
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Sec 2.2.3
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Sec 2.2.3
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Sec 2.2.8
Synthesis	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study	Sec 2.2.8

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
methods		intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Sec 2.2.8, sec A.4
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Sec 2.2.8
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Sec 2.2.8, Sec A.3, Sec 2.2.5.1 - 2.2.5.5
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Sec 2.2.8, Sec 2.2.6
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Sec 2.2.8, Sec 2.2.7
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Sec 2.2.8
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	n/a
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Sec 2.3.1, Fig 2.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Reasons given in Fig 2.1
Study characteristics	17	Cite each included study and present its characteristics.	Tab 2.3, Tab 2.4
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Sec 2.3.3, Tab 2.5, Fig 2.2
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Tab 2.3, Tab 2.4, Sec 2.3.4, Fig 2.3, Sec 2.3.5

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Sec 2.3.2, Sec 2.3.3
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Sec 2.3.4, Sec 2.3.5
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	n/a
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Sec 2.3.4, Fig 2.4
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Fig 2.5, Sec 2.3.4
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	n/a
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Sec 2.4
	23b	Discuss any limitations of the evidence included in the review.	Sec 2.4
	23c	Discuss any limitations of the review processes used.	Sec 2.4
	23d	Discuss implications of the results for practice, policy, and future research.	Sec 2.4, Sec 2.5
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Pg 15
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Pg 15
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Amendments can be seen by following the hyperlink on pg 15
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Pg v
Competing interests	26	Declare any competing interests of review authors.	Pg i

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Data extracted: Tab 2.3, Tab 2.4, Sec A.4. Stata code: Sec A.3. Data collection spreadsheets available on request.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71
For more information, visit: http://www.prisma-statement.org/

A.2 Systematic review search strategies

A.2.1 MEDLINE

Interface: OVID. Searched on December 2018.

1. ((shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff) adj3 (instability or bursitis or frozen or impinge* or tendonitis or tendinitis or pain* or osteoarthr* or periarthriti* or peri arthriti* or arthralgia)).ti,ab,kw.
2. Shoulder Impingement Syndrome/
3. exp Bursitis/
4. Rotator Cuff/
5. adhesive capsuliti*.ti,ab,kw.
6. Shoulder Pain/
7. or/1-6
8. exp Pain/
9. pain*.ti,ab,kw.
10. Arthralgia/
11. arthralgia.ti,ab,kw.
12. or/8-11
13. Shoulder/
14. Shoulder joint/
15. Acromioclavicular Joint/
16. (shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff).ti,ab,kw.
17. or/13-16
18. 12 and 17
19. 7 or 18
20. exp Diabetes Mellitus/
21. diabet*.ti,ab,kw.
22. (DMi or DM i).ti,ab,kw.
23. (DM1 or DM 1).ti,ab,kw.

24. (DM2 or DM 2).ti,ab,kw.
25. (DMii or DM ii).ti,ab,kw.
26. (DM adj2 type).ti,ab,kw.
27. or/20-26
28. 19 and 27
29. exp animals/ not humans/
30. 28 not 29

A.2.2 EMBASE

Interface: OVID. Searched on December 2018.

1. ((shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff) adj3 (instability or bursitis or frozen or impinge* or tendonitis or tendinitis or pain* or osteoarthr* or periarthriti* or peri arthriti* or arthralgia)).ti,ab,kw.
2. exp shoulder impingement syndrome/
3. exp bursitis/
4. exp rotator cuff/
5. exp humeroscapular periarthritis/
6. adhesive capsuliti*.ti,ab,kw.
7. exp shoulder pain/
8. or/1-7
9. exp pain/
10. pain*.ti,ab,kw.
11. exp arthralgia/
12. arthralgia.ti,ab,kw.
13. or/9-12
14. exp shoulder/
15. Acromioclavicular Joint/
16. (shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff).ti,ab,kw.
17. or/14-16
18. 13 and 17
19. 8 or 18
20. exp Diabetes Mellitus/
21. diabet*.ti,ab,kw.
22. (DMi or DM i).ti,ab,kw.
23. (DM1 or DM1).ti,ab,kw.
24. (DM2 or DM 2).ti,ab,kw.
25. (DMii or DM ii).ti,ab,kw.

26. (DM adj2 type).ti,ab,kw.
27. or/20-26
28. 19 and 27
29. exp animals/ not humans/
30. 28 not 29
31. limit 30 to embase

A.2.3 AMED

Interface: OVID. Searched on December 2018.

1. ((shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff) adj3 (instability or bursitis or frozen or impinge* or tendonitis or tendinitis or pain* or osteoarthr* or periarthriti* or peri arthriti* or arthralgia)).ti,ab.
2. exp Shoulder impingement syndrome/
3. exp Bursitis/
4. exp Rotator cuff/
5. adhesive capsuliti*.ti,ab.
6. exp shoulder pain/
7. or/1-6
8. exp Pain/
9. pain*.ti,ab.
10. exp Arthralgia/
11. arthralgia.ti,ab.
12. or/8-11
13. shoulder/
14. (shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff).ti,ab.
15. or/13-14
16. 12 and 15
17. 7 or 16
18. exp Diabetes mellitus/
19. diabet*.ti,ab.
20. (DMi or DM i).ti,ab.
21. (DM1 or DM 1).ti,ab.
22. (DM2 or DM 2).ti,ab.
23. (DMii or DM ii).ti,ab.
24. (DM adj2 type).ti,ab.
25. or/18-24

26. 17 and 25
27. exp animals/ not humans/
28. 26 not 27

A.2.4 PsycINFO

Interface: OVID. Searched on December 2018.

1. ((shoulder* or glenohumer* or subacromi* or acromi* or rotator cuff) adj3 (instability or bursitis or frozen or impinge* or tendonitis or tendinitis or pain* or osteoarthr* or periarthriti* or peri arthriti* or arthralgia)).ti,ab.
2. Shoulder Impingement Syndrome.ti,ab.
3. bursitis.ti,ab.
4. rotator cuff.ti,ab.
5. adhesive capsuliti*.ti,ab.
6. shoulder pain.ti,ab.
7. or/1-6
8. exp PAIN/
9. pain*.ti,ab.
10. arthralgia.ti,ab.
11. or/8-10
12. *“shoulder (anatomy)”/
13. shoulder*.ti,ab.
14. shoulder joint.ti,ab.
15. acromi*.ti,ab.
16. glenohumer*.ti,ab.
17. subacromi*.ti,ab.
18. or/12-17
19. 11 and 18
20. 7 or 19
21. exp DIABETES MELLITUS/
22. diabet*.ti,ab.
23. (DMi or DM i).ti,ab.
24. (DM1 or DM 1).ti,ab.
25. (DM2 or DM 2).ti,ab.

26. (DMii or DM ii).ti,ab.
27. (DM adj2 type).ti,ab.
28. or/21-27
29. 20 and 28

A.2.5 Web of Science

Science Citation Index Expanded and the Science Conference Proceedings Citation Index. Searched on December 2018.

((

TS=(Shoulder* NEAR/3 instability) OR TS=(Shoulder* NEAR/3 bursitis) OR TS=(Shoulder* NEAR/3 frozen) OR TS=(Shoulder* NEAR/3 impinge*) OR TS=(Shoulder* NEAR/3 tendonitis) OR TS=(Shoulder* NEAR/3 tendinitis) OR TS=(Shoulder* NEAR/3 pain) OR TS=(Shoulder* NEAR/3 osteoarthr*) OR TS=(Shoulder* NEAR/3 periarthriti*) OR TS=(Shoulder* NEAR/3 "peri arthriti*") OR TS=(Shoulder* NEAR/3 arthralgia)

OR

TS=(glenohumer* NEAR/3 instability) OR TS=(glenohumer* NEAR/3 bursitis) OR TS=(glenohumer* NEAR/3 frozen) OR TS=(glenohumer* NEAR/3 impinge*) OR TS=(glenohumer* NEAR/3 tendonitis) OR TS=(glenohumer* NEAR/3 tendinitis) OR TS=(glenohumer* NEAR/3 pain) OR TS=(glenohumer* NEAR/3 osteoarthr*) OR TS=(glenohumer* NEAR/3 periarthriti*) OR TS=(glenohumer* NEAR/3 "peri arthriti*") OR TS=(glenohumer* NEAR/3 arthralgia)

OR

TS=(subacromi* NEAR/3 instability) OR TS=(subacromi* NEAR/3 bursitis) OR TS=(subacromi* NEAR/3 frozen) OR TS=(subacromi* NEAR/3 impinge*) OR TS=(subacromi* NEAR/3 tendonitis) OR TS=(subacromi* NEAR/3 tendinitis) OR TS=(subacromi* NEAR/3 pain) OR TS=(subacromi* NEAR/3 osteoarthr*) OR TS=(subacromi* NEAR/3 periarthriti*) OR TS=(subacromi* NEAR/3 "peri arthriti*") OR TS=(subacromi* NEAR/3 arthralgia)

OR

TS=(acromi* NEAR/3 instability) OR TS=(acromi* NEAR/3 bursitis) OR TS=(acromi* NEAR/3 frozen) OR TS=(acromi* NEAR/3 impinge*) OR TS=(acromi* NEAR/3 tendonitis) OR TS=(acromi* NEAR/3 tendinitis) OR TS=(acromi* NEAR/3 pain) OR TS=(acromi* NEAR/3 osteoarthr*) OR TS=(acromi* NEAR/3 periarthriti*) OR TS=(acromi* NEAR/3 "peri arthriti*") OR TS=(acromi* NEAR/3 arthralgia)

OR

TS=("rotator cuff" NEAR/3 instability) OR TS=("rotator cuff" NEAR/3 bursitis) OR TS=("rotator

cuff" NEAR/3 frozen) OR TS=("rotator cuff" NEAR/3 impinge*) OR TS=("rotator cuff" NEAR/3 tendonitis) OR TS=("rotator cuff" NEAR/3 tendinitis) OR TS=("rotator cuff" NEAR/3 pain) OR TS=("rotator cuff" NEAR/3 osteoarthr*) OR TS=("rotator cuff" NEAR/3 periarthriti*) OR TS=("rotator cuff" NEAR/3 "peri arthriti*") OR TS=("rotator cuff" NEAR/3 arthralgia)

OR

TS=("Rotator cuff")

OR

TS=("Adhesive capsuliti*")

)

OR

TS=(arthralgia NEAR/3 shoulder* or arthralgia NEAR/3 glenohumer* or arthralgia NEAR/3 subacromi* or arthralgia NEAR/3 acromi* or arthralgia NEAR/3 "rotator cuff")

OR TS=(pain* NEAR/3 shoulder* or pain* NEAR/3 glenohumer* or pain* NEAR/3 subacromi* or pain* NEAR/3 acromi* or pain* NEAR/3 "rotator cuff")

)

And

TS=(diabet* or DM1 or "DM 1" or DM2 or "DM 2" or DMi or "DM i" or DMii or "DM ii" or DM NEAR/2 type)

A.2.6 CINAHL

Interface: EBSCO. Searched on December 2018. Filters: title or abstract

(

((shoulder* or glenohumer* or subacromi* or acromi* or "rotator cuff") N3 (instability or bursitis or frozen or impinge* or tendonitis or tendinitis or pain* or osteoarthr* or periarthriti* or "peri arthriti*" or arthralgia))

OR

(MH "Shoulder Impingement Syndrome") OR (MH "Bursitis+") OR (MH "Rotator Cuff+") OR (MH "Periarthritis") OR (MH "Adhesive Capsulitis+") OR (MH "Shoulder Pain")

OR

((MH "Pain+") or pain or (MH "Arthralgia+") or arthralgia) and ((MH "Shoulder") or (MH "Acromioclavicular Joint") or shoulder* or glenohumer* or subacromi* or acromi* or "rotator cuff")

)

AND

((MH "Diabetes Mellitus+") or diabet* or (DMi or "DM i") or (DM1 or "DM 1") or (DMii or "DM ii") or (DM2 or "DM 2") or (DM N2 type))

A.2.7 Epistemonikos

Searched on December 2018. Filters: title or abstract. Primary study. Not an RCT.

((“frozen shoulder” or “shoulder impinge*” or “shoulder bursitis” or “shoulder tendonitis” or “shoulder tendinitis” or “shoulder pain” or “pain in the shoulder” or “painful shoulder” or “shoulder osteoarthr*” or “shoulder joint arthr*” or “shoulder arthr”)

OR

(“glenohumeral impinge*” or “glenohumeral bursitis” or “glenohumeral tendonitis” or “glenohumeral tendinitis” or “glenohumeral pain” or “pain in the glenohumeral” or “glenohumeral osteoarthr*” or “glenohumeral arthr*” or “glenohumeral arthr”)

OR

(“subacromial impinge*” or “subacromial bursitis” or “subacromial tendonitis” or “subacromial tendinitis” or “subacromial pain” or “pain in the subacromial” or “subacromial osteoarthr*” or “subacromial arthr*” or “subacromial arthr”)

OR

“Rotator cuff”

OR

“periarthriti*”

OR

“peri arthriti*”

OR

“Adhesive capsuliti*”

)

AND

(diabet* or DM1 or DM2 or DMi or DMii or “type 1 DM” or “type 2 DM” or “type i DM” or “type ii DM”)

A.2.8 TRIP

Searched on December 2018.

("frozen shoulder" or "shoulder pain" or "periathriti*" or "peri arthriti*" or "adhesive capsuliti*" or "shoulder impingement" or "bursitis" or "rotator cuff") and "diabet*"

A.2.9 PEDro

Searched on December 2018. Filters: body part = upper arm, shoulder or shoulder girdle

Title and abstract search: diabet*

A.2.10 Open Grey

Searched on December 2018.

Search 1: Diabet* and shoulder*

Search 2: Diabet* and glenohumer*

Search 3: Diabet* and subacromi*

Search 4: Diabet* and acromi*

Search 5: Diabet* and "rotator cuff*"

Search 6: Diabet* and bursitis

Search 7: Diabet* and periarthriti*

Search 8: Diabet* and "peri arthriti*"

Search 9: Diabet* and "adhesive capsuliti*"

Search 10: Diabet* and arthralgia

A.2.11 Grey literature report

Searched on December 2018.

Diabet*

A.3 Fitting the random-effects model – Stata code

The random-effects model was fitted using the `admetan` package in Stata version 16.1 [175]. The following code was used to fit the model:

```
admetan lnOR SElnOR, re(reml, hk) eform effect(Odds Ratio)
   forestplot(lcols(Author) nonames leftjustify rcols(
   TotalSampleSize) xtitle(Odds Ratio, size(2.5)) xlabel(0.5 1
   2 5 10 30) cirange(0.5 30) range(0.5 30) spacing(3) boxscale
   (70))
```

The first argument `lnOR` and the second argument `SElnOR` specify the names of the variables containing the log odds ratios and the corresponding standard errors for each study. The argument `re(reml,hk)` specifies that a random-effects model with REML estimation is to be used to estimate parameters and the Hartung-Knapp-Sidik-Jonkman (HKSJ) correction is to be used to estimate the 95% confidence interval for the pooled odds ratio. The argument `eform` specifies that the output of the model should be given on the exponentiated log odds scale. `effect(Odds Ratio)` gives the title "Odds Ratio" for the "effect size" column in the output. `forest plot` plots the data and results in a forest plot. The other options are to edit the forest plots contents and appearance and can be found on the `forest plot` help page in Stata.

A.4 Meta-analysis data

Author, Year	Number of cases	Number of controls	Number of cases with diabetes	Number of controls with diabetes
Boyle-Walker, et al., 1997* [178]	32	31	7	0
Li, et al., 2014† [179]	182	196	44	18
Lee, et al., 2012 [180]	40	40	6	1
Milgrom, et al., 2008 [181]	126	98	37	11
Wang, et al., 2013 [182]	87	176	17	13
Kingston, et al., 2018 [183]	2190	2190	572	188

Table A.1: Meta-analysis data for the association between diabetes and the onset of frozen shoulder. *A continuity correction of 0.5 was added to all cells in this row to avoid dividing by the zero count in the last column. † (Li, et al. 2014) adjusted for history of minor shoulder trauma in a multivariable regression model and the adjusted odds ratio presented in (Li, et al. 2014) was used in the meta-analysis

Appendix B

Type 2 diabetes and the risk of developing frozen shoulder: a cohort study

B.1 STROBE checklist

The following pages contain a completed STROBE checklist corresponding to the work presented in Chapter 4.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	47
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	n/a
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Pg 47, Chapter 1
Objectives	3	State specific objectives, including any prespecified hypotheses	Section 3.2
Methods			
Study design	4	Present key elements of study design early in the paper	Sec 3.3.8, Sec 3.3.9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Sec 3.3.8, Sec 3.3.9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Sec 3.3.8, Sec 3.3.9
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	Sec 3.3.9
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Sec 3.3.11, Fig 3.10, Sec B.2, Sec B.3
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Sec 3.3.9, Sec B.2, Sec B.3
Bias	9	Describe any efforts to address potential sources of bias	Sec 3.3.10
Study size	10	Explain how the study size was arrived at	Sec 3.3.9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Sec 3.3.10, Sec B.3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Sec 3.3.10, Sec 3.3.1-Sec 3.3.7
		(*b*) Describe any methods used to examine subgroups and interactions	Sec 3.3.10
		(*c*) Explain how missing data were addressed	Sec 3.3.10
		(*d*) If applicable, explain how loss to follow-up was addressed	Sec 3.3.10, Sec 3.3.6-3.3.7
		(*e*) Describe any sensitivity analyses	Sec 3.3.7.4, Sec 3.3.10
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Sec 3.4.1, Fig 3.11
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential	Sec 3.4.1, Tab 3.1

		confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Tab 3.1
		(c) Summarise follow-up time (eg, average and total amount)	Sec 3.4.1, Sec 3.4.2
Outcome data	15*	Report numbers of outcome events or summary measures over time	Sec 3.4.2, Fig 3.11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Sec 3.4.2, Sec 3.3.11
		(*b*) Report category boundaries when continuous variables were categorized	Sec B.2.8
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Sec 3.4.2
Discussion			
Key results	18	Summarise key results with reference to study objectives	Sec 3.5
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Sec 3.5
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Sec 3.5
Generalisability	21	Discuss the generalisability (external validity) of the study results	Sec 3.5
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Pg v

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

B.2 Read code lists

B.2.1 Type 2 diabetes Read code list

The type 2 diabetes Read code list in Table B.1 was constructed by two general practitioners. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
66A..	Diabetes monitoring
9OL..	Diabetic monitoring admin
C10	Diabetes mellitus
C100	Diab.mell. - no complication
C1000	Diab.mell.no comp. - juvenile
C1000-1	Insulin dependent diab mellit.
C1001	Diab.mell.no comp. - adult
C1001-1	Maturity onset diabetes
C1001-2	Non-insulin depend.diabet.mell
C100z	Diab.mell.no comp. - onset NOS
C102	Diab.mell. + hyperosmolar coma
C104	Diabetic nephropathy with renal manifestation
C1041	Diab.mell.+nephropathy - adult
C104-1	Diabetic nephropathy
C104z	Diab.mell.+nephropathy NOS
C105	Diab.mell.+ eye manifestation
C1051	Diab.mell.+eye manif - adult
C106	Diabetes + neuropathy
C106-2	Diab.mell. with neuropathy
C1061	Diab.mell.+neuropathy - adult
C106-1	Diabetic amyotrophy
C106-3	Diabetic mellitus with polyneuropathy
C106-99	Diabetes + neuropathy
C106y	Oth specf diab mel+neuro comps
C106z	Diab.mell.+neuropathy NOS
C1071	Diab.+periph.circ.dis.-adult
C107	Diab.mell. With peripheral circulatory disorder

Continued on next page

Read code	Detail
C107-1	Diab.mell. With gangrene
C107-2	Diab. With gangrene
C1074	NIDDM periph circulat disord
C107z	Diab.+periph.circ.disease NOS
C108	IDDM-Insul depend diabet melit
C108-1	IDDM-Insul depend diabet melit
C1085	Insul depen diab mel+ulcer
C1085-2	Insul depen diab mel+ulcer
C1088	Insul dep diab mell-poor contr
C109	Non-insulin dependent diabetes mellitus
C1091	Non-ins-dp diab mel+ophth comp
C109-1	NIDDM - Non-insulin dependent diabetes mellitus
C1091-1	Type II diab mel+ophth comp
C1091-2	Type 2 diab mel+ophth comp
C109-2	Type 2 diabetes mellitus
C1092-2	Type 2 diab mell neurol comp
C109-3	Type II diabetes mellitus
C1094	Non insulin dependent diab mell with ulcer
C1094-1	Type II diab mell with ulcer
C1094-2	Type 2 diab mell with ulcer
C1097	Non-insul dep diab-poor contr
C1097-1	Type II diab-poor contr
C1097-2	Type 2 diab-poor contr
C1099	Non-insul-dep diab mel no comp
C1099-1	Type II diabetes mellitus without complication
C1099-2	Type 2 diabetes mellitus without complication
C109C	Non inslulin dependant diab mell nephropathy
C109C-1	Type II diab mell nephropathy
C109C-2	Type 2 diab mell nephropathy
C109J	Insul treated Type 2 diab mell

Continued on next page

Read code	**Detail**
C109J-1	Insul treated non-insulin dep diab mell
C109J-2	Insul treated Type 2 diab mell
C10B	Diabet mel induced by steroids
C10B0	Sterod ind diab mel w/out comp
C10E	Insulin dep diabetes mellitus
C10E-1	Insulin dep diabetes mellitus
C10E-2	Insulin dep diabetes mellitus
C10E8	Insul dep diab mell-poor contr
C10EC-2	Insulin dependent diabetes mellitus with polyneuropathy
C10F	Type 2 diabetes mellitus
C10F0	Type 2 diab mell + renal compl
C10F0-1	Type 2 diab mell + renal compl
C10F1	Type 2 diab mell+ophthal comp
C10F-1	Type II diabetes mellitus
C10F1-1	Type II diab mell+ophthal comp
C10F2	Type 2 diab mell + neurol comp
C10F2-1	Type 2 diab mell + neurol comp
C10F6	Type 2 diab mell + retinopathy
C10F6-1	Type II diab mell + retinopathy
C10F7	Type 2 diab mell+poor control
C10F7-1	Type II diab mell+poor control
C10F8	Reaven's syndrome
C10F9	Type 2 diab mell without comp
C10F9-1	Type II diab mell without comp
C10FB	Type 2 diab mell + polyneurop
C10FB-1	Type II diab mell + polyneurop
C10FJ	Insul treated Type 2 diab mell
C10FJ-1	Insul treated Type 2 diab mell
C10FK	Hyperos non-ket stat typ 2 d m
C10FK-1	Hyperos non-ket stat typ 2 d m
C10FL	Type 2 d m + persist proteinur
C10FL-1	Type II d m + persist proteinur
C10FM	Type 2 d m + persist microalb
C10FM-1	Type II d m + persist microalb
C10FQ	Type 2 d m + exudat maculopath

Continued on next page

Read code	Detail
C10FQ-1	Type II d m + exudat maculopath
C10FR	Type 2 dm with gastroparesis
C10y	Diab.mell.+other manifestation
C10zz	Diab.mell. + unspec comp NOS

Table B.1: Table detailing the type 2 diabetes Read code list. Note: ".." represents the inclusion of all daughter codes

B.2.2 Diabetes (all types) Read code list

The diabetes Read code list in Table B.2 was constructed by two general practitioners. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
66A	Diabetic monitoring
66A1	Initial diabetic assessment
66A2	Follow-up diabetic assessment
66A3	Diabetic on diet only
66A4	Diabetic on oral treatment
66A5	Diabetic on insulin
66A6	Last hypo. attack
66A71	Frequency GP/param hypoglycaem
66A8	Has seen dietician - diabetes
66A9	Understands diet - diabetes
66AA	Injection sites
66AA-1	Injection sites - diabetic
66Ab	Diabetic foot examination
66Ac	Diabetic periph neurop screen
66AD	Fundoscopy - diabetic check
66Ae	HBA1c target
66Ae0	HBA1c target level - IFCC standardised
66AG	Diabetic drug side effects
66AH	Diabetic treatment changed
66AH0	Conversion to insulin
66AI	Diabetic - good control
66AJ	Diabetic - poor control
66AJ0	Chronic hyperglycaemia
66AJ1	Brittle diabetes
66AJ-1	Unstable diabetes
66AJz	Diabetic - poor control NOS
66Am	Insulin dose changed
66AM	Diabetic - follow-up default
66Ao	Diabetes type 2 review
66Aq	Diabetic foot screen
66AQ	Diabetes: shared care programme
66AQ0	Unsuit diab year care program

Continued on next page

Read code	Detail
66AR	Diabetes management plan given
66AS	Diabetic annual review
66AS0	Diabetic annual review
66AT	Annual diabetic blood test
66AV	Diabetic on insulin+oral treat
66AW	Diabetic foot risk assessment
66AZ	Diabetic monitoring NOS
9OL1	Attends diabetes monitoring
9OL2	Refuses diabetes monitoring
9OL4	Diabetes monitoring 1st letter
9OL5	Diabetes monitoring 2nd letter
9OL6	Diabetes monitoring 3rd letter
9OL7	Diabetes monitor.verbal invite
9OL8	Diabetes monitor.phone invite
9OL9	Diabetes monitoring deleted
9OLA	Diabetes monitored
9OLA-1	Diabetes monitor. check done
9OLD	Diabet pt unsuit dig ret photo
C10	Diabetes mellitus
C100	Diab.mell. - no complication
C1000	Diab.mell.no comp. - juvenile
C1000-1	Insulin dependent diab mellit.
C1001	Diab.mell.no comp. - adult
C1001-1	Maturity onset diabetes
C1001-2	Non-insulin depend.diabet.mell
C100z	Diab.mell.no comp. - onset NOS
C101	Diab.mell.with ketoacidosis
C1011	Diab.mell.+ketoacid - adult
C101y	Oth specfd diab mel+ketoacidos
C101z	Diab.mell.+ketoacid -onset NOS
C102	Diab.mell. + hyperosmolar coma
C104	Diabetic nephropathy with renal manifestation
C1041	Diab.mell.+nephropathy - adult
C104-1	Diabetic nephropathy
C104z	Diab.mell.+nephropathy NOS

Continued on next page

Read code	Detail
C105	Diab.mell.+ eye manifestation
C1051	Diab.mell.+eye manif - adult
C106	Diabetes + neuropathy
C106-2	Diab.mell. with neuropathy
C1061	Diab.mell.+neuropathy - adult
C106-1	Diabetic amyotrophy
C106-3	Diabetic mellitus with polyneuropathy
C106-99	Diabetes + neuropathy
C106y	Oth specf diab mel+neuro comps
C106z	Diab.mell.+neuropathy NOS
C1071	Diab.+periph.circ.dis.-adult
C107	Diab.mell. With peripheral circulatory disorder
C107-1	Diab.mell. With gangrene
C107-2	Diab. With gangrene
C1074	NIDDM periph circulat disord
C107z	Diab.+periph.circ.disease NOS
C108	IDDM-Insul depend diabet melit
C108-1	IDDM-Insul depend diabet melit
C108-2	Type 1 diabetes mellitus
C108-3	Type I diabetes mellitus
C1085	Insul depen diab mel+ulcer
C1085-1	TypeI diabetes +ulcer
C1085-2	Insul depen diab mel+ulcer
C1088	Insul dep diab mell-poor contr
C1088-1	Type I diab mell-poor contr
C1088-2	Type 1 diab mell-poor contr
C109	Non-insulin dependent diabetes mellitus
C1091	Non-ins-dp diab mel+ophth comp
C109-1	NIDDM - Non-insulin dependent diabetes mellitus
C1091-1	Type II diab mel+ophth comp
C1091-2	Type 2 diab mel+ophth comp
C109-2	Type 2 diabetes mellitus
C1092-2	Type 2 diab mell neurol comp
C109-3	Type II diabetes mellitus

Continued on next page

Read code	Detail
C1094	Non insulin dependent diab mell with ulcer
C1094-1	Type II diab mell with ulcer
C1094-2	Type 2 diab mell with ulcer
C1097	Non-insul dep diab-poor contr
C1097-1	Type II diab-poor contr
C1097-2	Type 2 diab-poor contr
C1099	Non-insul-dep diab mel no comp
C1099-1	Type II diabetes mellitus without complication
C1099-2	Type 2 diabetes mellitus without complication
C109C	Non inslulin dependant diab mell nephropathy
C109C-1	Type II diab mell nephropathy
C109C-2	Type 2 diab mell nephropathy
C109J	Insul treated Type 2 diab mell
C109J-1	Insul treated non-insulin dep diab mell
C109J-2	Insul treated Type 2 diab mell
C10B	Diabet mel induced by steroids
C10B0	Sterod ind diab mel w/out comp
C10E	Insulin dep diabetes mellitus
C10E-1	Insulin dep diabetes mellitus
C10E-2	Insulin dep diabetes mellitus
C10E8	Insul dep diab mell-poor contr
C10E8-1	Type I diab mell-poor contr
C10E8-2	Type 1 diab mell-poor contr
C10EC	Type 1 diabetes mellitus with polyneuropathy
C10EC-1	Type I diabetes mellitus with polyneuropathy
C10EC-2	Insulin dependent diabetes mellitus with polyneuropathy
C10EK	Type 1 diabetes mellitus with persistent proteinuria
C10EL	Type 1 diabetes mellitus with persistent microalbuminuria
C10EL-1	Type I diabetes mellitus with persistent microalbuminuria

Continued on next page

Read code	Detail
C10F	Type 2 diabetes mellitus
C10F0	Type 2 diab mell + renal compl
C10F0-1	Type 2 diab mell + renal compl
C10F1	Type 2 diab mell+ophthal comp
C10F-1	Type II diabetes mellitus
C10F1-1	Type II diab mell+ophthal comp
C10F2	Type 2 diab mell + neurol comp
C10F2-1	Type 2 diab mell + neurol comp
C10F6	Type 2 diab mell + retinopathy
C10F6-1	Type II diab mell + retinopathy
C10F7	Type 2 diab mell+poor control
C10F7-1	Type II diab mell+poor control
C10F8	Reaven's syndrome
C10F9	Type 2 diab mell without comp
C10F9-1	Type II diab mell without comp
C10FB	Type 2 diab mell + polyneurop
C10FB-1	Type II diab mell + polyneurop
C10FJ	Insul treated Type 2 diab mell
C10FJ-1	Insul treated Type 2 diab mell
C10FK	Hyperos non-ket stat typ 2 d m
C10FK-1	Hyperos non-ket stat typ 2 d m
C10FL	Type 2 d m + persist proteinur
C10FL-1	Type II d m + persist proteinur
C10FM	Type 2 d m + persist microalb
C10FM-1	Type II d m + persist microalb
C10FQ	Type 2 d m + exudat maculopath
C10FQ-1	Type II d m + exudat maculopath
C10FR	Type 2 dm with gastroparesis
C10y	Diab.mell.+other manifestation
C10zz	Diab.mell. + unspec comp NOS

Table B.2: Table detailing the diabetes (all types) Read code list. Note: ".." represents the inclusion of all daughter codes

B.2.3 Frozen shoulder Read code list

The frozen shoulder Read code list in Table B.3 was constructed by two general practitioners. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
N210	Adhesive capsulitis – shoulder
N210-2	Frozen shoulder
EGTON131	Semi Frozen Shoulder
EGTON251	? Frozen Right Shoulder
N0951	Shoulder joint stiffness
N095A	Stiff shoulder NEC
N2120	Periarthritis of shoulder
N21z0	Capsulitis NOS
N0951	Stiff joint NEC, of the shoulder region
N095A	Stiff shoulder NEC
N2120	Periarthritis of shoulder

Table B.3: Table detailing the frozen shoulder Read code list

B.2.4 Shoulder pain Read code list

The shoulder pain Read code list in Table B.4 was obtained from the Prognostic AND Diagnostic Assessment of Shoulder Pain (PANDA-S) study conducted in the Primary Care Centre Versus Arthritis. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
N0511	Local.primary OA-shoulder regn
N0521	Local.secondary OA-shoulder
N0531	Local.OA unsp.-shoulder region
N0541	Oligoartic OA, unspec-shoulder
N05z1	Osteoarthritis -shoulder joint
N05z9	Osteoarthritis NOS of shoulder
N05zA	OA NOS-sternoclavicular joint
N05zB	OA NOS-acromioclavicular join
N06z1	Arthropathy NOS-shoulder
N0941	Arthralgia - shoulder
N094A	Arthralgia of shoulder
N094B	Arthralgia - sternoclav joint
N094C	Arthralgia - acromioclav joint
N0961	Other joint sympt.-shoulder

Continued on next page

Read code	Detail
N096B	Other symptoms - sternoclav jt
N210	Adhesive capsulitis - shoulde
N2110	Rotator cuff syndrome unspecif
N2113	Supraspinatus tendinitis
N211z	Painful arc syndrome
N2122	Subacromial impingement
N2124	Impingement syndr of shoulder
N2125	Shoulder tendonitis
N21z2	Supraspinatus tendonitis
N245	Shoulder pain
N2457	Shoulder pain
N210-2	Frozen shoulder
N211z-1	Painful arc syndrome
N245-7	Shoulder pain
N03x0	Arthr assoc oth dis-shoulder
N03x1	Arthr ass oth dis-sternoclav j
N03x2	Arthr ass oth dis-acromioclv j
N0631	Climacteric arthr.-shoulder
N0651	Unsp.polyarthr.-shoulder
N0661	Unsp.monoarthr.-shoulder
N06y1	Other spec.arthr.-shoulder
N0801	Artic.cart.dis.-shoulder
N080B	Artic cart disord oth j-should
N0811	Loose body joint-shoulder
N0819	Loose body in shoulder joint
N0841	Joint contracture-shoulder
N084B	Extension contracture-shoulder
N084C	Abduction contracture-shoulder
N084D	Adduction contracture-shoulder
N084E	Int rotat contracture-shoulder
N084F	Ext rotat contracture-shoulder
N0851	Joint ankylosis-shoulder
N085A	Ankylosis of shoulder joint
N08y1	Oth.joint deran.NEC-shoulder
N08z1	Joint derange.NOS-shoulder
N0901	Joint effusion-shoulder region

Continued on next page

Read code	Detail
N090A	Effusion of shoulder
N090B	Effusion of sternoclav joint
N090C	Effusion of acromioclav joint
N0951	Shoulder joint stiffness
N095A	Stiff shoulder NEC
N095B	Stiff sternoclavic joint NEC
N095C	Stiff acromioclavicular joint NEC
N096A	Other symptoms - shoulder
N096C	Other symptoms - acromioclav j
N0980	Synov osteochondromat-shoulder
N09y1	Other joint dis.-shoulder
N09z1	Joint disord.NOS-shoulder
N211	Rotator cuff shoulder syndrome
N2111	Calcifying tendinitis shoulder
N2114	Part thickn rotator cuff tear
N2115	Full thickn rotator cuff tear
N2118	Bursitis of shoulder
N212	Other shoulder affections NEC
N2120	Periarthritis of shoulder
N2121	Scapulohumeral fibrositis
N212z	Other shoulder affect.NEC NOS
N21z0	Capsulitis NOS
NyuAB	[X]Other shoulder lesions
S50	Sprained shoulder
S500	Sprain acromio-clav ligament
S501	Sprain, coraco-clav ligament
S504	Rotator cuff sprain
S505	Sprain subscapularis tendon
S507	Sprain shoulder joint
S5070	Sprain shoulder joint anterior
S5071	Sprain shoulder joint posterior
S50w	Other shoulder sprain
S50y	Shoulder sprain NOS
S5y41	Sternoclavicular sprain
Syu46	[X]Spr/str oth/un part shl gir
N0611	Traumatic arthropathy of the shoulder region

Continued on next page

Read code	Detail
N061A	Traumatic arthropathy of shoulder
N0641	Transient arthropathy of the shoulder region
N064A	Transient arthropathy of shoulder
N0661	Unspecified monoarthritis of the shoulder region
N06z1	Arthropathy NOS, of the shoulder region
N0831	Recurrent joint dislocation, of the shoulder region
N083A	Recurrent dislocation of shoulder - anterior
N083B	Recurrent dislocation of shoulder - posterior
N083C	Recurrent subluxation of shoulder - anterior
N083D	Recurrent subluxation of shoulder - posterior
N083E	Recurrent dislocation of shoulder - inferior
N083F	Recurrent subluxation of shoulder - inferior
N083G	Recurrent dislocation of shoulder - anterior
N083H	Recurrent subluxation of shoulder - anterior
N083J	Recurrent dislocation of shoulder - multidirectional
N083K	Recurrent subluxation of shoulder - multidirectional
N083L	Habitual dislocation of the shoulder
N083M	Habitual subluxation of the shoulder
N0878	Snapping shoulder
N08y1	Other joint derangement NEC, of the shoulder region
N08z1	Joint derangement NOS, of the shoulder region
N0911	Haemarthrosis of the shoulder region
N091A	Haemarthrosis of shoulder
N092B	Villonodular synovitis of sternoclavicular joint
N092C	Villonodular synovitis of acromioclavicular joint
N0951	Stiff joint NEC, of the shoulder region
N095A	Stiff shoulder NEC
N09y1	Other specified joint disorders of the shoulder region
N09z1	Joint disorder NOS, of shoulder region
N2120	Periarthritis of shoulder
S5Q6	Injury of tendon of the rotator cuff of shoulder
SD28	Multiple superficial injuries of shoulder and upper arm
SD2y	Superficial injury shoulder/upper arm NOS, without infection
SD2y0	Superficial injury shoulder NOS

Continued on next page

Read code	Detail
SD2yz	Superficial injury shoulder/upper arm, without infection NOS
SK12	Other shoulder and upper arm injuries
SK122	Other shoulder injuries
SK12z	Other shoulder and upper arm injury NOS
Syu4	[X]Injuries to the shoulder and upper arm
Syu40	[X]Other superficial injuries of shoulder and upper arm
Syu4D	[X]Other specified injuries of shoulder and upper arm
Syu4E	[X]Unspecified injury of shoulder and upper arm
1M02	Shoulder joint painful on movement
1M03	Shoulder joint painful on external rotation
N2457	Shoulder pain
7NC3	[SO]Ligament of shoulder or elbow
7NC30	[SO]Ligament of sternoclavicular joint
7NC31	[SO]Ligament of acromio-clavicular joint
7NC32	[SO]Coraco-clavicular ligament
7NC33	[SO]Coraco-acromial ligament
7NC34	[SO]Gleno-humeral ligament
7NC35	[SO]Glenoid labrum
7NC36	[SO]Ligament of elbow joint
7NC37	[SO]Ligament of superior radio-ulnar joint
7NC38	[SO]Annular ligament
7NC39	[SO]Ligament of inferior radio ulnar joint
S50	Sprain of shoulder and upper arm
S500	Sprain, acromio-clavicular ligament
S501	Sprain, coraco-clavicular ligament
S502	Coracohumeral sprain
S503	Sprain, infraspinatu tendon
S504	Rotator cuff sprain
S505	Sprain, subscapularis tendon
S506	Sprain, supraspinatus tendon
S507	Sprain, shoulder joint
S5070	Sprain, shoulder joint, anterior
S5071	Sprain, shoulder joint, posterior
S508	Sprain, biceps tendon
S509	Sprain, long head of biceps tendon

Continued on next page

Read code	Detail
S50A	Sprain, triceps tendon
S50X	Sprain and strain of other and unspecified parts of shoulder girdle
S50w	Other shoulder sprain
S50x	Other upper arm sprain
S50y	Shoulder sprain NOS
S50z	Upper arm sprain NOS
N2124	Impingement syndrome of shoulder
7N82	[SO]Muscle of shoulder or upper arm
7N820	[SO]Deltoid
7N821	[SO]Rotator cuff
7N823	[SO]Biceps brachii
7N827	[SO]Supraspinatus
7N828	[SO]Flexor of upper arm
7N829	[SO]Extensor of upper arm
7N82y	[SO]Specified muscle of shoulder or upper arm NEC
7N82z	[SO]Muscle of shoulder or upper arm NEC
7NAD	[SO]Joint of shoulder girdle or arm
7NAD0	[SO]Sternoclavicular joint
7NAD1	[SO]Acromioclavicular joint
7NAD2	[SO]Glenohumeral joint
7NAD3	[SO]Shoulder joint
7NADz	[SO]Joint of shoulder girdle or arm NEC
N211	Rotator cuff shoulder syndrome and allied disorders
N2110	Rotator cuff syndrome, unspecified
N2111	Calcifying tendinitis of the shoulder
N2112	Bicipital tenosynovitis
N2113	Supraspinatus tendinitis
N2114	Partial thickness rotator cuff tear
N2115	Full thickness rotator cuff tear
N2116	Subacromial bursitis
N2117	Subdeltoid bursitis
N2118	Bursitis of shoulder
N211z	Rotator cuff syndrome NOS

Table B.4: Table detailing shoulder pain Read code list

B.2.5 Hypertension Read code list

The hypertension Read code list in Table B.5 was obtained from another study (Blagojevic-Bucknall et al., 2017 [356]) conducted in the Primary Care Centre Versus Arthritis. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
G24z100	Hypertension secondary to drug
G202.00	Systolic hypertension
G21z.00	Hypertensive heart disease NOS
9OI..00	Hypertension monitoring admin.
G24z000	Secondary renovascular hypertension NOS
G210.00	Malignant hypertensive heart disease
9OI1.00	Attends hypertension monitor.
G22z.11	Renal hypertension
9OIA.11	Hypertension monitored
G241z00	Secondary benign hypertension NOS
G203.00	Diastolic hypertension
9OI..11	Hypertension clinic admin.
G20..00	Essential hypertension
662O.00	On treatment for hypertension
8HT5.00	Referral to hypertension clinic
G201.00	Benign essential hypertension
8CR4.00	Hypertension clinical management plan
662..12	Hypertension monitoring
G240z00	Secondary malignant hypertension NOS
6629	Hypertension:follow-up default
G240000	Secondary malignant renovascular hypertension
G22..00	Hypertensive renal disease
G222.00	Hypertensive renal disease with renal failure
Gyu2100	[X]Hypertension secondary to other renal disorders
662P.00	Hypertension monitoring
8I3N.00	Hypertension treatment refused
662c.00	Hypertension six month review
G20z.00	Essential hypertension NOS
G2...11	BP - hypertensive disease
G20..11	High blood pressure
662G.00	Hypertensive treatm.changed

Continued on next page

Read code	Detail
9N1y200	Seen in hypertension clinic
662H.00	Hypertension treatm.stopped
G244.00	Hypertension secondary to endocrine disorders
67H8.00	Lifestyle advice regarding hypertension
9h31.00	Excepted from hypertension qual indicators: Patient unsuit
6627	Good hypertension control
G24zz00	Secondary hypertension NOS
9OI2.00	Refuses hypertension monitor.
9OI4.00	Hypertens.monitor.1st letter
6628	Poor hypertension control
G240.00	Secondary malignant hypertension
662b.00	Moderate hypertension control
662d.00	Hypertension annual review
G220.00	Malignant hypertensive renal disease
14A2.00	H/O: hypertension
G20z.11	Hypertension NOS
G232.00	Hypertensive heart&renal dis wth (congestive) heart failure
G21zz00	Hypertensive heart disease NOS
G234.00	Hyperten heart&renal dis+both(congestv)heart and renal fail
662F.00	Hypertension treatm. started
Gyu2.00	[X]Hypertensive diseases
G210000	Malignant hypertensive heart disease without CCF
G21z011	Cardiomegaly - hypertensive
9OIA.00	Hypertension monitor.chck done
G200.00	Malignant essential hypertension
9h32.00	Excepted from hypertension qual indicators: Informed dissent
G2...00	Hypertensive disease
G24z.00	Secondary hypertension NOS
9OI6.00	Hypertens.monitor 3rd letter
G24..00	Secondary hypertension
9N03.00	Seen in hypertension clinic
G241.00	Secondary benign hypertension
G221.00	Benign hypertensive renal disease
G211.00	Benign hypertensive heart disease

Continued on next page

Read code	Detail
G2z..00	Hypertensive disease NOS
9OI5.00	Hypertens.monitor 2nd letter
662q.00	Trial reduction of antihypertensive therapy
9OI7.00	Hypertens.monitor verbal inv.
G241000	Secondary benign renovascular hypertension
G21..00	Hypertensive heart disease
9OI8.00	Hypertens.monitor phone invite
G2y..00	Other specified hypertensive disease
G211100	Benign hypertensive heart disease with CCF
G22z.00	Hypertensive renal disease NOS

Table B.5: Table detailing hypertension Read code list

B.2.6 Hyperlipidaemia Read code list

The hyperlipidaemia Read code list in Table B.6 was obtained from another study (Blagojevic-Bucknall et al., 2017 [356]) conducted in the Primary Care Centre Versus Arthritis. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
C320100	Hyperbetalipoproteinaemia
9N0J.00	Seen in cholesterol clinic
8I3J.00	Lipid lowering therapy declined
Cyu8D00	[X]Other hyperlipidaemia
C320200	Hyperlipidaemia, group A
1442	H/O: raised blood lipids
C321.00	Pure hyperglyceridaemia
ZV65317	[V]Dietary surveillance in hypercholesterolaemia
8B28.00	Lipid lowering therapy
C321000	Hypertriglyceridaemia
9Oc3.00	Lipid disorder monitoring second letter
9Oc..00	Lipid disorder monitoring administration
8BAG.00	Cholesterol reduction programme
C320300	Low
66X..00	Lipid disorder monitoring
8B6A.00	Statin prophylaxis
C320500	Familial defective apolipoprotein B-100

Continued on next page

Read code	Detail
C320y00	Other specified pure hypercholesterolaemia
8I76.00	Statin not tolerated
C322.11	Fredrickson type IIb lipidaemia
C322.12	Fredrickson type III lipidaemia
8CR3.00	Hyperlipidaemia clinical management plan
8BAG200	Cholesterol reduction program - declined
C324.00	Hyperlipidaemia NOS
C325000	High density lipoid deficiency
C327z00	Lipidoses NOS
44P3.00	Serum cholesterol raised
C323.12	Fredrickson type I lipaemia
8I3C.00	Statin declined
C325.00	Lipoprotein deficiencies
C322.00	Mixed hyperlipidaemia
C328.00	Dyslipidaemia
C320.12	Fredrickson type IIa lipidaemia
C320.13	Low density lipoproteinaemia
C320.11	Familial hypercholesterolaemia
8BG2.00	Lipid lowering therapy indicated
8BAG000	Cholesterol reduction programme - invited
C320z00	Pure hypercholesterolaemia NOS
C320000	Familial hypercholesterolaemia
44P4.00	Serum cholesterol very high
C321.11	Fredrickson type IV lipidaemia
C321.12	Very low density lipoprotinaemia
9Oc0.00	Attends lipid disorder monitoring
8HT1.00	Referral to lipid clinic
C320.00	Pure hypercholesterolaemia
9N0I.00	Seen in lipid clinic
9N4K.00	DNA - Did not attend cholesterol clinic
8BL1.00	Patient on maximal tolerated lipid lowering therapy
ZC2CJ00	Dietary advice for hyperlipidaemia
8BAG100	Cholesterol reduction program - attended

Table B.6: Table detailing hyperlipidaemia Read code list

B.2.7 Thyroid dysfunction (all types) Read code list

The thyroid dysfunction Read code list in Table B.7 was constructed by a general practitioner. The codes are recorded in the clinical file within CPRD GOLD.

Read code	Detail
C046.00	Autoimmune myxoedema
A175.00	Tuberculosis of thyroid gland
9O39100	Thyroid monitoring SMS text message first invitation
C02y200	Thyrotoxicosis factitia
7113y00	Other specified other operation on thyroid gland
C043z00	Iatrogenic hypothyroidism NOS
C051.11	De Quervain's thyroiditis
C043200	Hypothyroidism resulting from resorcinol
1433	H/O: thyroid disorder NOS
C04z100	Myxoedema coma
711z.00	Thyroid gland and parathyroid gland operations NOS
BB5fz00	[M]Thyroid adenoma or adenocarcinoma NOS
C000.13	Thyroid nodule
C03y100	Congenital hypothyroidism without goitre
U602100	[X]Thyroid horms + substits caus adverse eff in therap use
C02yz00	Thyrotoxicosis of other specified origin NOS
PK25z00	Anomaly of thyroid gland NEC NOS
C02..00	Thyrotoxicosis
C052.11	Autoimmune thyroiditis
C02y000	Thyrotoxicosis of other specified origin with no crisis
C050.00	Acute thyroiditis
F395300	Myopathy due to myxoedema
5A12.00	Thyroid tumour/metast irradiat
4423	Thyroid hormone tests low
F4G2000	Thyrotoxic exophthalmos
C044.00	Postinfectious hypothyroidism
7113z00	Other operation on thyroid gland NOS
ByuB.00	[X]Malignant neoplasm of thyroid and other endocrine glands
C03z.00	Congenital hypothyroidism NOS
7110000	Total thyroidectomy
C03..00	Congenital hypothyroidism

Continued on next page

Read code	Detail
7N10000	[SO]Thyroid gland
7111y00	Other specified operation on aberrant thyroid tissue
Cyu1400	[X]Other chronic thyroiditis
C062.00	Thyroid cyst
C04z.12	Thyroid insufficiency
C03y.00	Other specified congenital hypothyroidism
C022.00	Toxic multinodular goitre
TJ27000	Adverse reaction to liothyronine sodium
C024z00	Thyrotoxicosis from ectopic thyroid nodule NOS
C024000	Thyrotoxicosis from ectopic thyroid nodule with no crisis
C00z.00	Goitre NOS
L181z00	Thyroid dysfunction in pregnancy/childbirth/puerperium NOS
9O39200	Thyroid monitoring SMS text message 2nd invitation
C05y.00	Other and unspecified chronic thyroiditis
8B71.00	Iodine-goitre prophylaxis
C011.00	Nontoxic multinodular goitre
C03z.11	Congenital thyroid insufficiency
1431	H/O: hyperthyroidism
C061.00	Dyshormonogenic goitre
C041z00	Postablative hypothyroidism NOS
C052.00	Chronic lymphocytic thyroiditis
C052.12	Hashimoto's disease
7110300	Lobectomy of thyroid gland NEC
66B4.00	Thyroid eye disease
C04z000	Premature puberty due to hypothyroidism
Cyu1300	[X]Other thyrotoxicosis
212P.00	Hyperthyroidism resolved
TJ27200	Adverse reaction to thyroglobulin
C06y000	Thyroid-binding globulin abnormality
C03y000	Congenital hypothyroidism with diffuse goitre
C021000	Toxic uninodular goitre with no crisis
66B..00	Thyroid disease monitoring
C02z.00	Thyrotoxicosis without mention of goitre or other cause
B8yy000	Carcinoma in situ of thyroid gland
PK25.00	Anomalies of thyroid gland NEC

Continued on next page

Read code	Detail
C000.12	Substernal thyroid goitre
1431.11	H/O: thyrotoxicosis
Cyu1100	[X]Other sp cified hypothyroidism
C020200	Thyroid-associated dermopathy
FyuBD00	[X]Dysthyroid exophthalmos
C02y100	Thyrotoxicosis of other specified origin with crisis
Cyu1600	[X]Iodine-deficiency-related (endemic) goitre, unspecified
9h71.00	Excepted from thyroid quality indicators: Patient unsuitable
AC22.00	Thyroid echinococcus granulosus
C047.00	Subclinical hypothyroidism
7113100	Biopsy of lesion of thyroid gland
C134300	TSH - thyroid-stimulating hormone deficiency
1432	H/O: hypothyroidism
F144100	Cerebellar ataxia due to myxoedema
C0A1.00	Congenital iodine-deficiency syndrome, myxoedematous type
C0A5.00	Subclinical iodine-deficiency hypothyroidism
7N10100	[SO]Aberrant thyroid tissue
C020.00	Toxic diffuse goitre
C025.00	Subclinical hyperthyroidism
G557500	Thyrotoxic heart disease
Q44V.00	Neonatal goitre, not elsewhere classified
711..00	Thyroid gland and parathyroid gland operations
C000.14	Colloid goitre
C000.11	Retrosternal thyroid goitre
9O39300	Thyroid monitoring SMS text message third invitation
5A11.00	Thyroid gland ablat - irradiat
4422	Thyroid hormone tests high
C043000	Hypothyroidism resulting from para-aminosalicylic acid
C023z00	Toxic nodular goitre NOS
C0...00	Disorders of thyroid gland
7110200	Hemithyroidectomy
66BA.00	Thyroid dis.treatment stopped
7110z00	Thyroidectomy NOS
C04z.00	Hypothyroidism NOS

Continued on next page

Read code	Detail
Qyu6400	[X]Other transitory neonatal disorders/thyroid function,NEC
L181200	Thyroid dysfunction in puerperium - baby delivered
Q443.00	Neonatal thyrotoxicosis
442G.00	Thyroid hormone tests abnormal
C0AX.00	Iodine-deficiency-related (endemic) goitre, unspecified
C05z.00	Thyroiditis NOS
C010.00	Nontoxic uninodular goitre
687H.00	Congenital hypothyroidism screening related finding
SL27.00	Thyroid hormone and thyroid derivatives poisoning
TJ28z00	Adverse reaction to antithyroid agents NOS
7113200	Incision of lesion of thyroid gland
TJ27.00	Adverse reaction to thyroid and thyroid derivatives
66B8.00	Thyroid dis.treatment changed
9Oj..00	Hypothyroidism monitoring administration
9Oj1.00	Hypothyroidism monitoring second letter
9Oj3.00	Hypothyroidism monitoring verbal invite
66B5.00	Thyroid symptom change
C063000	Thyroid haemorrhage
9Oj4.00	Hypothyroidism monitoring telephone invitation
711..12	Thyroid gland operations
F381600	Myasthenic syndrome due to thyrotoxicosis
C063.00	Thyroid haemorrhage and infarction
C0A4.00	Iodine-deficiency-related multinodular (endemic) goitre
7110400	Isthmectomy of thyroid gland
SL27300	Thyroglobulin poisoning
C023000	Toxic nodular goitre unspecified with no crisis
C063100	Thyroid infarction
C01z.11	Adenomatous goitre
C0...11	Struma - goitre
66BA.11	Thyroxine Rx stopped
8CR5.00	Hypothyroidism clinical management plan
L181000	Thyroid dysfunction - unspec whether in pregnancy/puerperium
C045.00	Acquired atrophy of thyroid
Cyu1.00	[X]Disorders of thyroid gland

Continued on next page

Read code	Detail
C050200	Abscess of thyroid
66B3.00	Inactive thyroid disease
C053.00	Chronic fibrous thyroiditis
C05y400	Chronic thyroiditis with transient thyrotoxicosis
F395400	Myopathy due to thyrotoxicosis
Cyu1200	[X]Other specified nontoxic goitre
6762	Education about thyroid disease in pregnancy
22H3.00	O/E - thyroid swelling -bilat.
7113300	Exploration of thyroid gland
C0A3.00	Iodine-deficiency-related diffuse (endemic) goitre
C06yz00	Other specified thyroid disorder NOS
L181400	Thyroid dysfunction in puerperium- baby previously delivered
B53..00	Malignant neoplasm of thyroid gland
C020100	Toxic diffuse goitre with crisis
C06z.00	Thyroid disorder NOS
442I.00	Thyroid function tests abnormal
C00..00	Simple and unspecified goitre
7113000	Excision of lesion of thyroid gland
C050z00	Acute thyroiditis NOS
C02z100	Thyrotoxicosis without mention of goitre, cause with crisis
R145.00	[D]Thyroid function test abnormal
C041000	Irradiation hypothyroidism
PK25011	Retrosternal thyroid gland
C06y100	Thyroid atrophy
C06..00	Other disorders of thyroid
9h72.00	Excepted from thyroid quality indicators: Informed dissent
C023100	Toxic nodular goitre unspecified with crisis
SL27z00	Thyroid hormone and thyroid derivative poisoning NOS
C021.00	Toxic uninodular goitre
C023.00	Toxic nodular goitre unspecified
C053.11	Riedel's thyroiditis
R145000	[D]Thyroid scan abnormal
C021z00	Toxic uninodular goitre NOS
TJ28.00	Adverse reaction to antithyroid agents
R145z00	[D]Thyroid function tests abnormal NOS

Continued on next page

Read code	Detail
C04y.00	Other acquired hypothyroidism
66B9.00	Thyroid dis.treatment started
C04..13	Hypothyroidism
C04..12	Thyroid deficiency
C04..11	Myxoedema
L181100	Thyroid dysfunction during pregnancy - baby delivered
C04z.13	Hypothyroid goitre, acquired
C04z.11	Pretibial myxoedema - hypothyroid
C01z.00	Nontoxic nodular goitre NOS
Q433700	Neonatal jaundice with congenital hypothyroidism
9Oj0.00	Hypothyroidism monitoring first letter
9Oj2.00	Hypothyroidism monitoring third letter
66B2.00	Follow-up thyroid assessment
C02..11	Hyperthyroidism
C02..12	Toxic goitre
66B6.00	Thyroid drug side effects
C00z.11	Thyroid enlargement
7111000	Excision of substernal thyroid tissue
66BB.00	Hypothyroidism annual review
U602200	[X]Antithyroid drugs caus adverse effects in therapeut use
C040.00	Postsurgical hypothyroidism
7111z00	Operation on aberrant thyroid tissue NOS
1JM..00	Suspected hypothyroidism
C022100	Toxic multinodular goitre with crisis
B7G..11	Adenoma of thyroid gland
C02y.11	Factitia thyrotoxicosis
Cyu1500	[X]Other specified disorders of thyroid
C020000	Toxic diffuse goitre with no crisis
Cyu4J00	[X]Disorders of thyroid gland in diseases CE
N220411	De Quervain's disease
C02z000	Thyrotoxicosis without mention of goitre or cause no crisis
C02zz00	Thyrotoxicosis NOS
8BPG.00	Thyroid stimulating hormone suppression therapy
C054.00	Iatrogenic thyroiditis
22H2.00	O/E - thyroid swelling -unilat
7L1Z400	Oral delivery of radiotherapy for thyroid ablation

Continued on next page

Read code	Detail
711y.00	Thyroid gland or parathyroid gland operations OS
7110600	Thyroidectomy NEC
9N4T.00	DNA - Did not attend hyperthyroidism clinic
7110	Thyroidectomy operations
66BZ.00	Thyroid disease monitoring NOS
C022000	Toxic multinodular goitre with no crisis
PK25000	Aberrant thyroid gland
PK25100	Congenital absence of thyroid gland
BB5f.00	[M]Thyroid adenoma and adenocarcinoma
7110.11	Excision of thyroid gland operations
R145100	[D]Thyroid uptake abnormal
66B7.00	Thyroid-dubious diagn.criteria
44qV000	Congenital hypothyroidism screening, borderline result
C06y.00	Other specified thyroid disorders
143..11	H/O: thyroid disorder
C051.00	Subacute thyroiditis
7110500	Partial thyroidectomy NEC
L181500	Postpartum thyroiditis
C050000	Acute nonsuppurative thyroiditis
TJ27100	Adverse reaction to thyroxine sodium
C000.00	Simple goitre
B924000	Neoplasm of uncertain behaviour of thyroid gland
U602113	[X] Adverse reaction to levothyroxine sodium
7111	Operations on aberrant thyroid tissue
C041.00	Other postablative hypothyroidism
7111100	Excision of sublingual thyroid tissue
66B9.11	Thyroxine Rx started
C05..00	Thyroiditis
7110y00	Other specified thyroidectomy
SL28.00	Antithyroid agent poisoning
ZV10y15	[V]Personal history of malignant neoplasm of thyroid
7110111	Bilateral subtotal thyroidectomy
U602211	[X] Adverse reaction to antithyroid agents
C040.11	Post ablative hypothyroidism
442C.00	Thyroid horm tests borderline
B7G..00	Benign neoplasm of thyroid gland

Continued on next page

Read code	Detail
C02y.00	Thyrotoxicosis of other specified origin
9O39.00	Thyroid monitoring call
C042.00	Iodine hypothyroidism
L181.00	Thyroid dysfunction in pregnancy/childbirth/puerperium
22H4.00	O/E - thyroid lump
C043.00	Other iatrogenic hypothyroidism
7110100	Subtotal thyroidectomy
7113	Other operations on thyroid gland
1JM0.00	Suspected congenital hypothyroidism
F381400	Myasthenic syndrome due to hypothyroidism
L181300	Thyroid dysfunction in pregnancy - baby not yet delivered
9h7..00	Exception reporting: thyroid quality indicators
F11x500	Cerebral degeneration due to myxoedema
TJ27z00	Adverse reaction to thyroid and thyroid derivatives NOS
C020z00	Toxic diffuse goitre NOS
C022z00	Toxic multinodular goitre NOS
C01..00	Nontoxic nodular goitre
SL28z00	Antithyroid agent poisoning NOS
C02y300	Thyroid crisis
C024.00	Thyrotoxicosis from ectopic thyroid nodule
C04..00	Acquired hypothyroidism
C050100	Acute suppurative thyroiditis

Table B.7: Table detailing thyroid dysfunction Read code list

B.2.8 Ethnicity, deprivation, smoking, alcohol, weight, and height codes

Table B.8 contains the information required (data source, entity type/enttype [required for data contained in the additional CPRD files], data column) to identify the ethnicity, deprivation (IMD), smoking, alcohol, weight, and height data used in this thesis. A description of how height and weight data were cleaned and how BMI was calculated is given in Section B.2.9.

The records of smoking use and alcohol use that were recorded closest to the index date (but before the index date) were used.

Variable	Source	enttype	Data column
Ethnicity	HES	n/a	`gen_ethnicity`
Deprivation (IMD)	IMD	n/a	`imd2015_5`
Smoking	Additional CPRD files	4	`data1`
Alcohol	Additional CPRD files	5	`data1`
Weight	Additional CPRD files	13	`data1`
Height	Additional CPRD files	14	`data1`

Table B.8: Description of sources of data for ethnicity, deprivation, smoking, alcohol, weight, and height

B.2.9 Cleaning of height and weight data, and calculating BMI

All height values less than or equal to zero or possible missing values (“.” or “999”) were removed. (Note that CPRD height data should be entered in metres.) If the height value was between 3 and 7 then it was assumed that the data had been entered in feet and the values were converted to metres. Height values exceeding 2.3 m or less than 0.8 m were excluded. The height measurement recorded closest to the index date was used.

Again, all weight values less than or equal to zero or possible missing values (“.” or “999”) were removed. (Note that CPRD weight data should be entered in kilograms.) Weight values exceeding 450 kg were excluded. If weight values were less than 40 then it was assumed that the weight had been recorded in stones and the values were converted to kg. If the weight was still less than 40 then it was assumed that the entry was incorrect and the value was removed. The weight value that was used for the confounder weight value was the measurement closest

to the index date (but before the index date). The weight value that was used for the mediator weight value was the largest weight value post-index but before the individual was censored or developed the outcome of interest.

Confounder BMI values were calculated using the height and confounder weight values described in the previous two paragraphs (BMI = weight (kg) $\div$ [height (m)]^2). Patients were classed as obese if they had a BMI exceeding 30 but less than 250. Patients were classed as not obese if they had a BMI less than 30 but more than 10. Otherwise patients BMI values were classed as missing.

Mediator BMI values were calculated using the same height value as above but with the mediator weight value instead of the confounder weight value. Patients were said to have developed obesity during follow up if they were not classed as being obese pre-index and they had a mediator BMI value exceeding 30 but less than 250. (Note that the mediator was defined as 'number of metabolic factors developed during follow up' so it is important to ensure that the patient was not already classed as being obese pre-index.)

B.3 Collapsing categorical variables

Due to some factor levels of categorical variables having low cell counts, some levels were collapsed. Below are the original and collapsed factor levels for categorical variables.

Ethnicity original: Bangladeshi, black African, black Caribbean, black – other, Chinese, Indian, mixed, other Asian, other, Pakistani, missing, white.
Ethnicity after collapsing levels: white, not white, missing.

Alcohol original: currently drink alcohol, never drunk alcohol, have quit drinking alcohol, missing.
Alcohol after collapsing levels: do drink/have previously drunk alcohol, never drunk alcohol, missing.

All other categorical variables were left with their original categories. The categories for each variable are given below:

Gender: male, female.
Deprivation: least deprived IMD quintile, 2nd least deprived quintile, 3rd least deprived quintile, 4th least deprived quintile, most deprived quintile.[1]
Smoking: current smoker, ex smoker, never smoked, missing.
Thyroid dysfunction: diagnosed, not diagnosed.
Hyperlipidaemia: diagnosed, not diagnosed.
Hypertension: diagnosed, not diagnosed.
Obesity: obese, not obese, missing.

[1]The small proportion of people with missing deprivation data were excluded since there was no logical way to assign them to a group.

B.4 R code

B.4.1 Causal mediation analysis model

Firstly, the weights for the mediator need to be created. To achieve this, an ordinal logistic regression model is fitted for the mediator conditional on the exposure and confounders. Since there is a need to be able to use both the observed and counterfactual value of the exposure in the analysis, the variable `dm_index_temp` is created and will be used later in the code.

```
library("VGAM")
dta$dm_index_temp=dta$dm_index
fitM = vglm(ordered(NoNewMetSFactorsMediator) ~ factor(
   dm_index_temp) + age_index + factor(gender) + factor(alcohol
   ) + factor(smoking) + factor(thyroid_confounder) + factor(
   imd2015_5) + factor(ethnicity) + factor(
   hypertension_confounder) + factor(hyperlipidaemia_confounder
   ) + factor(obesityconfounder), family=propodds, id=dta$pair
   , data=dta)
```

Next, an ID variable is created so each copy of an individual can be matched to its counterfactual version. The variable `dm_indexStar` is also created to act as the variable X_i^* from Section 3.7.2. Also, the dataset is replicated with `dm_indexStar` being set to `dm_index` in one dataset and being set to `1-dm_index` in the other dataset.

```
N <- nrow(dta)
dta$id <- 1:N
data1 <- dta
data2 <- dta

data1$dm_indexStar <- dta$dm_index
data2$dm_indexStar <- 1-dta$dm_index
newMyData <- rbind(data1, data2)
```

Next, the weights are computed using the predicted values from the mediator model `fitM`. This is done once with exposure equal to `dm_index` (which allows the estimation of the denominator in Equation 3.9) and then repeated with exposure equal to `dm_indexStar` (which allows the estimation of the numerator in Equation 3.9).

```
newMyData$NoNewMetSFactorsMediator <-
   newMyData$NoNewMetSFactorsMediator + 1 #the mediator needs
   to be coded so that it starts from one

newMyData$dm_index_temp <- newMyData$dm_index
tempDir <- as.matrix(predict(fitM,type = "response", newdata=
   newMyData))[cbind(1:(2*N),newMyData$NoNewMetSFactorsMediator
   )]

newMyData$dm_index_temp <- newMyData$dm_indexStar
tempIndir <- as.matrix(predict(fitM,type = "response", newdata=
   newMyData))[cbind(1:(2*N),newMyData$NoNewMetSFactorsMediator
   )]

newMyData$weightM <- tempIndir/tempDir
```

Similarly, the stabilised weights for the exposure can be created using logistic regression.

```
dta$dm_index_temp=dta$dm_index

fitAdenom = geeglm(dm_index ~ age_index + factor(gender) +
   factor(alcohol) + factor(smoking) + factor(
   thyroid_confounder) + factor(imd2015_5) + factor(ethnicity)
   + factor(hypertension_confounder) + factor(
   hyperlipidaemia_confounder) + factor(obesityconfounder),
   family=binomial(), data=dta, id=dta$pair)
```

```
fitAnumer = glm(factor(dm_index) ~ 1, family=binomial(), data=
    dta)
newMyData$weightAnumer <- as.matrix(predict(fitAnumer,type = "
    response", newdata=newMyData))
newMyData$weightAdenom <- as.matrix(predict(fitAdenom,type = "
    response", newdata=newMyData))
newMyData$weightA = newMyData$weightAnumer/
    newMyData$weightAdenom
```

When conducting the sensitivity analysis with truncated weights, the following code was run at this stage.

```
weighttrunclimit=quantile(newMyData$weightA,probs=seq(0,1,0.95)
    )
newMyData$weightA[newMyData$weightA>weighttrunclimit]=
    weighttrunclimit
```

Then the final weights W_i can be made by multiplying the exposure weights W_i^X and mediator weights W_i^M together.

```
newMyData$weightAM = as.vector(newMyData$weightA *
    newMyData$weightM)
```

Now the Cox model can be run with the weights applied.

```
Cox=coxph(Surv(time_to_event_outcome, frozen_shoulder) ~ factor
    (dm_index) + factor(dm_indexStar), data=newMyData, weights =
     newMyData$weightAM)
```

The total effect, direct effect, indirect effect and proportion mediated can be extracted using

```
TE = exp(sum(coef(Cox)[c('factor(dm_index)1','factor(
    dm_indexStar)1')]))
DE = exp(unname(coef(Cox)['factor(dm_index)1']))
IE = exp(unname(coef(Cox)['factor(dm_indexStar)1']))
```

```
PM = log(IE) / log(TE)
```

To estimate bootstrap confidence intervals, the above code needs to be packed into the function `doEffectDecomp`

```
doEffectDecomp = function(dta)
{   dta$dm_index_temp=dta$dm_index
fitM = vglm(ordered(NoNewMetSFactorsMediator) ~ factor(
   dm_index_temp) + age_index + factor(gender) + factor(alcohol
   ) + factor(smoking) + factor(thyroid_confounder) + factor(
   imd2015_5) + factor(ethnicity) + factor(
   hypertension_confounder) + factor(hyperlipidaemia_confounder
   ) + factor(obesityconfounder), family=propodds, id=dta$pair,
    data=dta)

N <- nrow(dta)
dta$id <- 1:N
data1 <- dta
data2 <- dta
data2$pair <- data2$pair + max(dta$pair) #maybe not needed?
data1$dm_indexStar <- dta$dm_index
data2$dm_indexStar <- 1-dta$dm_index
newMyData <- rbind(data1, data2)
newMyData$NoNewMetSFactorsMediator <-
   newMyData$NoNewMetSFactorsMediator + 1

newMyData$dm_index_temp <- newMyData$dm_index
tempDir <- as.matrix(predict(fitM,type = "response",
newdata=newMyData))[cbind(1:(2*N),
   newMyData$NoNewMetSFactorsMediator)]
newMyData$dm_index_temp <- newMyData$dm_indexStar
```

```
tempIndir <- as.matrix(predict(fitM,type = "response",
newdata=newMyData))[cbind(1:(2*N),
   newMyData$NoNewMetSFactorsMediator)]

newMyData$weightM <- tempIndir/tempDir

fitAdenom = geeglm(dm_index ~ age_index + factor(gender) +
   factor(alcohol) + factor(smoking) + factor(
   thyroid_confounder) + factor(imd2015_5) + factor(ethnicity)
   + + factor(hypertension_confounder) + factor(
   hyperlipidaemia_confounder) + factor(obesityconfounder),
   family=binomial(), data=dta, id=dta$pair)
fitAnumer = glm(factor(dm_index) ~ 1, family=binomial(), data=
   dta)
newMyData$weightAnumer <- as.matrix(predict(fitAnumer,type ="
   response", newdata=newMyData))
newMyData$weightAdenom <- as.matrix(predict(fitAdenom,type ="
   response", newdata=newMyData))
newMyData$weightA = newMyData$weightAnumer/
   newMyData$weightAdenom

#remove the hash from the two lines below to run the analysis
   with truncated weights
#weighttrunclimit=quantile(newMyData$weightA,probs=seq
   (0,1,0.95))
#newMyData$weightA[newMyData$weightA>weighttrunclimit]=
   weighttrunclimit

newMyData$weightAM = as.vector(newMyData$weightA *
```

```
   newMyData$weightM)

Cox=coxph(Surv(time_to_event_outcome, frozen_shoulder) ~ factor
   (dm_index) + factor(dm_indexStar) , data=newMyData, weights
   = newMyData$weightAM)
TE = exp(sum(coef(Cox)[c('factor(dm_index)1','factor(
   dm_indexStar)1')]))
DE = exp(unname(coef(Cox)['factor(dm_index)1']))
IE = exp(unname(coef(Cox)['factor(dm_indexStar)1']))
PM = log(IE) / log(TE)
return(c(exp(coef(Cox)), TE=TE, DE=DE, IE=IE, PM=PM))
}
```

Bootstrap resampling can be achieved with the function `Samp` which draws matching pairs from the original dataset (drawing matching pairs preserves the age-, gender-, and practice-matching) with replacement and returns a new dataset. This is repeated 500 times, and for each of the 500 resampled datasets, the function `doEffectDecomp` is executed to return the estimates of the total effect, direct effect, indirect effect and proportion mediated. The 2.5th and 97.5th percentiles for each of the total effect, direct effect, indirect effect and proportion mediated can be extracted to provide 95% CIs for the total effect, direct effect, indirect effect and proportion mediated, respectively.

```
Samp = function(dta)
{
s = sample(unique(dta$pair), replace=TRUE)
return(do.call('rbind', lapply(s, function(x) dta[dta$pair == x
   , ])))
}

HRs = replicate(500, doEffectDecomp(Samp(dta)))
apply(HRs, 1, quantile, c(0.025, 0.975))
```

Note: the proportion mediated was not reported in the results since the indirect was slightly negative and the total effect was positive; thus, the proportion mediated has no sensible meaning.

Note: the code used to conduct this analysis was adapted from Rochon et al. [341].

B.4.2 Schoenfeld residual plots

The following code was used to create the Schoenfeld residual plots for the model `Cox` that was created in Section B.4.1.

```
library("survminer", "survival")
prophaz <- cox.zph(Cox)
ggcoxzph(prophaz, xlab="Time (years; recorded on log scale)",
   ylab="Scaled Schoenfeld residuals", resid=TRUE)
```

B.4.3 Kaplan-Meier Plot

To create a Kaplan-Meier plot for the patients with type 2 diabetes and the patients without diabetes in R the `survminer` package was used. The following code may be used to create a Kaplan-Meier plot using survival time variable `time_to_event_outcome`, event indicator `frozen_shoulder`, and the variable `dm_index` to indicate whether the patient had type 2 diabetes on the index date. The option `risk.table=TRUE` instructs R to include the percentage at risk table `ggsurv$table`.

```
library("survminer", "survival")
sfit <- sfit <- survfit(Surv(time_to_event_outcome,
   frozen_shoulder) ~ factor(dm_index) , data = dta)
ggsurv=ggsurvplot(sfit, data = dta, ylim=c(0.9,1), censor=FALSE
   , risk.table=TRUE, conf.int=TRUE, legend.labs=c("No diabetes
   ", "Type 2 diabetes")) + xlab("Time (years)")
ggsurv$table=ggsurvtable(sfit,data = dta, survtable = c("risk.
   table"), risk.table.type = c("percentage"), xlim=c(0,15),
```

```
   break.time.by=5, legend.labs=c("No diabetes", "Type 2
   diabetes")) + xlab("Time (years)")
ggsurv
```

B.5 Weight distributions

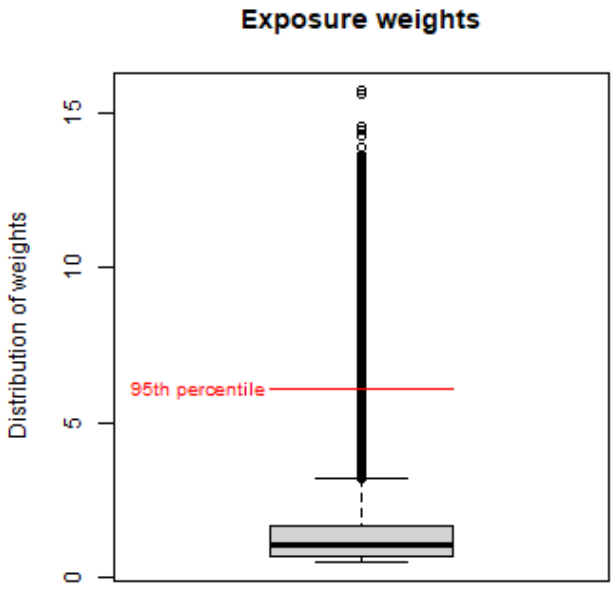

Figure B.1: Box plot of exposure weights, W_i^X

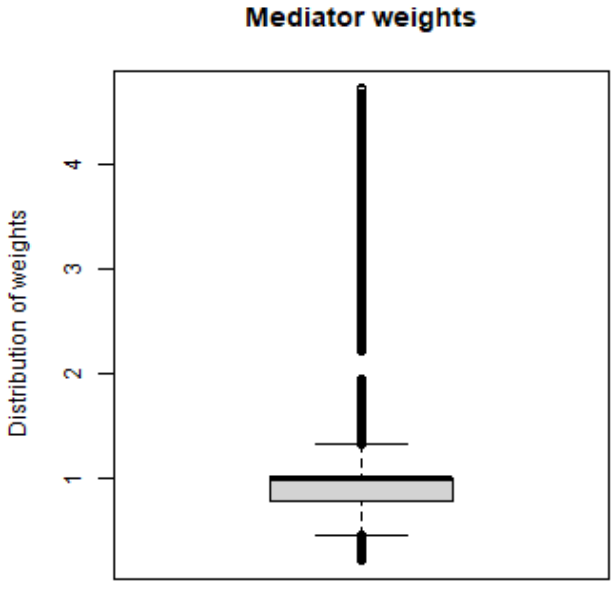

Figure B.2: Box plot of mediator weights, W_i^M

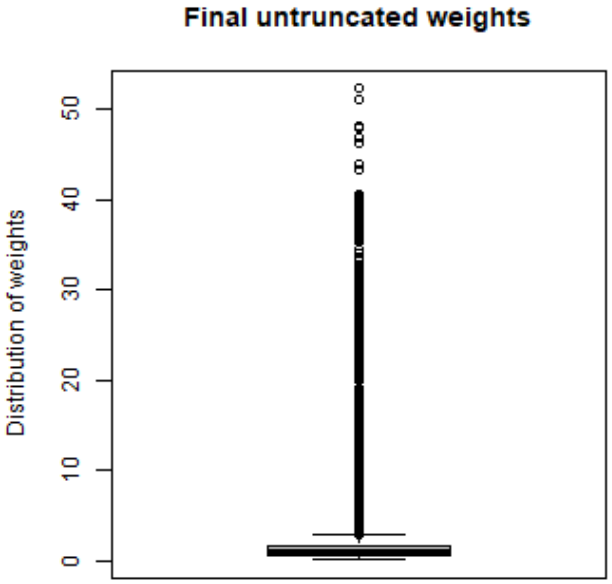

Figure B.3: Box plot of final untruncated weights, W_i

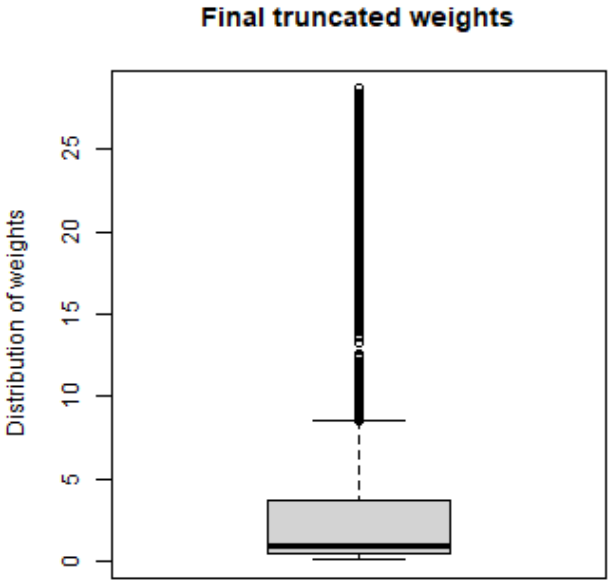

Figure B.4: Box plot of final truncated weights, $W_{i,trunc}$

Appendix C

Are patients with newly diagnosed frozen shoulder more likely to be diagnosed with type 2 diabetes? A cohort study

C.1 STROBE checklist

The following pages contain a completed STROBE checklist corresponding to the work presented in Chapter 5.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Pg 99
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	n/a
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Sec 4.1, Chapter 1
Objectives	3	State specific objectives, including any prespecified hypotheses	Section 4.2
Methods			
Study design	4	Present key elements of study design early in the paper	Sec 4.3.1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Sec 4.3.1
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Sec 4.3.1
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	Sec 4.3.1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Sec 4.3.1, Sec 4.3.2, Sec B.2, Sec B.3
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Sec 4.3.1, Sec B.2, Sec B.3
Bias	9	Describe any efforts to address potential sources of bias	Sec 4.3.2
Study size	10	Explain how the study size was arrived at	Sec 4.3.1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Sec 4.3.2, Sec B.3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Sec 4.3.2
		(*b*) Describe any methods used to examine subgroups and interactions	Sec n/a
		(*c*) Explain how missing data were addressed	Sec 4.3.2
		(*d*) If applicable, explain how loss to follow-up was addressed	Sec 4.3.2
		(*e*) Describe any sensitivity analyses	Sec 4.3.2
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Sec 4.4.1, Fig 4.1
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Sec 4.4.1, Tab 4.1
		(b) Indicate number of participants with missing data for each variable of interest	Tab 4.1
		(c) Summarise follow-up time (eg, average and total amount)	Sec 4.4.1, Sec

			4.4.2
Outcome data	15*	Report numbers of outcome events or summary measures over time	Sec 4.4.2, Fig 4.1
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Sec 4.4.2
		(*b*) Report category boundaries when continuous variables were categorized	Sec B.2.8
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Sec 4.4.2
Discussion			
Key results	18	Summarise key results with reference to study objectives	Sec 4.5
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Sec 4.5
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Sec 4.5
Generalisability	21	Discuss the generalisability (external validity) of the study results	Sec 4.5
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Pg v

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

C.2 R code

C.2.1 Kaplan-Meier Plot

To create a Kaplan-Meier plot for the frozen shoulder group and control group in R the `survminer` package was used. The following code may be used to create a Kaplan-Meier plot using survival time variable `time_to_event_outcome`, event indicator `T2dm_post_index`, and the variable `dta.case` to indicate whether the patient was in the frozen shoulder group or control group. The option `risk.table=TRUE` instructs R to include the percentage at risk table `ggsurv$table`.

```
library("survminer", "survival")
sfit <- survfit(Surv(time_to_event_outcome, T2dm_post_index) ~
   dta.case, data=dta)
ggsurv=ggsurvplot(sfit, data = dta, ylim=c(0.9,1), censor=FALSE
   , risk.table=TRUE, conf.int=TRUE, legend.labs=c("No frozen
   shoulder", "Frozen shoulder")) + xlab("Time (years)")
ggsurv$table=ggsurvtable(sfit,data = dta, survtable = c("risk.
   table"), risk.table.type = c("percentage"), xlim=c(0,15),
   break.time.by=5, legend.labs=c("No frozen shoulder", "Frozen
    shoulder")) + xlab("Time (years)")
ggsurv
```

C.2.2 Cox proportional hazards model

The age- and gender-adjusted Cox model was run using the `coxph` function from the `survival` package. The term `frailty(dta.pair)` introduces random effects terms to the Cox model to account for shared frailty between matching pairs which are identified by the `dta.pair` variable.

```
library("survival")
Cox=coxph(Surv(time_to_event_outcome, T2dm_post_index) ~ dta.
   case + + dta.gender + indexage + frailty(dta.pair), data=dta
```

```
   )
summary(Cox)
```

The fully adjusted model was run using the code:

```
library("survival")
Cox2=coxph(Surv(time_to_event_outcome, T2dm_post_index) ~ dta.
   case + dta.gender + indexage + NumberConsultationsPerYear +
   dta.imd2015_5 + dta.ethnicity + dta.obesity + dta.
   hyperlipidaemia + dta.hypertension + dta.thyroid + frailty(
   dta.pair) , data=dta)
summary(Cox2)
```

C.2.3 Schoenfeld residual plots

The following code was used to create the Schoenfeld residual plot for the model `Cox` that was created in Section C.2.2.

```
library("survminer", "survival")
prophaz <- cox.zph(Cox)
ggcoxzph(prophaz, xlab="Time (years; recorded on log scale)",
   ylab="Scaled Schoenfeld residuals", resid=TRUE, title="Age-
   and gender-adjusted model")
```

Appendix D

Diabetes as a prognostic factor in frozen shoulder: a systematic review

D.1 PRISMA checklist

The following pages contain a completed PRISMA checklist corresponding to the work presented in Chapter 6.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Pg 111
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	n/a
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Sec 5.1
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Sec 5.1.1
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Sec 5.2.1, Tab 5.1, Tab 5.1
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Sec 5.2.2, Sec 2.2.2.1, Sec 2.2.2.2, Sec A.2
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Sec A.2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Sec 5.2.2.2
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Sec 5.2.3
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Tab 5.1, Sec 5.2.3, Sec 5.2.5
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Sec 5.2.3
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Sec 5.2.4
Effect	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or	Sec 5.2.5

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
measures		presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Sec 5.2.5
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Sec 5.2.5
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Sec 5.2.5
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Sec 5.2.5
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Sec 5.2.5
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	n/a
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	n/a
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Sec 5.2.5
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Sec 5.3.1, Fig 5.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Reasons given in Fig 5.1
Study characteristics	17	Cite each included study and present its characteristics.	Sec 5.3.2, Tab 5.3
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Sec 5.3.3, Fig 5.2, Tab 5.4
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Tab D.1-Tab D.5
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Tab 5.3
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary	Sec 5.3.4,

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
		estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Fig 5.3-5.5
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Tab 5.6, Sec 5.3.4, Fig 5.3-5.5
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	n/a
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	n/a
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Tab 5.5, Tab 5.6, Sec 5.3.4.2
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Sec 5.4
	23b	Discuss any limitations of the evidence included in the review.	Sec 5.4
	23c	Discuss any limitations of the review processes used.	Sec 5.4
	23d	Discuss implications of the results for practice, policy, and future research.	Sec 5.4, Sec 5.5
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Pg 111
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Pg 111
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Amendments can be seen by following the hyperlink on pg 111
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Pg v
Competing interests	26	Declare any competing interests of review authors.	Pg i
Availability of data, code	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the	Data extracted:

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
and other materials		review.	Tab 2.3, Tab 2.4, Sec A.4. Stata code: Sec A.3. Data collection spreadsheets available on request.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71
For more information, visit: http://www.prisma-statement.org/

D.2 Results for diabetes as a prognostic factor in frozen shoulder

This appendix section contains tables which include summaries of results from studies investigating the association between diabetes and outcomes of follow-up. Tables D.1–D.4 include result summaries from studies reporting ROM results, pain scores, multi-dimensional clinical scores, and function and disability results, respectively. Results from studies reporting less common outcomes (<4 studies) are reported in Table D.5.

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
G. P. Nicholson, 2003	Diabetes group: 8 shoulders; Non-diabetes group: 17 shoulders	Pre-capsular release and at a mean 3 years post-arthroscopic capsular release (range 2-8 years)	**Active forward elevation**		High
			Pre-treatment mean:	Follow-up mean:	
			Diabetes: 85,	Diabetes: 154,	
			Non-diabetes: 83.	Non-diabetes: 170.	
			Active external rotation		
			Pre-treatment mean:	Follow-up mean:	
			Diabetes: 12,	Diabetes: 45	
			Non-diabetes: 10.	Non-diabetes: 57.	
			Active internal rotation (hand behind back)		
			Pre-treatment median:	Follow-up median:	
			Diabetes: Trochanter,	Diabetes: T12,	
			Non-diabetes: Buttock.	Non-diabetes: T10.	

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
G. L. Cvetanovich et al., 2018	Diabetes group: 8 shoulders; Non-diabetes group: 19 shoulders	Pre-capsular release and at an average 3.7 years post-capsular release (range 2-6 years)	**Forward flexion** Pre-treatment mean: Diabetes: 116.4 (SD=20.1), Non-diabetes: 114.3 (SD=23.1), p=0.80. **External rotation** Pre-treatment mean: Diabetes:30.9 (SD=14.5), Non-diabetes: 26.7 (SD=17.3), p=0.47.	Follow-up mean: Diabetes: 158.6 (SD=12.3), Non-diabetes: 155 (SD=17.9), p=0.50. Follow-up mean: Diabetes: 57.7 (SD=10.3), Non-diabetes: 56.4 (SD=18.0), p=0.79.	High

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
R. G. E. Clement et al., 2013	Diabetes group: 12 people; Non-diabetes group: 39 people	Pre-hydrodilatation and 1-month post-hydrodilatation	**Flexion**		High
			Pre-treatment mean: Diabetes: 117.1 (95% CI: 104.8-129.4), Non-diabetes: 108.8 (95% CI: 101.3-116.3), p=0.27.	Follow-up mean: Diabetes: 154.1 (95% CI: 142.7-165.4), Non-diabetes: 148.2 (95% CI: 140.1-156.3), p=0.42.	
			Abduction		
			Pre-treatment mean: Diabetes: 81.3 (95% CI: 60.6-101.9), Non-diabetes: 80.0 (95% CI: 70.2-89.8), p=0.92.	Follow-up mean: Diabetes: 141.8 (95% CI: 124.8-158.9), Non-diabetes: 132.6 (95% CI: 121.2-143.9), p=0.39.	
			External rotation		
			Pre-treatment mean: Diabetes: 11.7 (95% CI: 4.2-19.1), Non-diabetes: 12.2 (95% CI: 8.6-15.8), p=0.90.	Follow-up mean: Diabetes: 22.3 (95% CI: 13.2-31.4), Non-diabetes: 33.4 (95% CI: 28.7-38.1), p=0.049.	

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
S. Bell et al., 2003	Diabetes group: 15 people; Non-diabetes group: 94 people	Pre-hydrodilatation and 2 months post-hydrodilatation	**External rotation** Pre-treatment mean: Diabetes: 28, Non-diabetes: 25. **Passive gleno-humeral abduction** Pre-treatment mean: Diabetes: 60, Non-diabetes: 55. **Active forward elevation** Pre-treatment mean: Diabetes: 124, Non-diabetes: 113.	 Follow-up mean: Diabetes: 62, Non-diabetes: 56. Follow-up mean: Diabetes: 80, Non-diabetes: 81. Follow-up mean: Diabetes: 154, Non-diabetes: 152.	High
H. Vastamäki et al., 2013	Diabetes group: 4 people; Non-diabetes group: 11 people	Mean 23.1 years post-MUA (range 19-30 years)	**Flexion** Diabetes mean: 145, Non-diabetes mean: 145. **External Rotation** Diabetes mean: 40, Non-diabetes mean: 44.	**Abduction** Diabetes mean: 139, Non-diabetes mean: 156.	High

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
C-H Cho et al., 2016	Diabetes group: 17 shoulders pre-capsular release and final follow-up, 15 at 3 months, 9 at 6 months, 13 at 12 months; Non-diabetes group: 20 shoulders pre-capsular release, at 3 months and final follow-up, 17 at 6	Pre-capsular release and 3 months, 6 months, 12 months post-capsular release and at a final follow-up of mean 48.4 months (SD=15.8 months)	**Forward flexion** Pre-treatment mean: Diabetes: 90 (SD=23.2), Non-diabetes: 95 (SD=20.6), (SD=11.8), p=0.64. 6 months mean: Diabetes: 152.2 (SD=12.0), Non-diabetes: 161.8 (SD=6.1), p=0.045. Final follow-up mean: Diabetes: 168.8 (SD=4.9), Non-diabetes: 169.5 (SD=2.2), p=0.96. **External rotation** Pre-treatment mean: Diabetes: 15.0 (SD=11.9), Non-diabetes: 15.3 (SD=10.3), p=0.94. 6 months mean: Diabetes: 43.3 (SD=10.0), Non-diabetes: 55.0 (SD=11.3), p=0.021.	3 months mean: Diabetes: 140 (SD=13.1), Non-diabetes: 151.5 p=0.011. 12 months mean: Diabetes: 162.7 (SD=8.3), Non-diabetes: 168.9 (SD=4.5), p=0.06. 3 months mean: Diabetes: 34.3 (SD=13.5), Non-diabetes: 39.8 (SD=11.5), p=0.11. 12 months mean: Diabetes: 57.3 (SD=10.1), Non-diabetes: 63.7 (SD=9.9), p=0.10.	Moderate

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
	months, 15 at 12 months		Final follow-up mean: Diabetes: 65.9 (SD=6.2), Non-diabetes: 65.8 (SD=9.1), p=0.68.		
			Internal rotation (hand behind back)*		
			Pre-treatment mean: Diabetes: 17.2 (SD=1.9), Non-diabetes: 16.4 (SD=1.7), p=0.01.	3 months mean: Diabetes mean: 14.7 (SD=1.0), Non-diabetes: 14.4 (SD=1.5), p=0.42.	
			6 months mean: Diabetes: 13.7 (SD=0.9), Non-diabetes: 11.7 (SD=2.1), p=0.006	12 months mean: Diabetes: 12.2 (SD=2.6), Non-diabetes: 9.9 (SD=2.5), p=0.041.	
			Final follow-up mean: Diabetes: 9.8 (SD=3.1), Non-diabetes: 9.0 (SD=2.2), p=0.56.		
			*Cho et al. converted measurements to a continuous scale, with T1-T12 converted to 1-12, L1-L5 to 13-17, sacrum to 18, buttock to 19.		

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
A. Ando et al., 2018	Diabetes group: 10 shoulders; Non-diabetes group: 42 shoulders	Pre-treatment and at a mean follow-up of 4.8 years (SD=3.5 years) for the diabetes group, and at mean 5.1 years (SD=2.4 years) follow-up for the non-diabetes group.	**Forward flexion** Pre-treatment mean: Diabetes: 108.0, Non-diabetes: 104.9, p=0.43. **Abduction** Pre-treatment mean: Diabetes: 94.5, Non-diabetes: 90.4, p=0.34. **External rotation** Pre-treatment mean: Diabetes: 24, Non-diabetes: 20, p=0.37. **Internal rotation (hand behind back)** Pre-treatment median: Diabetes: Sacrum, Non-diabetes: Sacrum, p=0.87.	Follow-up mean: Diabetes: 140.5, Non-diabetes: 158.0, p=0.002. Follow-up mean: Diabetes: 121.5, Non-diabetes: 158.3, p=0.001. Follow-up mean: Diabetes: 41.0, Non-diabetes: 51.9, p=0.004. Follow-up median: Diabetes: T12, Non-diabetes: T8, p<0.001.	High

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
İ. Düzgün et al., 2012	Diabetes group: 12 people; Non-diabetes group: 38 people	Pre-treatment and after treatment protocol averaging 8 weeks.	**Flexion** Pre-treatment mean: Diabetes: 125.3, Non-diabetes: 120.4. **Abduction** Pre-treatment mean: Diabetes: 99.6, Non-diabetes: 91.5. **External rotation** Pre-treatment mean: Diabetes: 39.0, Non-diabetes: 30.8. **Internal rotation (hand behind back)** Pre-treatment mean: Diabetes: 41.9, Non-diabetes: 39.4.	Follow-up mean: Diabetes: 154.7, Non-diabetes: 160.6. Follow-up mean: Diabetes: 154.7, Non-diabetes: 141.1. Follow-up mean: Diabetes: 57.5, Non-diabetes: 61.1. Follow-up mean: Diabetes: 66.8, Non-diabetes: 67.2.	Moderate

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
H. Vastamäki et al., 2016	Diabetes group: 29 shoulders; Non-diabetes group: 169 shoulders	During frozen shoulder and at a follow-up of mean 10 years (SD=8 years) for the diabetes groups, and mean 9.7 years (SD=7 years) for the non-diabetes group.	**Flexion** Pre-treatment mean: Diabetes: 93 (SD=29), Non-diabetes: 101 (SD=18), p=0.20. **Abduction** Pre-treatment mean: Diabetes: 83 (SD=29), Non-diabetes: 86 (SD=24), p=0.60. **External rotation** Pre-treatment mean: Diabetes: 23 (SD=17), Non-diabetes: 22 (SD=16), p=0.88. **Internal rotation (hand behind back)** Pre-treatment median: Diabetes: Below buttock, Non-diabetes: Above buttock, p=0.18.	Follow-up mean: Diabetes: 144 (SD=19), Non-diabetes: 157 (SD=14), p=0.001. Follow-up mean: Diabetes: 154 (SD=36), Non-diabetes: 173 (SD=16), p=0.008. Follow-up mean: Diabetes: 41 (SD=19), Non-diabetes: 53 (SD=14), p=0.002. Follow-up median: Diabetes: LIII, Non-diabetes: ThXII, p<0.001.	High

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
C-H Cho et al., 2020	Diabetes group: 32 shoulders; Non-diabetes group: 110 shoulders	Pre-treatment and 3 weeks, 6 weeks, 12 weeks post-treatment	**Forward flexion** Pre-treatment mean: Diabetes: 118.9 (SD=20.9), Non-diabetes: 112.3 (SD=23.1), (SD=118.3). 6 weeks mean: Diabetes: 151.4 (SD=15.0), Non-diabetes: 157.3 (SD=14.4), (SD=13.6). **Abduction** Pre-treatment mean: Diabetes: 104.7 (SD=24.1), Non-diabetes: 103.3 (SD=24.7), (SD=23.6). 6 weeks mean: Diabetes: 143.1 (SD=19.3), Non-diabetes: 148.9 (SD=21.0), (SD=21.4). **External rotation** Pre-treatment mean: Diabetes: 37.7 (SD=12.0), Non-diabetes: 37.8 (SD=16.8), 6 weeks mean: Diabetes: 58.4 (SD=12.2), Non-diabetes: 65.7 (SD=11.6),	3 weeks mean: Diabetes: 146.7 (SD=17.1), Non-diabetes: 149.3 12 weeks mean: Diabetes:149.5 (SD=21.5), Non-diabetes: 159.3 3 weeks mean: Diabetes: 134.8 (SD=22.5), Non-diabetes: 140.3 12 weeks mean: Diabetes: 139.4 (SD=28.2), Non-diabetes: 149.9 3 weeks mean: Diabetes: 54.4 (SD=11.1), Non-diabetes: 58.3 (SD=13.4). 12 weeks mean: Diabetes: 58.1 (SD=15.1), Non-diabetes: 65.9 (SD=11.4).	High

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
			Internal rotation*		
			Pre-treatment mean:	3 weeks mean:	
			Diabetes: 16.2 (SD=3.0),	Diabetes: 12.3 (SD=2.8),	
			Non-diabetes: 16.5 (SD=2.6),	Non-diabetes: 11.7 (SD=2.7).	
			6 weeks mean:	12 weeks mean:	
			Diabetes: 11.8 (SD=3.0),	Diabetes:11.7 (SD=3.3),	
			Non-diabetes: 10.1 (SD=2.5),	Non-diabetes: 9.7 (SD=2.8).	
			* (T1–T12 maps to 1–12, L1–L5 to 13–17, sacrum to 18; coccyx to 19; and buttocks to 20)		

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Y. W. Ko et al., 2021	Diabetes group: 32 shoulders; Non-diabetes group: 203 shoulders	Pre-treatment and 6 weeks, 3 months post-treatment	**Forward flexion** Pre-treatment mean: Diabetes: 92 (SD=10.7), Non-diabetes: 92 (SD=9.9), 3 months mean: Diabetes: 141 (SD=20.7), Non-diabetes: 151 (SD=22.0). **External rotation** Pre-treatment mean: Diabetes: 15 (SD=3.9), Non-diabetes: 15 (SD=5.9), 3 months mean: Diabetes: 29 (SD=24.1), Non-diabetes: 42 (SD=15.1). **Internal rotation*** Pre-treatment mean: Diabetes: 6 (SD=4.9), Non-diabetes: 7 (SD=4.5), 3 months mean: Diabetes: 24 (SD=16.4), Non-diabetes: 35 (SD=20.7).	6 weeks mean: Diabetes: 130 (SD=17.7), Non-diabetes: 147 (SD=19.8). 6 weeks mean: Diabetes: 24 (SD=15.2), Non-diabetes: 36 (SD=16.1). 6 weeks mean: Diabetes: 21 (SD=14.5), Non-diabetes: 31 (SD=20.7).	Moderate

Continued on next page

Author, Year	Sample size	Follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
G. L. Yanlei et al., 2019	Diabetes group: 32 shoulders; Non-diabetes group: 24 shoulders	Pre-treatment and 12 months post-treatment	**Forward flexion**		High
			Pre-treatment mean:	12 months mean:	
			Diabetes: 83.1,	Diabetes: 123.0,	
			Non-diabetes: 82.5,	Non-diabetes: 132.5.	
			Abduction		
			Pre-treatment mean:	12 months mean:	
			Diabetes: 57.3,	Diabetes: 117.0,	
			Non-diabetes: 60.0,	Non-diabetes: 127.7.	
			External rotation		
			Pre-treatment mean:	12 months mean:	
			Diabetes: 2.31,	Diabetes: 8.06,	
			Non-diabetes: 0.96,	Non-diabetes: 9.36.	
			Internal rotation (hand behind back)		
			Pre-treatment mean:	12 months mean:	
			Diabetes: 2.81,	Diabetes: 6.19,	
			Non-diabetes: 3.17,	Non-diabetes: 8.50.	

Continued on next page

Author, Year	**Sample size**	**Follow-up measurements taken**	**Brief summary of evidence**		**QUIPS risk of bias score**
F. Barbosa et al., 2019	Diabetes group: 46 shoulders; Non-diabetes group: 164 shoulders	Pre-treatment and 3, 6, 12 months post-treatment	**External rotation** Pre-treatment mean: Diabetes: 3.9, Non-diabetes: 14.5. 6 months mean: Diabetes: 25.8, Non-diabetes: 50.6, p=0.017.	3 months mean: Diabetes: 58.7, Non-diabetes: 62.8, p=0.509. 12 months mean: Diabetes: 29.7, Non-diabetes: 48.3, p=0.8.	High

Table D.1: Summary of results for studies reporting results for the association between diabetes and ROM in patients with frozen shoulder

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**		**QUIPS risk of bias score**
R. G. E. Clement et al., 2013	Diabetes group: 12 people; Non-diabetes group: 39 people	Visual Analog Scores (VAS) pre-hydrodilatation and mean 14 months post-hydrodilatation	**Pre-treatment** **Follow-up**	Diabetes mean: 8.1, 95% CI: 7.1–9.1, Non-diabetes mean: 6.8, 95% CI: 6.2 – 7.5, p=0.048. Diabetes mean: 5.4, 95% CI: 3.3–7.5, Non-diabetes mean: 3.0, 95% CI: 1.9-4.1, p=0.065.	High
S. Bell et al., 2003	Diabetes group: 12 people; Diabetes group: 15 people; Non-diabetes group: 94 people	VAS pre-hydrodilatation and 2 months post-hydrodilatation	**Pre-treatment** **Follow-up**	Diabetes: 33% severe pain, 66% moderate pain. Non-diabetes:22% severe pain, 47% moderate pain, 30% mild pain, 1% no pain. Diabetes: 13% severe pain, 20% moderate pain, 33% mild pain, 33% no pain. , Non-diabetes: 0% severe pain, 7% moderate pain, 25% mild pain, 62% no pain.	High
G. P. Nicholson, 2003	Diabetes group: 8 shoulders; Non-diabetes group: 17 shoulders	VAS pre-capsular release and at an average 3 years post-capsular release (range 2-8 years)	**Pre-treatment** **Follow-up**	Diabetes median: 4.5, Non-diabetes median: 7. Diabetes median: 1, Non-diabetes median: 0.	High

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
H. Vastamäki et al., 2013	Diabetes group: 4 people; Non-diabetes group: 11 people	VAS at mean 23.1 years post-MUA (range 19-30 years)	**During exertion**	Diabetes mean: 2.2, Non-diabetes mean: 1.3.	High
			At rest	Diabetes mean: 0.5, Non-diabetes mean: 0.3.	
			At night	Diabetes mean: 0.7, Non-diabetes mean: 0.8. p>0.45 in each case. (Exact p-values not reported)	
H. Vastamäki et al., 2016	Diabetes group: 29 shoulders; Non-diabetes group: 169 shoulders	VAS at final follow-up, which was at 10 years (SD=8 years) for the diabetes group, and 9.7 years (SD=7 years) for the non-diabetes group.	**During exertion**	Diabetes mean: 2.5 (SD=0.9), Non-diabetes mean: 1.4 (SD=2.3), p=0.034.	High
			At rest	Diabetes mean: 0.9 (SD=1.3), Non-diabetes mean: 0.6 (SD=1.3), p=0.40.	
			At night	Diabetes mean: 1.5 (SD=2.2), Non-diabetes mean: 0.9 (SD=1.8), p=0.20.	

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
C-H Cho et al., 2016	Diabetes group: 17 shoulders pre-capsular release and final follow-up, 15 at 3 months, 9 at 6 months, 13 at 12 months; Non-diabetes group: 20 shoulders pre-capsular release, at 3 months and final follow-up, 17 at 6 months, 15 at 12 months.	VAS pre-capsular release and 3 months, 6 months, 12 months post-capsular release and at a final follow-up of mean 48.4 months (SD=15.8 months)	**Pre-treatment**	Diabetes mean: 7.0 (SD=1.8), Non-diabetes mean: 7.4 (SD=1.5), p=0.78.	Moderate
			3 months	Diabetes mean: 3.5 (SD=1.5), Non-diabetes mean: 3.2 (SD=2.1), p=0.56.	
			6 months	Diabetes mean: 2.4 (SD=1.4), Non-diabetes mean: 1.8 (SD=1.5), p=0.24.	
			12 months	Diabetes mean: 2.2 (SD=1.3), Non-diabetes mean: 1.4 (SD=2.0), p=0.11.	
			Final follow-up	Diabetes mean: 0.5 (SD=1.3), Non-diabetes 0.5 (SD=1.0), p=1.00.	

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
A. Ando et al., 2018	Diabetes group: 10 shoulders; Non-diabetes group: 42 shoulders	VAS pre-manipulation under ultrasound-guided brachial plexus block and at follow-up of mean 4.8 years (SD=3.5 years for the diabetes group, and mean 5.1 years (SD=2.4 years) for the non-diabetes group.	**Pre-treatment** **Follow-up**	Diabetes mean: 6.3, Non-diabetes mean: 6.7, p=0.51. Diabetes mean: 2.1, Non-diabetes mean: 0.7, p=0.007.	Moderate
Cho et al., 2020	Diabetes group: 32 shoulders; Non-diabetes group: 110 shoulders	Pre-treatment and 3, 6, 12 weeks post-treatment	**Pre-treatment** **3 weeks** **6 weeks** **12 weeks**	Diabetes mean: 7.7 (SD=1.8), Non-diabetes mean: 7.3 (SD=1.7) Diabetes mean: 2.8 (SD=1.5), Non-diabetes mean: 2.4 (SD=1.3). Diabetes mean: 2.8 (SD=1.6), Non-diabetes mean: 1.9 (SD=1.3). Diabetes mean: 3.3 (SD=1.9), Non-diabetes mean: 2.0 (SD=1.4).	High

Continued on next page

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**		**QUIPS risk of bias score**
Y. W. Ko et al., 2021	Diabetes group: 32 shoulders; Non-diabetes group: 203 shoulders	Pre-treatment and 6 weeks, 3 months post-treatment	**At rest**		Moderate
			Pre-treatment	Diabetes mean: 2 (SD=1.7), Non-diabetes mean: 2 (SD=1.6).	
			6 weeks	Diabetes mean: 1 (SD=1.6), Non-diabetes mean: 1 (SD=1.5).	
			3 months	Diabetes mean: 1 (SD=1.1), Non-diabetes mean: 1 (SD=1.4).	
			During motion		
			Pre-treatment	Diabetes mean: 6 (SD=1.8), Non-diabetes mean: 5 (SD=1.8).	
			6 weeks	Diabetes mean: 3 (SD=2.1), Non-diabetes mean: 3 (SD=1.8).	
			3 months	Diabetes mean: 3 (SD=2.1), Non-diabetes mean: 2 (SD=1.9).	
G. L. Yanlei et al., 2019	Diabetes group: 32 shoulders; Non-diabetes group: 24 shoulders	Pre-treatment and 12 months post-treatment	**Pre-treatment**	Diabetes mean: 6.47, Non-diabetes mean: 6.96.	High
			12 months	Diabetes mean: 2.09, Non-diabetes mean: 1.20.	

Table D.2: Summary of results for studies reporting results for the association between diabetes and pain in patients with frozen shoulder

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Mehta et al., 2014	Diabetes group: 21 people; Non-diabetes group: 21 people	Constant scores pre-treatment, 6 weeks, 6 months and 2 years post-capsular release.	**Pre-treatment**	Diabetes: Mean 36.6 (SD=4.6), Non-diabetes: Mean 38.4 (SD=5.7).	High
			6 weeks	Diabetes: Mean 55.6 (SD=4.7), Non-diabetes: Mean 66.8 (SD=4.5), $p<0.01$.	
			6 months	Diabetes: Mean 67.4 (SD=5.6), Non-diabetes: Mean 79.6 (SD=3.8), $p<0.01$.	
			24 months	Diabetes: Mean 67.4 (SD=5.6), Non-diabetes: Mean 79.6 (SD=3.8), $p<0.01$.	
Çinar et al., 2009	Diabetes group: 15 shoulders; Non-diabetes group: 13 shoulders	Constant scores pre-treatment, and at follow-up, which was at mean 48.5 months for the diabetes group and mean 60.2 months for the non-diabetes group	**Pre-treatment**	Diabetes: Mean 30.4 (SD=6.2), Non-diabetes: Mean 29.6 (SD=5.8), $p>0.05$.	High
			Follow-up	Diabetes: Mean 82.0 (SD=18.2), Non-diabetes: Mean 93.6 (SD=10.2), $p<0.05$.	

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Wang et al., 2010	Diabetes group: 21 shoulders; No-diabetes group: 42 shoulders	Modified constant scores (with the 25 points for assessment of muscle strength excluded) pre-MUA, 3 weeks post-MUA and an average of 95 months (range 18-189 months) post-MUA	**Pre-treatment** **3 weeks** **Final follow-up**	Diabetes: Mean 24.04 (SD=3.7), Non-diabetes: Mean 23.66 (SD=3.36), p=0.34. Diabetes: Mean 55.9 (SD=5.29), Non-diabetes: Mean 55.78 (SD=3.46), p=0.46. Diabetes: Mean 72.14 (SD=4.3), Non-diabetes: Mean 72.38 (SD=4.28), p=0.42.	High
Düzgün et al., 2012	Diabetes group: 12 people; Non-diabetes group: 38 people	Constant scores pre-treatment and after treatment protocol averaging 8 weeks.	**Pre-treatment** **Follow-up**	Diabetes: Mean 42 (SD=11), Non-diabetes: Mean 41 (SD=11), p>0.05. Diabetes: Mean 68 (SD=11), Non-diabetes: Mean 73 (SD=12), p>0.05.	Moderate
Celik et al., 2017	Diabetes group: 12 shoulders; Non-diabetes group: 20 shoulders	Constant scores pre-treatment and at follow-up, which was at mean 49.5 months (range: 24–90 months)	**Pre-treatment** **Follow-up**	Diabetes: Mean 39.24 (SD=5.72), Non-diabetes: Mean 38.23 (SD=4.30), p=0.69. Diabetes: Mean 91.24 (SD=15.99), Non-diabetes: Mean 79.69 (SD=25.85), p=0.066.	High

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Vastamäki et al., 2016	Diabetes group: 29 shoulders; Non-diabetes group: 169 shoulders	Constant scores at final follow-up, which was at 10 years (SD=8 years) for the diabetes group, and 9.7 years (SD=7 years) for the non-diabetes group.	**Diabetes** **Non-diabetes**	Mean 76 (SD=16), Mean 82 (SD=12), p=0.055.	High
Ando et al., 2018	Diabetes group: 10 shoulders; Non-diabetes group: 42 shoulders	Constant scores pre-treatment and at follow-up of mean 4.8 years (SD=3.5 years for the diabetes group, and mean 5.1 years (SD=2.4 years) for the non-diabetes group.	**Pre-treatment** **Follow-up**	Diabetes: Mean 53.1, Non-diabetes: Mean 51.8 p=0.82. Diabetes: Mean 81.5, Non-diabetes: Mean 92.4, p=0.002.	Moderate

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Y. W. Ko et al., 2021	Diabetes group: 32 shoulders; Non-diabetes group: 203 shoulders	Constant scores pre-treatment and 6 weeks, 3 months post-treatment	**Pre-treatment** **6 weeks** **3 months**	Diabetes mean: 42 (SD=9.0), Non-diabetes mean: 44 (SD=8.1). Diabetes mean: 50 (SD=6.1), Non-diabetes mean: 50 (SD=8.2). Diabetes mean: 50 (SD=5.1), Non-diabetes mean: 52 (SD=6.7).	Moderate
G. L. Yanlei et al., 2019	Diabetes group: 32 shoulders; Non-diabetes group: 24 shoulders	Constant scores pre-treatment and 12 months post-treatment	**Pre-treatment** **12 months**	Diabetes mean: 29.2, Non-diabetes mean: 24.4. Diabetes mean: 64.3, Non-diabetes mean: 74.2.	High
Sinha et al., 2017	Diabetes group: 26 people; Non-diabetes group: 90 people	Regression models with outcomes: pre-procedure Oxford Shoulder Score (OSS); improvement in OSS at 4 weeks post-procedure.	Multivariable regression of OSS scores, adjusted for covariates: prior physiotherapy (yes/no), predominant symptom, stage of disease (plateau/worse), prior steroid injections (yes/no), age and symptom duration. **Pre-treatment** Diabetes present: β=1.77, SE=2.38, p=0.46. **4 weeks post-treatment** Diabetes present: β=-2.99, SE=2.71, p=0.27.		Moderate

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence	QUIPS risk of bias score
Theodorides et al., 2014	Diabetes group: 39 people, Non-diabetes group: 256 people	OSS at mean follow-up 28 days post-MUA and at mean follow-up 3.6 years post-MUA (IQR 1.7 – 5.0 years)	Mean change in OSS, adjusted for gender, age, aetiology, symptom duration, follow-up: **Mean Change 28-days post-treatment** Diabetes: 14.44 (95% CI: 12.08 – 16.80), Non-diabetes: 15.19 (95% CI: 14.18 – 16.19), p=0.55. **Mean Change 3.6-years post-treatment** Diabetes: 7.24 (95% CI: 5.11-9.37), Non-diabetes: 8.11 (95% CI: 7.38-8.84), p=0.43. Note: No baseline OSS scores were reported.	Moderate
Lyhne et al., 2018	Diabetes group: 18 people; Non-diabetes group: 75 people	Improvement between pre-procedure and 6-month post-op OSS's.	**Diabetes** Mean improvement 11.5 (95% CI: 6.2-16.4), **Non-Diabetes** Mean improvement 15.8 (95% CI: 13.6-17.9), p=0.09. Note: differences in improvement between groups could be partly due to the diabetes group's better mean pre-operative score (although exact scores were not reported), and hence be due to regression to the mean.	High

Continued on next page

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**		**QUIPS risk of bias score**
Clement et al., 2013	Diabetes group: 12 people; Non-diabetes group: 39 people	OSS pre-hydrodilatation and mean 14 months post-hydrodilatation	**Pre-treatment**	Diabetes: 19.2 (95% CI: 13.9 – 24.5), Non-diabetes: 23.3 (95% CI: 21.0 – 25.6), p=0.182.	High
			Follow-up	Diabetes: 33.9 (95% CI: 26.9 – 40.9), Non-diabetes: 41.1 (95% CI: 37.9 – 44.3), p=0.090.	
F. Barbosa et al., 2019	Diabetes group: 46 shoulders; Non-diabetes group: 164 shoulders	OSS pre-treatment and 3, 6, 12 months post-treatment	**Pre-treatment**	Diabetes mean: 19, Non-diabetes mean: 17.5.	High
			3 months	Diabetes mean: 26.6, Non-diabetes mean: 34.3, $p<0.01$.	
			6 months	Diabetes mean: 30.9, Non-diabetes mean: 28.9, $p<0.01$.	
			12 months	Diabetes mean: 33.4, Non-diabetes mean: 30.0, p=0.603.	

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Lamplot, et al., 2018	Diabetes group: 9 people, Non-diabetes group: 51 people	American Shoulder and Elbow Surgeons shoulder score (ASES) at baseline and a minimum 2-year follow-up (mean, 3.4 years)	**Baseline** **Follow-up**	Diabetes: Mean 33.3 (95% CI: 24.9, 41.7), Non-diabetes: 42.7 (95% CI: 34.5, 50.9), p=0.44. Diabetes: Mean 85.1 (95% CI: 81.3, 88.9), Non-diabetes: 93.3 (95% CI: 89.8, 96.8), p=0.06.	High
Nicholson, 2003	Diabetes group: 8 shoulders; Non-diabetes group: 17 shoulders	ASES pre-capsular release and at an average 3 years post-capsular release (range 2-8 years)	**Pre-treatment** **Post-treatment**	Diabetes: Mean 39.2, Non-diabetes: Mean 36.9. Diabetes: Mean 88, Non-diabetes: Mean 94.4.	High

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence		QUIPS risk of bias score
Cho et al., 2016	Diabetes group: 17 shoulders pre-capsular release and final follow-up, 15 at 3 months, 9 at 6 months, 13 at 12 months; Non-diabetes group: 20 shoulders pre-capsular release, at 3 months and final follow-up, 17 at 6 months, 15 at 12 months	ASES pre-capsular release and 3 months, 6 months, 12 months post-capsular release and at a final follow-up of mean 48.4 months (SD=15.8 months)	**Pre-treatment**	Diabetes: Mean 28.1 (SD=14.9), Non-diabetes: 30.0 (SD=15.4), p=0.66.	Moderate
			3 months	Diabetes: Mean 62.6 (SD=13.9), Non-diabetes: 69.8 (SD=18.4), p=0.20.	
			6 months	Diabetes: Mean 74.8 (SD=14.9), Non-diabetes: 80.6 (SD=14.7), p=0.31.	
			12 months	Diabetes: Mean 77.7 (SD=15.1), Non-diabetes: 88.8 (SD=13.8), p=0.025.	
			Final follow-up	Diabetes: Mean 95.0 (SD=8.2), Non-diabetes: 96.7 (SD=6.1), p=0.48.	

Continued on next page

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**		**QUIPS risk of bias score**
Cho et al., 2020	Diabetes group: 32 shoulders; Non-diabetes group: 110 shoulders	ASES pre-treatment and 3, 6, 12 weeks post-treatment	**Pre-treatment**	Diabetes mean: 33.8 (SD=14.9), Non-diabetes mean: 34.4 (SD=14.3).	High
			3 weeks	Diabetes mean: 71.6 (SD=11.8), Non-diabetes mean: 77.2 (SD=12.2).	
			6 weeks	Diabetes mean: 74.7 (SD=12.8), Non-diabetes mean: 83.1 (SD=11.6).	
			12 weeks	Diabetes mean: 70.8 (SD=16.2), Non-diabetes mean: 82.8 (SD=13.3).	

Table D.3: Summary of results for studies reporting results for the association between diabetes and clinical multi-dimensional scores in patients with frozen shoulder

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**	**QUIPS risk of bias score**
Nicholson, 2003	Diabetes group: 8 shoulders; Non-diabetes group: 17 shoulders	Simple Shoulder Test (SST) pre-capsular release and at an average 3 years post-capsular release (range 2-8 years)	**Pre-treatment:** Diabetes: Median 2.5, Non-diabetes: Median 4. **Follow-up:** Diabetes: Median 9, Non-diabetes: Median 11. Note: a score closer to zero indicates worse function. The SST is scored from 0-12.	High
Rill et al., 2011	Diabetes nonoperative group: 19 patients; Non-diabetes nonoperative group: 49 shoulders. Diabetes surgery group: 9 shoulders; Non-diabetes surgery group: 15 shoulders.	SST measured at follow-up of minimum 2 years, mean 40 months, range 24-68 months.	**After nonoperative treatment:** Diabetes: Mean 8.8 (SD=3.4), Non-diabetes: Mean 10.6 (SD=2.3). **After surgery:** Diabetes: Mean 9.7, Non-diabetes: Mean 9.7. (No measure of spread reported)	High

Continued on next page

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**	**QUIPS risk of bias score**
Lyhne et al., 2018	Diabetes group: 18 people; Non-diabetes group: 75 people	Improvement between pre-procedure and 6-month post-op Visual Quality Scale scores.	**Diabetes:** Mean improvement 39.6, **Non-diabetes:** Mean improvement 44.5, p=0.50. Note: Pre-operative scores and measures of spread were not reported. The Visual Quality Scale is a 0-100 score with 100 indicating higher levels of satisfaction.	High
Lamplot, et al, 2018	Diabetes group: 9 people, Non-diabetes group: 51 people	Shoulder activity scale scores at baseline and a minimum 2-year follow-up (mean, 3.4 years)	**At baseline:** Diabetes: 11.5 (95% CI: 10.5 – 12.5), Non-diabetes: 10.5 (95% CI: 9.4 – 11.6), p=0.84. **Follow-up:** Diabetes: 8.6 (95% CI: 7.2 – 10.0), Non-diabetes: 9.3 (95% CI: 8.2 – 10.4), p=0.44. Note: The shoulder activity scale is a 0-20 scale with a higher score representing higher shoulder activity levels.	High

Table D.4: Summary of results for studies reporting results for the association between diabetes and function and disability outcome measures in patients with frozen shoulder

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence	QUIPS risk of bias score
W. N. Levine, et al., 2007	Diabetes group: 19 shoulders, Non-diabetes group: 86 shoulders.	Requiring surgery after a nonoperative treatment programme averaging 4.7 months (range 0.2-43.9 months).	Percentage of shoulders requiring operative management: Diabetes group: 10.5%, Non-diabetes group: 10.5%.	High
K. Kingston et al., 2018	Diabetes group: 572 patients, Non-diabetes group: 1618 patients.	Requiring surgery, follow-up not reported.	Percentage of shoulders requiring operative management: Diabetes group: 14.0%, Non-diabetes group: 17.4%.	High
F. Barbosa et al., 2019	Diabetes group: 46 shoulders; Non-diabetes group: 164 shoulders	Requiring surgery, follow-up unclear	Percentage of shoulders requiring operative management: Diabetes group: 70%, Non-diabetes group: 44%.	High
Gundtoft et al., 2020	Diabetes group: 34 shoulders; Non-diabetes group: 201 shoulders	Requiring arthroscopic capsular release, 2 year follow-up	Percentage of shoulders requiring arthroscopic capsular release: Diabetes group: 14.7%, Non-diabetes group: 5.5%, $p<0.05$.	High

Continued on next page

Author, Year	Sample size	Measurement system used & follow-up measurements taken	Brief summary of evidence	QUIPS risk of bias score
J. D. Lamplot et al., 2018	Diabetes group: 9 people, Non-diabetes group: 51 people	i) A diagnosis of contralateral frozen shoulder. ii) Requiring a second glenohumeral joint injection. Follow-up was a minimum of 2-years (mean, 3.4 years)	Percentage of shoulders developing contralateral frozen shoulder: Diabetes group: 77.8%, Non-diabetes group: 29.4%, p=0.009. Percentage of patients requiring a second glenohumeral joint injection: Diabetes group: 55.6%, Non-diabetes group: 29.4%, p=0.15.	High
D. A. Woods et al., 2017.	Diabetes group: 96 shoulders (56 type 1, 40 type 2), Non-diabetes group: 696 shoulders.	Requiring a second MUA. Follow-up length unclear.	Percentage of shoulders requiring a second MUA: Type 1 diabetes: 37.9%, Type 2 diabetes: 25%, Non-diabetes: 15.7%.	High
E. F. Jenkins et al., 2012	Diabetes group: 39 shoulders, Non-diabetes group: 274 shoulders.	Requiring a second MUA. Follow-up length unclear.	Percentage of shoulders requiring a second MUA: Diabetes: 36%, Non-diabetes: 15%.	High

Continued on next page

Author, Year	**Sample size**	**Measurement system used & follow-up measurements taken**	**Brief summary of evidence**	**QUIPS risk of bias score**
A. Ando et al., 2013	Diabetes group: 61 shoulders, Non-diabetes group: 356 shoulders.	Failure to recover from frozen shoulder. 30-month follow-up period.	Cox proportional hazards model adjusted for age, gender, onset to visit time interval, external rotation; with outcome defined as recovery from frozen shoulder. Diabetes group hazard ratio (compare to the non-diabetes group, defined as the reference category) was 0.54 (95% CI: 0.36-0.96), p=0.007.	Moderate

Table D.5: Summary of results for studies reporting results for the association between diabetes and less common outcome measures in patients with frozen shoulder

D.3 Peer-reviewed journal publication

The following pages contain the peer-reviewed journal publication corresponding to the work presented in Chapter 6.

Archives of Rehabilitation Research and Clinical Translation (2021) 3, 100141

ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Rehabilitation Research and Clinical Translation

Archives of Rehabilitation Research and Clinical Translation 2021;3: 100141

Available online at www.sciencedirect.com

Systematic Review

Diabetes as a Prognostic Factor in Frozen Shoulder: A Systematic Review

Brett P. Dyer, MSc , Claire Burton, PhD, Trishna Rathod-Mistry, PhD, Milisa Blagojevic-Bucknall, PhD, Danielle A. van der Windt, PhD

Primary Care Centre Versus Arthritis, School of Medicine, Keele University, Staffordshire, United Kingdom

KEYWORDS
Adhesive capsulitis;
Diabetes;
Frozen shoulder;
Prognosis;
Rehabilitation

Abstract *Objective:* To summarize evidence from longitudinal observational studies to determine whether diabetes (types 1 and 2) is associated with the course of symptoms in people with frozen shoulder.
Data Sources: A systematic literature search of 11 bibliographic databases (published through June 2021), reference screening, and emailing professional contacts.
Study Selection: Studies were selected if they had a longitudinal observational design that included people diagnosed with frozen shoulder at baseline and compared outcomes at follow-up (>2wk) among those with and without diabetes at baseline.
Data Extraction: Data extraction was completed by 1 reviewer using a predefined extraction sheet and was checked by another reviewer. Two reviewers independently judged risk of bias using the Quality in Prognostic Factor Studies tool.
Data Synthesis: A narrative synthesis, including inspection of forest plots and use of the prognostic factor Grading of Recommendations, Assessment, Development and Evaluations framework. Twenty-eight studies satisfied the inclusion criteria. Seven studies were judged to be at a moderate risk of bias and 21 at a high risk of bias. Diabetes was associated with worse multidimensional clinical scores (moderate certainty in evidence), worse pain (low certainty in evidence), and worse range of motion (very low certainty in evidence).
Conclusions: This review provides preliminary evidence to suggest that people with diabetes may experience worse outcomes from frozen shoulder than those without diabetes. If high-quality studies can confirm the findings of this review, then clinicians should monitor patients with

List of abbreviations: GRADE, Grading of Recommendations, Assessment, Development and Evaluations; MUA, manipulation under anesthesia; QUIPS, Quality in Prognostic Factor Studies; ROM, range of motion; VAS, visual analog scale.

Supported by the Versus Arthritis PhD scholarship scheme (grant no. 21899). Claire Burton is funded by an NIHR (National Institute for Health Research) Clinical Lectureship. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care.
PROSPERO Registration No.: CRD42019122963.
Disclosures: None.
Cite this article as: Arch Rehabil Res Clin Transl. 2021;3:100141

https://doi.org/10.1016/j.arrct.2021.100141

frozen shoulder with diabetes more closely and offer further treatment if pain or lack of function persists long-term.

Frozen shoulder, also known as adhesive capsulitis, is a painful condition that can cause prolonged disability.[1] Stiffness of the capsule surrounding the glenohumeral joint reduces both active and passive range of motion (ROM), particularly external rotation.[2] Frozen shoulder is commonly, but incorrectly, said to be a self-limiting condition (meaning that, in time, the condition will resolve without intervention).[3-5] However, there is an abundance of evidence to suggest that many patients with frozen shoulder suffer from long-term pain and restricted movement.[6-9] Frozen shoulder is initially treated using conservative (nonsurgical) methods including analgesics, local corticosteroid injection, and gentle mobilization and exercise.[5] Cases that are resistant to conservative management may be treated surgically with manipulation under anesthesia, arthroscopic capsular release, or arthrographic distention/hydrodilatation.[10] Currently there is no clear consensus as to which management strategy is the most effective way to treat frozen shoulder.[10-12]

The onset of frozen shoulder most commonly occurs between 40 and 70 years of age, with patients rarely presenting before the age of 40.[13] Fifty-eight percent of people with frozen shoulder are women.[13] In 6%-17% of patients the contralateral shoulder is also affected, usually within 5 years of the first shoulder recovering.[5,14] The prevalence of frozen shoulder in the general population has often been stated in the literature to be around 2%,[14] although any estimates of the incidence or prevalence of frozen shoulder will be inconsistent owing to the variability in diagnostic criteria for frozen shoulder.[15]

People with diabetes are 5 times more likely to have frozen shoulder than people without diabetes, and the prevalence of frozen shoulder in people with diabetes has been estimated to be 13.4%.[14] Although it is currently unclear why diabetes is associated with frozen shoulder, it has been hypothesized that glycation processes may cause changes in capsule tissues and consequently lead to the development of frozen shoulder.[16] People with diabetes make up around 30% of the frozen shoulder population[14]; therefore, it is important to understand whether and how the outcomes of frozen shoulder may differ for people with diabetes compared with those without diabetes.

This review summarizes evidence from longitudinal observational studies to investigate whether diabetes is a prognostic factor in people with frozen shoulder.

Methods

Search strategy

The protocol for this systematic review was registered on PROSPERO (CRD42019122963), and the review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses statement.[17] MEDLINE, EMBASE, AMED, PsycINFO, Web of Science core collection, Cumulative Index to Nursing and Allied Health Literature, Epistemonikos, Trip, PEDro, OpenGrey, and The Grey Literature Report were searched from inception to June 2021. Reference lists of included studies were screened and a professional contact of 1 author (D.vdW.) was contacted. The search for MEDLINE, using Medical Subject Headings and free-text words related to shoulder pain and diabetes, can be found in supplemental appendix S1 (available online only at http://www.archives-pmr.org/). The search strategy was constructed (with the help of a health information expert) to identify studies about shoulder pain in general, rather than frozen shoulder, to maximize the sensitivity of the search.

Study selection

Titles and abstracts were screened by 1 reviewer (B.P.D.) and a 20% random sample was independently checked by 2 reviewers (M.B-B., C.B.) using predefined inclusion and exclusion criteria. Disagreements were resolved by discussion with another reviewer (D.vdW.). Full-text articles were screened by B.P.D. and were independently checked by 3 reviewers (M.B-B., C.B., T.R-M.) using predefined inclusion and exclusion criteria. Any discrepancies were reviewed by and discussed with D.vdW.

To be eligible for inclusion, studies were required to have a longitudinal observational design (prospective or retrospective), include people diagnosed with frozen shoulder at baseline, establish self-reported or clinically diagnosed type 1 or type 2 diabetes at baseline, and compare outcomes between those with and without diabetes at follow-up (>2wk). Cross-sectional studies, case studies, and trials were excluded. We included population-based studies as well as clinical cohorts, with no limitations in terms of treatment received for frozen shoulder. When a full-text article could not be obtained, the study was excluded. All outcome variables related to frozen shoulder, including ROM, pain, and functionality, were eligible for inclusion. Non—English-language papers were assessed by a reviewer with appropriate language skills.

Data extraction and risk of bias assessment

Data were extracted using a predefined extraction sheet by 1 reviewer (B.P.D.) and independently checked by 3 reviewers (M.B-B., C.B., T.R-M.). Extracted data included details of study design, setting, sample characteristics,

exposure/outcome/covariate measurement, sample size, treatment type, attrition, inclusion and exclusion criteria, statistical analysis, and association estimates and their corresponding raw data (if presented). Risk of bias for all studies was independently assessed by pairs of reviewers. The Quality In Prognosis Studies (QUIPS) tool[18] was used to judge risk of bias. The QUIPS tool covers 6 domains: (1) study participation; (2) study attrition; (3) prognostic factor measurement; (4) outcome measurement; (5) study confounding; and (6) statistical analysis and reporting. Each domain is scored as either low, medium, or high risk of bias, and each domain contains numerous prompting items to help guide decision making. The bias scores for each domain and potential effect of biases on the overall study risk of bias were then used to judge the overall risk of bias. Disagreements regarding data extraction and risk of bias were resolved by discussion.

Analysis

There was a high level of variation between studies in terms of outcome measures and length of follow-up; therefore, pooled estimates of associations between diabetes and outcome were not calculated. Generally, results were reported as continuous data, so forest plots of mean differences in outcome scores between people with and without diabetes were plotted. This allowed for the results from the primary studies to be visualized and helped inspect the magnitude, direction, and consistency of possible associations. Where studies provided sufficient raw data, confidence intervals were calculated and included in the forest plot; otherwise, only the point estimate was used. Forest plots were plotted using R version 4.0.2.[19,a]

The synthesis and grading of evidence were conducted using an adapted version of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework, which is tailored specifically for the use of prognostic factor research.[20] The GRADE framework scores prognostic evidence on 6 factors (phase of investigation, study limitations, inconsistency, indirectness, imprecision, publication bias) that may decrease the quality of evidence and 2 factors that may increase the certainty in evidence (effect size, exposure-response gradient). Evidence for the association of diabetes with outcomes in people with frozen shoulder was graded after considering all QUIPS scores, tallies of association direction, raw data, and forest plots.

Some studies reported results for the same outcome at multiple follow-up points. To ensure that these cohorts only contributed once to the results but were still incorporated into the evidence synthesis and GRADE assessment, the direction of association reported in the evidence synthesis for these studies was the direction of association observed most frequently throughout follow-up points. When multiple ROM movements were reported within a single article, the most common direction of association observed for those movements was used for the scoring of the GRADE direction of association between diabetes and ROM for that study. Additionally, some studies used multiple instruments to measure the same domain, so the measure that was used most frequently by papers in the review was included in the evidence synthesis.

Results

Search results

The searches identified 1784 unique citations; 46 studies were selected for full-text screening and 28 studies were selected for the final review (fig 1). Twenty-one studies reported results about the outcome domains ROM (abduction, forward flexion, external and internal rotation), pain (eg, 0-10 visual analog score[21]) and/or multidimensional clinical scores (eg, Constant score) and are summarized in this article, and 13 studies that reported less common (<5 studies) outcomes are reported in supplemental appendix S2, tables SB1 and SB2 (available online only at http://www.archives-pmr.org/). Thirteen studies investigated the association between diabetes and ROM, 10 with pain, and 19 with multidimensional clinical scores.

Study characteristics

The 21 articles reporting ROM, pain, or multidimensional clinical scores in this review were cohort studies. Patients received arthroscopic capsular release in 7 studies, hydrodilatation in 3, manipulation under anesthesia in 4, physiotherapy alone in 2, physiotherapy and ultrasound-guided intraarticular corticosteroid injection and exercise in 1, manipulation and arthroscopic capsular release in 1, manipulation under ultrasound-guided brachial plexus block in 1, and a mixture of surgical and conservative treatments in 3. Nine studies were from Europe, 9 from Asia, 3 from North America, and 1 from Oceania. Thirteen studies were hospital based, 2 were based in medical centers, 1 was based in a physiotherapy clinic, 1 was based in a sports medicine clinic, and the remaining 5 did not specify the setting. Sample size ranged from 15-295 shoulders, with a median sample size of 56 shoulders. The percentage of shoulders from people with diabetes ranged from 13%-57% with a mean of 25%±13%.

Risk of bias

QUIPS risk of bias assessments can be found in supplementarl table SC1 (available online only at http://www.archives-pmr.org/) and overall risk of bias scores for each study can be found in table 1. Across studies, the reviewers agreed on risk of bias scores for 82% of bias domains and agreed on 26 of the 28 overall risk of bias scores. Twenty-one studies were judged to be at a high risk of bias and 7 were judged to be at a moderate risk of bias. In general, the methods used to account for potential confounders, prognostic factors/diabetes measurements, and statistical analysis and reporting were poor (fig 2). Studies often used basic univariate tests to compare outcomes between the diabetes and non-diabetes groups and thus did not adjust for confounders and did not present or compare the characteristics of the 2 groups. Reporting of results was often incomplete and confidence intervals and measures of spread were not always reported (supplemental tables SB1-SB5, available online only at http://www.archives-pmr.org/). It was not clear whether the decision to compare outcomes between those with and without diabetes was based on a priori hypotheses or a posteriori hypotheses, meaning that there was potential

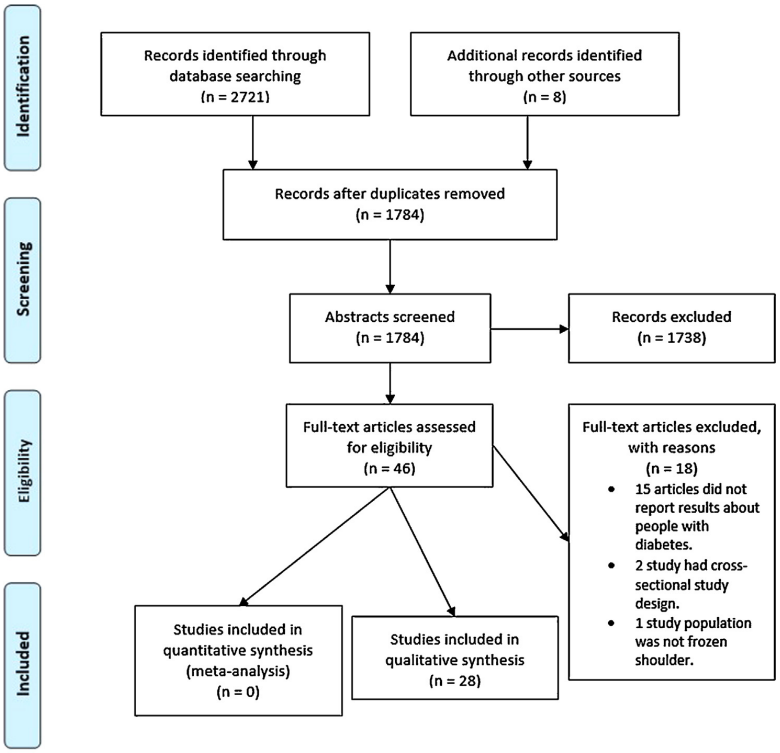

Fig 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram summarizing study identification and selection.

reporting bias present. Studies rarely defined diabetes or reported how diabetes was established (self-reported, tested, or from medical records).

Results for diabetes as a prognostic factor in frozen shoulder

Studies that analyzed the relationship between the presence of diabetes at baseline with either range of motion, pain, or multidimensional clinical scores are summarized in table 1, and full results for these studies can be found in supplemental appendix S2, tables SB3-SB5 (available online only at http://www.archives-pmr.org/). For all 3 outcome domains, there was very little evidence to suggest that people with diabetes had worse baseline scores (see supplemental tables SB3-SB5, available online only at http://www.archives-pmr.org/).

The forest plot of mean differences in ROM (fig 3) shows that generally people with diabetes had worse ROM at follow-up than those without diabetes, although association sizes were inconsistent. The strength of evidence was downgraded in the GRADE framework for the inconsistency in the direction of association because 3 of 13 studies suggested that diabetes is associated with better ROM (table 2). The forest plot contains results on abduction for 673 people from 8 studies, results on external rotation for 1581 people across 13 studies, and results on flexion for 997 people across 12 studies.

The forest plot of mean differences in pain/visual analog scale scored on a 0-10 scale (fig 4) includes 920 people across 10 studies. The plot suggests a possible association between diabetes and worse pain that is consistent across studies but small in magnitude (see fig 4; table 2). The differences often did not meet the statistical significance threshold defined in the respective article (see supplemental table SB4, available online only at http://www.archives-pmr.org/), which could be partly owing to small association sizes but also owing to small sample sizes.

Articles reporting multidimensional clinical scores consistently demonstrated results suggesting that people with diabetes had worse outcomes from frozen shoulder (see supplemental table SB5, available online only at http://www.archives-pmr.org/; fig 5). The forest plot includes 2315 people, including 9 studies based on 758 people measured using Constant scores,[43] 2 studies of 148 people measured with Oxford Shoulder Scores,[44] and 4 studies consisting of a total of 264 people measured with American Shoulder and Elbow Surgeons Shoulder Scores.[45] In some smaller cohorts the difference did not meet statistical significance, but, in general, studies showed associations of similar magnitude and direction, in which people with diabetes had worse outcomes (see table 2).

Table 1 Summary of study characteristics for studies reporting ROM, pain, or multidimensional clinical scores

Study	Study Design and Setting	Treatment Type	Outcomes Measured and Tools Used	Follow-Up Measurements Taken	Sample Size	QUIPS Risk of Bias Score
Nicholson[22]	Cohort study Hospital, US	Arthroscopic capsular release	ROM, pain (VAS), multidimensional score (ASES)	Mean 3 y post-capsular release (range, 2-8 y)	Diabetes: 8 shoulders; non-diabetes: 17 shoulders	High
Cvetanovich et al[23]	Cohort study Medical center, US	Arthroscopic capsular release	ROM	Mean 3.7 y post-capsular release (range, 2-6 y)	Diabetes: 8 shoulders; non-diabetes: 19 shoulders	High
Clement et al[24]	Cohort study Hospital, UK	Hydrodilatation.	ROM, pain (VAS), multidimensional score (OSS)	1 mo post-hydrodilatation	Diabetes: 12 people; non-diabetes: 39 people	High
Bell et al[25]	Cohort study Setting unclear, Australia	Hydrodilatation	ROM, pain (VAS) scored as nil, mild, moderate, or severe	2 mo post-hydrodilatation	Diabetes: 15 people; non-diabetes: 94 people	High
Vastamäki and Vastamäki[26]	Cohort study Hospital, Finland	MUA	ROM, pain (VAS)	Mean 23.1 y post-MUA (range, 19-30 y)	Diabetes: 4 people; non-diabetes: 11 people	High
Cho et al[27]	Cohort study Setting unclear, Republic of Korea	Arthroscopic capsular release	ROM, Pain (VAS), multidimensional score (ASES)	3 mo, 6 mo, 12 mo post-capsular release and a final follow-up of mean 48.4±15.8 mo	Diabetes: 17 shoulders pre-capsular release and final follow-up, 15 at 3 mo, 9 at 6 mo, 13 at 12 mo; non-diabetes: 20 shoulders pre-capsular release, at 3 mo and final follow-up, 17 at 6 mo, 15 at 12 mo	Moderate
Ando et al[28]	Cohort study Setting unclear, Japan	Manipulation under ultrasound-guided brachial plexus block	ROM, pain (VAS), multidimensional score (Constant score)	Mean 4.8±3.5 y for the diabetes group; mean 5.1±2.4 y for the non-diabetes group	Diabetes: 10 shoulders; non-diabetes: 42 shoulders	High
Düzgün et al[29]	Cohort study Physiotherapy center, Turkey	Physiotherapy	ROM, multidimensional score (Constant score)	Following the treatment protocol averaging 8 wk duration	Diabetes: 12 people; non-diabetes: 38 people	Moderate
Vastamäki et al[30]	Cohort study Hospital, Finland	Diabetes group: 69% underwent MUA and 31% underwent conservative treatment; non-diabetes group: 53.3% underwent MUA and 37.3% underwent conservative treatment	ROM, pain (VAS), multidimensional score (Constant score)	Mean 10±8 y for the diabetes group and mean 9.7±7 y for the non-diabetes group	Diabetes: 29 shoulders; non-diabetes: 169 shoulders	High
Mehta et al[31]	Cohort study Hospital, UK	Arthroscopic capsular release	multidimensional score (Constant score)	6 weeks, 6 mo, and 2 y post-capsular release	Diabetes: 21 people; non-diabetes: 21 people	High
Çınar et al[32]	Cohort study Setting unclear, Turkey	Arthroscopic capsular release	Multidimensional score (Constant score)	Mean 48.5 mo for the diabetes group and mean 60.2 mo for the non-diabetes group	Diabetes: 15 shoulders; non-diabetes: 13 shoulders	High
Wang et al[33]	Cohort study Medical center, Taiwan	MUA	Multidimensional score (Adjusted constant score, excluding the 25 points for assessment of muscle strength)	3 wk post-MUA and an average of 95 mo (range, 18-189 mo) post-MUA	Diabetes: 21 shoulders; non-diabetes: 42 shoulders	High

(*continued*)

Table 1 (*Continued*)

Study	Study Design and Setting	Treatment Type	Outcomes Measured and Tools Used	Follow-Up Measurements Taken	Sample Size	QUIPS Risk of Bias Score
Celik et al[34]	Cohort study Setting unclear, Turkey	Manipulation and arthroscopic capsular release	Multidimensional score (Constant score)	Mean 49.5 mo (range, 24-90 mo)	Diabetes: 12 shoulders; non-diabetes: 20 shoulders	High
Sinha et al[35]	Cohort study Hospital, UK	Hydrodilatation	Multidimensional score (OSS)	Improvement in OSS between pre-procedure and 4 wk post-procedure	Diabetes: 26 people; non-diabetes: 90 people	Moderate
Lyhne et al[36]	Cohort study Hospital, Denmark	Arthroscopic capsular release	Multidimensional score (OSS)	Improvement between pre-procedure and 6-mo post-op OSSs	Diabetes: 18 people; non-diabetes: 75 people	High
Theodorides et al[37]	Cohort study Hospital, UK	MUA	Multidimensional score (OSS)	Mean follow-up 28 d post-MUA and at mean follow-up 3.6 y post-MUA (IQR, 1.7-5.0 y)	Diabetes: 39 people; non-diabetes: 256 people	Moderate
Lamplot et al[38]	Cohort study Sports medicine clinic, USA	Conservative treatment	Multidimensional score (ASES)	Minimum 2-y follow-up (mean, 3.4 y)	Diabetes: 9 people; non-diabetes: 51 people	High
Cho et al[39]	Cohort study Hospital, Republic of Korea	Ultrasound-guided intraarticular corticosteroid injection	ROM, pain (VAS), multidimensional score (ASES)	3 wk, 6 wk, 12 wk posttreatment	Diabetes group: 32 shoulders; non-diabetes group: 110 shoulders	High
Ko et al[40]	Cohort study Hospital, Republic of Korea	MUA	ROM, pain (VAS), multidimensional score (Constant score)	6 wk, 3 mo posttreatment	Diabetes group: 32 shoulders; non-diabetes group: 203 shoulders	Moderate
Yanlei et al[41]	Cohort study Hospital, Singapore	Arthroscopic capsular release	ROM, pain (VAS), multidimensional scores (Constant score)	12 mo posttreatment	Diabetes group: 32 shoulders; non-diabetes group: 24 shoulders	High
Barbosa et al[42]	Cohort study Hospital, UK	Mixture of conservative or surgical treatment	ROM, multidimensional score (OSS)	3, 6, 12 mo follow-up	Diabetes group: 46 shoulders; non-diabetes group: 164 shoulders	High

NOTE. Unless otherwise stated, ROM is measured in degrees, pain is measured on a 0-10 visual analog scale with 10 being the worst pain, Constant scores and ASES scores range from 0-100, OSS scores range from 0-48, and for each score reported, a higher score represents a better patient outcome.
Abbreviations: ASES, American Shoulder and Elbow Surgeons Shoulder Score; MUA, manipulation under anesthesia; OSS, Oxford Shoulder Score; UK, United Kingdom; US, United States; VAS, visual analog scale.

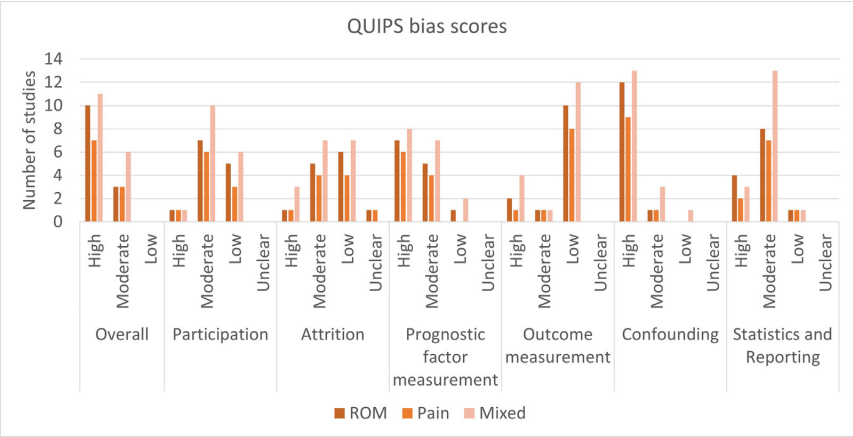

Fig 2 Bar graph of QUIPS scores for each of the 6 bias domains: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, and statistical analysis and reporting.

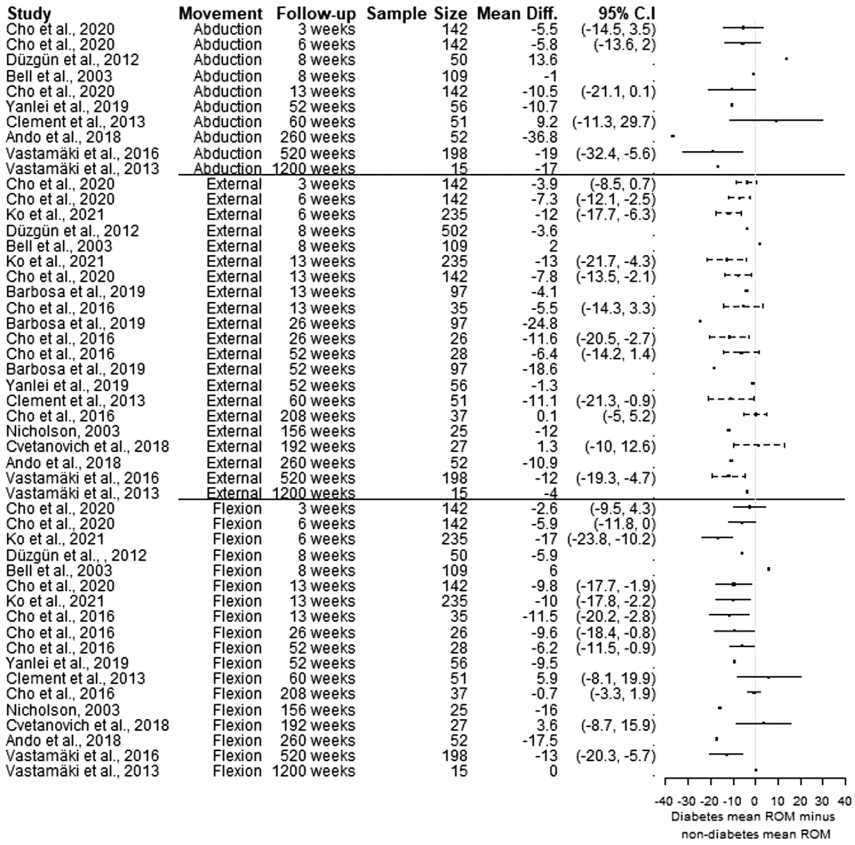

Fig 3 Forest plots of mean differences in ROM scores (degrees) between those with diabetes vs those without diabetes.

Table 2 Summary of GRADE results

Outcome Domain	Number of Participants	Number of Studies	Direction of Association			GRADE Factors							Overall Certainty in Evidence
			Diabetes Group Generally Had Better Outcomes	Tie in Direction of Association	Diabetes Group Generally Had Worse Outcomes	Phase of Investigation	Study Limitations	Inconsistency	Indirectness	Imprecision	Publication Bias	Effect Size	Exposure-Response Gradient	
ROM	2107	13	3	0	10	Phase 2	X	X	✓	X	X	X	X	Very low
Pain	920	10	0	0	10	Phase 2	X	✓	✓	X	X	X	X	Low
Multidimensional scores	1785	18	2	1	15	Phase 2	X	✓	✓	X	X	✓	X	Moderate

NOTE. GRADE factor scoring: ✓ = no serious limitations (or present for moderate/large association size, exposure-response gradient); X = serious limitations (or absent for moderate/large effect size, exposure-response gradient). Within the GRADE framework, a study is classed as phase 2 if it is "a cohort study that seeks to confirm independent associations between the prognostic factor and the outcome."[20(p71)]

For all 3 outcome domains (ROM, pain, multidimensional scores), certainty in evidence was downgraded/not upgraded on the GRADE factors limitations, imprecision, publication bias, and exposure-response gradient (see table 2) with respective reasoning being risk of bias was often high; no rationale for sample sizes was given and some studies produced imprecise estimates; some studies reported associations between diabetes and the outcome without corresponding hypotheses defined at the onset of the study; and diabetes was measured as a binary variable in all studies so there was no exposure-response gradient.

Final GRADE certainty in evidence for diabetes being associated with worse frozen shoulder outcomes was very low for ROM outcome scores, low for pain outcome scores, and moderate for multidimensional clinical outcome scores.

Results from articles reporting less common outcomes can be found in supplemental tables SB1 and SB2(available online only at http://www.archives-pmr.org/). These studies contained 11 results suggesting that people with diabetes had worse outcomes at follow-up, 1 result suggesting that people with diabetes had better outcomes at follow-up, and 3 results where there were no differences in outcomes between those with and without diabetes.

Discussion

This review demonstrates evidence of moderate to very low strength that people with diabetes are likely to experience poorer outcome after a diagnosis of frozen shoulder than those who do not have diabetes. The quality of evidence of diabetes as a prognostic factor in frozen shoulder was very low for ROM outcomes, low for pain outcomes, and moderate for mult-dimensional clinical scores. Through inspection of the forest plots it appears that many studies may have been underpowered, with wide confidence intervals including 0 despite an apparent association. Twenty-one of the 28 studies were at a high risk of bias, meaning that any conclusions based on the results need to be taken with caution. However, a general trend observed suggested that people with diabetes had worse outcomes at follow-up than people without diabetes.

The results of this systematic review are consistent with existing reviews on the topic. Whelton and Peach[46] reported the results of 23 studies but lacked any evidence synthesis. The authors of the review concluded that people with diabetes had a more severe and intractable form of the condition. Boutefnouchet et al[47] conducted a systematic review comparing the outcomes of patients with and without diabetes after arthroscopic capsular release. Again, the review lacked any evidence synthesis strategy, but after reporting the results of 6 studies, the authors concluded that patients with diabetes have more pain, reduced ROM, and inferior function compared with patients without diabetes. Boutefnouchet et al[48] also suggested that the reason patients with diabetes have worse ROM and function could be because they experience more pain and the pain inhibits their ability to do the exercises that are recommended as treatment. The systematic review presented in this article provides a clear evidence synthesis approach using the GRADE framework, along with forest plots to assist data visualization and the use of the QUIPS tool to assess risk of bias.

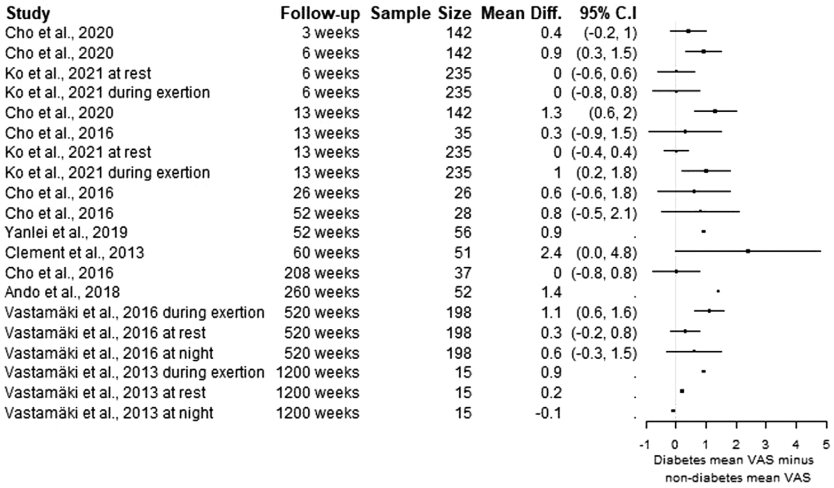

Fig 4 Forest plots of mean differences in visual analog scale scores (0-10 scale) between those with diabetes vs those without diabetes.

Study Limitations

A transparent GRADE approach to evidence synthesis was used and raw data were analyzed, using forest plots to help visualize data from primary studies. High variation in definitions of outcome measures and length of follow-up meant that quantitative pooling of the results was not appropriate. The GRADE synthesis method uses vote counting of statistically significant associations in primary studies. A limitation of this approach is that, if interpreted alone, the vote counting of statistically significant associations does not take sample sizes, association sizes, and measures of spread into account.[48,49] We attempted to overcome this issue by using forest plots to visualize association size and direction and

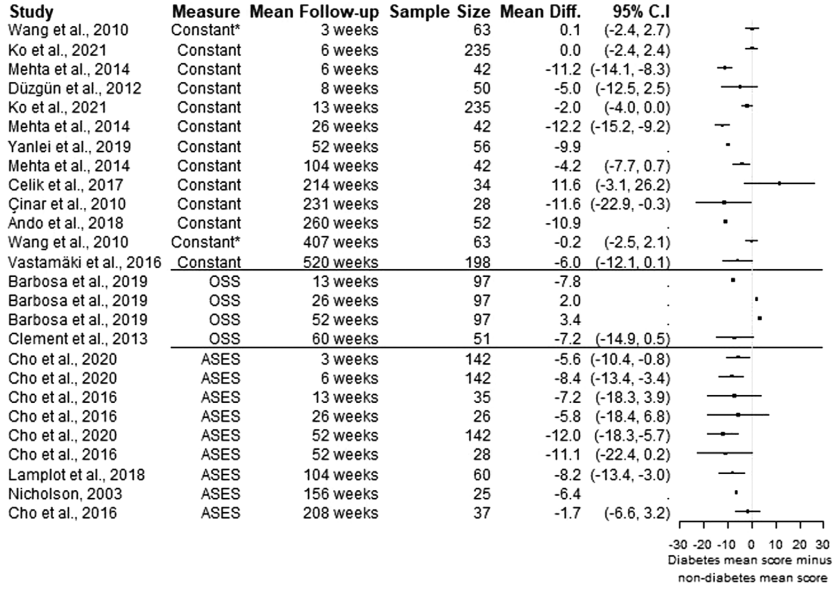

Fig 5 Forest plots of mean differences in multidimensional clinical scores between those with diabetes vs those without diabetes. Constant scores and American Shoulder and Elbow Surgeons Shoulder Scores are on a 0-100 scale; Oxford Shoulder Scores are scored from 0-48. For each measure, a higher score represents a better patient outcome. *Wang et al[33] used an adjusted Constant score, excluding the 25 points for assessment of muscle strength.

the precision of estimates, alongside the presentation of raw data and tallies of association direction to score the GRADE factors. The tallies of association direction, along with presentation of raw data, were used to enable the summarization of results measured categorically that could not be included in the forest plot of mean differences, which uses continuously measured outcomes.

Another limitation of the review is that the GRADE synthesis approach relies on the judgment of the reviewer to score the GRADE factors. This approach is therefore less transparent than methods such as meta-analysis. Through the reporting of results in forest plots and supplemental tables, we have attempted to be transparent in showing the data that guided the scoring of GRADE factors and ultimately the conclusions drawn from this review.

A limitation of current available evidence is that many studies were judged to be at a high risk of bias. Few studies adjusted for confounders or compared baseline characteristics between diabetes and non-diabetes groups, and the type of diabetes that participants had was often not reported. Additionally, future research should clarify whether the decision to compare outcomes in people with and without diabetes is based on an a priori hypothesis or an a posteriori hypothesis. Furthermore, patients in most cohorts received treatments that are generally offered in secondary care settings to patients who have not responded to other treatments (pain relief, mobilization, exercise). This means that it is still unclear how diabetes may affect overall prognosis and treatment outcome in most patients with frozen shoulder managed in primary care, who may have a milder form or may be in an earlier phase of the condition.

Conclusions

To conclude, this review provides preliminary evidence to suggest that people with diabetes may have worse outcomes from frozen shoulder, but high-quality studies are required before more firm conclusions can be made. Nevertheless, given the evidence in this review, clinicians should monitor patients with frozen shoulder with diabetes and recommend further treatment if pain or lack of function persists long-term. Further work is warranted to determine whether patients with diabetes do indeed experience a less favorable outcome from frozen shoulder treated with conservative management in primary care, compared with patients who do not have diabetes. Additionally, further research may explore whether diabetes influences the effectiveness of specific treatments for frozen shoulder, which would need to be investigated in appropriately powered randomized controlled trials or using individual participant data from multiple smaller trials.

Supplier

a. R software, version 4.0.2; The R Foundation.

Corresponding author

Brett P. Dyer, MSc, Primary Care Centre Versus Arthritis, School of Medicine, David Weatherall building, University Rd, Keele University, Staffordshire, ST5 5BG, UK. *E-mail address:* b.p.dyer@keele.ac.uk.

Acknowledgments

We thank the Keele University Research Institute Versus Arthritis systematic review team for their support throughout this project and Elaine Willmore, PhD, for taking time to check the results of our searches for this review. Further thanks are given to Linda Chesterton, PhD, who was involved in the initial conceptualization of this study.

References

1. Bagheri F, Ebrahimzadeh MH, Moradi A, Bidgoli HF. Factors associated with pain, disability and quality of life in patients suffering from frozen shoulder. Arch Bone Jt Surg 2016;4:243–7.
2. Bunker T. Frozen shoulder. J Orthop Trauma 2011;25:11–8.
3. Nagy MT, Macfarlane RJ, Khan Y, Waseem M. The frozen shoulder: myths and realities. Open Orthop J 2013;7:352–5.
4. Manske RC, Prohaska D. Diagnosis and management of adhesive capsulitis. Curr Rev Musculoskelet Med 2008;1:180–9.
5. Dias R, Cutts S, Massoud S. Frozen shoulder. BMJ 2005;331:1453–6.
6. Manske RC, Prohaska D. Clinical commentary and literature review: diagnosis, conservative and surgical management of adhesive capsulitis. Shoulder Elbow 2010;2:238–54.
7. Lewis J. Frozen shoulder contracture syndrome—aetiology, diagnosis and management. Man Ther 2015;20:2–9.
8. Le HV, Lee SJ, Nazarian A, Rodriguez EK. Adhesive capsulitis of the shoulder: review of pathophysiology and current clinical treatments. Shoulder Elbow 2017;9:75–84.
9. Wong CK, Levine WN, Deo K, et al. Natural history of frozen shoulder: fact or fiction? A systematic review. Physiotherapy 2017;103:40–7.
10. Uppal HS, Evans JP, Smith C. Frozen shoulder: a systematic review of therapeutic options. World J Orthop 2015;6:263–8.
11. Maund E, Craig D, Suekarran S, et al. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. Health Technol Assess 2012;16:1–264.
12. Rangan A, Brealey SD, Keding A, et al. Management of adults with primary frozen shoulder in secondary care (UK FROST): a multicentre, pragmatic, three-arm, superiority randomised clinical trial. Lancet 2020;396:977–89.
13. Cho C-H, Koo TW, Cho N-S, et al. Demographic and clinical characteristics of primary frozen shoulder in a Korean population. Clin Shoulder Elbow 2015;18:133–7.
14. Zreik NH, Malik RA. Adhesive capsulitis of the shoulder and diabetes: a meta-analysis of prevalence. Muscles Ligaments Tendons J 2016;6:26–34.
15. van der Windt DAWM, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. Ann Rheum Dis 1995;54:959–64.
16. Hsu C-L, Sheu WH-H. Diabetes and shoulder disorders. J Diabetes Investig 2016;7:649–51.
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.
18. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280–6.
19. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available at: https://www.R-project.org/. Accessed June 15, 2021.

20. Huguet A, Hayden JA, Stinson J, et al. Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. Syst Rev 2013;2:71.
21. Delgado DA, Lambert BS, Boutris N, et al. Validation of digital visual analog scale pain scoring with a traditional paper-based visual analog scale in adults. J Am Acad Orthop Surg Glob Res Rev 2018;2:e088.
22. Nicholson GP. Arthroscopic capsular release for stiff shoulders: effect of etiology on outcomes. Arthroscopy 2003;19:40–9.
23. Cvetanovich GL, Leroux TS, Bernardoni ED, et al. Clinical outcomes of arthroscopic 360° capsular release for idiopathic adhesive capsulitis in the lateral decubitus position. Arthroscopy 2018;34:764–70.
24. Clement RGE, Ray AG, Davidson C, Robinson CM, Perks FJ. Frozen shoulder: long-term outcome following arthrographic distension. Acta Orthop Belg 2013;79:368–74.
25. Bell S, Coghlan J, Richardson M. Hydrodilatation in the management of shoulder capsulitis. Australas Radiol 2003;47:247–51.
26. Vastamäki H, Vastamäki M. Motion and pain relief remain 23 years after manipulation under anesthesia for frozen shoulder. Clin Orthop Relat Res 2013;471:1245–50.
27. Cho C-H, Kim D-H, Lee Y-K. Serial comparison of clinical outcomes after arthroscopic capsular release for refractory frozen shoulder with and without diabetes. Arthroscopy 2016;32:1515–20.
28. Ando A, Hamada J, Hagiwara Y, Sekiguchi T, Koide M, Itoi E. Short-term clinical results of manipulation under ultrasound-guided brachial plexus block in patients with idiopathic frozen shoulder and diabetic secondary frozen shoulder. Open Orthop J 2018;12:99–104.
29. Düzgün İ, Baltacı G, Atay ÖA. Manual therapy is an effective treatment for frozen shoulder in diabetics: an observational study. Jt Dis Relat Surg 2012;23:94–9.
30. Vastamäki H, Ristolainen L, Vastamäki M. Range of motion of diabetic frozen shoulder recovers to the contralateral level. J Int Med Res 2016;44:1191–9.
31. Mehta SS, Singh HP, Pandey R. Comparative outcome of arthroscopic release for frozen shoulder in patients with and without diabetes. Bone Joint J 2014. 96-B:1355-8.
32. Çınar M, Akpınar S, Derincek A, Circi E, Uysal M. Comparison of arthroscopic capsular release in diabetic and idiopathic frozen shoulder patients. Arch Orthop Trauma Surg 2010;130:401–6.
33. Wang J-P, Huang T-F, Ma H-L, Hung S-C, Chen T-H, Liu C-L. Manipulation under anaesthesia for frozen shoulder in patients with and without non-insulin dependent diabetes mellitus. Int Orthop 2010;34:1227–32.
34. Celik H, Seckin MF, Akcal MA, Kara A, Kilinc BE, Akman S. Mid-long term results of manipulation and arthroscopic release in frozen shoulder. Acta Ortop Bras 2017;25:270–4.
35. Sinha R, Patel P, Rose N, et al. Analysis of hydrodilatation as part of a combined service for stiff shoulder. Shoulder Elbow 2017;9:169–77.
36. Lyhne JM, Jacobsen JR, Hansen SJ, Jensen CM, Deutch SR. Diabetic and non-diabetic patients report equal symptom relief after arthroscopic capsular release of frozen shoulder. J Clin Orthop Trauma 2019;10:261–4.
37. Theodorides AA, Owen JM, Sayers AE, Woods DA. Factors affecting short- and long-term outcomes of manipulation under anaesthesia in patients with adhesive capsulitis of the shoulder. Shoulder Elbow 2014;6:245–56.
38. Lamplot JD, Lillegraven O, Brophy RH. Outcomes from conservative treatment of shoulder idiopathic adhesive capsulitis and factors associated with developing contralateral disease. Orthop J Sports Med 2018;6:1–8.
39. Cho CH, Jin H-J, Kim D-H. Comparison of clinical outcomes between idiopathic frozen shoulder and diabetic frozen shoulder after a single ultrasound-guided intra-articular corticosteroid injection. Diagnostics 2020;10:370.
40. Ko YW, Park JH, Youn S-M, Rhee YG, Rhee S-M. Effects of comorbidities on the outcomes of manipulation under anesthesia for primary stiff shoulder. J Shoulder Elbow Surg 2021;30:e482–92.
41. Yanlei GL, Keong MW, Tjoen DLT. Do diabetic patients have different outcomes after arthroscopic capsular release for frozen shoulder? J Orthop 2019;16:211–5.
42. Barbosa F, Swamy G, Salem H, et al. Chronic adhesive capsulitis (frozen shoulder): comparative outcomes of treatment in patients with diabetes and obesity. J Clin Orthop Trauma 2019;10:265–8.
43. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. Clin Orthop Relat Res 1987;214:160–4.
44. Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford Shoulder Score revisited. Arch Orthop Trauma Surg 2009;129:119–23.
45. Richards RR, An K-N, Bigliani LU, et al. A standardized method for the assessment of shoulder function. J Shoulder Elbow Surg 1994;3:347–52.
46. Whelton C, Peach CA. Review of diabetic frozen shoulder. Eur J Orthop Surg Traumatol 2018;28:363–71.
47. Boutefnouchet T, Jordan R, Bhabra G, ModiC Saithna A. Comparison of outcomes following arthroscopic capsular release foridiopathic, diabetic and secondary shoulder adhesive capsulitis: a systematic review. Orthop Traumatol 2019;105:839–46.
48. Wasserstein RL, Lazar NA. The ASA statement on *P*-values: context, process, and purpose. Am Stat 2016;70:129–33.
49. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 2016;31:337–50.

Appendix E

Is diabetes a predictor of surgery in people with frozen shoulder? A cohort study

E.1 STROBE checklist

The following pages contain a completed STROBE checklist corresponding to the work presented in Chapter 7.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Pg 143
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	n/a
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Sec 6.1, Chapter 1
Objectives	3	State specific objectives, including any prespecified hypotheses	Section 6.2
Methods			
Study design	4	Present key elements of study design early in the paper	Sec 6.3.1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Sec 6.3.1, 6.3.2
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Sec 6.3.1
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Sec 6.3.2, Sec 6.3.3, Sec 6.3.4, Sec B.2, Sec B.3, Sec E.2
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Sec 6.3.1-6.3.4, Sec B.2, Sec B.3
Bias	9	Describe any efforts to address potential sources of bias	Sec 6.3.5
Study size	10	Explain how the study size was arrived at	Sec 6.3.1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Sec B.3
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Sec 6.3.5
		(*b*) Describe any methods used to examine subgroups and interactions	Sec n/a
		(*c*) Explain how missing data were addressed	Sec 6.3.5
		(*d*) If applicable, explain how loss to follow-up was addressed	Sec 6.3.5
		(*e*) Describe any sensitivity analyses	Sec 6.3.5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Sec 6.4.1, Fig 6.1
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential	Sec 6.4.1, Tab 6.1

		confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Tab 6.1
		(c) Summarise follow-up time (eg, average and total amount)	Sec 6.4.1, Sec 6.4.2
Outcome data	15*	Report numbers of outcome events or summary measures over time	Sec 6.4.2, Fig 6.1
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Sec 6.4.2
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Sec 6.4.2
Discussion			
Key results	18	Summarise key results with reference to study objectives	Sec 6.5
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Sec 6.5
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Sec 6.5
Generalisability	21	Discuss the generalisability (external validity) of the study results	Sec 6.5
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Pg v

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

E.2 Frozen shoulder surgery Read code list

The frozen shoulder surgery code list in Table E.1 was constructed by two general practitioners. The codes are recorded in the HES dataset.

Read code	Detail
7K6a9	Manipulation of shoulder joint
7K6a9	Manipulation of shoulder joint under anaesthetic
7K6T2	Release of contracture of shoulder joint
7K46	Diagnostic arthroscopy of shoulder joint
7K4y	Other specified operations on shoulder joint
7K4z	Shoulder joint operations NOS
7K6Z5	Injection of steroid into shoulder joint
7K6ZC	Injection of Lederspan into shoulder joint
7K6Z9	Injection of hydrocortisone acetate into shoulder joint
54P6	Shoulder arthrogram

Table E.1: Table detailing the frozen shoulder surgery codes

E.3 R code

E.3.1 Kaplan-Meier Plot

To create a Kaplan-Meier plot for the people with diabetes vs. the people without diabetes in R the `survminer` package was used. The following code may be used to create a Kaplan-Meier plot using survival time variable `time_to_event_outcome`, event indicator `surgery`, and the variable `dta.case` to indicate whether the patient had diabetes. The option `risk.table=TRUE` instructs R to include the percentage at risk table `ggsurv$table`.

```
library("survminer", "survival")
sfit <- survfit(Surv(time_to_event_outcome, surgery) ~ dta.case
   , data=dta)
ggsurv=ggsurvplot(sfit, data = dta, ylim=c(0.85,1), censor=
   FALSE, risk.table=TRUE, conf.int=TRUE, legend.labs=c("No
   Diabetes", "Diabetes")) + xlab("Time (years)") + xlab("Time
   (years)")
ggsurv$table=ggsurvtable(sfit,data = dta, survtable = c("risk.
   table"), risk.table.type = c("percentage"), xlim=c(0,15),
   break.time.by=5, legend.labs=c("No Diabetes", "Diabetes")) +
    xlab("Time (years)")
ggsurv
```

E.3.2 Cox proportional hazards model

The Cox model was run using the `coxph` function from the `survival` package.

```
library("survival")
Cox=coxph(Surv(time_to_event_outcome, surgery) ~ dta.case + dta
   .gender + indexage + dta.imd2015_5 + dta.ethnicity, data=dta
   )
summary(Cox)
```

E.3.3 Schoenfeld residual plots

The following code was used to create the Schoenfeld residual plot for the model `Cox` that was created in Section E.3.2.

```
library("survminer", "survival")
prophaz <- cox.zph(Cox)
ggcoxzph(prophaz, xlab="Time (years; recorded on log scale)",
   ylab="Scaled Schoenfeld residuals", resid=TRUE, title="Age-
   and gender-adjusted model", resid=TRUE, var=c("dta.case"))
```

E.4 Three- and five-year follow-up Schoenfeld residual plots

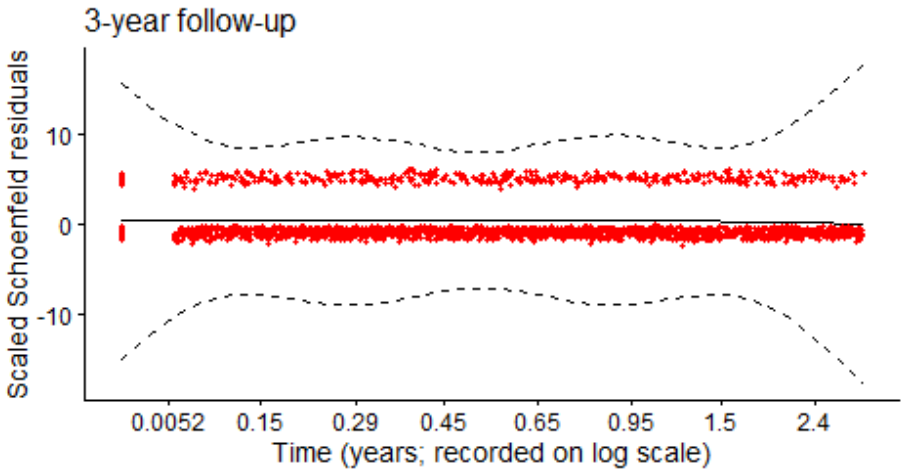

Figure E.1: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the diabetes coefficient in the Cox model with a maximum of 3 years follow-up

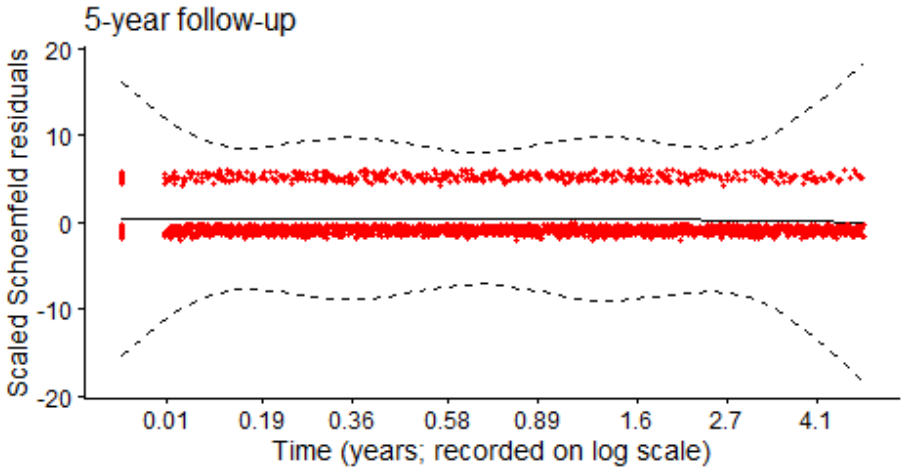

Figure E.2: Residual plot of scaled Schoenfeld residuals, with an added smoothing spline and 95% confidence bands, for the diabetes coefficient in the Cox model with a maximum of 5 years follow-up

Chapter 9

References

[1] C. M. Robinson, K. T. M. Seah, et al. Frozen shoulder. *The Journal of Bone & Joint Surgery*, 94(1):1–9, 2012.

[2] F. Bagheri, M. H. Ebrahimzadeh, et al. Factors associated with pain, disability and quality of life in patients suffering from frozen shoulder. *The Archives of Bone and Joint Surgery*, 4(3):243–7, 2016.

[3] S. Jones, N. Hanchard, et al. A qualitative study of patients' perceptions and priorities when living with primary frozen shoulder. *BMJ Open*, 3:e003452, 2013.

[4] A. Ewald. Adhesive capsulitis: A review. *American Family Physician*, 83(4):417–22, 2011.

[5] T. Bunker. Frozen shoulder: unravelling the enigma. *Annals of The Royal College of Surgeons of England*, 79:210–3, 1997.

[6] E. A. Codman . *The shoulder: rupture of the supraspinatus tendon and other lesions in or about the subacromial bursa (2nd ed.)*. T. Todd Company, Boston, 1934.

[7] J. S. Neviaser. Adhesive capsulitis of the shoulder. Study of pathological findings in periarthritis of the shoulder. *The Journal of Bone and Joint Surgery*, 27:211–22, 1945.

[8] R. Dias, S. Cutts, et al. Frozen shoulder. *BMJ*, 331(7530):1453–6, 2005.

[9] T. Bunker. Time for a new name for frozen shoulder—contracture of the shoulder. *Shoulder & Elbow*, 1:4–9, 2009.

[10] T. Bunker. Frozen shoulder. *Orthopaedics and Trauma*, 25(1):11–8, 2011.

[11] Udaix. Shoulder joint of human body anatomy infographic diagram with all parts including bones ligaments muscles bursa cavity capsule cartilage membrane for medical science education and health care. Photograph. Shutterstock. Accessed: 19/08/2021.

[12] ScientificStock. Shoulder movement: flexion, extension, abduction, adduction, external rotation and internal rotation. Photograph. Shutterstock. Accessed: 19/08/2021.

[13] M. Rookmoneea, L. Dennis, et al. The effectiveness of interventions in the management of patients with primary frozen shoulder. *The Journal of Bone & Joint Surgery*, 92(9):1267–72, 2010.

[14] N. Grubbs. Frozen Shoulder Syndrome: A Review of Literature. *Journal of Orthopaedic & Sports Physical Therapy*, 18(3):479–87, 1993.

[15] J. D. Zuckerman, A. Rokito. Frozen shoulder: a consensus definition. *Journal of Shoulder and Elbow Surgery*, 20(2):322–5, 2011.

[16] B. J. Lundberg. The frozen shoulder: Clinical and radiographical observations the effect of manipulation under general anesthesia structure and glycosaminoglycan content of the joint capsule local bone metabolism. *Acta Orthopaedica Scandinavica*, 40:1–59, 1969.

[17] A. Rangan, L. Goodchild, et al. Frozen shoulder. *Journal of Shoulder and Elbow Surgery*, 7(4):299–307, 2015.

[18] M. J. Kelley, P. W. Mcclure, et al. Frozen shoulder: Evidence and a proposed model guiding rehabilitation. *Journal of orthopaedic & sports physical therapy*, 39(2):135–48, 2009.

[19] K. Walker-Bone, K. T. Palmer, et al. Prevalence and Impact of Musculoskeletal Disorders of the Upper Limb in the General Population. *Arthritis & Rheumatism*, 51(4):642–51, 2004.

[20] D. A. W. M. van der Windt, B. W. Koes, et al. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Annals of the Rheumatic Diseases*, 54:959–64, 1995.

[21] C-H. Cho, T. W. Koo, et al. Demographic and clinical characteristics of primary frozen shoulder in a korean population. *Clinics in Shoulder and Elbow*, 18(3):133–7, 2015.

[22] C. T. Wadsworth. Frozen shoulder. *Physical Therapy*, 66(12):1878–83, 1986.

[23] M. T. Nagy, R. J. MacFarlane, et al. The frozen shoulder: Myths and realities. *The Open Orthopaedics Journal*, 7:352–5, 2013.

[24] D. Cucchi, A. Marmotti, et al. Risk factors for shoulder stiffness: Current concepts. *Joints*, 5(4):517–223, 2017.

[25] N. H. Zreik, R. A. Malik, et al . Adhesive capsulitis of the shoulder and diabetes: a meta-analysis of prevalence. *Muscle, Ligament and Tendons*, 6(1):26–34, 2016.

[26] C. M Sung, T.S Sung, et al. Are serum lipids involved in primary frozen shoulder? a case-control study. *The Journal of Bone and Joint Surgery*, 96(21):1828–33, 2014.

[27] S. P. Nayak, C. K. Panda. Is hyperlipidaemia a cause of primary frozen shoulder? A case-controlled study. *Journal of Evidence Based Medicine and Healthcare*, 4(13):697–701, 2017.

[28] D. C. Austin, I. Gans, et al. The association of metabolic syndrome markerswith adhesive capsulitis. *Journal of shoulder and elbow surgery*, 23:1043–51, 2014.

[29] S-W. Huang, J-W Lin, et al. Hyperthyroidism is a risk factor for developing adhesive capsulitis of the shoulder: a nationwide longitudinal population-based study. *Scientific Reports*, 4:4183, 2014.

[30] C. Cohen, S. Tortato, et al. Association between Frozen Shoulder and Thyroid Diseases: Strengthening the Evidence. *Revista Brasileira de Ortopedia*, 55(4):483–9, 2020.

[31] S-F. Lo, S-W. Chu, et al. Diabetes mellitus and accompanying hyperlipidemia are independent risk factors for adhesive capsulitis: a nationwide population-based cohort study. *Rheumatology International*, 34:67–74, 2014.

[32] D. T Loveday, P. Johnston, et al. Frozen shoulder: a clinical review. *British Journal of Hospital Medicine*, 70(5):276–8, 2009.

[33] S. P. Smith, V. S. Devaraj, et al. The association between frozen shoulder anddupuytren's disease. *Journal of Shoulder and Elbow Surgery*, 10(2):149–51, 2001.

[34] T. D. Bunker, P. P. Anthony. The pathology of frozen shoulder. A Dupuytren-like disease. *The Journal of Bone & Joint Surgery*, 77(5):677–83, 1995.

[35] J. A. Dickson, E. H. Crosby. Periarthritis of the shoulder. *JAMA*, 99:2252–7, 1932.

[36] A. Omari, T. D. Bunker. Open surgical release for frozen shoulder: Surgical findings and results of the release. *Journal of Shoulder & Elbow Surgery*, 10(4):353–7, 2001.

[37] O. Kilian, U. Pfeil, et al. Enhanced alpha 1(I) mRNA expression in frozen shoulder and dupuytren tissue. *European Journal of Medical Research*, 12(12):585–90, 2007.

[38] E. Itoi, G. Arce, et al. Shoulder Stiffness: Current Concepts and Concerns. *Arthroscopy*, 32(7):1402–14, 2016.

[39] V. Ryan, H. Brown, et al. The pathophysiology associated with primary (idiopathic) frozen shoulder: A systematic review. *BMC Musculoskeletal Disorders*, 17:340, 2016.

[40] Y. M. Lho, E. Ha, et al. Inflammatory cytokines are overexpressed in the subacromial bursa of frozen shoulder. *Journal of Shoulder & Elbow Surgery*, 22(5):666–72, 2013.

[41] B. Kabbabe, S. Ramkumar, M. Richardson. Cytogenetic analysis of the pathology of frozen shoulder. *International Journal of Shoulder Surgery*, 4(3):75–8, 2010.

[42] C-H. Cho, K-S. Song, et al. Biological Aspect of Pathophysiology for Frozen Shoulder. *BioMed Research International*, page 7274517, 2018.

[43] S. A. Rodeo, J. A. Hannafin, et al. Immunolocalization of cytokines and their receptors in adhesive capsulitis of the shoulder. *Journal of Orthopaedic Research*, 15(3):427–36, 1997.

[44] R. S. Watson, E. Gouze, et al. Gene delivery of TGF-b1 induces arthrofibrosis and chondrometaplasia of synovium in vivo. *Labaratory Investigation*, 90:1615–27, 2010.

[45] J. Hartupee, D. L. Mann. Role of Inflammatory Cells in Fibroblast Activation. *Journal of Molecular and Cellular Cardiology*, 93:143–8, 2016.

[46] K. Tamai, M. Akutsu, Y. Yano. Primary frozen shoulder: brief review of pathology and imaging abnormalities. *Journal of Orthopaedic Science*, 19(1):1–5, 2014.

[47] P. J. Rundquist, P. M. Ludewig. Patterns of motion loss in subjects with idiopathic lossof shoulder range of motion. *Clinical Biomechanics*, 19:810–8, 2004.

[48] R. C. Manske, D. Prohaska. Clinical commentary and literature review: diagnosis, conservative and surgical management of adhesive capsulitis. *Shoulder and Elbow*, 2:238–54, 2010.

[49] J. Lewis. Frozen shoulder contracture syndrome - aetiology, diagnosis and management. *Manual Therapy*, 20(1):2–9, 2015.

[50] P. H. Laubscher, T. G. Rösch. Frozen Shoulder: A review. *The South African Orthopaedic Journal*, 8(3):24–9, 2009.

[51] C. K. Wong, W. N. Levine, et al. Natural history of frozen shoulder: fact or fiction? a systematic review. *Physiotherapy*, 103:40–7, 2017.

[52] M. H. Khan, S. Nuhmani. Frozen Shoulder—A Review of Current Concepts. *Journal of Musculoskeletal Pain*, 22(3):308–13, 2014.

[53] S. Abrassart, F. Kolo, et al. 'Frozen shoulder' is ill-defined. How can it be described better? *Shoulder & Elbow*, 5:273–9, 2020.

[54] R. G. Grey. The natural history of 'idiopathic' frozen shoulder. *Journal of bone and joint surgery*, 60:564, 1978.

[55] H. V. Le, S. J. Lee, et al. Adhesive capsulitis of the shoulder: review of pathophysiology and current clinical treatments. *Shoulder and Elbow*, 9(2):75–84, 2017.

[56] A. I. Binder, D. Y. Bulgen, et al. Frozen shoulder: a long-term prospective study. *Annals of the Rheumatic Diseases*, 43:361–4, 1984.

[57] B. Hazleman. The painful stiff shoulder. *Rheumatology and Physical Medicine*, 11(8):413–21, 1972.

[58] B. Reeves. The natural history of the frozen shoulder syndrome. *Scandinavian journal of rheumatology*, 4(4):193–6, 1975.

[59] C. Hand, K. Clipsham, et al. Long-term outcome of frozen shoulder. *Journal of Shoulder and Elbow Surgery*, 17(2):231–6, 2008.

[60] I. Dudkiewicz. Idiopathic adhesive capsulitis: long-term results of conservative treatment. *The Israel Medical Association Journal*, 6(9):524–6, 2004.

[61] B. Shaffer, J. E. Tibone, R. K. Kerlan. Frozen shoulder: a long-term follow-up. *Journal of Bone and Joint Surgery*, 74A:738–46, 1992.

[62] L. Dennis, S. Brealey, et al. Managing idiopathic frozen shoulder: a survey of health professionals' current practice and research priorities. *Shoulder & Elbow*, 2:294–300, 2010.

[63] S. M. van de Laar, P. van der Zwaal. Management of the frozen shoulder. *Orthopedic Research and Reviews*, 6:81–90, 2014.

[64] E. Maund, D. Craig, et al. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. *Health technology assessment*, 16(11):1–264, 2012.

[65] D. A. W. M. van der Windt, B. W. Koes, et al. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. *BMJ*, 317:1292–6, 1998.

[66] S. Carette, H. Moffet, et al. Intraarticular corticosteroids, supervised physiotherapy, or a combination of the two in the treatment of adhesive capsulitis of the shoulder: a placebo-controlled trial. *Arthritis and Rheumatism*, 48(3):829–38, 2003.

[67] J. C. Winters, J. S. Sobel, et al. Comparison of physiotherapy, manipulation, and corticosteroid injection for treating shoulder complaints in general practice: randomised, single blind study. *BMJ*, 314:1320–8, 1997.

[68] H. S. Uppal, J. P. Evans, et al. Frozen shoulder: A systematic review of therapeutic options. *World Journal of Orthopedics*, 6(2):263–8, 2015.

[69] R. J. Neviaser, T. J. Neviaser. The frozen shoulder. Diagnosis and management. *Clinical Orthopaedics and Related Research*, 223:59–64, 1987.

[70] K. Kanbe. Clinical outcome of arthroscopic capsular release for frozen shoulder: essential technical points in 255 patients. *Journal of Orthopaedic Surgery and Research*, 13:56, 2018.

[71] A. Rangan, S. D. Breadley, et al. Management of adults with primary frozen shoulder in secondary care (UK FROST): a multicentre, pragmatic, three-arm, superiority randomised clinical trial. *Lancet*, 396:977–89, 2020.

[72] S. Jones, N. Hanchard, S. Hamilton, et al. A qualitative study of patients' perceptions and priorities when living with primary frozen shoulder. *BMJ Open*, 3:e003452, 2013.

[73] C. Srikesavan, F. Toye, S. Brealey. Experiences and perceptions of trial participants and healthcare professionals in the uk frozen shoulder trial (uk frost): a nested qualitative study. *BMJ Open*, 11:e040829, 2021.

[74] American Diabetes Association. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes 2020. *Diabetes Care*, 43(Suppl. 1):S14–S31, 2020.

[75] International Diabetes Federation. *IDF Diabetes Atlas, Second Edition*. Brussels, Belgium: International Diabetes Federation, 2019.

[76] NHS Digital. National Diabetes Audit 2017/18 - Report 1 Care Processes and Treatment Targets. 2020.

[77] N. G. Forouhi, N. J. Wareham. Epidemiology of diabetes. *Medicine*, 47(1):22–7, 2019.

[78] K. Rett, U. Gottwald-Hostalek. Understanding prediabetes: definition, prevalence, burden and treatment options for an emerging disease . *Current Medical Research and Opinion*, 35(9):1529–34, 2019.

[79] W. T. Cefalu. "Prediabetes": are there problems with this label? No, we need heightened awareness of this condition! *Diabetes Care*, 39:1472–7, 2016.

[80] Y. Lawal, F. Bello, et al. Prediabetes Deserves More Attention: A Review . *Clinical Diabetes*, 38(4):328–38, 2020.

[81] N. Bansal. Prediabetes diagnosis and treatment: A review . *World Journal of Diabetes*, 6(2):296–303, 2015.

[82] England NHS. NHS Diabetes Prevention Programme (NHS DPP) . https://www.england.nhs.uk/ourwork/qual-clin-lead/diabetes-prevention/. Accessed: 28-08-2021.

[83] J-W. Yoon, H-S. Jun. Autoimmune destruction of pancreatic beta cells. *American Journal of Therapeutics*, 12(6):580–91, 2005.

[84] M. A. Atkinson, G. S. Eisenbarth, A. W. Michels. Type 1 diabetes. *Lancet*, 383(9911):69–82, 2014.

[85] American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*, 37 (Suppl. 1):S81–S90, 2014.

[86] S. E. Khan, M. E. Cooper, S. Del Prato. Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present and future. *Lancet*, 383(9922):1068–83, 2014.

[87] K. Kaku. Pathophysiology of Type 2 Diabetes and Its Treatment Policy. *Research and Reviews*, 53(1):41–6, 2010.

[88] J. S. Sklyer, G. L. Bakris, et al. Differentiation of Diabetes by Pathophysiology, Natural History, and Prognosis. *Diabetes*, 66:241–55, 2017.

[89] A. V. Matveyenko, P. C. Butler. Relationship between beta-cell mass and diabetes onset. *Diabetes, obesity & metabolism*, 10 (Suppl. 4):23–31, 2008.

[90] Lifestyles Team, NHS Digital. Health Survey for England 2018. 2019.

[91] NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. *Lancet*, 387:1513–30, 2016.

[92] M. Rewers, J. Ludvigsson. Environmental risk factors for type 1 diabetes. *Lancet*, 387(10035):2340–8, 2016.

[93] F. Pociot, Å. Lernmark. Genetic risk factors for type 1 diabetes. *Lancet*, 387(10035):2331–9, 2016.

[94] A. B. Olokoba, O. A. Obateru, L. B. Olokoba. Type 2 Diabetes Mellitus: A Review of Current Trends. *Oman Medical Journal*, 27(4):269–73, 2012.

[95] B. Fletcher, M. Gulanick, C. Lamendola. Risk Factors for Type 2 Diabetes Mellitus. *The Journal of Cardiovascular Nursing*, 16(2):17–23, 2002.

[96] S. S. Zghebi, D. T. Steinke, et al. Examining trends in type 2 diabetes incidence, prevalence and mortality in the UK between 2004 and 2014. *Diabetes, Obesity and Metabolism*, 19:1537–45, 2017.

[97] D. E. Moller, K. D. Kaufman. Metabolic Syndrome: A Clinical and Molecular Perspective. *Annual Review of Medicine*, 56:45–62, 2005.

[98] S. O'Neill, L. O'Driscoll. Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies. *Obesity reviews*, 16:1–12, 2015.

[99] E. Kassi, P. Pervanidou, et al. Metabolic Syndrome: Definitions and Controversies. *BMC Medicine*, 9:48, 2011.

[100] K. G. M. M. Alberti, R. H. Eckel, et al. Harmonizing the Metabolic Syndrome. *Circulation*, 120:1640–5, 2009.

[101] K. D. Bruce, C. D. Byrne. The metabolic syndrome: common origins of a multifactorial disorder. *Postgraduate Medical Journal*, 85:614–21, 2009.

[102] K. H. Collins, W. Herzog, et al. Obesity, Metabolic Syndrome, and Musculoskeletal Disease: Common Inflammatory Pathways Suggest a Central Role for Loss of Muscle Integrity. *Frontiers in Physiology*, 9:112, 2018.

[103] P. Reddy, D. Lent-Schochet, et al. Metabolic syndrome is an inflammatory disorder: A conspiracy between adipose tissue and phagocytes. *Clinica Chimica Acta*, 496:35–44, 2019.

[104] S. Ghosal, A. Ghosal. Diabetes and musculoskeletal disorders-a review. *Journal of Diabetes, Metabolic Disorders & Control*, 7(2):63–71, 2020.

[105] D. R. Seaman. Body mass index and musculoskeletal pain: is there a connection? *Chiropractic & Manual Therapies*, 21:15, 2013.

[106] M. Pietrzak. Adhesive capsulitis: An age related symptom of metabolic syndrome and chronic low-grade inflammation? *Medical Hypotheses*, 88:12–7, 2016.

[107] Diabetes UK. Facts and stats January - 2019. 2019, Available from: https:// www.diabetes.org.uk/ resources-s3/ 2019-02/ 1362B_Facts %20and %20stats%20 Update% 20Jan %202019_LOW% 20RES_ EXTERNAL.pdf [accessed 9 November 2020].

[108] M. J. Fowler. Microvascular and Macrovascular Complications of Diabetes. *Clinical Diabetes*, 26(2):77–82, 2008.

[109] T Sözen , N. Ç. Başaran, et al. Musculoskeletal problems in diabetes mellitus. *European Journal of Rheumatology*, 5(4):258–65, 2018.

[110] A. Goldin, J. A. Beckman, et al. Advanced glycation end products. *Circulation*, 114:597–605, 2006.

[111] A. Negre-Salvayre, R. Salvayre, et al. Hyperglycemia and Glycation in Diabetic Complications. *Antioxidants & Redox Signaling*, 11(12):3073–96, 2009.

[112] R. G. Paul, A. J. Bailey. Glycation of collagen: the basis of its central role in the late complications of ageing and diabetes. *International Journal of Biochemistry & Cell Biology*, 28(12):1297–310, 1996.

[113] K. R. Hwang, G. A. Murrell, et al. Advanced glycation end products in idiopathic frozen shoulders. *Journal of Shoulder and Elbow Surgery*, 25(6):981–8, 2016.

[114] J. E. Hsu, O. A. Anakwenze, et al. Current review of adhesive capsulitis. *Journal of Shoulder and Elbow Surgery*, 20(3):502–14, 2011.

[115] N. Shanmugam, M. A. Reddy, et al. High glucose-induced expression of proinflammatory cytokine and chemokine genes in monocytic cells. *Diabetes*, 52(2):1256–64, 2003.

[116] C. Whelton, C. A. Peach. Review of diabetic frozen shoulder. *The European Journal of Orthopaedic Surgery and Traumatology*, 28(3):363–71, 2018.

[117] T. Boutefnouchet, R. Jordan, G. Bhabra, C. Modi, A. Saithna. Comparison of outcomes following arthroscopic capsular release for idiopathic, diabetic and secondary shoulder adhesive capsulitis: A systematic review. *Orthopaedics & Traumatology: Surgery & Research*, 105:839–46, 2019.

[118] E. Herrett, A. M. Gallagher, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology*, 44(3):827–36, 2015.

[119] Clinical Practice Research Datalink. 2021, Available from: https://www.cprd.com [accessed 30th January 2021].

[120] M. J. Carr, S. Steeg, et al. Effects of the COVID-19 pandemic on primary care-recorded mental illness and self-harm episodes in the UK: a population based cohort study. *Lancet Public Health*, 6:e124–35, 2021.

[121] CPRD. CPRD Aurum Frequently asked questions (FAQs) Version 2.2 .

https://www.cprd.com/ sites/ default/ files/ CPRD%20Aurum%20FAQs%20v2.2.pdf. Accessed: 18-07-2021.

[122] NHS Digital. SNOMED CT . https://digital.nhs.uk/ services/ terminology-and-classifications/ snomed-ct. Accessed: 18-07-2021.

[123] M. Roland. Linking physicians' pay to the quality of care—a major experiment in the united kingdom. *New England Journal of Medicine*, 351(14):1448–54, 2004.

[124] S. J. Gillam, A. N. Siriwardena, N. Steel. Pay-for-performance in the united kingdom: Impact of the quality and outcomes framework—a systematic review. *Annals of Family Medicine*, 10(5):461–8, 2012.

[125] E. Kontopantelis, D. Reeves, et al. Recorded quality of primary care for patients with diabetes in England before and after the introduction of a financial incentive scheme: a longitudinal observational study. *BMJ quality & safety*, pages 53–64, 2013.

[126] K. Khunti, R. Gadsby, et al. Original article quality of diabetes care in the uk: comparison of published quality-of-care reports with results of the quality and outcomes framework for diabetes. *Diabetic Medicine*, 24:1436–41, 2007.

[127] E. Kontopantelis, R. Parisi, et al. Longitudinal multiple imputation approaches for body mass index or other variables with very low individual-level variability: the mibmi command in stata. *BMC Research Notes*, 10:41, 2017.

[128] L. Marston, J. R. Carpenter, et al. Issues in multiple imputation of missing data for large general practice clinical databases. *Pharmacoepidemiology and Drug Safety*, 19(6):618–26, 2010.

[129] E. Herrett, A. M. Gallagher, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD) . *International Journal of Epidemiology*, 44(3):827–36, 2015.

[130] C. B. Tighe, W. S. Oakley. The prevalence of a diabetic condition and adhesive capsulitis of the shoulder. *Southern Medical Journal*, 101(6):591–5, 2008.

[131] S. K. Rai, M. Kashid, et al. Is it necessary to screen patient with adhesive capsulitis of shoulder for diabetes mellitus? *Journal of Family Medicine and Primary Care*, 8(9):2927–32, 2019.

[132] D. Moher, A. Liberati, et al. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *BMJ*, 339:b2535, 2009.

[133] U.S. National Library of Medicine. MEDLINE: Description of the Database . https://www.nlm.nih.gov/bsd/medline.html. Accessed: 19-05-2019.

[134] Ovid. Embase Factsheet. http://www.ovid.com/ site/ catalog/ databases/ pdfs/ 903.pdf. Accessed: 19-05-2019.

[135] Ovid. AMED Factsheet. http://www.ovid.com/ site/ catalog/ databases/ pdfs/ 12.pdf. Accessed: 19-05-2019.

[136] Ovid. PsycINFO Factsheet. http://www.ovid.com/ site/ catalog /databases/ pdfs/ 139.pdf. Accessed: 19-05-2019.

[137] Keele University Library. Web of Science. https://www.keele.ac.uk/ library/ find/ databases/ webofscience/. Accessed: 19-05-2019.

[138] EBSCO. CINAHL Complete. https://www.ebscohost.com/ nursing/ products/ cinahl-databases/ cinahl-complete. Accessed: 19-05-2019.

[139] University of Leicester Library. Epistemonikos. https://www2.le.ac.uk/ library/ find/ databases/ e/ epistemonikos. Accessed: 19-05-2019.

[140] Trip. https://www.tripdatabase.com/. Accessed: 19-05-2019.

[141] P.Blobaum . Physiotherapy evidence database (pedro). *Journal of the Medical Library Association*, 94(4):477–8, 2006.

[142] OpenGrey. http://www.opengrey.eu/. Accessed: 20-05-2019.

[143] The New York Academy of Medicine. http://www.greylit.org/home. Accessed: 20-05-2019.

[144] Elsevier. *Mendeley Desktop*. 2020.

[145] K. J. Rothman, S. Greenland, et al. *Modern Epidemiology: Third Edition*. Lippincott Williams & Wilkins, 2008.

[146] M. A. Hernán, J. A. Robins. *Causal Inference: What if*. Chapman & Hall/CRC, 2020.

[147] J. A. Hayden, D. A. van der Windt, et al. . Assessing bias in studies of prognostic factors. *Annals of Internal Medicine*, 158(4):280–6, 2019.

[148] T. J. VanderWeele. On the definition of a confounder. *Annals of Statistics*, 41(1):196–220, 2013.

[149] E. Bareinboim, J. Pearl. Controlling Selection Bias in Causal Inference. *JMLR Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, 22:100–8, 2012.

[150] M. Szklo, F. J. Nieto. *Epidemiology: Beyond the Basics 4th Ed.* Jones and Bartlett Publishers, Inc, 2018.

[151] D. E. Lilienfeld, P. D. Stolley. *Foundations of Epidemiology 3rd Ed.* New York, NY: Oxford University Press, 1994.

[152] L. Gordis. *Epidemiology 5th Ed.* Philadelphia, PA: Elsevier Saunders, 2014.

[153] M. Egger. Meta-analysis: Potentials and promise. *BMJ*, 315:1371, 1997.

[154] M. Egger. Meta-analysis: Principles and procedures. *BMJ*, 315:1533, 1997.

[155] M. Borenstein, L. V. Hedges, et al. . *Introduction to Meta-Analysis*. Wiley, 2011.

[156] A. A. Veroniki, D. Jackson, et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods*, 7(1):55–79, 2016.

[157] J. Sánchez-Meca, F. Marín-Martínez. Confidence intervals for the overall effect size in random-effects meta-analysis. *Psychology methods*, 13(1):31–28, 2008.

[158] J. IntHout, J. P. A. Ioannidis, et al. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology*, 14(25):1–12, 2014.

[159] J. Hartung, G. Knapp. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine*, 20(12):1771–82, 2001.

[160] J. Hartung, G. Knapp. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine*, 20(24):3875–89, 2001.

[161] K. Sidik, J. N. Jonkman. A simple confidence interval for meta-analysis. *Statistics in Medicine*, 21(21):3153–9, 2002.

[162] R. D. Riley, S. G. Gates, et al. Statistical methods can be improved within Cochrane pregnancy and childbirth reviews. *Journal of Clinical Epidemiology*, 64(6):608–18, 2011.

[163] M. Mueller, M. D'Addario, et al. Methods to systematically review and meta-analyse observational studies: a systematic scoping review of recommendations. *BMC Medical Research Methodology*, 18(1):44, 2018.

[164] O. M. Dekkers, J. P. Vandenbroucke, et al. COSMOS-E: Guidance on conducting systematic reviews and meta-analyses of observational studies of etiology. *PLoS Med*, 16(2):e1002742, 2019.

[165] J. E. Hunter, F. L. Schmidt. Fixed Effects vs. Random Effects MetaAnalysis Models: Implications for Cumulative Research Knowledge. *International Journal of Selection and Assessment*, 8(4):275–92, 2000.

[166] R. J. Hardy, S. G. Thompson. Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine*, 17:841–56, 1998.

[167] J. P. T. Higgins, S. G. Thompson, et al. Measuring inconsistency in meta-analyses. *BMJ*, 327:557–60, 2003.

[168] M. Borenstein, J. P. T. Higgins, et al. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, 8:5–18, 2017.

[169] W. Viechtbauer, M. W. L. Cheung. Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2):112–25, 2010.

[170] D. Jackson, R. Turner. Power analysis for random-effects meta-analysis. *Research Synthesis Methods*, 8:290–302, 2017.

[171] T. Friede, C. Röver, et al. Meta-analysis of few small studies in orphan diseases. *Research Synthesis Methods*, 8(1):79–91, 2017.

[172] C. Röver, G. Knapp, T. Friede. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. *BMC Medical Research Methodology*, 15:99, 2015.

[173] J. IntHout, J. P. A. Ioannidis, et al. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, 6:e010247, 2016.

[174] C. Partlett, R. D. Riley. Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following reml estimation. *Statistics in Medicine*, 36:301–17, 2017.

[175] StataCorp. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC, 2019.

[176] J. Sterne, R. Harbord. Funnel plots in meta-analysis. *Stata Journal*, 4(2):127–41, 2004.

[177] J. P. T. Higgins, S. Green . *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2011.

[178] K. L. Boyle-Walker, D. L. Gabard, et al. A profile of patients with adhesive capsulitis. *Journal of Hand Therapy*, 10(3):222–8, 1997.

[179] W. Li, N. Lu, et al. Case control study of risk factors for frozen shoulder in china. *International Journal of Rheumatic Diseases*, 18(5):508–13, 2014.

[180] S-Y. Lee, J. Park, et al. Correlation of mr arthrographic findings and range of shoulder motions in patients with frozen shoulder. *Musculoskeletal Imaging*, 198:173–9, 2012.

[181] C. Milgrom, V. Novack, et al. Risk factors for idiopathic frozen shoulder. *Israel Medical Association Journal*, 10(5):361–4, 2008.

[182] K. Wang, V. Ho, et al. Risk factors in idiopathic adhesive capsulitis: a case control study. *Journal of Shoulder and Elbow Surgery*, 22(7):24–9, 2013.

[183] K. Kingston, E. J. Curry, et al. Shoulder adhesive capsulitis: epidemiology and predictors of surgery. *Journal of Shoulder and Elbow Surgery*, 27(8):1437–43, 2018.

[184] Y-P. Huang, C-Y. Fann, et al. Association of diabetes mellitus with the risk of developing adhesive capsulitis of the shoulder: A longitudinal population-based followup study. *Arthritis Care & Research*, 65(7):1197–1202, 2013.

[185] S. J. Thomas, C. McDougall, et al. Prevalence of symptoms and signs of shoulder problems in people with diabetes mellitus. *Journal of Shoulder and Elbow Surgery*, 16(6):748–51, 2007.

[186] J. F. Bridgeman. Periarthritis of the shoulder and diabetes mellitus. *Annals of the Rheumatic Diseases*, 31:69–71, 1972.

[187] G. Heinze, D. Dunkler. Five myths about variable selection. *Transplant International*, 30(1):6–10, 2017.

[188] S. Greenland, R. Neutra. Control of confounding in the assessment of medical technology. *International Journal of Epidemiology*, 9(4):361–7, 1978.

[189] L. G. Dales, H. K. Ury. An improper use of statistical significance testing in studying covariables. *International Journal of Epidemiology*, 7(4):373–5, 1978.

[190] J. Pearl. *Causality, Second Edition*. Cambridge University Press, 2009.

[191] S. A. E. Peters, R. R. Huxley, M. Woodward. Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes. *Lancet*, 383(9933):1973–80, 2014.

[192] L. Chien-Chang, S. Chun-Chuan, et al. Impact of Diabetes on Stroke Risk and Outcomes Two Nationwide Retrospective Cohort Studies. *Medicine*, 94(52):e2282, 2015.

[193] M. Janghorbani, F. B. Hu, et al. Prospective Study of Type 1 and Type 2 Diabetes and Risk of Stroke Subtypes. *Diabetes Care*, 30(7):1730–5, 2007.

[194] R. Chen, B. Ovbiagele, W. Feng. Diabetes and Stroke: Epidemiology, Pathophysiology, Pharmaceuticals and Outcomes. *American Journal of the Medical Sciences*, 351(4):380–6, 2016.

[195] J. Pearl, D. Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Penguin Books, 2018.

[196] M. M. Glymour, S. Greenland. Causal diagrams 183. *Modern Epidemiology, Third Edition*, pages 183–209, 2008.

[197] E. F. Schisterman, S. R. Cole, R. W. Platt. Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies. *Epidemiology*, 20(4):488–95, 2009.

[198] P. W. G. Tennant, E. J. Murray, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *International Journal of Epidemiology*, pages 620–32, 2020.

[199] D. Westreich, S. Greenland. The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*, 177(4):292–8, 2013.

[200] J. H. Chan, B. S. Ho, et al. The relationship between the incidence of adhesive capsulitis and hemoglobin a1c. *Journal of Shoulder and Elbow Surgery*, 26(10):1834–7, 2017.

[201] E. H. Yian, R. Contreras, J. F. Sodl. Effects of glycemic control on prevalence of diabetic frozen shoulder. *Journal of Bone and Joint Surgery*, 94(10):919–23, 2012.

[202] Y. Saisho. Metformin and Inflammation: Its Potential Beyond Glucose-lowering Effect. *Endocrine, metabolic & immune disorders drug targets*, 15(3):196–205, 2015.

[203] Y. Wang, S. M. Hussain, et al. Association between metformin use and disease progression in obese people with knee osteoarthritis: data from the Osteoarthritis Initiative-a prospective cohort study. *Arthritis research & therapy*, 21(1):127, 2019.

[204] F. Bonnet, A. J. Scheen. Effects of SGLT2 inhibitors on systemic and tissue low-grade inflammation: The potential contribution to diabetes complications and cardiovascular disease. *Diabetes & metabolism*, 44(6):457–464, 2018.

[205] M. A. Hernán, J. Hsu, et al. A second chance to get causal inference right: A classification of data science tasks. *CHANCE*, 32(1):42–9, 2019.

[206] M. A. Hernán. The c-word: The more we discuss it, the less dirty it sounds. *The American Journal of Public Health*, 108(5):625–6, 2018.

[207] K. F. Arnold, V. Davies, et al. Generalised linear models for prognosis and intervention: Theory, practice, and implications for machine learning. *TBC*, 1:1, 2020.

[208] M. A. Hernán. The c-word: Scientific euphemisms do not improve causal inference from observational data. *The American Journal of Public Health*, 108:616–9, 2018.

[209] C. M. Schooling, H. E. Jones. Clarifying questions about "risk factors": predictors versus explanation. *Emerging Themes in Epidemiology*, 15(10), 2018.

[210] M. van Diepen, C. L. Ramspek, et al. Prediction versus aetiology: common pitfalls and how to avoid them. *Nephrology Dialysis Transplantation*, 32(2):ii1–ii5, 2017.

[211] P. Royston, K. G. M. Moons, et al. Prognosis and prognostic research: Developing a prognostic model. *BMJ*, 338:604, 2009.

[212] C. F. Dormann, J. Elith, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1):27–46, 2013.

[213] D. L. Cassell, P. L. Flom. Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. *Proceedings of the North East SAS Users Group (NESUG) 2007 Conference*, https://lexjansen.com/pnwsug/2008/DavidCassell-StoppingStepwise.pdf.

[214] F. E. Harrell. *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. Springer, New York, 2001.

[215] K. Han, K. Song, et al. How to develop, validate, and compare clinical prediction models involving radiological parameters: Study design and statistical methods. *Korean Journal of Radiology*, 17(3):339–50, 2016.

[216] D. G. Altman, Y. Vergouwe, et al. Prognosis and prognostic research: validating a prognostic model. *BMJ*, 338:605, 2009.

[217] E. W. Steyerberg, F. E. Harrell. Prediction models need appropriate internal, internal-external, and external validation. *Journal of Clinical Epidemiology*, 69:245–7, 2016.

[218] M. A. Hernán. A definition of causal effect for epidemiological research. *Journal of Epidemiology and Community Health*, 58:265–71, 2004.

[219] M. A. Hernán, S. Hernández-Díaz, et al. A structural approach to selection bias. *Epidemiology*, 15(5):615–25, 2004.

[220] S. Greenland. Quantifying biases in causal models: Classical confounding vs collider-stratification bias. *Epidemiology*, 14:300–6, 2003.

[221] J. Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–710, 1995.

[222] S. Greenland, J. Pearl, et al. Causal diagrams for epidemiologic reseach. *Epidemiology*, 10:37–48, 1999.

[223] S. Greenland, B. Brumback. An overview of relations among causal modelling methods. *International Journal of Epidemiology*, 31(5):1030–7, 2002.

[224] M. M. Glymour. Using causal diagrams to understand common problems in social epidemiology 393. *Methods in Social Epidemiology*, pages 393–428, 2006.

[225] J. Pearl, M. Glymour, et al. . *Causal Inference in Statistics - A Primer*. Wiley, 2016.

[226] S. Hernández-Díaz, E. F. Schisterman, et al. The birth weight "paradox" uncovered? *American Journal of Epidemiology*, 164(11):1115–20, 2006.

[227] Y-K Tu, D. Gunnel, et al. Simpson's paradox, lord's paradox, and suppression effects are the same phenomenon – the reversal paradox. *Emerging Themes in Epidemiology*, 5(2), 2008 https://doi.org/10.1186/1742-7622-5-2.

[228] D. Westreich. Berkson's bias, selection bias, and missing data. *Epidemiology*, 23(1):159–64, 2013.

[229] J. Pearl. Graphical models, causality, and intervention. *Statistical Science*, 8(3):266–9, 1993.

[230] T. J. VanderWeele, N. Staudt. Causal diagrams for empirical legal research: a methodology for identifying causation, avoiding bias and interpreting results. *Law, Probability and Risk*, 10(4):329–354, 2011.

[231] D. Westreich. *Epidemiology by Design: A Causal Approach to the Health Sciences*. Oxford University Press, 2020.

[232] G. W. Imbens, D. B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, 2015.

[233] P. W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–60, 1986.

[234] M. M. Garrido, B. Dowd, et al. Understanding treatment effect terminology in pain and symptom management research. *Journal of Pain and Symptom Management*, 52(3):446–52, 2016.

[235] D. H. Rehkopf, M. M. Glymour, T. L. Osypuk. The consistency assumption for causal inference in social epidemiology: When a rose is not a rose. *Current Epidemiology Reports*, 3(1):63–71, 2016.

[236] M. A. Hernán, J. M. Robins. Estimating causal effects from epidemiological data. *Journal of Epidemiology & Community Health*, 60(7):578–86, 2006.

[237] D. B. Rubin . Inference and missing data. *Biometrika*, 63(3):581–92, 1976.

[238] M. A. Hernán. Beyond exchangeability: The other conditions for causal inference in medical research. *Statistical Methods in Medical Research*, 21(1):3–5, 2011.

[239] D. Westreich, S. R. Cole. Invited commentary: Positivity in practice. *American Journal of Epidemiology*, 171(6):674–7, 2010.

[240] E. J. Tchetgen Tchetgen, T. J. VanderWeele. On causal inference in the presence of interference. *Statistical methods in medical research*, 21(1):55–75, 2012.

[241] T. J. VanderWeele, E. J. Tchetgen Tchetgen, M. E. Halloran. Interference and sensitivity analysis. *Statistical Science*, 29(4):687–706, 2014.

[242] T. J. VanderWeele. Concerning the consistency assumption in causal inference. *Epidemiology*, 20:880–3, 2009.

[243] P. R. Rosenbaum, D. B. Rubin. Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome. *Journal of the Royal Statistical Society. Series B*, 45(2):212–8, 1983.

[244] T. J. VanderWeele. Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health*, 37:17–32, 2016.

[245] D. P. MacKinnon, J. Cheong, A. G. Prirlott. *In Cooper, H., Camic, P. M., Long, D. L., Panter, A. T., Rindskopf, D., Sher, K. J. (Eds.) (2012). APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological*. Washington, DC, US: American Psychological Association, 2012.

[246] D. F. Alwin, R. M. Hauser. The Decomposition of Effects in Path Analysis. *American Sociological Review*, 40(1):37–47, 1975.

[247] R. M. Baron, D. A. Kenny. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–82, 1986.

[248] T. J. VanderWeele . *Explanation in Causal Inference*. Oxford University Press, 2015.

[249] T. J. VanderWeele. A unification of mediation and interaction: a four-way decomposition. *Epidemiology*, 25(5):749–61, 2014.

[250] J. Pearl. Direct and Indirect Effects. *Proceedings of the Seventeenth Conference on Uncertainty in Artifi*

cial Intelligence, pages 411–20, 2001.

[251] D. P. MacKinnon, G. Warsi, J. H. Dwyer. A simulation study of mediated effect measures. *Multivariate behavioural research*, 30(1):41–62, 1995.

[252] D. P. MacKinnon, J. H. Dwyer. Estimating mediated effects in prevention studies. *Evaluation Review*, 17:144–58, 1993.

[253] T. J. VanderWeele, S. Vansteelandt. Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *American Journal of Epidemiology*, 172(12):1339–48, 2010.

[254] K-M. Leung, R. M. Elashoff, A. A. Afifi. Censoring issues in survival analysis. *Annual Review of Public Health*, 18:83–104, 1997.

[255] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187–220, 1972.

[256] T. J. VanderWeele. Causal mediation analysis with survival data. *Epidemiology*, 22(4):582–5, 2011.

[257] G. Hong. Ratio of mediator probability weighting for estimating natural direct and indirect effects. *JSM Proceedings, Biometrics Section*, pages 2401–2415, 2010.

[258] G. Hong, J. Deutsch, H. D. Hill. Ratio-of-mediator-probability weighting for causal mediation analysis in the presence of treatment-by-mediator interaction. *Journal of Educational and Behavioral Statistics*, 40(3):307–40, 2015.

[259] T. Lange, S. Vansteelandt, M. Bekaert. A simple unified approach for estimating natural direct and indirect effects. *American Journal of Epidemiology*, 176(3):190–5, 2012.

[260] J. M. Robins, M. A. Hernán, B. Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–60, 2000.

[261] T. J. VanderWeele, D. Peng. Sensitivity analysis in observational research: Introducing the e-value. *Annals of Internal Medicine*, 167(4):268–74, 2017.

[262] E. von Elm, D. G. Altman, et al. Strengthening the reporting of observational studies in epidemiology (strobe) statement: guidelines for reporting observational studies. *BMJ*, 335:806, 2007.

[263] P. Reddy, D. Lent-Schochet, et al. Metabolic syndrome is an inflammatory disorder: A conspiracy between adipose tissue and phagocytes. *Clinica Chimica Acta*, 496:35–44, 2019.

[264] J. P. Sutherland, B. McKinley, et al. The metabolic syndrome and inflammation. *Metabolic disorder and related disorders*, 2(2):82–104, 2004.

[265] K. Esposito, D. Giugliano. The metabolic syndrome and inflammation: association or causation? *Nutrition, metabolism, and cardiovascular diseases*, 14(5):228–32, 2004.

[266] K-Y. Liang, S. L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, pages 13–22, 1986.

[267] M. N. Mittinty, S. Vansteelandt. Longitudinal mediation analysis using natural effect models. *American Journal of Epidemiology*, 189(11):1427–35, 2020.

[268] S. R. Cole, M. A. Hernán . Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, 168(6):656–64, 2008.

[269] A. Sjölander, S. Greenland et al. Ignoring the matching variables incohort studies – when is it validand why? *Statistics in Medicine*, 32:4696–708, 2013.

[270] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2019.

[271] R-H. Chen, H-Y. Chen, et al. Thyroid diseases increased the risk of type 2 diabetes mellitus: a nationwide cohort study. *Medicine*, 98(20):e15631, 2019.

[272] L. Chaker, S. Ligthart, et al. Thyroid function and risk of type 2 diabetes: a population-based prospective cohort study. *BMC Medicine*, 14:e150, 2016.

[273] N. Gronich, S. N. Deftereos, et al. Hypothyroidism is a risk factor for new-onset diabetes: A cohort study. *Diabetes Care*, 38(9):1657–64, 2015.

[274] M. Ladan, A. Atieh, A. Fereidoun. Thyroid disease and the metabolic syndrome. *Current Opinion in Endocrinology & Diabetes and Obesity*, 26(5):256–65, 2019.

[275] L. Yang, X. Lv, et al. Subclinical hypothyroidism and the risk of metabolic syndrome: A meta-analysis of observational studies. *Endocrine Research*, 41(2):158–65, 2015.

[276] K. A. Iwen, E. Schröder, G. Brabant. Thyroid hormones and the metabolic syndrome. *European Thyroid Journal*, 2:83–92, 2013.

[277] Y. Y. Liu, G. A. Brent. Thyroid hormone crosstalk with nuclear receptor signaling in metabolic regulation. *Trends in Endocrinology and Metabolism*, 21(3):166–73, 2010.

[278] M. Schiefer, P. F. Santos Teixeira, et al. Prevalence of hypothyroidism in patients with frozen shoulder. *Journal of Shoulder and Elbow Surgery*, 26(1):49–55, 2017.

[279] K. Mansfield, E. Crellin, et al. Completeness and validity of alcohol recording in general practice within the UK: a cross-sectional study. *BMJ Open*, 9:e031537, 2019.

[280] L. Linsell, J. Dawson, et al. Prevalence and incidence of adults consulting for shoulder conditions in UK primary care; patterns of diagnosis and referral. *Rheumatology*, 45:215–21, 2006.

[281] O. Dorrestijn, K. Greving. Patients with shoulder complaints in general practice: consumption of medical care. *Rheumatology*, 50(2):389–95, 2011.

[282] T. J. VanderWeele. On the distinction between interaction and effect modification. *Epidemiology*, 20:863–71, 2009.

[283] S. Maini, A. Bansal, et al. Prevalence of diabetes in patients with frozen shoulder. *Journal of Evolution of Medical and Dental Sciences*, 8(11):804, 2019.

[284] R. H. Withrington, F. L. Girgis, et al. A comparative study of the aetiological factors in shoulder pain. *British Journal of Rheumatology*, 24:24–6, 1985.

[285] C. A. Whicher, S. O'Neill, R. I. G. Holt. Diabetes in the uk: 2019. *Diabetic Medicine*, 37:242–7, 2020.

[286] M. I. Harris, R. Klein, et al. Onset of niddm occurs at least 4–7 yr before clinical diagnosis. *Diabetes Care*, 15(7):815–9, 1992.

[287] T. J. Thompson, M. M. Engelgau, et al. The onset of niddm and its relationship to clinical diagnosis in egyptian adults. *Diabetic Medicine*, 13(4):337–40, 1996.

[288] J. L. Gross, M. J. de Azevedo, et al. Diabetic nephropathy: Diagnosis, prevention, and treatment. *Diabetes Care*, 28(1):164–76, 2005.

[289] A. K. H. Lim. Diabetic nephropathy – complications and treatment. *International Journal of Nephrology and Renovascular Disease*, 7:361–81, 2014.

[290] E. J. Duh, J. K. Sun, A. W. Stitt. Diabetic retinopathy: current understanding, mechanisms, and treatment strategies. *JCI Insight*, 2(14):e93751, 2017.

[291] E. L. Feldman, B. C. Callaghan et al. Diabetic neuropathy. *Nature Reviews Disease Primers*, 5:41, 2019.

[292] G. Said. Diabetic neuropathy - a review. *Nature Reviews Neurology*, 3:331–40, 2007.

[293] B. C. Callaghan, H. Cheng, et al. Diabetic neuropathy: Clinical manifestations and current treatments. *Lancet Neurology*, 11(6):521–34, 2012.

[294] C. C. Low Wang, C. N. Hess, et al. Atherosclerotic cardiovascular disease and heart failure in type 2 diabetes mellitus – mechanisms, management, and clinical considerations. *Circulation*, 133:2459–502, 2016.

[295] J. R. Petrie, T. J. Guzik, et al. Diabetes, hypertension, and cardiovascular disease: Clinical insights and vascular mechanisms. *The Canadian Journal of Cardiology*, 34(5):575–84, 2018.

[296] P. M. Grambsch, T. M. Therneau. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3):515–26, 1994.

[297] D. Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69(1):239–41, 1982.

[298] K. Hirsch, A. Wienke. Software for semiparametric shared gamma and log-normal frailty models: An overview. *Computer Methods and Programs in Biomedicine*, 107(3):582–97, 2012.

[299] P. H. Gundtoft, A. K. Kristensen, et al. Prevalence and impact of diabetes mellitus on the frozen shoulder. *Southern Medical Journal*, 111(11):654–9, 2018.

[300] M. Artus, D. A. van der Windt, et al. . Management of shoulder pain by UK general practitioners (GPs): a national survey. *BMJ open*, 7(6):e015711, 2017.

[301] H. Hemingway, P. Croft, et al. Prognosis research strategy (progress) 1: A framework for researching clinical outcomes. *BMJ*, 346:e5595, 2013.

[302] R. D. Riley, J. A. Hayden, et al. Prognosis research strategy (PROGRESS) 2: Prognostic Factor Research. *PLoS Med*, 10(2):e1001380, 2013.

[303] D. A. Delgado, B. S. Lambert, et al. Validation of digital visual analog scale pain scoring with a traditional paper-based visual analog scale in adults. *Journal of the American Academy of Orthopaedic Surgeons Global Research & Reviews*, 2(3):e088, 2018.

[304] R Core Team: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2013.

[305] A. Huguet, J. A. Hayden, et al. Judging the quality of evidence in reviews of prognostic factor research: adapting the grade framework. *Systematic Reviews*, 2(1):71, 2013.

[306] R. R. Richards, K-N. An, et al. A standardized method for the assessment of shoulder function. *Journal of Shoulder and Elbow Surgery*, 3(6):347–52, 1994.

[307] J. Dawson, K. Rogers, et al. The oxford shoulder score revisited. *Archives of Orthopaedic and Trauma Surgery*, 129:119–23, 2009.

[308] G. P. Nicholson. Arthroscopic capsular release for stiff shoulders: effect of etiology on outcomes. *Arthroscopy*, 19(1):20–9, 2003.

[309] G. L. Cvetanovich, T. S. Leroux, et al. Clinical outcomes of arthroscopic 360° capsular release for idiopathic adhesive capsulitis in the lateral decubitus position. *Arthroscopy*, 34(3):764–70, 2018.

[310] R. G. E. Clement, A. G. Ray, et al. Frozen shoulder: long-term outcome following arthrographic distension. *Acta Orthopaedica Belgica*, 79(4):368–74, 2013.

[311] S. Bell, J. Coghlan, M. Richardson. Hydrodilatation in the management of shoulder capsulitis. *Australasian Radiology*, 47(3):247–51, 2003.

[312] H. Vastamäki, M. Vastamäki. Motion and pain relief remain 23 years after manipulation under anesthesia for frozen shoulder. *Clinical Orthopaedics and Related Research*, 471(4):1245–50, 2013.

[313] C-H. Cho, D-H. Kim, Y-K. Lee. Serial comparison of clinical outcomes after arthroscopic capsular release for refractory frozen shoulder with and without diabetes. *Arthroscopy*, 32(8):1515–20, 2016.

[314] A. Ando, J. Hamada, et al. Short-term clinical results of manipulation under ultrasound-guided brachial plexus block in patients with idiopathic frozen shoulder and diabetic secondary frozen shoulder. *The Open Orthopaedics Journal*, 12(1):99–104, 2018.

[315] İ. Düzgün, G. Baltacı, ÖA. Atay. Manual therapy is an effective treatment for frozen shoulder in diabetics: An observational study. *Joint Diseases and Related Surgery*, 23(2):94–9, 2012.

[316] H. Vastamäki, L. Ristolainen, M. Vastamäki. Range of motion of diabetic frozen shoulder recovers to the contralateral level. *Journal of International Medical Research*, 44(6):1191–9, 2016.

[317] C-H. Cho, H-J Jin, D-H Kim. Comparison of Clinical Outcomes between Idiopathic Frozen Shoulder and Diabetic Frozen Shoulder After a Single Ultrasound-Guided Intra-Articular Corticosteroid Injection. *Diagnostics*, 10:370, 2020.

[318] Y. W. Ko, J. H. Park, et al. Effects of comorbidities on the outcomes of manipulation under anesthesia for primary stiff shoulder. *Journal of Shoulder & Elbow Surgery*, 1:11, 2021.

[319] G. L. Yanlei, M. W. Keong, et al. Do diabetic patients have different outcomes after arthroscopic capsular release for frozen shoulder? *Journal of Orthopaedics*, 16:211–5, 2019.

[320] F. Barbosa, G. Swamy, et al. Chronic adhesive capsulitis (Frozen shoulder): Comparative outcomes of treatment in patients with diabetes and obesity. *Journal of Cinical Orthopaedics and Trauma*, 10:265–8, 2019.

[321] S. S. Mehta, H. P. Singh, R. Pandey. Comparative outcome of arthroscopic release for frozen shoulder in patients with and without diabetes. *The Bone & Joint Journal*, 96-B(10):1355–8, 2014.

[322] M. Çınar, S. Akpınar, et al. Comparison of arthroscopic capsular release in diabetic and idiopathic frozen shoulder patients. *Archives of Orthopaedic and Trauma Surgery*, 130(3):401–6, 2010.

[323] J-P. Wang, T-F. Huang, et al. Manipulation under anaesthesia for frozen shoulder in patients with and without non-insulin dependent diabetes mellitus. *International Orthopaedics*, 34(8):1227–32, 2010.

[324] H. Celik, M. F. Seckin, et al. Mid-long term results of manipulation and arthroscopic release in frozen shoulder. *Acta Ortopédica Brasileira*, 25(6):270–4, 2017.

[325] R. Sinha, P. Patel, et al. Analysis of hydrodilatation as part of a combined service for stiff shoulder. *Shoulder & Elbow*, 9(3):169–77, 2017.

[326] J. M. Lyhne, J. R. Jacobsen, et al. Diabetic and non-diabetic patients report equal symptom relief after arthroscopic capsular release of frozen shoulder. *Journal of Clinical Orthopaedics and Trauma*, 10(2):261–4, 2019.

[327] A. A. Theodorides, J. M. Owen, et al. Factors affecting short- and long-term outcomes of manipulation under anaesthesia in patients with adhesive capsulitis of the shoulder. *Shoulder & Elbow*, 6(4):245–56, 2014.

[328] J. D. Lamplot, O. Lillegraven, R. H. Brophy. Outcomes from conservative treatment of shoulder idiopathic adhesive capsulitis and factors associated with developing contralateral disease. *Orthopaedic Journal of Sports Medicine*, 6(7):1–8, 2018.

[329] B. K. Rill, C. M. Fleckenstein, et al. Predictors of outcome after nonoperative and operative treatment of adhesive capsulitis. *American Journal of Sports Medicine*, 39(3):567–74, 2011.

[330] W. N. Levine, C. P. Kashyap, et al. Nonoperative management of idiopathic adhesive capsulitis. *Journal of Shoulder and Elbow Surgery*, 16(5):569–73, 2007.

[331] P. H. Gundtoft, M. L. Attrup, et al. Diabetes mellitus affects the prognosis of frozen shoulder. *Danish Medical Journal*, 67(10):A02200071, 2020.

[332] D. A. Woods, K. Loganathan. Recurrence of frozen shoulder after manipulation under anaesthetic (mua): the results of repeating the mua. *The Bone & Joint Journal*, 99-B(6):812–7, 2017.

[333] E. F. Jenkins, W. J. C. Thomas, et al. The outcome of manipulation under general anesthesia for the management of frozen shoulder in patients with diabetes mellitus. *Journal of Shoulder and Elbow Surgery*, 21(11):1492–8, 2012.

[334] A. Ando, H. Sugaya, et al. Identification of prognostic factors for the nonoperative treatment of stiff shoulder. *International Orthopaedics*, 37(5):859–64, 2013.

[335] C. R. Constant, A. H. Murley. A clinical method of functional assessment of the shoulder. *Clinical Orthopaedics and Related Research*, 214:160–4, 1987.

[336] F. Barbosa, G. Swamy, et al. Chronic adhesive capsulitis (frozen shoulder): Comparative outcomes of treatment in patients with diabetes and obesity. *Journal of Clinical Orthopaedics and Trauma*, 10:265–8, 2019.

[337] P. H. Gundtoft, M. L. Attrup, et al. . Diabetes mellitus affects the prognosis of frozen shoulder. *Danish Medical Journal*, 67(10):A02200071, 2020.

[338] G. Peat, R. D. Riley, et al. Improving the Transparency of Prognosis Research: The Role of Reporting, Data Sharing, Registration, and Protocols. *PLoS Med*, 11(7):e1001671, 2014.

[339] M. A. Hernán, J. M. Robins. Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, 183(8):758–64, 2016.

[340] J. A. C. Sterne, M. A. Hernán, et al. Robins-i: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355:i4919, 2016.

[341] J. Rochon, A. du Bois, T. Lange. Mediation analysis of the relationship between institutional research activity and patient survival. *BMC Medical Research Methodology*, 14:9, 2014.

[342] R. Persson, C. Vasilakis-Scaramozza, et al. Cprd aurum database: Assessment of data quality and completeness of three important comorbidities. *Pharmascoepidemiology & Drug Safety*, 29(11):1456–64, 2020.

[343] I. E Højskov, L. C. Thygsesen, et al. The challenge of non-adherence to early rehabilitation after coronary artery bypass surgery: Secondary results from the SheppHeartCABG trial. *European Journal of Cardiovascular Nursing*, 19(3):238–47, 2019.

[344] M. J. Armstrong, R. J. Sigal, et al. Cardiac rehabilitation completion is associated with reduced mortality in patients with diabetes and coronary artery disease. *Diabetologia*, 58:691–8, 2015.

[345] M. U. C. Worcester, B. M. Murphy, et al. Cardiac rehabilitation programmes: predictors of non-attendance and drop-out. *Diabetologia*, 11(4):328–335, 2004.

[346] M. Wittmer, M. Volpatti, et al. Expectation, satisfaction, and predictors of dropout in cardiac rehabilitation. *European journal of preventative cardiology*, 19(5):1082–8, 2012.

[347] M. Forhan, B. M. Zagorski, et al. Predicting Exercise Adherence for Patients with Obesity and Diabetes Referred to a Cardiac Rehabilitation and Secondary Prevention Program. *Canadian Journal of Diabetes*, 37(3):189–94, 2013.

[348] W. Pan, H. Lu, et al. Prognostic value of HbA1c for in-hospital and short-term mortality in patients with acute coronary syndrome: a systematic review and meta-analysis. *Cardiovascular Diabetology*, 18:169, 2019.

[349] J. Gu, J. Pan, et al. Prognostic impact of HbA1c variability on long-term outcomes in patients with heart failure and type 2 diabetes mellitus. *Cardiovascular Diabetology*, 17:96, 2018.

[350] M. Motoishi, S. Sawai, et al. The preoperative HbA1c level is an independent prognostic factor for the postoperative survival after resection of non-small cell lung cancer in elderly patients. *Surgery Today*, 48(5):517–24, 2018.

[351] M. Abate, C. Schiavone, et al. Limited joint mobility in diabetes and ageing: Recent advances in pathogenesis and therapy. *INTERNATIONAL JOURNAL OF IMMUNOPATHOLOGY AND PHARMACOLOGY*, 23(4):997–1003, 2010.

[352] T. Sözen, N. C. Başaran, et al. Musculoskeletal problems in diabetes mellitus. *European Journal of Rheumatology*, 5(4):258–65, 2018.

[353] F. Takase, Y. Mifune, et al. Association of advanced glycation end products in dupuytren disease. *Journal of Orthopaedic Surgery and Research*, 13:143, 2018.

[354] J. P. Löfgren, M. Zimmerman, et al. Diabetes mellitus as a risk factor for trigger finger –a longitudinal cohort study over more than 20 years. *Frontiers in Clinical Diabetes and Healthcare*, 2:708721, 2021.

[355] R. H. Miller, S. J. Anderson, et al. Hemoglobin a1c level and cardiovascular disease incidence in persons with type 1 diabetes: An application of joint modeling of longitudinal and time-toevent data in the pittsburgh epidemiology of diabetes complications study. *American Journal of Epidemiology*, 187(7):1520–9, 2018.

[356] M. Blagojevic-Bucknall, C. Mallen, et al. Gout as a consequence of sleep apnoea: a matched cohort study. *Rheumatology*, 56(2):183, 2017.