Outcomes of shoulder replacement surgery in routine care

A thesis submitted to the University of Oxford in partial fulfilment of the requirements for the degree of

Doctor of Philosophy



Richard Stephen Craig

St Edmund Hall
Michaelmas Term 2019

Outcomes of shoulder replacement surgery in routine care

Doctor of Philosophy in Musculoskeletal Sciences: Biomedical Data Science
Richard Stephen Craig, St Edmund Hall, Michaelmas 2019

Better evidence is needed on the expected outcomes of shoulder replacement surgery to help patients to decide whether it is the right treatment for them. This thesis systematically evaluates the potential benefits and risks of shoulder replacement surgery observed in routine UK practice, and explores the patient factors associated with good and poor outcomes.

A Cochrane review of randomised controlled trials (RCTs) was performed. The available studies were of low quality and produced few reliable findings to inform practice. RCTs provided very little useful data on the rates of revision surgery and adverse events.

Despite the lack of high-quality evidence, temporal trend analysis revealed that the standardised incidence of primary shoulder replacement surgery recorded by NHS England increased from $3.9/10^5$ in 1999 to $13.6/10^5$ in 2017. These trends were marked by an increase in the number of conventional and reverse polarity total shoulder replacements performed in patients aged 65 years and older, and an increase in recorded patient comorbidities.

Hospital Episode Statistics data were used to calculate the risks of long-term revision and early serious adverse events. The lifetime risk of implant revision surgery following an elective shoulder replacement was less than 1 in 20 for patients aged 85 years or older, but as high as 1 in 4 for some patients aged less than 60 years. These rates were approximately 1.5 times higher if other types of reoperation were included. Serious adverse events were recorded after 1 in 30 elective procedures within the first 30 days. Patients aged over 85 years were far more likely to experience a serious adverse event (1 in 5 men and 1 in 10 women). Following procedures performed for acute trauma, the risk of needing revision surgery was considerably lower, but the likelihood of experiencing a serious adverse event was three times higher.

National Joint Registry data for England, Wales, Northern Ireland and the Isle of Man were used to predict postoperative patient reported outcomes with modest accuracy. Most patients experienced considerable improvement in pain and function within the first year following an elective shoulder replacement, but 1 in 10 patients experienced no benefit at all. Patients aged 70 to 80 years with a diagnosis of primary osteoarthritis or rotator cuff-tear arthropathy had the best predicted outcomes. Younger age and a history of previous trauma and/or surgery were the most important predictors of no improvement. Patient reported outcomes were better following total shoulder replacement compared with hemiarthroplasty in patients without a tear of the rotator cuff. This agreed with the low-quality evidence from RCTs. In patients with an attenuated rotator cuff, there was no clear difference between conventional and reverse polarity total shoulder replacements.

This thesis estimated lifetime risks of revision surgery following a primary shoulder replacement for the first time. This new information, together with the predicted patient reported outcomes and adverse event data, represents the best available evidence from routine practice for surgeons and patients considering shoulder replacement surgery. Further research needs to identify treatment strategies that optimise long-term outcomes for younger patients and the safety of surgery for older patients with comorbidities. The choice of procedure type for patients with incomplete rotator cuff tears is an ongoing uncertainty and may be best addressed with a future well-designed trial.

Acknowledgements

I am fortunate to have had excellent mentors throughout my clinical and academic training. From my first experiences as a medical student and house officer at St Mary's Hospital, Paddington, I have worked with surgeons who have embraced and challenged the uncertainty of what we do. I am now proud to be involved in and contribute to a department that provides some of the best current evidence to share with patients about their orthopaedic care, supporting them to make well-informed decisions. I am grateful to everyone within NDORMS who has supported me over the past three years.

I would first like to thank my supervisors, most importantly Jonny Rees, without whom this thesis would not have been possible. Since we first met, Jonny has guided me expertly, both as a direct supervisor, as a clinical academic mentor, and by creating an excellent support network. Gary Collins challenged me to understand and independently implement the analyses to high methodological standards. Sally Hopewell provided steadfast support throughout the Cochrane review process, and Amar Rangan provided valuable insights on the most important study questions and clinical utility of the results.

I am indebted to the National Joint Registry and Royal College of Surgeons for appointing me to their two-year fellowship programme. This extends not only to the NJR/RCS staff, but also to the hard-working hospital staff who input data into the NJR dataset and most importantly to patients for allowing researchers like myself to use their data. I am grateful to the NIHR Oxford Biomedical Research Centre for supporting the final year of my studies.

Thank you to Linda Hunt and particularly Adrian Sayers at the NJR statistical support unit at the University of Bristol for their help developing the data extract, insights on the analysis of NJR data, and encouragement at the start of the process. I look forward to future collaborations. Thank you to Jenny Lane, my fellow upper limb data enthusiast, for her work at the start of the project to plan and obtain our HES data extract. Thank you to Jenny, Dom Furniss and Andy Carr for their helpful manuscript comments.

For the Cochrane review chapter I would like to thank Henry Goodier for his hard work as the second reviewer for study screening and risk of bias. Thank you to Jas Singh for his comments on the final submitted manuscript, to Eli Harriss and Jordi Pardo for their advice on the search strategy, and to Renea Johnston for editorial support from the Cochrane Musculoskeletal Group.

Thank you to Steve Gwilym, Dani Prieto-Alhambra and Andy Judge for their comments at transfer and confirmation assessments that have helped to improve the final thesis. Thank you to Abtin Alvand, without whose encouragement I may never have stepped through the door of the Botnar.

Finally, to Becs and Isla, thank you for your ongoing love and support. I owe you a trip Down Under.

Contents

Abstract **i**

Acknowledgements **ii**

List of Figures **xv**

List of Tables **xix**

Abbreviations **xxi**

1 Introduction **1**

1.1 Background 1

1.2 Aims 3

1.3 Indications for shoulder replacement surgery 3

1.4 Interventions for glenohumeral arthroses 6

1.4.1 Non-surgical interventions 6

1.4.2 Non-arthroplasty surgical interventions 8

1.5 History and development of prosthetic shoulder replacements 11

1.6 Classification of contemporary shoulder replacements 17

1.7 Evaluation of surgical interventions and innovations 18

1.8 Measuring outcomes in shoulder surgery 20

1.8.1 Core outcome sets 20

1.8.2 Commonly used shoulder-specific outcome measures 21

1.8.3 Mortality 22

1.8.4 Revision surgery 22

1.8.5 Adverse events 23

1.9 Published clinical prediction models in shoulder replacement surgery 24

1.10 Factors influencing the outcome of shoulder replacement surgery 26

1.11 Preoperative decision making 26

1.12 Scope and structure of the thesis 27

2 Cochrane systematic review of shoulder replacement surgery for osteoarthritis and cuff tear arthropathy 29

2.1 Introduction 29

2.2 Methods 30

2.2.1 Inclusion/exclusion criteria 30

2.2.2 Electronic searches 31

2.2.3 Data collection and analysis 32

2.3 Results 36

2.3.1 Results of the searches 36

2.3.2 Characteristics of included studies 37

2.3.3 Participants 38
2.3.4 Interventions 38
2.3.5 Outcomes 41
2.3.6 Ongoing and excluded studies 41
2.3.7 Risk of bias in the included studies 41
2.3.8 Effects of interventions 45
2.4 Discussion 55
2.4.1 Summary of findings 55
2.4.2 Overall completeness and applicability of the evidence 57
2.4.3 Quality of the evidence 57
2.4.4 Potential for bias in the review process 58
2.4.5 Agreements and disagreements with other studies or reviews 58
2.4.6 Conclusions 60

3 Data sources and data management **61**

3.1 Summary 61
3.2 Hospital Episode Statistics 61
3.2.1 Data quality 62
3.2.2 Data collection procedures and description of the dataset 62
3.2.3 Data linkage and patient consent 63
3.2.4 Inclusion criteria 63

3.2.5 Episode linkage to generate dataset for analysis 64
3.2.6 Exclusion criteria 64
3.2.7 HES data flowchart 65
3.2.8 Calculation of comorbidity indices 66
3.2.9 Handling missing data 66
3.3 National Joint Registry 66
3.3.1 Data quality 67
3.3.2 Data collection procedures and description of the dataset 67
3.3.3 Data linkage and patient consent 68
3.3.4 Inclusion criteria 69
3.3.5 Exclusion criteria 69
3.3.6 NJR data flowchart 70
3.3.7 Database reformatting 72
3.3.8 Handling missing data 72
3.3.9 Reclassification of the NJR component database 74
3.4 Unit of analysis and patient grouping issues 75
3.5 Approvals and pre-registration 75
3.6 Software 76

4 Temporal trends in the incidence of primary shoulder replacement surgery recorded in Hospital Episode Statistics for NHS England, 1998-2017 77

4.1 Introduction 77

4.2 Methods . . . 78
4.2.1 Data source and participants . . . 78
4.2.2 Calculation and standardisation of incidence rates . . . 78
4.3 Results . . . 79
4.3.1 Standardised incidence rates . . . 79
4.3.2 Age- and sex-specific incidence rates . . . 79
4.3.3 Types of shoulder replacement performed . . . 80
4.4 Discussion . . . 84
4.4.1 Key findings . . . 84
4.4.2 Strengths and limitations . . . 84
4.4.3 Comparison to other studies . . . 85
4.4.4 Meaning of the study . . . 86
4.4.5 Unanswered questions and future research . . . 87

5 What are the long-term and lifetime risks of revision surgery following a primary shoulder replacement recorded in Hospital Episode Statistics for NHS England? 88

5.1 Introduction . . . 88
5.2 Methods . . . 89
5.2.1 Data source and participants . . . 89
5.2.2 Data processing . . . 90
5.2.3 Statistical analysis . . . 90
5.3 Results . . . 92

5.3.1 Demographics . . . 92
5.3.2 Lifetime risk of revision or reoperation following elective surgery . . . 93
5.3.3 Unadjusted estimates of revision and reoperation risk . . . 95
5.3.4 Age- and sex-specific estimates of revision risk . . . 98
5.3.5 Revision risk by indication for surgery . . . 101
5.3.6 Revision risk by procedure type . . . 102
5.3.7 Temporal trends in revision risk by year of surgery . . . 104
5.4 Discussion . . . 107
5.4.1 Key findings . . . 107
5.4.2 Strengths and limitations . . . 108
5.4.3 Comparison to other studies . . . 109
5.4.4 Meaning of the study . . . 111
5.4.5 Unanswered questions and future research . . . 111

6 What are the risks of serious adverse events following shoulder replacement surgery recorded in Hospital Episode Statistics for NHS England? 113

6.1 Introduction . . . 113
6.2 Methods . . . 114
6.2.1 Data source and participants . . . 114
6.2.2 Adverse event outcome definitions . . . 115
6.2.3 Data processing . . . 116
6.2.4 Statistical analysis . . . 116

6.3 Results . . . 117

6.3.1 Demographics . . . 117

6.3.2 Mortality after shoulder replacement surgery . . . 117

6.3.3 Total serious adverse events at 30 and 90 days following surgery . . . 120

6.3.4 Timing of serious adverse events over first year following surgery . . . 122

6.3.5 Risk of serious adverse events within 30 days adjusted for age, sex and comorbidities . . . 125

6.3.6 Trends in serious adverse events and patient demographics over time . . . 126

6.4 Discussion . . . 130

6.4.1 Key findings . . . 130

6.4.2 Strengths and limitations . . . 130

6.4.3 Comparison to other studies . . . 132

6.4.4 Meaning of the study . . . 134

6.4.5 Unanswered questions and future research . . . 135

7 Development and validation of models to predict the outcomes following shoulder replacement surgery using National Joint Registry data 137

7.1 Introduction . . . 137

7.2 Methods . . . 141

7.2.1 Data source and participants . . . 141

7.2.2 Definition of outcome variables . . . 141

7.2.3 Independent variables . . . 142

7.2.4 Model fitting procedures . . . 143
7.2.5 Sample size considerations . . . 144
7.2.6 Model performance and internal validation . . . 144
7.3 Results . . . 146
7.3.1 Demographics . . . 146
7.3.2 Model 1: Prediction of postoperative Oxford Shoulder Scores . . . 148
7.3.3 Model 2: Prediction of early revision risk . . . 150
7.3.4 Model 3: Prediction of patients with a poor outcome . . . 156
7.3.5 Example scenarios . . . 159
7.4 Discussion . . . 159
7.4.1 Key findings . . . 159
7.4.2 Strengths and limitations . . . 160
7.4.3 Comparison to other studies . . . 162
7.4.4 Meaning of the study . . . 163
7.4.5 Unanswered questions and future research . . . 164

8 Variation in patient reported outcomes by implant type, age and rotator cuff status recorded in the National Joint Registry 166

8.1 Introduction . . . 166
8.2 Methods . . . 167
8.2.1 Data source and participants . . . 167
8.2.2 Interventions/comparisons . . . 168

8.2.3 Outcome measures . . . 168
8.2.4 Statistical analysis . . . 168
8.2.5 Sample size . . . 169
8.3 Results . . . 169
8.3.1 Demographics . . . 169
8.3.2 Implant choice by age and rotator cuff status . . . 171
8.3.3 PROMs . . . 171
8.3.4 Revision risk at two years . . . 174
8.4 Discussion . . . 176
8.4.1 Key findings . . . 176
8.4.2 Strengths and limitations . . . 176
8.4.3 Comparison to other studies . . . 178
8.4.4 Meaning of the study . . . 179
8.4.5 Unanswered questions and future research . . . 180

9 Conclusions **182**

9.1 Summary of findings . . . 183
9.1.1 Evidence from trials . . . 183
9.1.2 Trends in practice . . . 183
9.1.3 Prediction of patient reported outcomes . . . 184
9.1.4 Long-term revision risk . . . 185

9.1.5 Serious adverse events . . . 185

9.2 Implications for practice . . . 186

9.3 Limitations and future work . . . 187

9.3.1 Scope for future trials . . . 187

9.3.2 Reducing serious adverse events . . . 188

9.3.3 Improving data quality . . . 188

9.3.4 Dissemination of findings . . . 189

A Appendices **190**

A.1 Grants, Awards, Publications, Presentations . . . 190

A.1.1 Related to this thesis . . . 190

A.1.2 Additional projects and collaborations . . . 192

A.2 Cochrane review additional materials . . . 194

A.2.1 Hierarchy of outcome measures . . . 194

A.2.2 Characteristics of excluded studies . . . 195

A.2.3 Characteristics of ongoing studies . . . 195

A.2.4 Additional study data . . . 199

A.2.5 Search strategy . . . 202

A.3 Approvals . . . 202

A.3.1 Approvals: HES studies sponsorship letter . . . 203

A.3.2 Approvals: HES studies ethics waiver . . . 204

A.3.3 Approvals: NJR studies sponsorship letter 205

A.3.4 Approvals: NJR studies ethics waiver 206

A.3.5 Acknowledgement and disclaimer statement 207

A.4 HES database code lists 208

A.5 NJR database: MDS forms, component classification and PROMs collection . . 209

A.5.1 NJR MDS form, version 6: Primary procedures 210

A.5.2 NJR MDS form, version 7: Primary procedures 214

A.5.3 Component database reclassification 219

A.5.4 PROMs completeness and timing of collection 222

A.6 Revision: additional materials 225

A.6.1 Calculation of lifetime risk of revision 225

A.6.2 Table of lifetime risk estimates 226

A.6.3 Sensitivity analyses accounting for death as a competing risk 227

A.7 Serious adverse events: additional data 228

A.7.1 Age and Charlson score distributions 228

A.7.2 Risks of serious adverse events by age and sex 230

A.7.3 Model equation for risks of any serious advent within 30 days 232

A.8 Prediction model equations 234

A.9 PROMS: additional materials 236

A.9.1 Distribution of Q1 OSS by rotator cuff status and implant type 236

Bibliography **237**

List of Figures

1.1 Published trends in incidence of shoulder replacement surgery across seven national registries 2

1.2 Replica of original Péan prosthesis 11

1.3 Plain radiograph showing glenoid loosening 13

1.4 Different designs of glenoid component 14

1.5 Stemless TSR with modern bearing components 16

1.6 Classification of main shoulder replacement constructs 18

2.1 PRISMA flowchart for study screening and inclusion 37

2.2 Summary of risk of bias judgements 44

2.3 Forest plot of pain for TSR vs hemiarthroplasty 47

2.4 Forest plot of disability/function for TSR vs hemiarthroplasty 47

2.5 Forest plot of the risk of revision for TSR vs hemiarthroplasty 48

2.6 Forest plot of the risk of revision for TSR using pegged vs keeled glenoid components 51

3.1 Data flow chart for HES studies 65

3.2 Data flow chart for NJR studies 70

3.3 Procedure types listed on the NJR minimum dataset forms with intended mapping between versions 75

4.1 Crude and standardised incidence of primary shoulder replacement surgery . . 81

4.2 Standardised incidence of primary shoulder replacement surgery: elective vs trauma 81

4.3 Age- and sex-specific incidence of primary elective shoulder replacement surgery 82

4.4 Age- and sex-specific incidence of primary shoulder replacement surgery for acute trauma 82

4.5 Annual volume of primary elective shoulder replacements by procedure type . . 83

4.6 Annual volume of primary shoulder replacements for acute trauma by procedure type 83

5.1 Age- and sex-specific lifetime risks of revision or reoperation following primary elective shoulder replacement surgery 94

5.2 Unadjusted risk of revision or reoperation following primary shoulder replacement 97

5.3 Age- and sex-specific risk of revision following elective shoulder replacement . 99

5.4 Age- and sex-specific risk of revision following shoulder replacement for acute trauma 100

5.5 Risk of revision by indication for primary shoulder replacement 101

5.6 Risk of revision by type of primary shoulder replacement 103

5.7 Temporal trends in revision risk by year of primary surgery 105

5.8 Temporal trends in revision risk by year of primary surgery and procedure type 106

6.1 Cumulative risk of death up to 15 years following primary shoulder replacement 118

6.2 Cumulative risk of death up to one year following primary shoulder replacement 119

6.3 Rates of serious adverse events within one year following elective shoulder replacement surgery 123

6.4 Rates of serious adverse events within one year following shoulder replacement surgery for acute trauma 124

6.5 Predicted probability of any adverse event 126

6.6 Temporal trends in 30-day mortality 128

6.7 Temporal trends in serious adverse events within 30 days 128

6.8 Temporal trends in mean age at time of surgery 129

6.9 Temporal trends in mean Charlson Index for patients undergoing primary shoulder replacement surgery 129

7.1 Model 1: Performance plots 149

7.2 Model 1: Partial-effects plot for predicted Q2 OSS 150

7.3 Model 2A: Performance plots 153

7.4 Model 2B: Performance plots 154

7.5 Model 2C: Performance plots 155

7.6 Model 3: Partial-effects plot for predicted probability of a poor outcome 157

7.7 Model 3: Performance plots 158

8.1 Age distributions by state of rotator cuff and type of surgery performed 173

8.2 Predicted Q2 OSS by rotator cuff status, age and procedure type 173

A.1 Proportion of elective procedures with a valid linked Q1 OSS 222

A.2 Proportion of acute trauma procedures with a valid linked Q1 OSS 222

A.3 Proportion of elective procedures with a valid linked Q2 OSS 223

A.4 Proportion of acute trauma procedures with a valid linked Q2 OSS 223

A.5 Timing of Q2 score completion in relation to date of surgery 224

A.6 Histogram of reported Q1 OSS scores for the acute trauma cohort 224

A.7 Density plot for age of patients in the HES cohort 228

A.8 Density plot for Charlson Index of patients in the HES cohort 229

A.9 Q1 OSS distributions by state of rotator cuff and type of surgery performed . . 236

List of Tables

2.1 Characteristics of the included studies 39

2.2 Study participant demographics 40

2.3 Summary of findings: Conventional TSR versus hemiarthroplasty 46

3.1 Description of variables recorded on the NJR minimum dataset forms 71

5.1 Demographics of HES cohort 93

5.2 Risk of revision surgery up to 19 years 96

5.3 Risk of reoperation for infection 102

6.1 Association of serious adverse events with 90-day mortality 120

6.2 Rates of serious adverse events within 30 and 90 days following surgery 121

7.1 Summary of previously published prediction models 139

7.2 Description of the variables used to fit prediction models 143

7.3 Model fitting procedures and performance measures 144

7.4 Modelling sample demographics 147

8.1 Demographics by rotator cuff status 170

8.2 Risk of revision at two years by implant type and rotator cuff status 175

A.4 List of OPCS-4 codes used to identify shoulder replacements in HES database . 208

A.5 List of OPCS-4 codes used to identify reoperations in HES database 209

A.6 List of ICD-10 codes used to identify serious adverse events in HES database . 209

A.7 Main component groups for classification 219

A.8 Examples of valid/invalid constructs 220

A.9 Procedure reclassification matrix 221

A.10 Lifetime risk of revision and reoperation by age and sex 226

A.11 Risk of revision surgery up to 19 years - KM vs CI 227

A.12 Risks of each serious adverse event for elective cohort 230

A.13 Risks of each serious adverse event for acute trauma cohort 231

A.14 Rates of complications at 30 days from NSQIP studies 233

Abbreviations

ACSNSQIP	American College of Surgeons National Surgical Quality Improvement Program
AOANJRR	Australian Orthopaedic Association National Joint Replacement Registry
APC	Admitted Patient Care
ASA	American Society of Anaesthesiologists Grade
ASES Score	American Shoulder and Elbow Surgeons Score
AUC	Area Under a Receiver Operating Characteristics Curve
AVN	Avascular necrosis
BESS	The British Elbow and Shoulder Society
CTA	Rotator cuff-tear arthropathy
DASH	Disabilities of the Arm, Shoulder and Hand
FDA	US Food and Drug Administration
HES	Hospital Episode Statistics
HQIP	The Healthcare Quality Improvement Partnership
ICD-10	WHO International Statistical Classification of Diseases and Related Health Problems, 10th revision
JLA	James Lind Alliance
MAR	Missing at Random
MCAR	Missing Completely at Random
MCID	Minimum clinically important difference
MDS	Minimum Dataset Form
NDORMS	Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences

NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NIS	Nationwide Inpatient Sample
NJR	National Joint Registry of England, Wales, Northern Ireland and the Isle of Man
NSAID	Non-steroidal anti-inflammatory drug
NZJR	New Zealand Joint Registry
OA	Osteoarthritis
ONS	Office for National Statistics
OPCS-4	Office for Population Censuses and Surveys Classification of Interventions and Procedures
OSS	Oxford Shoulder Score
PROM	Patient-reported outcome measure
Q1 OSS	Preoperative Oxford Shoulder Score
Q2 OSS	Postoperative Oxford Shoulder Score
RA	Rheumatoid arthritis
RCT	Randomised controlled trial
RTSR	Reverse polarity total shoulder replacement
SF-12	12-Item Short Form Health Survey
SST	Simple Shoulder Test
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
TSR	Conventional/anatomic total shoulder replacement
UCLA Score	University of California Los Angeles Shoulder Score
VAS	Visual Analogue Scale
WOOS Index	Western Ontario Osteoarthritis of the Shoulder Index

Chapter 1

Introduction

1.1 Background

For patients suffering with severe shoulder arthritis, there is not enough high-quality evidence to guide the decisions made with healthcare professionals.[1–3] Shoulder replacement surgery can result in considerable gains in both physical and mental quality of life.[4–6] It is rapidly becoming the treatment of choice for many of these patients but there is limited information on the individual expected benefits and risks to help them decide whether to undergo surgery. Discussion of the available evidence is an important component of any shared decision making process and a duty for healthcare professionals.[7,8] Increasingly, patients will expect that treatments are tailored to their individual circumstances.[9]

A nationally commissioned priority setting partnership published in 2016 identified the top ten uncertainties for shoulder conditions most in need of further research.[10] Within the list, several key questions relevant to shoulder replacement surgery were identified.

1. "For the main shoulder conditions of arthritis, frozen shoulder, impingement, rotator cuff tears and instability, can you predict which patients will do well with surgery to help them decide on whether to have surgery or not?"

2. "In patients with 3 and 4 part proximal humeral fractures, what is the long-term outcome of reverse total shoulder replacement compared to hemiarthroplasty?"

3. "In patients with shoulder arthritis, is a hemiarthroplasty or a total shoulder replacement or a reverse replacement most effective?"

Whilst they are performed far less frequently than joint replacements of the hip and knee,[11] worldwide trends universally demonstrate a rapid and ongoing annual increase in the absolute number and incidence of shoulder replacements performed (shown in figure 1.1).[12–14] Trends from UK registry data match these patterns but are only available for a relatively short period since 2012. Rates of surgery per capita have been considerably lower in the UK than in other developed nations.[13]

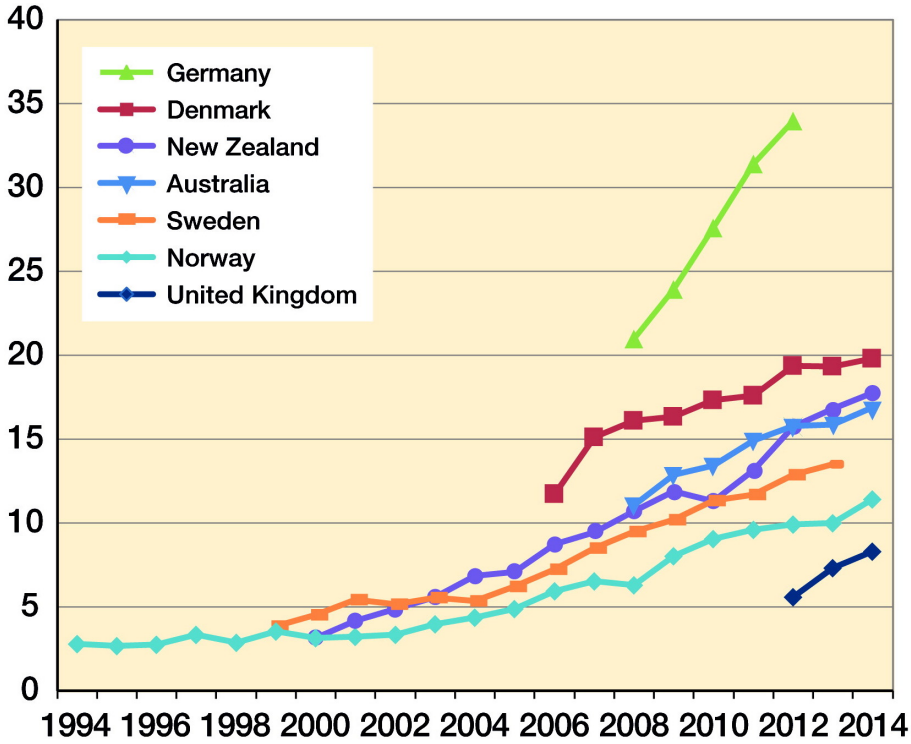


Figure 1.1: Trends in incidence of shoulder replacement surgery across seven national registries (/10^5 population). Reproduced from Lubbeke et al under CC BY-NC licence.[13]

Multiple factors may plausibly contribute to the rapid increase in annual caseload including an increasing population of older adults at risk, availability of innovative new surgical implant designs and larger numbers of fellowship trained shoulder surgeons. The apparent increased confidence in shoulder replacement as a treatment modality has occurred in the absence of robust level-one evidence. The studies identified for inclusion in a 2010 Cochrane review of shoulder surgery for osteoarthritis did not compare shoulder replacement to any other treatment modalities.[15] Nonetheless, shoulder replacement has become the standard treatment

for advanced glenohumeral joint pathologies.

1.2 Aims

The aim of this thesis is to provide better evidence to patients and clinicians on the benefits and risks of shoulder replacement surgery. The following key objectives address this aim.

1. To produce an updated appraisal of the current state of high-level evidence supporting shoulder replacement surgery practice.
2. To document current patterns and trends of practice, revision rates, adverse events, and patient reported outcomes in the UK using routinely collected national datasets.
3. To develop and validate models to predict postoperative outcomes following shoulder replacement surgery.

1.3 Indications for shoulder replacement surgery

Shoulder replacement surgery is an accepted treatment for end-stage glenohumeral arthropathies including osteoarthritis (OA), rotator cuff-tear arthropathy (CTA), inflammatory arthritis and displaced proximal humeral fractures.[16,17] The most common indications reported to the UK registry are OA and CTA, accounting for approximately 80% of procedures undertaken.[11] Symptomatic glenohumeral OA is thought to be relatively rare, however there is no population-based epidemiology to support this. Estimates are frequently inferred from rates of intervention which can introduce bias from socioeconomic influences, healthcare system factors and variable thresholds for the decision to treat. OA is the commonest indication for shoulder replacement reported internationally.[13]

CTA is a specific pattern of glenohumeral arthritis that occurs in the presence of a massive rotator cuff tear and was first described by Neer in 1983.[18] The rotator cuff plays an important role in maintaining a stable glenohumeral fulcrum.[19] In the presence of a full-thickness tear of

the supraspinatus with marked extension into the posterior cuff and/or the subscapularis, the anterior-posterior force couple is disrupted.[20] The humeral head migrates superiorly resulting in eccentric superior glenoid erosion and eventually articulation between the humeral head and the undersurface of the acromion.

The management and prognosis for many patients with inflammatory arthritides, particularly rheumatoid arthritis (RA), have been greatly improved by the advent of biological disease modifying therapies.[21] Shoulder replacement remains an important option for patients with end-stage disease although there is already some evidence from Danish population studies that the frequency has fallen in this patient group.[22] Conversely, a large American study using the Nationwide Inpatient Sample found that rates of primary shoulder replacement may still be rising for patients with RA.[23] Notably, the concomitant recording of rotator cuff tears in their cohort had also increased and it is possible that cuff disease is the primary reason for surgery in a group of patients with otherwise well-controlled RA.

Avascular necrosis (AVN) of the humeral head develops in the presence of abnormal microcirculation in the subchondral bone, resulting in loss of cartilage, and progressive deformity and incongruity of the humeral head. It is the second most common site of AVN after the femoral head and the list of contributing causes is largely the same: post-traumatic, steroid use, alcohol misuse, smoking, haemoglobinopathies, vasculitis, Gaucher's disease and dysbarism.[24,25] AVN complicating proximal humeral fractures has a high conversion rate to shoulder replacement surgery.[26] Shoulder replacement provides considerable improvements in pain and range of movement for these patients.[27]

AVN is just one of the situations where shoulder replacement has a role following trauma. Post-traumatic sequelae cover a range of anatomical distortions which have have been classified by Pascal Boileau into four types: 1) humeral head collapse or necrosis with minimal tuberosity malunion, 2) locked dislocations or fracture-dislocations, 3) nonunions of the surgical neck, and 4) severe malunions of the tuberosities.[28] Results in this early series of conventional conventional/anatomic total shoulder replacements (TSRs) and hemiarthroplasties were disappointing for patients with type 3 and 4 sequelae. These cases may now be more reliably treated using reverse polarity total shoulder replacements (RTSRs), however the results

are are still inferior to type 1 sequelae treated with conventional TSR.[29]

There is a lack of consensus on the management of acute proximal humeral fractures.[30] The 2015 Cochrane review found no clinically important differences between surgery and non-operative management.[31] The surgical interventions tested against non-operative treatments were osteosynthesis or humeral hemiarthroplasty. One small randomised controlled trial (RCT) has demonstrated a clinically important benefit from RTSR compared to humeral hemiarthroplasty, but there was no non-operative arm in the study.[32] The PROFHER-2 multicentre RCT comparing hemiarthroplasty to RTSR to non-operative management of acute proximal humeral fractures is currently recruiting in the UK and is due to report in 2024 (ISRCTN76296703). Current pooled evidence can support a "watch and wait" policy in older patients where there may be little difference between the outcomes achieved with RTSR in either situation.[33] Some patterns of injury, including split humeral heads and fracture dislocations are generally considered clear indications for surgery and are excluded from published and ongoing RCTs.[30,34]

Patients with primary and secondary bone tumours present unique challenges in shoulder arthroplasty and further discussion is beyond the scope of this thesis. Both conventional implants and endoprostheses are are used with reasonable functional results, and reverse polarity designs are important for restoring function where the rotator cuff is resected or functionally incompetent.[35,36]

Tumours, post-trauma reconstruction, instability arthropathy, AVN, and glenoid dysplasia are all complex situations. The challenges include glenoid bone loss, tuberosity malunion, rotator cuff insufficiency, inadequacy of the soft tissue envelope, capsular contraction and scarring. A history of previous failed internal fixation may increase the risks of reoperation considerably.[37] The risks of intraoperative complications, including fracture, may be higher.[38] These complex situations often require specific expertise and shoulder prostheses with specific design features to accommodate the altered anatomy.[39] Randomised trials typically exclude these more complex patients. Due to the heterogeneity of both the patients and pathoanatomical problems, there is limited high-quality evidence to support treatment recommendations.

1.4 Interventions for glenohumeral arthroses

Whilst shoulder replacement surgery is one of the standard treatments for most of these end-stage glenohumeral arthropathies, many reasonable alternatives exist both for less advanced disease, and for those unfit for or keen to avoid major surgery.

1.4.1 Non-surgical interventions

US and UK guidelines on glenohumeral osteoarthritis are not proscriptive about the use of non-surgical treatments due to the low quality supporting evidence available.[1,3] In the absence of red-flag criteria for urgent referral, most patients treated in the National Health Service (NHS) will be offered low-risk, lower-cost interventions first, only being referred to secondary care if symptoms fail to improve. The inclusion criteria for the majority of clinical trials in shoulder surgery require a period of non-operative management to have occurred. The options include adequate analgesia, activity modification, physical therapy, therapeutic intra-articular injections, and nerve blocks.[3]

Analgesia

The WHO three-step pain ladder has become the ubiquitous foundation for most analgesia strategies.[40] The initial principles of this were sound for the primary purpose of improving palliation in cancer pain, however we are now seeing some of the problems of widening the use of opioid analgesia in chronic pain conditions including arthritides.[41] The benefits achieved from oral paracetamol for osteoarthritic pain are small, however due to its favourable adverse-event profile it remains a first-line analgesic.[42] Non-steroidal anti-inflammatory drugs (NSAIDs) have been shown to be more efficacious and preferred by more patients for osteoarthritic pain, but they remain second-line due to the increased risk of serious side effects including gastroduodenal bleeding and acute kidney injury.[43–45]

Weak and strong opioid analgesia is well established for the treatment of acute postoperative

and cancer pain.[46,47] In studies of tramadol in the management of osteoarthritic pain, there may be no important benefits compared to placebo.[48] The is no clear evidence for the efficacy of high-dose opioids in chronic non-cancer pain but serious adverse events occur in approximately 7.5% of patients on medium- to long-term therapy.[49,50] The 2016 Centers for Disease Control and Prevention (CDC) guidelines provide wide-ranging recommendations on the introduction, monitoring and alternatives to opioid treatment.[51]

Gabapentinoids (e.g. pregabalin, gabapentin) have an important role in the management of neuropathic pain and can be used in the short-term as part of multimodal postoperative pain strategies.[52] The evidence is strongest in the treatment of post-herpetic neuralgia and peripheral diabetic neuropathy, but more limited in other areas, including musculoskeletal neuropathic pain.[53] Like opioid analgesia, this class of drug has seen a rapid expansion in its use. Concerns regarding dependency and side effects prompted the the UK government to upgrade their status to controlled-class C.[54]

It is apparent that patients with short-lived mild-to-moderate pain may be appropriately managed with non-opioid analgesia to good effect. In patients with more severe, chronic and complex pain there needs to be an emphasis on identification of the root cause and, where possible, targeted treatment of that cause.

Physical therapy

Physical therapy is widely used as an umbrella term encompassing a number of different treatment approaches including simple advice and education regarding exercise/physical activity, supervised or structured exercise, passive mobilisation, laser treatments, pulsed electromagnetic field therapy and ultrasound. Some of these are effective for some shoulder conditions.[55] Measures to address physical activity and manual therapies are recommended in clinical practice guidelines for musculoskeletal pain.[56] Combined therapies may provide stable symptom control in the medium term for some patients.[57] Thresholds for pain, functional impairment and anatomical abnormality beyond which physical therapies may be less effective have not been described for shoulder arthritis.

Intra-articular injections

Intra-articular injections are typically considered in two classes: combined glucocorticoid steroids with local anaesthetic agents, or viscosupplements. A Cochrane review of corticosteroid injections for shoulder pain found that improvements in pain may be small and short-lived. The review did not include any RCTs in patients with arthritis.[58] Randomised controlled trials in knee osteoarthritis have shown that benefits at six weeks are not maintained beyond six months.[59] There is no shoulder-specific evidence to support or advise against these commonly used injections for osteoarthritis or cuff tear arthropathy.[60]

Viscosupplementation with intra-articular hyaluronic acid injections may have beneficial effects on the synovial fluid properties, chondrocyte turnover and pain relief in patients with osteoarthritis.[61] A recent systematic review of both randomised and non-randomised studies found that patients with glenohumeral OA treated with hyaluronic acid may have substantial improvements in pain. However, the effect attributable to hyaluronic acid is questionable since patients in the control arms also experienced comparable improvements.[62]

1.4.2 Non-arthroplasty surgical interventions

As with non-operative treatments, there is a lack of high-quality evidence to guide the use of non-arthroplasty surgery for arthritis.[15] The results of shoulder replacement surgery in younger patients aged less than 50 years can be unreliable, which may reflect that the indication for surgery in this group is rarely primary osteoarthritis.[63] One common cause of early secondary arthritis is shoulder instability. Younger more active patients may also place higher demands on their new shoulder replacement. Rates of revision arthroplasty surgery in this group are high (39% at ten years).[64]

Arthroscopic debridement, biological resurfacing, microfracture, and osteochondral transfer or allograft are amongst the options for cartilage defects. The choice depends on the size and location of the defect, and patient factors.[65] There is mounting evidence that the arthroscopic management for degenerative arthritis of other large joints may be less effective

than it has been believed to be. A number of randomised controlled studies of arthroscopic debridement versus sham surgery or physical therapy for knee OA or degenerative meniscal tears have demonstrated no clinically important benefits.[66–70] Millett and colleagues have introduced the concept of "Comprehensive Arthroscopic Management (CAM)" for patients with Kellgren-Lawrence grade 2, 3 or 4 glenohumeral OA.[71] Their technique emphasises the importance of systematically dealing with a number of key pathologies, including humeral osteoplasty, capsular release, axillary nerve decompression and biceps tenodesis. In a study of 47 younger patients undergoing CAM (mean age 52 years), 26% converted to a TSR within five years.[72] Those that did not convert to a TSR achieved substantial improvements in pain and functional outcome scores. The authors concluded that their method has a role in the treatment of younger patients and those wishing to avoid or delay joint replacement surgery. Their approach is somewhat more aggressive than that of other arthroscopic debridement studies.[73–75] In the absence of any level-one evidence, there is no clear consensus on the role of shoulder arthroscopy in the management of arthritis.[1,3]

For patients with bipolar arthritis (affecting both the humeral head and glenoid), arthroscopic interposition arthroplasty with either meniscal allograft or dermal collagen grafts is an option.[76] When biological treatments for the glenoid were combined with prosthetic humeral hemiarthroplasty or humeral head resurfacing procedures the results were poor and rates of reoperation unacceptable, regardless of whether meniscus, dermis or capsule was used as the interposition.[77] Better results may be achieved with hemiarthroplasty alone.[78] Reaming to achieve a concentric glenoid surface ("ream and run") is a more frequently performed companion procedure for hemiarthroplasty in the young patient with a biconcave glenoid.[79]

The management of massive irreparable cuff tears is controversial and worth mention since this is an evolving indication for RTSR, particularly in older patients.[80] A number of promising novel methods exist to manage massive rotator cuff tears in the absence of arthritic change, but the the best approach amongst these is not yet clear.[81] Amongst the primary goals of soft-tissue surgery is to depress and recentre the humeral head to maximise the mechanical advantage of the deltoid, achieving improved active forward flexion and abduction. The options include superior capsular reconstruction and subacromial balloon spacers. Initial case

series of subacromial balloon spacers show improved pain and function.[82] A multicentre UK RCT comparing arthroscopic debridement with or without balloon spacer is recruiting and is due to report at the end of 2022 (trial registration: ISRCTN17825590). Superior capsular reconstruction uses synthetic or allograft tissue to bridge between the superior glenoid and greater tuberosity of the humerus. The results from the technique designer and early adopting experts are encouraging and show some ability to reverse pseudoparalysis.[83,84] However, there are no high quality studies with a control arm.

For palliation of pain in patients with CTA who are unsuitable or opt against joint replacement, suprascapular nerve ablation can be performed under radiological guidance. The suprascapular nerve has sensory innervation to a large proportion of the glenohumeral joint caspsule. In a small series of twelve patients, pain improved from 62 to 39 on a 100mm Visual Analogue Scale (VAS) at six months.[85] Some patients who had a response at three months had begun to deteriorate again, therefore the effects may not be permanent. Notably, the suprascapular nerve also innervates the supraspinatus and infraspinatus muscles. Ablation is contraindicated in patients with an intact and functioning rotator cuff. The efficacy and safety of this method needs to be proven in a larger cohort of patients.

Arthrodesis may be used as a salvage procedure. The most common indication is limb paralysis from a brachial plexus injury.[86] Historically it had a role for arthroses prior to the development of reliable prosthetic replacements. Whilst the results in patients with upper limb paralysis were good, results in patients with glenohumeral arthroses were poor and persistent pain was a problem.[87] A contemporary series of 29 arthrodeses for mixed indications reported poor patient-reported outcome scores and high rates of moderate-to-severe persistent pain.[88] Excision arthroplasty may rarely be needed to manage a failed arthroplasty in the presence of poor bone stock and/or infection refractory to treatment. Functional results are poor and it has no role as a primary surgical procedure for shoulder pain except for the treatment of local severe infection.[89,90]

1.5 History and development of prosthetic shoulder replacements

Early shoulder prostheses

The first total joint replacement of any kind is attributed to the German surgeon Thermosticles Gluck, who pioneered the use of ivory implants in the knee, hip, ankle, wrist, elbow and shoulder in 1890.[91] However, the earliest available description of a joint replacement using man-made materials followed shortly afterwards in 1893, when Jules Emile Péan used a platinum and rubber device in a patient with gross destruction of the glenohumeral joint due to a severe tuberculous infection (figure 1.2).[92] This crude device provided good initial pain relief and some restoration of function before ultimately failing after two years due to the presence of persistent draining infection. A number of polymer humeral prostheses were tried in the 1950s but these largely failed due to high wear rates, foreign body reactions and breakages.[93–97] Silastic cup interposition suffered from high rates of dislocation and fragmentation.[98]

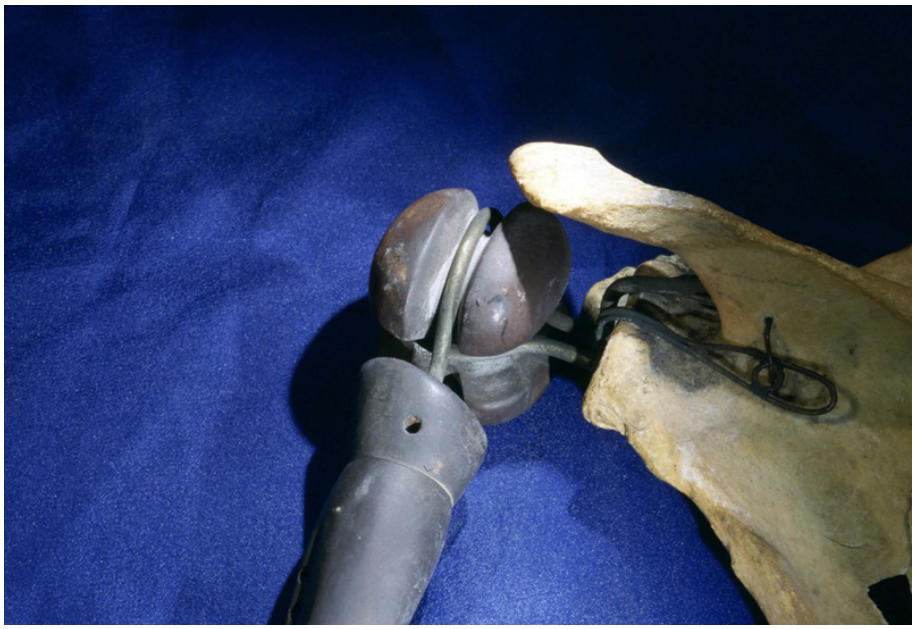


Figure 1.2: Replica of the original Péan prosthesis. (US Government photograph in public domain, via Wikimedia Commons)

Neer's vitallium prostheses

Substantial progress was made with the introduction of Krueger's vitallium prosthesis, which was the first modern humeral hemiarthroplasty moulded to the shape of a cadaveric humeral head and preserving the rotator cuff insertions.[99] Around the same time, Charles Neer was working to find a solution for the poor results of open reduction, excision arthroplasty or primary fusion for patients with proximal humeral fracture-dislocations.[100] His Neer I Vitallium hemiarthroplasty included an attachment for the tuberosities. Excellent or satisfactory results were achieved in 11 of 12 patients with acute and delayed treatment for proximal humeral fracture-dislocations or avascular necrosis.[101] Neer's prosthesis was iteratively improved and is the forerunner to the subsequent age of modern shoulder replacement surgery.

Neer's later series of 48 patients treated for glenohumeral OA only included one patient with replacement of the glenoid (total shoulder replacement) in a series of humeral hemiarthroplasties.[102] His observations were of good results with hemiarthroplasty alone in most patients, but he identified a subgroup of patients with flatter, more eroded glenoids where a total shoulder replacement could have a role to improve the fulcrum of the joint. The unconstrained device he used was not dissimilar to the common cemented implants that have followed. In 1982, he reported on 273 TSRs using five different designs of glenoid component including two with an extended articular surface, termed "semi-constrained", to reduce the extent of superior humeral head migration in patients with rotator cuff insufficiency.[103] Surgery and rehabilitation for such patients with rotator cuff insufficiency was described as "limited-goals", reflecting the challenge of restoring function in this setting when there is no fulcrum.

Glenoid component design

Glenoid component failure through aseptic loosening has been one of the primary weaknesses of shoulder replacement surgery (shown by the radiograph in figure 1.3). In a conventional polarity design, the compressive forces across the glenoid bearing and its interface with the bone are weak compared to lower limb arthroplasties loaded with body weight. These forces are also

more eccentrically directed. Conventional total shoulder replacements fail particularly rapidly in rotator cuff deficient shoulders as eccentric loading of the glenoid results in early accelerated wear and loosening, tipping it into a superior inclination (termed 'rocking horse' glenoid).[104] Radiological lucency around glenoid components has been extensively studied in the literature and may be present in upto 60% of cases. Estimates vary widely and standard radiological methods may not identify many at risk implants. Absolute failure rates with revision for component loosening are much lower, and so the clinical importance is not clear. However, a recent study by Schoch et al of 492 patients has identified a clear association between the presence of radiological lucencies and poorer patient-reported outcome measures (PROMs).[105]

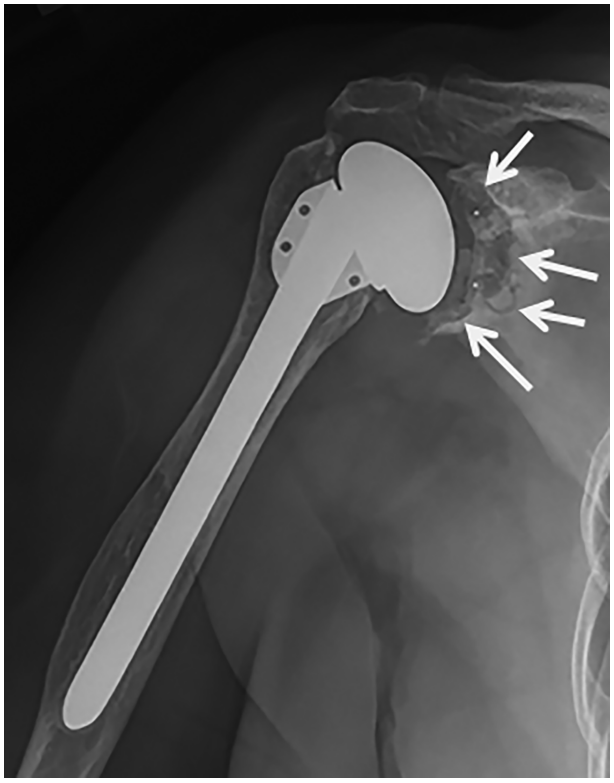


Figure 1.3: Plain radiograph showing advanced glenoid loosening with lucent areas around the component. Reproduced from McLendon et al under CC BY licence.
[106]

Neer's conventional polarity all-polyethylene glenoid bearing with a keel became widely used and copied. Figure 1.4 shows a range of contemporary glenoid components. Biomechanical and revision studies have shown that the cemented pegged polyethylene implants are superior to keeled implants, and they have become the standard of care.[107] There has been no high-level evidence to demonstrate an important patient reported difference between the two designs. Early uncemented metal-backed components suffered failures due to early loosening and

overstuffing of the joint.[108]

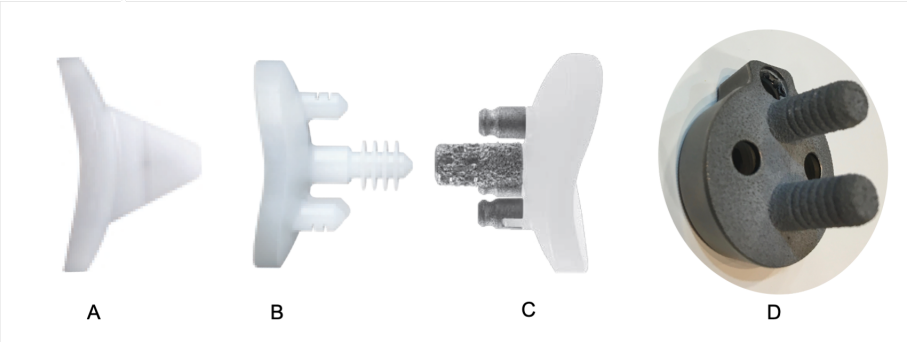


Figure 1.4: Different designs of glenoid component: A - All-polyethylene keeled, B - All-polyethylene pegged, C - Polyethylene with integrated metal pegs and cage, D - Calcium phosphate coated metal baseplate with pegs and screw holes (this example is used with a reverse bearing)

There is now renewed interest in the use of modular metal base plates. These offer the option for future conversion from a conventional to a reverse polarity replacement without the destruction required in a full glenoid revision. New coatings and highly porous materials, combined with powerful compression screws may improve the longevity of these implants. There remains little long term data to support this yet. Figure 1.4D shows a Calcium phosphate coated base plate with screw holes for use in reverse arthroplasty. The design principles are similar for newer conventional and universal base plates.

Fixed-fulcrum glenoids

Several early designers explored a fixed-fulcrum design philosophy to address the problem of rotator cuff deficiency.[109] Unfortunately, the increased constraint of these devices greatly increased the stresses at the bone-implant interface, particularly on the glenoid side, and these implants failed frequently by early glenoid loosening.[103,110,111]

Kessel and Bayley designed one of a number of replacements with a new reversed polarity philosophy, where a metal ball on a substantial central screw was placed on the glenoid side and a high-density polyethylene socket on the humeral side.[112] This, and other early constrained reverses, did not solve the problem of glenoid-side loosening.[113] The addition of a hydroxyapatite coating to the glenoid screw reduced the incidence of loosening somewhat but the range of active forward elevation achieved was modest.[114] The constrained reverse

bearing couple still has a clearly defined role to play in endoprosthetic replacement for salvage arthroplasty and tumour reconstruction[115]

Grammont's reverse prosthesis

A stepwise change in reverse shoulder replacement came when Grammont proposed the concept of medialisation and distalisation of the centre of rotation with an increased humeral neck-shaft angle. Grammont's "Delta" prosthesis is the prototype for most of the new reverse shoulder replacements that have followed.[116] Large single- and multi-centre series have demonstrated good functional results with Grammont-style prostheses.[117,118] Complications in these early series, including dislocation and infection, were common and the functional benefits seemed to wane over time. Some authors, therefore, recommended strict selection criteria and restriction of the procedure to older patients with lower functional expectations.[117–119]

The altered biomechanics of RTSRs may have largely circumvented the problems of early loosening associated with conventional repalcements. Such is the success of RTSR that some surgeons now advocate its use in patients with gross functional impairment due to a massive irreparable rotator cuff tear ('pseudoparalysis') even if there is no evidence of a painful arthropathy.[80] However, few recommend RTSR for younger high demand patients.

Humeral fixation

One of the other main areas where implant design diverges is in the extent of humeral resection and fixation in conventional polarity total shoulder replacement or hemiarthroplasty. Implants derived from Neer's gold standard typically require resection of the humeral head at the anatomical neck of the humerus, insertion of a humeral component with a stem entering the diaphysis of the humerus, and an integrated head with no modularity (monobloc) or a stem with a separate modular head. In resurfacing arthroplasties, minimal bone is resected from the diseased humeral head, maintaining the convex hemispherical profile and covering it with a metal cup. Similar results were reported using resurfacing hemiarthroplasties with the potential benefit of preserved humeral bone stock in the the event of further surgery.[120,121]

The minimal humeral resection performed as part of a resurfacing procedure provides limited access to the glenoid when undertaking a total shoulder replacement. Resurfacings also require reasonably well-preserved humeral head anatomy and may not be suitable in avascular necrosis or malunion cases. The introduction of "stemless" shoulder replacement offers a solution to some of these problems.[122] These implants require resection of the whole humeral head at the anatomic neck, as for a conventional stemmed humeral implant, and are fixed to the metaphyseal bone with one of a number of different short flanged or caged bodies. An example is shown in figure 1.5. Removal of the humeral head affords much better access to the glenoid whilst still avoiding violation of the medullary canal. The Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) refers to stemless implants as "mid-head" resection procedures.



Figure 1.5: Stemless TSR with ceramic humeral bearing and vitamin E enriched polyethylene glenoid component

Current innovations

The newest major innovation reported is the use of a pyrocarbon ball as a simple interposition arthroplasty. Its use is being explored as a bone-preserving option in younger patients with arthritis. It is currently only being used at a small number of centres until longer-term results are known.[123]

Overall, the number of shoulder replacement options marketed today is vast. Extensive modularity of both humeral and glenoid components provides the ability for surgeons to recreate individual anatomy more closely and to perform increasingly complex reconstructions in both primary and revision operations. Whilst not covered in detail in this section, shoulder replacements offer a wide range of options for cemented and uncemented fixation, and different

types of bearing surface material including metal, polyethylene, ceramic and pyrocarbon.

1.6 Classification of contemporary shoulder replacements

Replacements may be either a humeral hemiarthroplasty if just the humeral head is replaced, or a total shoulder replacement if both the humeral head and the glenoid are replaced. Whilst there are a number of interposition and biological resurfacing procedures described for the treatment of glenoid pathologies,[76] isolated replacement of the glenoid articulation with a prosthesis fixed to bone is not feasible due to our inability to access the glenoid face in the presence of an intact humeral head. Further key features are whether the bearing is conventional or reversed and whether the humeral component is stemmed, stemless or a resurfacing. The more recently developed prosthetic (pyrocarbon) ball interposition arthroplasty falls into its own category.

A simple schematic grid is shown in figure 1.6 and describes the major subgroups for different shoulder replacement constructs based on the bearing couple and the humeral component type. To be clear, a "construct" refers to shoulder replacement as a whole including both glenoid and humeral components. The construct is the basis for analyses within this thesis.

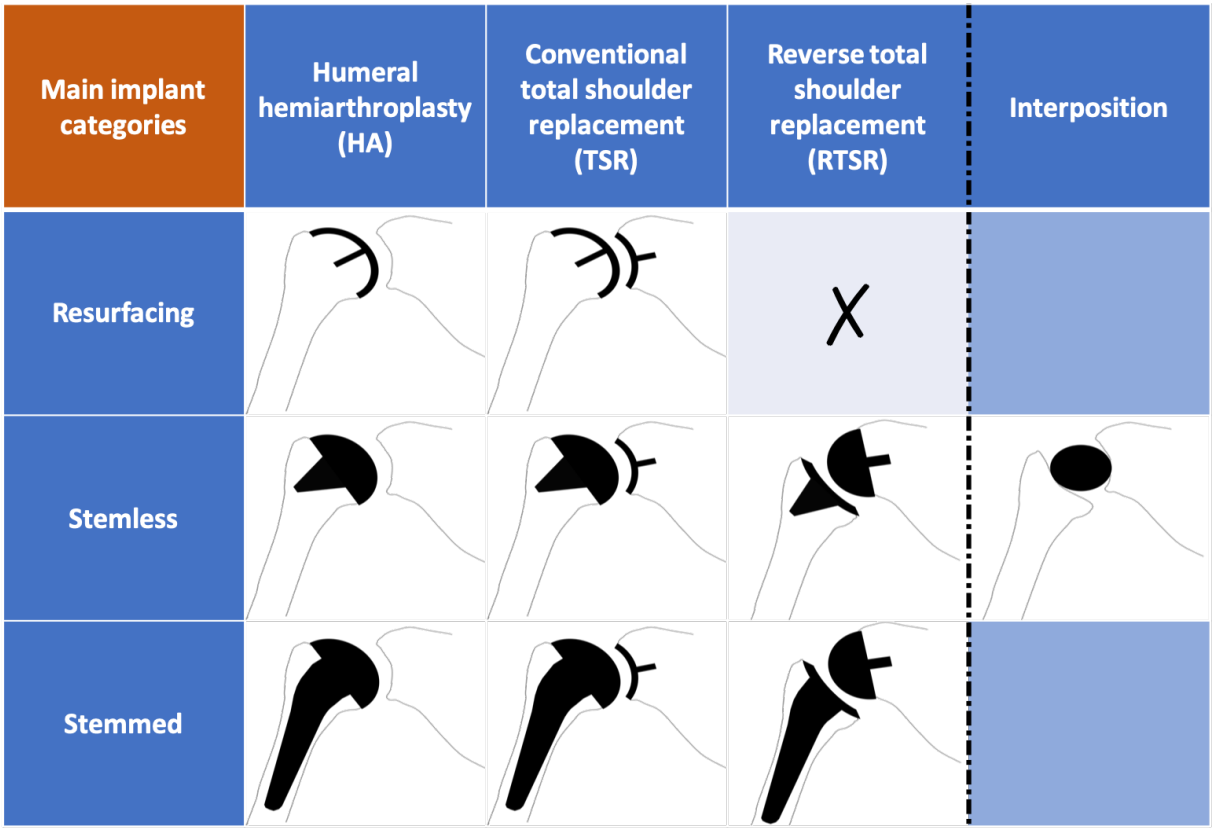


Figure 1.6: Classification of main shoulder replacement constructs

1.7 Evaluation of surgical interventions and innovations

Shoulder replacement surgery is an area of constant innovation and has led to the advances described already (section 1.5). Well-intended innovations can sometimes be unsuccessful and have the potential to cause considerable harm, illustrated clearly in orthopaedic surgery by the problems with metal-on-metal hip replacements.[125] The IDEAL framework (Idea, Development, Exploration, Assessment, Long-term study) describes a best-practice approach for the introduction of new technology and new techniques.[126] A crucial concept is that innovations should be adequately assessed and reported, and that eligible patient groups should be clearly defined before widespread use.

The levels of evidence described by Wright in 2003 are now widely used in orthopaedic

Some of the concepts in this section regarding evidence levels, RCTs and the role of registries were previously published as a BMJ opinion piece entitled: Providing better evidence in orthopaedic surgery: RCTs, registries, or both?[124]

journals.[127] The gold standard for proof of efficacy and effectiveness remains RCTs. Level-four (case series) and level-five (expert opinion) studies continue to be disproportionately represented in the literature on shoulder surgery. The UK surgical community is now at the centre of bucking this trend and has demonstrated the ability to deliver high-quality multicentre RCTs.[128–130]

Observational studies are conventionally considered inferior to RCTs and may produce a biased estimate of the true treatment effect, a stance based initially on a review by Sacks et al of RCTs versus studies with historical controls.[131] This largely holds to be true, but has been challenged by subsequent comparative analysis that found the results of RCTs were not systematically different to those from observational studies across five different treatment comparisons.[132] RCTs addressing the same comparison may report conflicting outcomes.[133] This may be due to small sample sizes or differences in the inclusion criteria and study interventions. RCTs should have robust internal validity but may lack generalisability. Observational studies do not necessarily avoid this.[134] Well-designed registries and databases of routine care are also essential to monitor longer-term outcomes, incidence of rare events and generalisability of outcomes from early studies to the population at large.

Arthroplasty registries are now well established across Europe, North America and Australasia. They may occupy a higher position than some other observational study designs since the datasets are prospectively managed, the data collection is defined at the outset, and they include large numbers of unselected patients treated in a wide range of settings by many different surgeons.[135] Registries play an important role in prognosis and risk factor studies, in the comparison of implants and surgeons, and in the long-term monitoring of outcomes not commonly reported by trials.[136] Ideally, we would know every planned question and every required data field at the inception of the registry to maximise the number of answerable research questions. There are practical restrictions on the quantity of data collection that is feasible and registries have minimum dataset lists to ensure that the most important variables are reliably collected.[135]

Large-scale administrative (claims) datasets are distinct from registries in that their primary purpose is typically not the study of outcomes, rather the collection of data for remuneration

or resource management.[137] As such, the quality and granularity of data is usually considered to be inferior.[138] However, some important but rare outcomes can only be studied by looking at these datasets where the volume and coverage is sufficient to capture rare events relevant to patient safety.[139] The signal from registry and administrative data can also help to inform areas for further prospectively designed research.

1.8 Measuring outcomes in shoulder surgery

Outcomes of medical and surgical interventions are typically considered in terms of their potential benefits and possible risks. In joint replacement surgery, the standard measure of effect has long been the rate of implant revision/survival, measured using time-to-event survival analysis techniques.[140] Subjective surgeon-perceived success is no longer an acceptable or useful metric.[141] There has rightly been an increasing emphasis placed on outcomes relevant to patients, including validated condition-specific PROMs.

1.8.1 Core outcome sets

The OMERACT group (Outcome Measures in Rheumatology, https://omeract.org) set up a Shoulder Core Outcome Set Special Interest Group in 2016 to establish a core outcome set for clinical trials of treatments for shoulder disorders.[142] The inner domain of most important outcomes comprised 1) pain, 2) physical function/activity, 3) "patient global - shoulder", and 4) adverse events including death.[143] Whilst agreed upon for RCTs, there is no reason why observational studies should not also report outcomes within these domains where available. The supporting systematic review for the process identified 319 different outcome measurement instruments from 409 trials. There were 47 different instruments to measure pain and 49 different instruments to measure function.[144] Future steps in the core outcome process will identify which measurement tools meet OMERACT criteria for endorsement.

1.8.2 Commonly used shoulder-specific outcome measures

The tools reported within the international literature show some geographic preferences related towards the tools developed locally and/or endorsed by national bodies. The most frequently cited score is the Constant-Murley score.[145] The Constant-Murley score combines patient-reported scores, and physician-evaluated measures of strength and range of motion. The University of California Los Angeles Shoulder Score (UCLA Score) score combines scores for patient satisfaction, range of movement, strength, pain and function.[146] Whilst both have been validated and are widely used, there is no clear rationale for the weighting of patient-reported and physician-evaluated components of these scores. The inclusion of a physician-evaluated component may increase the potential for assessors to introduce bias into the recording process.

In the UK, the most commonly used shoulder outcome measure is the Oxford Shoulder Score (OSS).[147] The OSS is a validated twelve-item questionnaire comprising pain and function domains with a minimum score of 0 points and maximum of 48 points, which equates to no pain and normal function.[148,149] It may be separated into pain and function subscales, although there is little precedent for this.[150] The OSS was developed for use in a range of shoulder conditions including impingement, rotator cuff disorders and arthritis. Both the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man (NJR) and New Zealand Joint Registry (NZJR) already collect the OSS, and the AOANJRR has also selected the OSS for future collection, together with additional quality-of-life measures.

The Western Ontario Osteoarthritis of the Shoulder Index (WOOS Index) was specifically developed for use in patients with shoulder osteoarthritis and hence advocated for use in shoulder replacement studies.[151] The original 19-question scale is usually translated to a 0 to 100 scale, where a score of 100 equates to normal shoulder function and no pain. The WOOS Index is widely used in North America. Numerous validated translations exist and it is the instrument of choice for both the Danish and Swedish shoulder registries.

Most studies from the United States use the American Shoulder and Elbow Surgeons Score (ASES Score), which is also a scale from 0 to 100 (100 equals normal function and no pain). Whilst it is reliable, responsive and has high construct validity, there are some items that may

be less relevant to the target population for shoulder replacement surgery, particularly those regarding throwing and sports participation.[152,153] There is a physician-evaluated section, but this does not contribute to the calculation of the 100 point score and so the ASES Score can still be considered a patient-reported score. The Simple Shoulder Test (SST) shares some of the limitations of the ASES Score due to the inclusion of North American specific activities.[154]

The Disabilities of the Arm, Shoulder and Hand (DASH) score considers the function of the upper limb as a whole.[155] It is rarely used as the primary outcome measure in shoulder replacement studies but sometimes appears as a secondary outcome measure.

Although the quality of the score development and validation processes vary, these scores all broadly try to measure the same domain of pain and function. The OSS and ASES Score are strongly correlated and one can be reliably predicted by the other.[156] The Constant-Murley score and OSS are also correlated, shown by a slightly less reliable association in a population of patients with proximal humeral fractures.[157]

1.8.3 Mortality

Mortality rates reported by the NJR are low. Within 90 days the rate after elective surgery is 2.7 per 1000 (95% CI 2.1 to 3.3), which is lower than the rate for elective total hip replacements. However, the rate following trauma is 17 per 1000 (95% CI 12 to 22). This much higher rate probably reflects different age, sex and comorbidity profiles, but these details are not provided in the NJR report.[11] Similar rates have been reported by the Danish Shoulder Arthroplasty Registry.[158]

1.8.4 Revision surgery

Failure of joint replacement surgery, including shoulder replacement, is usually identified by the occurrence of a revision operation and reported using a time-to-event survival analysis.[140] It is important to identify clearly what is meant by the endpoint of "revision", which has not always been consistently applied by international registries.[159] For the purposes of shoulder

replacement surgery, the UK NJR considers a revision procedure to be "the further operation to remove, change, add or reduce any of the components".[160] Revision rates are now widely reported by international registries. The NJR currently has data up to six years following surgery.[161] Studies of long-term outcomes are limited to large institutional series with 20-year follow-up from the Mayo clinic.[162,163] A systematic review and meta-analysis to collate ten-year revision rates from all available published and registry sources is currently in progress (PROSPERO registration: CRD42019140221).

There is a need for estimates that are better tailored to individual patient circumstances. The published series are not large enough to reliably inform patients on simple age- and sex-specific subgroup differences.

1.8.5 Adverse events

Adverse events following surgery are usually classified as serious or specific, and systemic or local. For the purposes of the US Food and Drug Administration (FDA), a serious adverse event is reportable if it results in a threat to life or death, admission or increased length of hospital stay, permanent disability or damage, or treatment to prevent such damage.[164] This is the standard applied to clinical trials and is adopted by the National Institute for Health Research (NIHR).[165] This is a very broad definition when applied to outcomes in routine practice and encompasses a number of causes for extended hospital stays including common hospital acquired infections.

Most large adverse event studies have reported on rates of venous thromboembolism and myocardial infarction. The most recent estimates from large US-based database studies range from 0.1% to 1.4% for pulmonary embolism and 0.1% to 1% for myocardial infarction.[166–171] Transferring these statistics to UK practice is not straightforward. The period for event capture in these studies varies and mostly includes only those occurring prior to initial discharge. Rates are also pooled for elective and trauma cases, and calculation of an estimate for absolute risk to share with individual patients from the published material is not user-friendly. The highest rates were reported in the study by Cancienne et al.[170] Over 40% of these patients had a BMI of >30kg/m^2 and the mean Charlson score was 4.85, which is alarmingly high and almost certainly

not representative of current secondary care shoulder referrals in the NHS.

A study using Hospital Episode Statistics (HES) data from 2005 to 2008 reported 90-day rates for pulmonary embolism of 0.20% following TSR, 0.06% following elective hemiarthroplasty and 0.52% following hemiarthroplasty for trauma.[172] For these rare events, the study was somewhat underpowered for these differences to be reliable and no stratification was published for other patient factors more plausibly related to incidence.

In a more recent cohort study of 125 766 patients undergoing TSR from the US Nationwide Inpatient Sample, a broader range of systemic complications was captured, including common hospital-acquired infections, pulmonary emboli, cardiac, renal and gastrointestinal complications. The overall estimate for any complication was 6.7%.[173] An association of complications with length of stay and subsequent mortality was noted. With no follow-up beyond discharge for short-stay procedures, it is likely that the reported rates underestimate the true frequencies of some conditions.

Local specific adverse events are also poorly defined. A Delphi consensus has been performed for the complications of rotator cuff repair but no such consensus exists for shoulder arthroplasty.[174] Thresholds and categories in for reporting specific adverse events are very variable and highly subjective. The NJR captures a number of intraoperative complications including fracture of the humerus or glenoid and vascular injury (see the Minimum Dataset Form (MDS) forms in appendix A.5. Nerve injuries were previously collected, however these are rarely identifiable at the time of surgery and have now been removed from the MDS forms (version 7). The rate of intraoperative fracture in the NJR data has been reported to be 1.6%.[175]

1.9 Published clinical prediction models in shoulder replacement surgery

Whilst several authors have used regression modelling to describe factors "predictive" of postoperative outcomes, there are few useful published models designed, validated and calibrated with the specific purpose of predicting outcomes on an individual per-patient basis.

Those available have been developed on small, homogeneous, single-institution datasets and are unlikely to perform as well when applied in other settings. Full model equations, including the intercept, are rarely provided.

Simmen et al developed a logistic regression model on a group of 140 patients to predict the likelihood of achieving an arbitrary threshold of 80 points on the Constant-Murley Score, which identifies approximately the top third of patients.[176] The predictions were reported to be accurate (Area Under a Receiver Operating Characteristics Curve (AUC)=0.79). Matsen at al developed a logistic model to predict an arbitrary improvement of at least 30% of maximum improvement on the Simple Shoulder Test.[177] Similarly, predictions were shown to be accurate (AUC=0.79) but the model development process included multiple steps likely to introduce considerable overfitting to the characteristics of their small development sample. Patients with lower starting PROMs may achieve lower absolute postoperative scores but gain greater benefit. One study has suggested thresholds for the ASES Score and 12-Item Short Form Health Survey (SF-12) that may predict the likelihood of achieving a clinically important improvement.[178]

Machine learning has been heralded by some as the answer to many of our prediction and classification questions. Multiple machine learning algorithms were trialled in one paper to identify those most at risk of adverse events, but these only achieved an AUC of 0.71 at best.[179] High-throughput PROMs such as the Patient Reported Outcome Measurement Information System (PROMIS) are not yet widely used. Models using PROMIS within multivariable models to predict physical functioning and pain interference had remarkably good AUCs (0.70 and 0.87 respectively).[180] However, the endpoint used to validate these measures was a "90% chance of achieving the MCID" which is not intuitive and the AUC is certainly overoptimistic and biased due to an inadequate sample size and through the flawed approach of using the preoperative score to predict a change score.

1.10 Factors influencing the outcome of shoulder replacement surgery

Many more studies have reported factors found to influence outcomes. These vary from one study to another depending on the study inclusion criteria, patient population characteristics, data availability and thresholds chosen to determine statistical or clinical importance.

Postoperative outcome scores or satisfaction may be correlated with preoperative PROMs, sex, comorbidities, prior surgery, glenoid morphology, indication for surgery, humeral head migration, rotator cuff status, range of motion, dominance of operated arm, diabetes, depression and preoperative expectations.[176,181–185]

The likelihood of revision surgery may be correlated with age, sex, BMI, rotator cuff disease, comorbidities, indication and implant choice.[162,163,186–192]

The likelihood of experiencing an adverse event or requiring unplanned readmission may be correlated with age, sex, comorbidities, hospital case volume, insurance status, workers compensation, depression, prolonged surgery, steroid use and implant choice.[39,168,169,187–189,193–202]

1.11 Preoperative decision making

Shoulder replacement surgery, and joint replacement surgery more widely, is a preference-sensitive treatment influenced by both patient and surgeon factors.[203] A balance must be struck between the likely benefits and risks of the proposed surgery. There has been a considerable shift away from paternalistic decision making by surgeons towards a shared decision making model with patients. Offering an individualised approach to treatment is an expected standard of NHS care.[204] The process of obtaining written consent for surgery is just one step in what should be a continuous process of information sharing and discussion to reach a decision agreed by both clinician and patient. Clinicians have a duty to inform these discussions

with relevant evidence on the risks, benefits and alternatives to surgery in terms that are easily understood.[7,8]

The context in which risk and benefit are framed is an important source of bias on patient decision making. In a randomised study of patients with rotator cuff tears, patients were either told that there was a 71% chance that the cuff would remain healed after surgery or that there was a 29% chance that the cuff would retear within two years of surgery. 84% of patients in the first ("benefit") group accepted surgery compared to only 46% in the second ("risk") group.[205] Presentation of relative risks rather than absolute risks can also change how patients perceive those risks, and absolute risks may be preferred.[206,207] Regardless of how risks are presented, they should ultimately be underpinned by reliable, generalisable estimates from a population that is adequately matched to the patient considering treatment. Where information permits, those estimates should be tailored closely to their specific characteristics.

1.12 Scope and structure of the thesis

Joint replacement surgery is generally successful and results in considerable quality-of-life benefits, but we need to appreciate that not all treatments work in all patients.[208] Identification of those most and least likely to benefit from shoulder joint replacement surgery allows us to manage patient expectations better and to channel available resources into the most appropriate strategies for that individual. The focus of this thesis is to improve our understanding of the benefits and risks of surgery in an elective context, and to provide information that may inform and enrich discussions between patients and clinicians at an individual level. Since the primary data sources are routinely collected datasets for patients in the UK, the estimates will be most relevant and reliable for NHS treated patients. Where data was available for patients undergoing surgery following acute trauma, a parallel analysis is included to provide useful contrast with the outcomes of elective surgery.

Chapter 2 is a Cochrane systematic review appraising the current state of evidence from randomised controlled trials informing the use of shoulder replacement surgery for osteoarthritis and rotator cuff tear arthropathy.

Chapter 3 appraises the quality of the data sources and outlines the data cleaning methods.

Chapters 4, 5 and 6 use HES data to explore long-term trends and outcomes following shoulder replacement surgery. Chapter 4 presents the incidence of shoulder replacement surgery in England over a 19-year period. Age-, sex- and implant-specific rates are explored and explanations for trends over time are discussed. Chapter 5 provides age- and sex-stratified estimates for the long-term risk of both revision surgery and other additional reoperations following a primary shoulder replacement. Revision risk is also presented in the context of patient life-expectancy (lifetime risk) using methods adapted from cancer epidemiology. Chapter 6 provides estimates of the incidence of serious systemic adverse events within 30 and 90 days of surgery. Estimates are adjusted for age, sex and preoperative comorbidities using multiple logistic regression.

Chapters 7 and 8 use data from the NJR including a pilot sample of postoperative PROMs. In Chapter 7, a series of prediction models are designed and internally validated to separately predict the likely postoperative OSS, the risk of needing early revision surgery, and the likelihood of having a poor overall outcome defined by a worse postoperative OSS and/or needing revision surgery within one year. Methods are employed to robustly account for missing PROMs data and to reduce the risk of model overfitting and bias. Finally, Chapter 8 explores the influence of implant type on PROMs in patients with OA or CTA. Postoperative and change scores are modelled across the full age-range, taking into account the effect of rotator cuff status on implant choice and outcomes.

RCTs, registry and administrative data sources all have a role to play in addressing specific questions and their relative merits and limitations will be considered. Implant registries and administrative records are living datasets subject to regular internal and independent analyses. Insights from this thesis relevant to these future analyses will be discussed.

Chapter 2

Cochrane systematic review of shoulder replacement surgery for osteoarthritis and cuff tear arthropathy

2.1 Introduction

Nationally commissioned priority setting partnerships and guideline setting committees have previously noted a lack of adequate evidence to guide treatment decisions in shoulder osteoarthritis.[10,209] Cochrane systematic reviews uphold high levels of methodological rigour and are recognised by many to be the gold standard in evidence-based healthcare.[210] The last Cochrane review in this field was published in 2010 and considered all surgical interventions for shoulder osteoarthritis.[15] Studies discussed in chapter 1 demonstrated an increase in the number of shoulder replacements performed internationally, despite this lack of evidence.

Much of the expansion in shoulder replacement surgery may be driven by the use of RTSRs. None of the studies identified by Singh et al included patients treated with RTSRs.[15] To address this gap in the previous review process, this chapter uses Cochrane systematic review

This chapter has been conducted as a Cochrane Systematic Review. A full, unabridged version of the manuscript has been peer-reviewed and submitted for final editorial review by the Cochrane Musculoskeletal Group. Some sections contains standardised Cochrane text.

methodology to produce an updated appraisal of the best available evidence in shoulder replacement surgery. The inclusion criteria were explicitly expanded to capture patients undergoing reverse total shoulder replacement and patients with a primary diagnosis of rotator cuff tear arthropathy.

2.2 Methods

This systematic review was performed and reported according to the Methodological Expectations of Cochrane Intervention Reviews.[211] The full study protocol was approved by the Cochrane Musculoskeletal group and is published in the Cochrane Database of Systematic Reviews.[212]

2.2.1 Inclusion/exclusion criteria

Types of studies

Only RCTs were considered for inclusion. Non-randomised and quasi-randomised studies were excluded. No articles required translation. Published and unpublished data were considered.

Participants

Study participants were adults (aged 18 years and over) with primary glenohumeral OA or arthritis secondary to CTA. Studies of adults undergoing surgery for inflammatory arthritis, benign or malignant tumours, adhesive capsulitis, shoulder instability or fractures were excluded.

Interventions and comparisons

Studies with any type of shoulder replacement in one study arm were included. The comparisons of interest included placebo, no treatment, non-surgical treatment, other types of

surgery, other types of shoulder replacement surgery or technical modifications.

Types of outcome measures

Based on the preliminary core domain set described by the OMERACT (Outcome Measures in Rheumatology) Special Interest Group, the following major outcomes were measured: 1. Pain, 2. Disability/function, 3. Serious adverse events (death, or those requiring hospitalisation), 4. Specific (including shoulder stiffness, instability, infection, and nerve damage), 5. Quality of life, 6. Revision or other reoperation, including treatment failure, 7. Patient-perceived success of treatment, 8. Physician-evaluated outcomes, including radiographic assessment of lucency.[142,143]

Several functional outcome scores (e.g. Constant-Murley Score) are often reported as sub-domains including pain. The validity and interpretation of individual subscales are not well defined and they have not been included in any meta-analyses. Measured range of motion and strength are considered to be of low utility, and these outcomes were not analysed separately beyond their contribution to a validated disability/function score.[142] Treatment failure was reported in some studies as patients who required revision surgery but either declined it or were unfit.

Outcome data were collected for the following time points: short-term (less than one year), intermediate (one to three years) and long-term (more than three years). The intermediate time point was used as the primary time point for comparisons.

2.2.2 Electronic searches

The following databases were searched from inception to 31st January 2019 with no language restrictions: The Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, CINAHL, SportDiscus, Web of Science, ClinicalTrials.gov (www.ClinicalTrials.gov) and the WHO International Clinical Trials Registry Platform (www.who.int/ictrp/en/). National and international society websites were also checked for

relevant conference abstracts. Detailed search strategies are included in appendix A.2.

2.2.3 Data collection and analysis

Selection of studies and data extraction

Studies identified by the searches were uploaded into the online review manager software, Covidence (https://www.covidence.org/).[213] Duplicate entries were removed and multiple reports of the same study were collated so that each study, rather than each report, was the unit of interest in the review. All steps of study selection, risk of bias assessment and data extraction were carried out independently by two reviewers. Disagreements were resolved by consensus and where consensus was not achieved initially, a third reviewer acted as an adjudicator. Titles and abstracts of studies were checked and those not relevant were discarded. Relevant full-text versions were obtained and fully reviewed to determine which studies should be included and have data extracted. A study data collection form was used to extract data relating to the study design, participants, interventions, comparisons, outcomes, and potential conflicts of interest.

Assessment of risk of bias

The risk of bias for each study was assessed using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions.[214] Risk of bias was summarised as high, unclear or low across nine domains and presented as summary figures.

Handling of outcome data

The number of participants per treatment group was extracted for all outcomes. The number of events was extracted for dichotomous outcomes, and the mean and standard deviation for continuous outcomes. Numerical data presented only in figures or graphs was first requested from the study authors, and when not available it was extracted from the plots using WebPlotDigitizer (https://automeris.io/WebPlotDigitizer/).[215] Final values were preferred

to change from baseline values, and unadjusted values were preferred to adjusted values. If more than one outcome measure was reported in a trial, they were prioritised based on the hierarchy of major outcomes listed in appendix A.2. Where possible, data were extracted based on intention-to-treat analyses.

Main planned comparisons

The main planned comparisons were:

1. Any type of shoulder replacement surgery versus placebo (sham-surgery)
2. Any type of shoulder replacement surgery versus any other type of surgery
3. Any type of shoulder replacement versus any type of non-surgical treatment
4. Any one type of shoulder replacement surgery versus any other type of shoulder replacement surgery
5. Any one type of shoulder replacement surgical technique versus any other type of shoulder replacement surgical technique (e.g. cemented versus uncemented implants)

Measures of treatment effect

Dichotomous data were analysed as risk ratios (RR), or Peto odds ratios (OR) when the outcome was a rare event (approximately less than 10%), and used 95% CIs. Continuous data were analysed as mean difference (MD) or standardised mean difference (SMD), with 95% CIs, depending on whether the outcome was measured using the same scale or different scales. When different scales were used to measure the same conceptual outcome, the SMD was back-translated to a typical scale (e.g. 0 to 10 for pain) as described in Chapter 12 of the Cochrane Handbook for Systematic Reviews of Interventions.[216] The minimum clinically important differences (MCIDs) used were 1.5 points on a 10-point VAS for pain, 1 and 4 points respectively for the physical and mental components of the SF-12, 10 points on a 100-point scale for disability/function measured by the WOOS Index, and 13.4 points for the ASES Score.[217–219]

Unit of analysis issues

The unit of analysis was each shoulder rather than each participant. None of the included trials reported more than two study arms. If a single trial reported multiple time points within one of the planned time points, then data that related to the latest time point was extracted.

Dealing with missing data

Study authors were contacted to obtain missing study characteristics and numerical outcome data. There was no missing data that would influence planned meta-analyses therefore no sensitivity analyses were necessary. For dichotomous outcomes (e.g. number of revision operations), event rates were calculated using the number of participants randomised in the group as the denominator, unless the number at risk was otherwise clearly stated. Where possible, standard deviations were calculated from other statistics such as standard errors, confidence intervals or p-values, according to the methods recommended in the Cochrane Handbook for Systematic Reviews of Interventions.[220]

Assessment of heterogeneity

Clinical and methodological diversity were considered in terms of participants, interventions, outcomes and study characteristics. Statistical heterogeneity was assessed by visual inspection of forest plots to assess for obvious differences in results between the studies, and using the I^2 and Chi-squared statistical tests, referring to the Cochrane Handbook for Systematic Reviews of Interventions to guide interpretation.[221]

Assessment of reporting biases

To assess for publication and outcome-reporting bias, trial protocols and registry entries were checked against their published reports.

Data synthesis

Meta-analyses were only undertaken where it was meaningful to do so. That is, if the treatments, participants and the underlying clinical question were similar enough for pooling to make sense. For clinically homogeneous studies, we pooled outcomes in a meta-analysis using the random-effects model as a default. All the included studies were small, hence sensitivity analysis to assess for small study bias was not relevant.

Quality of the evidence (GRADE)

The quality of the evidence was assess using the five "Grading of Recommendations Assessment, Development and Evaluation" (GRADE) considerations. GRADE considers study limitations, consistency of effect, imprecision, indirectness and publication bias, reporting the quality as high, moderate, low, or very low.[222] The methods and recommendations were applied as described in chapters 8, 11 and 12, of the Cochrane Handbook for Systematic Reviews of Interventions.[214,216,223]

Subgroup analysis and investigation of heterogeneity

Insufficient data were available to carry out the planned subgroup analyses for the following factors: age of participant, presence or absence of significant rotator cuff tear.

Sensitivity analysis

Insufficient data were available to carry out the following sensitivity analyses to investigate the robustness of the treatment effect of pain, disability and function: inclusion of missing data, inclusion of trials identified at risk of selection bias, inclusion of trials with unclear or inadequate blinding of the outcome assessor, choice of statistical method for pooled data (fixed-effect versus random-effects model).

2.3 Results

2.3.1 Results of the searches

Figure 2.1 details the studies screened and passed to each stage, reasons for exclusion and the studies included in the review. The initial searches performed on 2nd August 2018 yielded 2081 records with an additional 41 records from the updated search on 31st January 2019. After exclusion of duplicates and screening of abstracts and titles for eligibility 43 studies were identified for full-text review. Of the eight studies included in the previous version of this review,[15] one was excluded due to concerns regarding the study design.[224] Thirteen new studies met the inclusion criteria, providing a total of 20 studies for inclusion.

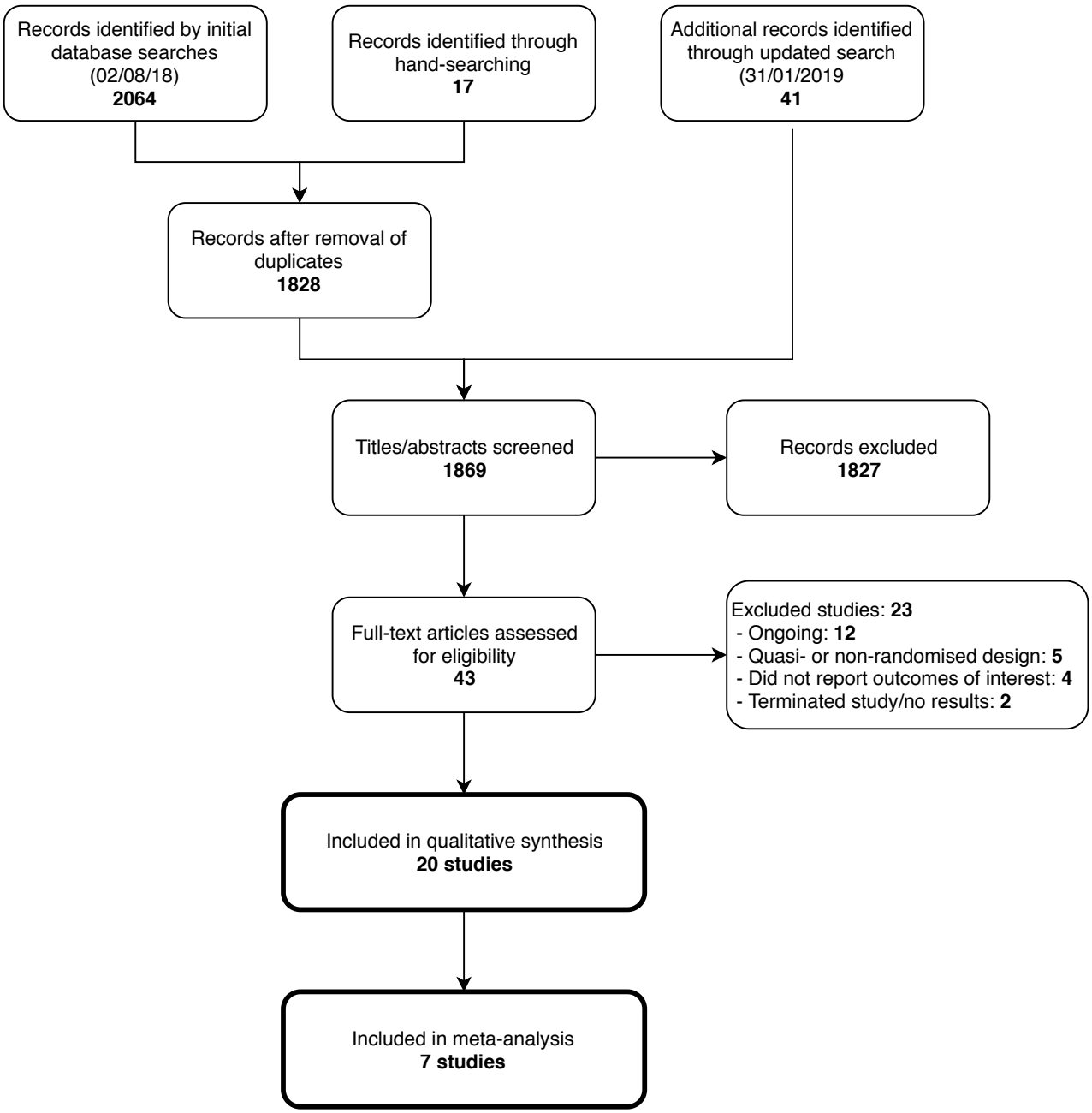


Figure 2.1: PRISMA flowchart for study screening and inclusion

2.3.2 Characteristics of included studies

The characteristics and citations for the 20 included studies are summarised in table 2.1. Hereafter, the studies will be referred to by the first author name and year of publication. All 20 studies were parallel-group RCTs with two arms. Only five studies recruited from more than one site. The median length of follow-up was two years (range: six weeks to 10 years). Eleven studies were performed in North America, seven in Europe, and one each in Australia and New Zealand.

A total of 1083 participants (1105 shoulders) were randomised. The median number of shoulders analysed in each study was 42 (range 16 to 161). In six studies, 10 patients underwent staged bilateral procedures where each shoulder was randomised and analysed independently. In one study, one participant underwent a simultaneous bilateral procedure. It is unclear whether the shoulders were randomised independently.

2.3.3 Participants

Sixteen studies reported on patients with glenohumeral OA (excluding CTA). In these studies, the mean age of participants ranged from 63 to 70 years and the median proportion of female participants was 55%. Four studies included patients with CTA (**Edwards 2012, Gobezie 2019, Greiner 2015, Poon 2014**). In these studies, the mean age of participants ranged from 69 to 81 years and the median proportion of female participants was 63%. Participant demographics for each study are summarised in table 2.2.

2.3.4 Interventions

The included studies addressed 13 different comparisons, which are summarised in the characteristics of studies table (table 2.1). All of the comparisons are of one type of shoulder replacement compared to another, or of one type of technique to another. No studies were found comparing any type of shoulder replacement surgery to placebo (sham-surgery), nor to any other type of surgery, nor to any type of non-surgical treatment.

Table 2.1: Characteristics of the included studies

Study identifier	Country	Number of centres	Intervention	Comparator
Gartsman 2000[225]	USA	1	Stemmed humeral hemiarthroplasty	Stemmed TSR
Lo 2005[226]	Canada	1	Stemmed humeral hemiarthroplasty	Stemmed TSR
Sandow 2013[227]	Australia	1	Stemmed humeral hemiarthroplasty	Stemmed TSR
Rasmussen 2015[228]	Denmark	1	Stemmed humeral hemiarthroplasty	Humeral head resurfacing
Uschok 2017[229]	Germany	2	Stemmed TSR	Stemless TSR
Boileau 2002[230]	France	1	TSR with cemented polyethylene glenoid	TSR with uncemented metal-backed glenoid
Edwards 2010[231]	USA	1	TSR with cemented keeled glenoid	TSR with cemented pegged glenoid
Gartsman 2005[232]	USA	1	TSR with cemented keeled glenoid	TSR with cemented pegged glenoid
Gascoyne 2017[233]	Canada	1	TSR with cemented keeled glenoid	TSR with cemented pegged glenoid
Nuttall 2007[234]	UK	1	TSR with cemented keeled glenoid	TSR with cemented pegged glenoid
Rahme 2009[235]	Sweden	2	TSR with cemented keeled glenoid	TSR with cemented pegged glenoid
Litchfield 2011[236]	Canada	7	TSR with cemented stem	TSR with uncemented stem
Kwon 2019[237]	USA	1	TSR via standard approach	TSR via subscapularis sparing approach
Lapner 2012[238]	Canada	2	TSR via lesser tuberosity osteotomy	TSR via subscapularis tenotomy
Levine 2019[239]	USA	1	TSR via lesser tuberosity osteotomy	TSR via subscapularis peel
Mechlenburg 2014[240]	Denmark	2	Humeral head resurfacing (Copeland)	Humeral head resurfacing (Global CAP)
Edwards 2012[241]	USA	1	RTSR with neutral glenosphere	RTSR with 10 degree inferior-tilted glenosphere
Greiner 2015[242]	Germany	1	RTSR with standard offset glenosphere	RTSR with increased offset glenosphere
Poon 2014[243]	NZ	1	RTSR with concentric glenosphere position	RTSR with eccentric glenosphere position
Gobezie 2019[244]	USA	1	RTSR with 155 degree neck-shaft angle	RTSR with 135 degree neck-shaft angle

Table 2.2: Study participant demographics

Study identifier	Number of participants	Shoulders randomised	Mean age	Percentage female	Diagnosis	Length of follow-up
Gartsman 2000[225]	51	55	65	40%	OA	36 months
Lo 2005[226]	42	42	70	56%	OA	24 months
Sandow 2013[227]	33	33	68/72[a]	67%	OA	120 months
Rasmussen 2015[228]	35	40	67	68%	OA	12 months
Uschok 2017[229]	40	40	67	58%	OA	60 months
Boileau 2002[230]	39	40	69	75%	OA	36 months
Edwards 2010[231]	50	59	69	47%	OA	24 and 60 months[b]
Gartsman 2005[232]	47	47	69	37%	OA	6 weeks
Gascoyne 2017[233]	15	16	64	33%	OA	24 months
Nuttall 2007[234]	20	20	67	30%	OA	24 months
Rahme 2009[235]	28	30	64	67%	OA	24 months
Litchfield 2011[236]	161	161	69	55%	OA	24 months
Kwon 2019[237]	107	107	69	53%	OA	24 months
Lapner 2012[238]	87	87	68	61%	OA	12 and 24 months[c]
Levine 2019[239]	60	60	NR	NR	OA	24 months
Mechlenburg 2014[240]	32	32	63	42%	OA	24 months
Edwards 2012[241]	52	52	69	56%	CTA	12 months
Greiner 2015[242]	34	34	75	65%	CTA	24 months
Poon 2014[243]	50	50	81	63%	CTA	24 months
Gobezie 2019[244]	100	100	73	66%	CTA	24 months

NR: Not reported. [a]Within group medians
[b]Some results reported in Kilian 2017[245] [c]Some results reported in Lapner 2013[246]

2.3.5 Outcomes

No one study assessed all of the major outcomes. Pain was reported by 12 studies, disability/function was reported on a number of validated scales by 19 studies, quality of life by three studies, specific adverse events by 12 studies, serious adverse events by three studies, and revision/reoperation/treatment failure by 14 studies. Patient reported success was not reported by any studies. The commonest additional physician reported outcomes were radiographic measures of lucency/loosening/notching etc reported by 10 studies, radiostereometric analysis of micromotion by five studies, range of motion by 15 studies, strength by nine studies, operating time in two studies and humeral head bone mineral density in one study.

2.3.6 Ongoing and excluded studies

Twelve ongoing trials were identified in clinical trials registers. Six are listed as actively recruiting, two are active but not recruiting, two have recently been completed but no published results are available yet, one study is not yet recruiting, and the status of one is unknown.

Eleven studies were excluded following full text review. Of these, four studies did not report on the outcomes of interest and four studies were either quasi- or non-randomised studies. One study was described as a randomised study, however significant concerns regarding the allocation process were identified by the reviewers and this study has also been excluded. Two studies identified in clinical trials registers were excluded: one had been terminated due to slow recruitment and poor follow-up, and one was listed in 2006 but no results have been published and the inclusion criteria did not match this review.

Details of the ongoing and excluded studies are summarised in appendix A.2.

2.3.7 Risk of bias in the included studies

The risk of bias judgements are summarised in figure 2.2.

Selection bias may be introduced in the sequence generation process and allocation concealment. Sequence generation was generally adequate and was most frequently completed using some form of random numbers table or list. The allocation-concealment process was poorly described in half of the studies. In **Kwon 2019**, group allocation was revealed to the surgeon prior to the intraoperative decision confirming suitability, therefore it was deemed high risk.

The operating surgeons could not be blinded to the treatment allocation. The possible effect of this on subsequent outcomes is unclear and these studies have all been assessed to be at unclear risk of performance bias, at best. Two studies were rated at high risk of performance bias. In **Gascoyne 2017**, the procedures were all performed by a single surgeon who terminated the trial early in response to published studies in the literature. In **Sandow 2013**, the procedures and follow-up were all performed by a single surgeon. The trial was terminated early due to a perceived higher early revision risk.

Most studies used patient-reported outcome tools that could be administered without supervision. However, the follow-up for two studies was performed open label (**Levine 2019, Sandow 2013**) and these were rated at high risk of bias in this domain. In ten studies, it was unclear how well and how long the participants were blinded for.

Only four studies were rated at low risk of bias for physician-evaluated outcomes. For the 16 studies that were judged to be at high risk of bias for physician-evaluated outcomes, the outcomes were mostly radiographic and/or radiostereometric, where it is not possible to blind assessors for the radiological appearances of different types of shoulder replacement implants. In **Sandow 2013**, one participant undergoing a revision surgery was excluded from analysis of failures since the surgeon attributed this failure to their own technical error. It is not possible to appreciate whether technical deficiencies contributed to failures in the other arm of the study, nor whether the procedure method was more likely to result in technical error and subsequent early failure. The contribution of surgical technique to failures was not discussed by any other papers. Therefore, for consistency, this technical failure was included in the analysis of revisions for this review.

The risk of bias due to attrition (incomplete outcome data) was unclear in five studies and high

in five. Participant flow and loss of follow up through studies was often poorly reported within the text of studies.

The risk of bias in trials due to selective reporting was unclear in five studies and high in three studies. No trials described or referenced a study protocol and clinical trials register entries were found for only five studies (**Kwon 2019, Lapner 2012, Levine 2019, Litchfield 2011, Mechlenburg 2014**).

Five studies were at high risk of bias from baseline imbalance. These imbalances were either in participants' sex or baseline disability/function scores. **Litchfield 2011** performed a sensitivity analysis and found that a significant effect in the primary study outcome was highly likely to be attributable to sex imbalance.

Sixteen studies described standardisation of their rehabilitation protocols. In **Kwon 2019**, different postoperative restrictions were applied to the two study groups and this domain was rated at high risk of bias.

The studies are generally small and a high proportion report industry funding or financial conflicts of interest. Eight studies reported funding from industry sources. A further seven studies reported a financial conflict of interest for study authors relating to study implants. Only three studies (**Gartsman 2000, Levine 2019, Rasmussen 2015**) reported freedom from any personal financial relationships or research funding that could be perceived as a conflict of interests.

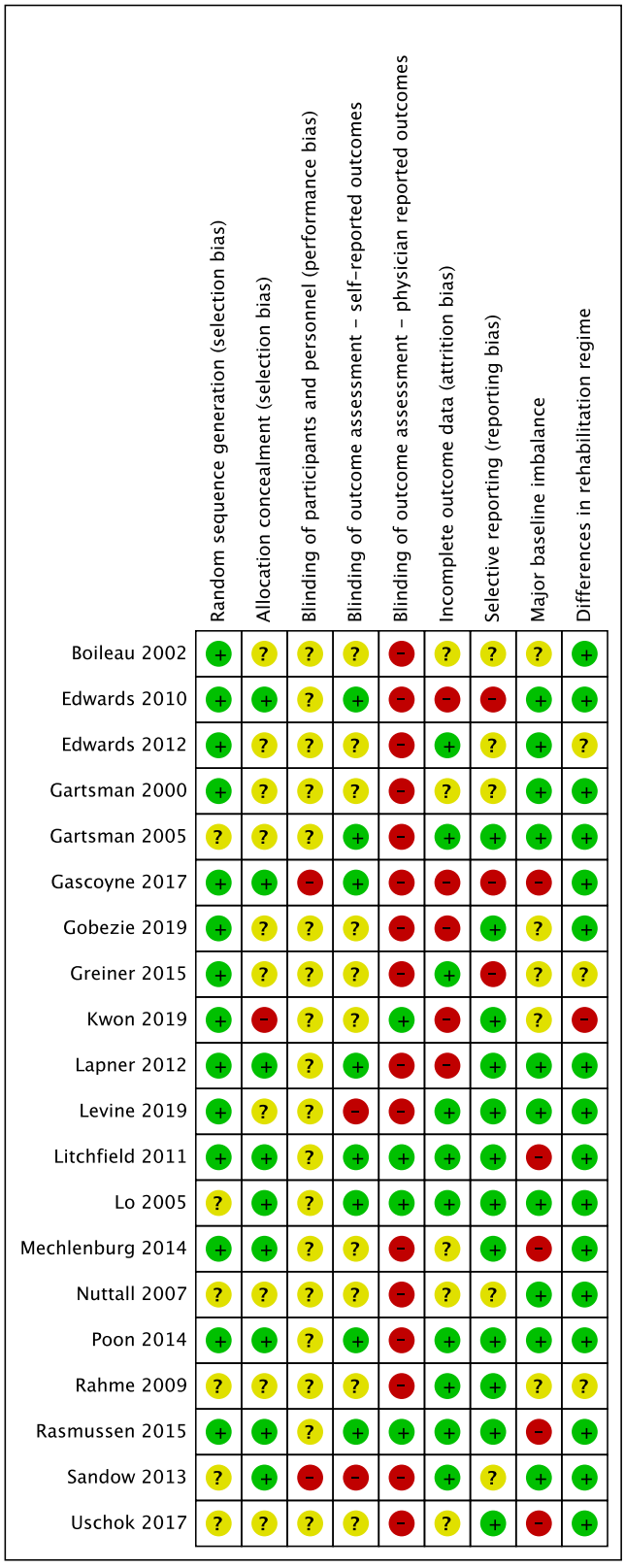


Figure 2.2: Summary of risk of bias judgements. + = Low, ? = Unclear, − = High

2.3.8 Effects of interventions

There were no studies identified comparing shoulder replacement surgery to placebo, no treatment, any non-operative treatment or other types of shoulder surgery. All studies compared one type of shoulder replacement (or technical modification) to another. Additional study data, not otherwise described in the text, tables and figures, can be found in appendix A.2.4.

Stemmed conventional TSR versus stemmed humeral hemiarthroplasty

The summary of findings for this comparison is reported in table 2.3. TSR may result in less pain and improved disability/function when compared to hemiarthroplasty at two years (figures 2.3 and 2.4). Disability/function was reported by **Lo 2005** on the WOOS Index and by **Gartsman 2000** on the ASES Score. The results were pooled since the scores are conceptually similar and both scaled 0-100. The difference in pain and disability/function were close to the predetermined thresholds for clinical importance. The quality of the evidence was assessed as low due to imprecision and risk of bias. For the same reasons, there may be no difference in quality of life but this is uncertain.

Disability/function outcomes reported by **Sandow 2013** were consistent with the direction and magnitude of the effect in the other studies but they were not presented in a format suitable for meta-analysis. The same study reported a trend towards lower revision rates at two years and ten years (RR 0.33, 95% CI 0.07 to 1.53) favouring TSRs, but there were serious concerns regarding risk of bias for this physician-determined outcome in this study and these results were also excluded from meta-analysis.

No difference was demonstrated in revision/reoperation/treatment failure rates or specific adverse events reported at two years and up to six years in **Gartsman 2000** and **Lo 2005** (figure 2.5). One fatal pulmonary embolism at two days post surgery was reported following a TSR in **Lo 2005**. The studies and pooled estimates for these binary outcomes were considerably underpowered such that no meaningful conclusions can be drawn. No additional outcomes meeting the eligibility criteria were reported.

Table 2.3: Summary of findings: Conventional TSR versus hemiarthroplasty

Outcome	Number of shoulders	Result	Favours	GRADE	Studies included	Comments
Pain (0-10; VAS)	92	MD -1.49 (-2.88 to -0.10)	TSR	Low	Gartsman 2000; Lo 2005	I^2: 55%
Disability/function (0-100; WOOS Index)	92	MD 10.57 (2.11 to 19.02)	TSR	Low	Gartsman 2000; Lo 2005	I^2: 0%
Quality of life: physical (0-100; SF-12)	41	MD -0.80 (-8.2 to 6.6)	HA	Low	Lo 2005	
Quality of life: mental (0-100; SF-12)	41	MD 1.00 (-5.1 to 7.1)	TSR	Low	Lo 2005	
Adverse events (specific)	41	RR 0.35 (0.08 to 1.54)	TSR	Very low	Lo 2005	
Adverse events (serious)	42	See text		Very low	Lo 2005	
Revision/reoperation/failure (intermediate)	92	RR 1.29 (0.30 to 5.53)	HA	Very low	Gartsman 2000; Lo 2005	I^2: 0%
Revision/reoperation/failure (long-term)	51	RR 0.89 (0.20 to 4.00)	TSR	Very low	Gartsman 2000	

MD: Mean Difference, RR: Relative Risk. Outcomes reported at intermediate endpoint (1-3 years) unless stated.

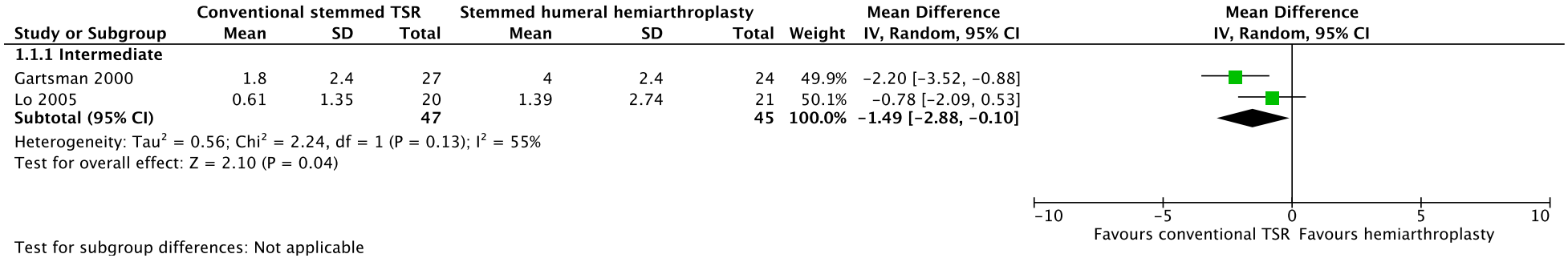


Figure 2.3: Forest plot of pain for TSR vs hemiarthroplasty (Visual Analogue Scale, 0-10 scale, lower=better)

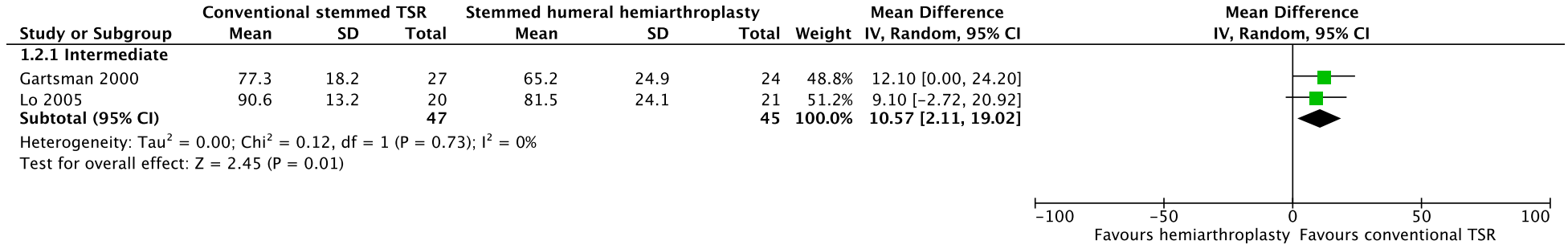


Figure 2.4: Forest plot of disability/function for TSR vs hemiarthroplasty (WOOS Index, 0-100 scale, higher=better)

Study or Subgroup	Conventional stemmed TSR Events	Total	Stemmed humeral hemiarthroplasty Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
1.7.1 Intermediate						
Gartsman 2000	2	27	1	24	38.8%	1.78 [0.17, 18.39]
Lo 2005	2	20	2	21	61.2%	1.05 [0.16, 6.76]
Subtotal (95% CI)		**47**		**45**	**100.0%**	**1.29 [0.30, 5.53]**
Total events	4		3			
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.12$, df = 1 (P = 0.73); $I^2 = 0\%$						
Test for overall effect: Z = 0.34 (P = 0.73)						
1.7.2 Long-term						
Gartsman 2000	3	27	3	24	100.0%	0.89 [0.20, 4.00]
Subtotal (95% CI)		**27**		**24**	**100.0%**	**0.89 [0.20, 4.00]**
Total events	3		3			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.15 (P = 0.88)						

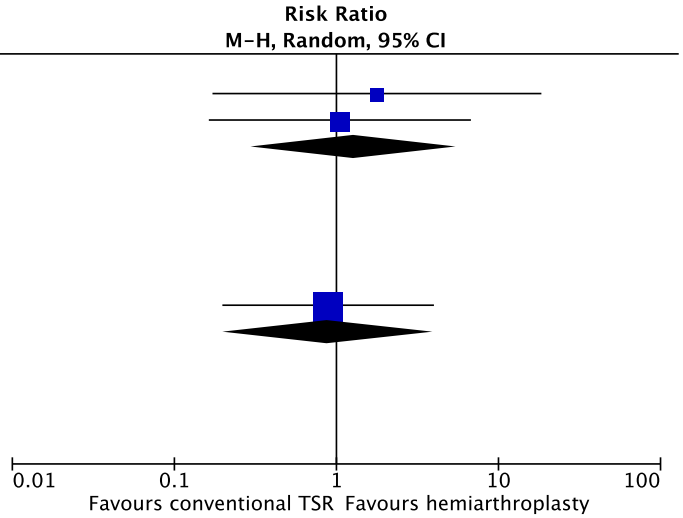


Figure 2.5: Forest plot of the risk of revision/reoperation/treatment failure for TSR vs hemiarthroplasty (Cumulative events, intermediate= 1-3 years, long-term= >3 years)

Resurfacing humeral hemiarthroplasty compared to stemmed humeral hemiarthroplasty

Rasmussen 2015 reported on 35 participants (40 shoulders) for this comparison. It is uncertain whether there is any difference between these two treatments across all of the domains because the evidence was rated as very low quality, downgraded for risks of bias and serious imprecision. Pain was recorded on a non-validated subscale of the Constant-Murley Score. Disability/function was better at one year following stemmed hemiarthroplasty, however there was a major baseline imbalance between the groups and the confidence interval for the estimate was wide. Quality of life was not reported. The study was underpowered to reliably capture adverse events and revision/reoperation/treatment failure. The study reported that operating time was 28 minutes shorter with resurfacing hemiarthroplasty (95% CI 19 to 37).

Conventional stemless TSR compared to conventional stemmed TSR

Uschock 2017 reported on 40 participants (40 shoulders). No difference was found between the two treatments across all of the domains, but it is uncertain whether a difference may exist because the evidence was rated as very low quality, downgraded for serious risks of bias and imprecision. Several measures of periprosthetic lucency were reported from plain radiographs. The results were confounded by the use of non-comparable scales and imbalances between the types of glenoid components used in each arm of the study.

Conventional stemmed TSR using metal-backed uncemented glenoid component compared to all-polyethylene keeled cemented glenoid component

Boileau 2002 studied 39 participants (40 shoulders) for this comparison. There was no difference reported between the groups for the domains of pain, disability/function and adverse events, but it is uncertain whether a difference may exist because the evidence was rated as very low-quality, downgraded for risk of bias and serious imprecision. There may be a clinically important increased risk of revision at a mean follow-up of 38 months for metal-backed uncemented glenoid implants compared to all-polyethylene keeled cemented components, based on 5/20 events in the metal-backed group versus 0/20 events in the cemented group (RR

11.00, 95% CI 0.65 to 186.62). This finding is uncertain since the evidence is very low-quality, downgraded due risks of bias and serious imprecision. The study reported the incidence of periprosthetic glenoid lucency seen on plain radiographs. Lucency was more frequently seen with cemented components compared to uncemented components (17/20 vs 5/20). However, none were observed to progress in the cemented group, compared to 4/5 progressions with clinical deterioration in the uncemented group. The reliability and validity of the assessment of glenoid lucency was not reported in the study.

Conventional stemmed TSR using pegged compared to keeled all-polyethylene cemented glenoid components

Five studies comprising a total of 160 participants (172 shoulders) provided data for this comparison (**Edwards 2010, Gartsman 2005, Gascoyne 2017, Nuttall 2007, Rahme 2009**). Data for **Edwards 2010** were obtained from the primary study and a secondary publication.[231,245] The secondary publication by Kilian et al included additional patients.[245] The reporting of patient flow and outcomes lacked transparency.

Only **Nuttall 2007** reported on pain, finding no difference, but only mean values were presented by this small study with no measures of spread, therefore no conclusions can be drawn. For disability/function, four studies independently reported no difference at between one and five years of follow-up, but the studies used different outcome scores, different measures of central tendency (mean/medians, exact p-values/range/no measure of spread), and different time points. No pooling of results for meta-analysis was possible and the overall outcome is uncertain. No studies reported any adverse events over and above those requiring revision surgery.

At two and five years of follow-up, studies reported fewer revisions of pegged glenoids compared to keeled glenoids, but this is uncertain due to serious imprecision and risk of bias (figure 2.6).

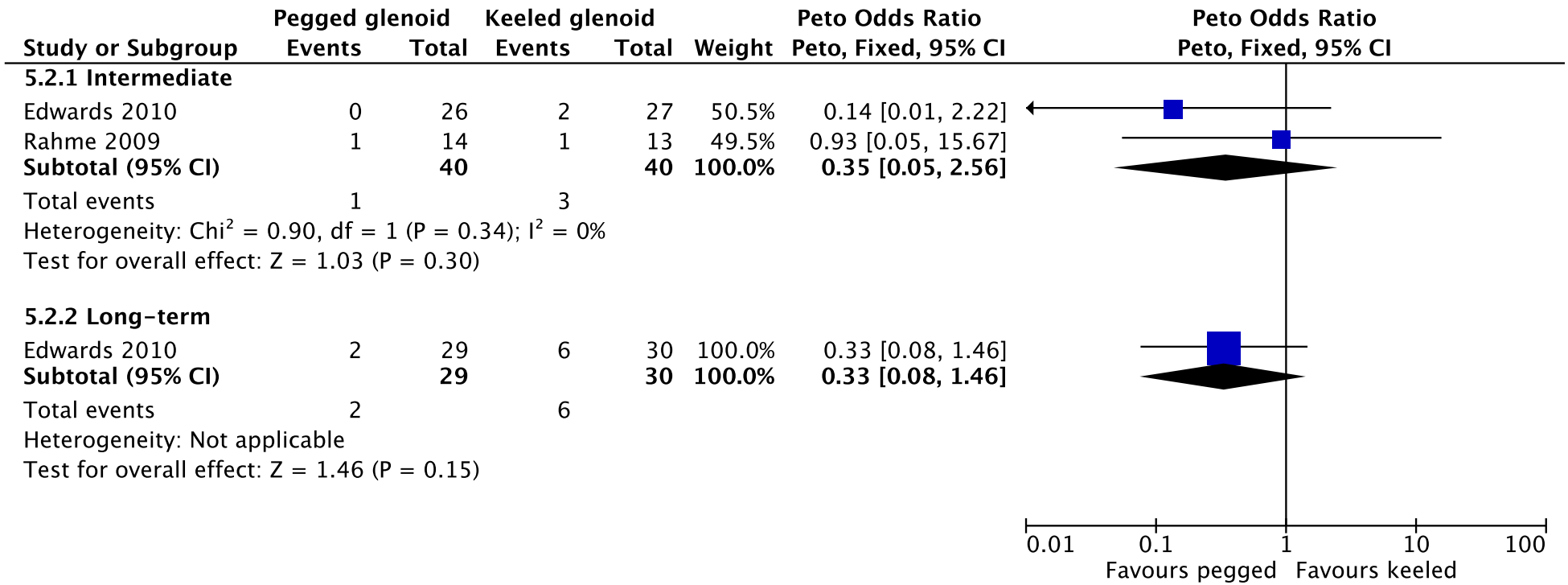


Figure 2.6: Forest plot of the risk of revision/reoperation/treatment failure for TSR using pegged vs keeled glenoid components (Cumulative events, intermediate= 1-3 years, long-term= >3 years)

The quality of evidence for radiological and radiostereometric outcomes was very low. Using a score of $\geq$4 (1-5 scale) as a cut-off on the scales described by Franklin et al and Lazarus et al,[104,247] there may be little to no difference in the rate of substantial periprosthetic radiolucency comparing pegged to keeled components at two-year follow-up (RR 0.38, 95% CI 0.02 to 8.83; 71 participants in two studies (**Edwards 2010, Rahme 2009**)). Radiostereometric analysis of glenoid component micromotion was performed in three studies. No important difference was found by **Nuttall 2007** and **Rahme 2009**. **Gascoyne 2017** found a possible difference in favour of pegged implants, but the study was at high risk of bias across several domains and the measurement method was changed partway through the study.

Conventional stemmed TSR using uncemented compared to cemented fixation of the humeral stem

One large multicentre study of 161 participants (161 shoulders) provided the data for this comparison (**Litchfield 2011**). Pain was not reported. The study found a small but clinically unimportant difference in disability/function, measured on the WOOS Index, favouring TSR with a cemented stem (MD 8.6, 95% CI 2.4 to 14.8). Due to risks of bias including a considerable imbalance of participant sex between the groups, this was rated as very low-quality evidence. Sensitivity analysis by the study authors confirmed that baseline imbalance may have inflated the treatment effect seen. It is uncertain whether there is any difference in quality of life, adverse events and revision/reoperation/treatment failure rates. No differences were found, but the evidence was of very low quality due to serious risks of bias and imprecision.

Conventional stemmed TSR using a subscapularis sparing approach compared to standard subscapularis tenotomy

Kwon 2019 randomised 107 patients but only reported on 70 for this comparison. No clinically important differences were observed for pain, disability/function and revision/reoperation/treatment failure. The evidence was very low-quality due to serious risks of bias and imprecision. Adverse events were not reported and no additional outcomes met the inclusion criteria for this review.

Conventional stemmed TSR using a lesser tuberosity osteotomy approach compared to subscapularis tenotomy/peel

Two studies of 147 participants (147 shoulders) provided the data for this comparison. **Lapner 2012** compared osteotomy to subscapularis peel, and **Levine 2019** compared osteotomy to subscapularis tenotomy. Results for **Lapner 2012** were obtained from both the primary study and a secondary publication in 2013.[238,246] **Levine 2019** reported pain and disability/function as mean values only, therefore it is uncertain whether there is any difference between osteotomy and tenotomy. In addition, there were inconsistencies between the values reported in the tables and the text; the authors did not respond to requests for clarification. **Lapner 2012** found that there is probably little to no difference in disability/function measured using the WOOS Index but it is uncertain (MD -1.70, 95% CI -9.16 to 5.76, low quality evidence). There were two revisions/reoperations/failures in the peel group and none in the osteotomy group within two years in **Lapner 2012**. **Levine 2019** reported one in each arm at one year.

Both studies reported on satisfactory radiological healing at one year and there may be little or no difference in this outcome (RR 0.99, 95% CI 0.87 to 1.13, low-quality evidence downgraded for risk of bias and imprecision). **Lapner 2012** found no difference in the rates of fatty infiltration of the rotator cuff and subscapularis strength. **Levine 2019** reported that operative time was shorter for subscapularis tenotomy (129 minutes versus 153 minutes), but this outcome is at serious risk of bias and was reported without any statistical measure of uncertainty.

Resurfacing humeral hemiarthroplasty using one brand compared to another brand

One study of 32 participants (32 shoulders) provided the data for this comparison (**Mechlenburg 2014**) and compared the newer Global C.A.P. shoulder implant with the well-established Copeland shoulder implant.

No difference was shown for disability/function and revision/reoperation/treatment failure. The evidence was very low-quality. Pain was not reported. The study also measured component micromotion using radiostereometry, and periprosthetic bone mineral density. No differences were reported in either outcome but the evidence was very low-quality.

Reverse stemmed TSR using a 10-degree inferior inclination of the glenosphere compared to neutral inclination

Edwards 2012 studied 52 patients (52 shoulders). There is low-quality evidence to suggest that may be no clinically important difference in disability/function at one year measured using the ASES (MD 7.60 in favour of an inferior tilted glenosphere, 95% CI 0.83 to 14.37, 0-100 scale, higher=better). Only one adverse event (prosthesis dislocation) was reported in the neutral group at up to one year of follow-up, and no revisions/reoperations/treatment failures were reported in either group. Scapular notching was measured according to the Nerot classification.[248] No difference was found between the groups for cases with a grade of notching $\geq$1 at one year (RR 1.15, 95% CI 0.85 to 1.56). The follow-up was short and this evidence was of very low quality.

Reverse stemmed TSR using a bony increased offset (BIO) of the glenosphere compared to standard offset

One study of 34 participants (34 patients) provided the data for this comparison (**Greiner 2015**). It is uncertain whether there is a difference in any of the domains reported due to serious risks of bias and imprecision.

At two years, no difference was found in pain recorded on a non-validated subscale of the Constant-Murley Score, and no difference was found in disability/function on the full Constant-Murley Score (MD 2.60, 95% CI -9.52 to 14.72). Two acromial stress fractures were reported in each arm, identified on routine CT scans at one year.

Reverse stemmed TSR using eccentric placement of the glenosphere compared to concentric placement

Poon 2014 included 50 participants (50 shoulders) and provided moderate-quality evidence that there is no clinically important difference in pain (MD 0.20, 95% CI -0.63 to 1.03, VAS, 0-10 scale, lower=better) and no clinically important difference in disability/function (MD -2.00,

95% CI -5.17 to 1.17, ASES, 0-100 scale, higher=better), both measured at two years. However, for specific adverse events and risk of revision/reoperation/treatment failure the evidence was of very low quality due to serious imprecision and risk of bias and it is uncertain whether there is any difference between the treatments. The same is true for glenoid notching.

Reverse stemmed TSR using a 135-degree humeral neck-shaft angle compared to a 155-degree neck-shaft angle

For the final comparison, **Gobezie 2019** studied 100 participants (100 shoulders). It is uncertain whether there is a difference in most of the domains reported due to serious risks of bias and imprecision. At two years, no difference was found in pain, disability/function, adverse events and revision/reoperation/treatment failure.

Use of a 135-degree neck shaft angle humeral component for reverse TSR may be associated with lower rates of scapular notching than 155-degree components (RR 0.37, 95% CI 0.37 to 0.74) based on low-quality evidence, downgraded two levels for risk of bias.

2.4 Discussion

2.4.1 Summary of findings

There were no published RCTs and no registered RCTs in trial registers comparing shoulder replacement surgery to placebo, non-operative management, or any other type of surgical treatment. The trials identified by this review were almost universally small and addressed diverse comparisons of different shoulder replacement types and technical modifications (20 trials, 13 different comparisons). Pooling of data for meta-analysis was only possible for a small selection of outcomes and comparisons. The overall quality of evidence for most comparisons was either low or very low and as such, few useful conclusions can be drawn. For trials reporting a power calculation, this was performed for their planned primary outcome, usually disability/function. For dichotomous outcomes (adverse events and risk of revision, reoperation

or treatment failure), the trials were all too small to reliably detect and compare these rare events. In addition, adverse event definitions were not predefined and follow-up periods were generally too short to monitor revision risks.

For the main comparison of conventional stemmed TSR compared to stemmed humeral hemiarthroplasty for osteoarthritis, there was low-quality evidence to suggest that there may be a small but clinically uncertain improvement in pain in favouring TSR (MD -1.49, 95% CI -2.88 to -0.10, VAS, 0-10 scale, lower=better) and there may be a clinically important improvement in patient reported disability/function favouring TSR (MD 10.57, 95% CI 2.11 to 19.02, WOOS Index, 0-100 scale, higher=better). There is probably no clinically important difference in overall quality of life measures. No conclusions could be drawn regarding risks of adverse events, revision, reoperation and treatment failure.

The one trial comparing metal-backed uncemented glenoid components to cemented all-polyethylene components for TSR suggested that there may be a clinically important higher risk of implant revision or reoperation at a mean of 38 months post surgery (RR 11.00, 95% CI 0.65 to 186.62), but confidence in this estimate is very low based on the serious imprecision in the estimate.

For reverse shoulder replacement, there may be a small but unimportant improvement in patient reported disability/function at one year from using a 10-degree inferior inclination glenosphere position compared to a neutral glenosphere position (MD 7.60 points, ASES, 0-100 scale, higher=better, low-quality evidence from one trial). There was moderate-quality evidence from one trial to suggest that there is probably no important difference in patient-reported pain and disability/function from using an eccentric compared to a concentric glenosphere position.

There were no other clinically important differences identified. This is primarily due to an absence of high-quality evidence, and is not based on any proof of no difference informed by high-quality studies.

2.4.2 Overall completeness and applicability of the evidence

No randomised trials in this field have adequately addressed the fundamental question of the effectiveness and risks of one type of shoulder replacement over another, or the effectiveness and risks of shoulder replacement surgery compared to no treatment, placebo or to any other form of treatment for OA or CTA. Important uncertainties regarding the major classes of shoulder replacement remain unanswered by this review, specifically the choice between humeral hemiarthroplasty, conventional TSR and reverse TSR, which are far more fundamental questions than those addressed in more specific narrow subtype studies. The importance and generalisability of some of the comparisons reviewed is not apparent. Many of the comparisons were supported by only one study, and the outcomes for comparisons with more than one study were mostly inadequately reported to allow meta-analysis. Results were therefore mostly inconclusive.

With regard to serious adverse events and risks of revision surgery, none of the included studies were of sufficient size and length of follow-up to be powered to identify these events reliably. Whilst three studies reported on revision risk for the main comparison of conventional stemmed TSR versus stemmed hemiarthroplasty, the quality of the evidence was very low. No firm conclusions on these outcomes can be made.

2.4.3 Quality of the evidence

The evidence identified was of low- or very low-quality. Studies were short, small and suffered from serious risks of bias. Only one study (**Lo 2005**) was free from high risk of bias across all domains. Of the 180 risk of bias assessments made, 46% were low, 34% were unclear and 20% were high. Due to the nature of the treatment, there was a risk of performance bias in all 20 of the included studies. The implications are not entirely clear but the potential for surgeons to influence the patient outcomes at multiple points within the studies needs to be considered. Further biases were common for radiological outcome measures, and imbalances between groups at baseline.

2.4.4 Potential for bias in the review process

The study was conducted according to the previously published protocol.[212] Whilst this is an update of a previous review on the topic,[15] there was a significant change of scope to explicitly include trials in participants with CTA and to restrict the included studies to only those with any type of shoulder replacement as one of the study arms. To reflect this change, all searches were redesigned and run without date limits and all studies were screened for this review by two independent reviewers. These reviewers made their decisions to include studies and their assessments of risk of bias independently of the previous review process with reference to updated Methodological Expectations for Cochrane Intervention Reviews (MECIR). The risk of bias from the review methodology process is therefore low. However, the analysis of revision risk and adverse events is a major limitation of this review, largely due to the inclusion criteria requiring only randomised controlled trials. Much larger studies are needed to identify events that occur at a frequency of between 1 in 1000 and 1 in 100. This information may be better provided by well-designed studies from large registry-based or routinely-collected datasets, or future trials designed to be nested in national registries to monitor this longer-term follow-up.

2.4.5 Agreements and disagreements with other studies or reviews

Two systematic reviews compared TSR to hemiarthroplasty for OA.[249,250] They performed meta-analyses based on the same four studies (including **Gartsman 2000**, **Lo 2005** and unpublished results from **Sandow 2013**, plus another unpublished study with mixed indications for surgery). Both concluded that function was superior following TSR compared to humeral hemiarthroplasty. Singh et al included the two published papers focused purely on osteoarthritis in the previous version of this Cochrane review.[15] They found that TSR may offer superior function to humeral hemiarthroplasty but, like this current review, identified that the supporting evidence is of low quality. They included seven RCTs overall for participants with OA, but none with CTA. One of those seven studies, by Kircher et al, was excluded from this version of the review.[224] Whilst this current review has identified an additional 13 studies for inclusion, including nine studies in patients with OA, the new studies are heterogeneous and low quality.

These studies did not support any reliable new conclusions.

The remaining published systematic reviews are largely based on low-quality evidence from non-randomised studies and pooled estimates from single arm studies. Radnay et al analysed 1952 patients from 23 studies (only one randomised) and felt able to make stronger conclusions in favour of TSR compared to humeral hemiarthroplasty for the outcomes of pain, function, satisfaction, range of motion and revision surgery.[251] Van den Bekerom et al performed a systematic review comparing long-term outcomes of total shoulder replacement to humeral hemiarthroplasty based on 1958 patients from 19 non-randomised studies.[252] They found that revision rates were higher after humeral hemiarthroplasty, however complication rates may be higher after TSR.

In a systematic review of 14 studies (both randomised and non-randomised), Choate et al failed to identify that any one method of subscapularis management in shoulder replacement surgery was superior to another.[253] Papadonikolakis et al reviewed 43 studies (only one comparative) and found a higher revision rate following TSR using metal-backed glenoid components versus all-polyethylene components.[254] Vavken et al analysed 1460 patients from eight comparative studies of pegged versus keeled glenoid components for TSR (including four RCTs).[255] No difference in rates of glenoid lucency were identified. The main conclusion was a slightly lower revision rate in favour of pegged components, weighted by the results of one large non-randomised study.

Erickson et al included 3302 patients from 65 studies to determine the effect of humeral inclination on range of movement achieved after RTSR.[256] The study found greater external rotation if a 135-degree inclination was used compared to 155 degrees, but no other differences. The same group used an overlapping set of 2222 shoulders in 38 studies to conclude that there is a higher rate of scapular notching with the 155-degree prosthesis, which is supported by the RCT included in this review.[244,256]

2.4.6 Conclusions

TSR may provide better function/less disability at two years when compared to humeral hemiarthroplasty. Whilst this finding may support the use of TSR in OA, this review lacks data on longer-term outcomes, revision risk and adverse events. The main findings are unchanged from the Cochrane review of 2010 and therefore the review generates no new guidance or implications for practice.

There remains a serious lack of high-quality evidence to determine the benefits and risks of shoulder replacement surgery. Investigators, commissioners and funders of research need to align research studies to address areas of uncertainty and to answer questions that have the potential to make significant changes to practice and patient care. Investigators should engage with methodologists to design studies that are suitable to reliably answer important questions.

The study of adverse events and revision risk remains difficult and may not be feasible within the context of a conventional surgical RCT with short follow-up. The numbers needed to identify less frequent events are prohibitive for normal trial designs. Our understanding of these events may be improved in the short-term by analysis of routinely collected and registry datasets. However, these studies may still be criticised due to the risks of selection bias. The *a priori* protocol for this Cochrane review specifically excluded observational studies. For future updates to this review, the inclusion criteria should be relaxed to include large well-conducted observational studies on the harms of shoulder replacement surgery. In the longer term, the possibility of conducting nationally coordinated, large-scale RCTs nested within a prospectively collected national registry should be explored.

Subsequent chapters in this thesis will explore whether trends in the use of shoulder replacement surgery in the UK match those seen internationally. The studies will examine whether large-scale datasets of the observed outcomes of shoulder replacement surgery support its increased use and whether patient groups most likely to benefit and most likely to suffer harm from surgery can be predicted.

Chapter 3

Data sources and data management

3.1 Summary

The following chapters in this thesis are based on analyses of data from two sources: Hospital Episode Statistics (HES) and the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man (NJR). A specific statistical analysis plan is described for each study within the relevant chapter. This chapter describes the fundamental properties of the data sources, data cleaning methods and statistical software used.

3.2 Hospital Episode Statistics

HES is a vast database storing information for all activity carried out in NHS hospitals or funded by NHS England. It is managed by NHS Digital.[257] This thesis uses data from the Admitted Patient Care (APC) database. Submission of data by NHS providers is mandatory to inform accurate remuneration for activity performed and it provides universal coverage of NHS England. The APC database captures all inpatient and day-case activity, and as such is a good resource for studying surgical interventions.[258]

3.2.1 Data quality

As a primarily administrative dataset, HES data does not contain the same level of detail regarding procedures that is available from joint replacement registries.[259] However, HES data has been used as a reference standard by the NJR to determine rates of procedure completeness. It has been shown that HES data may actually be at less risk of underreporting bias for revision procedures than NJR records, and is therefore less likely to underestimate true revision or reoperation rates.[260]

The quality of HES APC coding for shoulder replacement surgery has been audited by the North East Quality Observatory in 2015.[261] For 440 procedures, across eight different NHS providers, there was 96% agreement between the HES record and the hospital notes on the side of surgery. No procedures were incorrectly identified as shoulder replacements. There was 87% agreement on the category of shoulder replacement implanted, 80% agreement on the primary diagnosis/indication for surgery and 84% agreement on diagnosis when secondary codes were also used. The commonest procedure-type mismatches were for resurfacing shoulder replacements, and for incorrect polarity of RTSRs and TSRs. The accuracy of the implant fixation method (cemented/uncemented/hybrid) was less than 70% and is not used further in this thesis.

3.2.2 Data collection procedures and description of the dataset

The data within HES APC is extracted from hospital notes by local clinical coders applying nationally approved coding standards. Up to 20 diagnoses may be recorded using the WHO International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10) codes, and up to 24 procedures may be recorded using Office for Population Censuses and Surveys Classification of Interventions and Procedures (OPCS-4) codes. Procedures typically require the combination of a number of separate codes including the main procedure, the side of surgery, and the approach or type of anaesthetic. Diagnoses can be recorded with varying levels of specificity depending whether 3, 4, 5, or 6 digit codes are provided. The first listed diagnosis is considered to be the primary diagnosis for the admission.[262] A

large number of additional basic demographic fields are available (e.g. age, sex, ethnicity), together with details of the hospital providing care and registered GP practice. NHS Digital derive additional fields from the patient's home postcode and GP practice details to facilitate geographical analyses, typically using Clinical Commissioning Groups or Super Output Areas to illustrate variation. Full postcodes and dates of birth are restricted from research data releases to avoid identification of individuals in the dataset. Further steps are taken to desensitise data extracts prior to release, including pseudonymisation of consultant and GP identifiers.

HES data is stored as individual "finished consultant episodes" of care. Each time a patient is transferred to a different consultant-led team, a new episode is created. A number of completed episodes may be combined to form a "spell" of care, or what would be conventionally considered to be the full duration of a hospital admission. The HES dataset is organised by UK financial year running from April 1st to March 31st.

3.2.3 Data linkage and patient consent

Data were provided by NHS Digital together with linked mortality data from the Office for National Statistics (ONS), including date and cause of death. Death data is only available for deaths occurring since 1st January 1998. Since 2001, cause of death has been recorded using the ICD-10 system. Prior to that, ICD-9 codes were used. All linkage was performed by NHS Digital and provided in a non-identifiable format. Whilst no formal consent is required for the provision of non-identifiable HES data by NHS Digital, individuals have the right to request that their data is not released for purposes beyond their own individual care. This was known as a "Type 2 opt-out", now replaced by the National Data Opt-out.[263] As of March 2017, the end of the study dataset, 2.3% of patients across England had requested a type 2 opt-out.[264]

3.2.4 Inclusion criteria

Any patient undergoing any type of shoulder replacement procedure between 1st April 1998 and 31st March 2017 was selected based on the presence of a relevant OPCS-4 code listed in appendix A.4. NHS digital searched the HES database using the terms supplied. Patients were

included for analysis of outcomes if they were aged 18 years or older at the time of their primary shoulder replacement procedure. For the purposes of incidence trends, no age restrictions were applied.

3.2.5 Episode linkage to generate dataset for analysis

For each of the patients identified by the inclusion criteria, every linked episode of care was extracted to form a complete history of past and follow-up inpatient care. Multiple episodes for the same patient were cross-referenced using the pseudonymised HES ID and ordered by the episode start date. Overall lengths of stay were calculated using the continuous inpatient spell methodology.[265] For each patient, the first primary shoulder replacement per side for each patient was identified and labelled as the index case. All subsequent activity was then searched for medical complications of interest and further operations on the affected shoulder, including revisions. The event dates were then converted into time-to-event variables for the outcomes of interest. The sequence of shoulder replacement procedures was constructed per side for each patient. The data was collapsed into a one-row-per-procedure database for final analyses. Further details regarding specific complications and further surgery are described in the relevant chapters.

3.2.6 Exclusion criteria

Duplicate episodes and episodes occurring after the date of death were removed. Some index procedures were performed before 1st April 1998 but were found in the datacut because the episode was completed in the new financial year; these were excluded. Patients with missing age or sex data, or implausible ages were removed. Some pre-release data quality measures are in place to remove infeasible data. The upper limit applied by NHS digital is 120 years. In fact, for the extracted shoulder replacement cases the observed upper age limit was 109 years. Patients undergoing surgery with a diagnosis of tumour were excluded. Patients with an inappropriate operation sequence were excluded from the analysis: i.e. revision procedure as the first recorded shoulder replacement for that side. Any shoulder replacement procedure that

followed after the first primary shoulder replacement was treated as a revision procedure, even if it was not explicitly identified with a revision code. The list of OPCS-4 codes in appendix A.4 includes those explicitly identified as revisions.

3.2.7 HES data flowchart

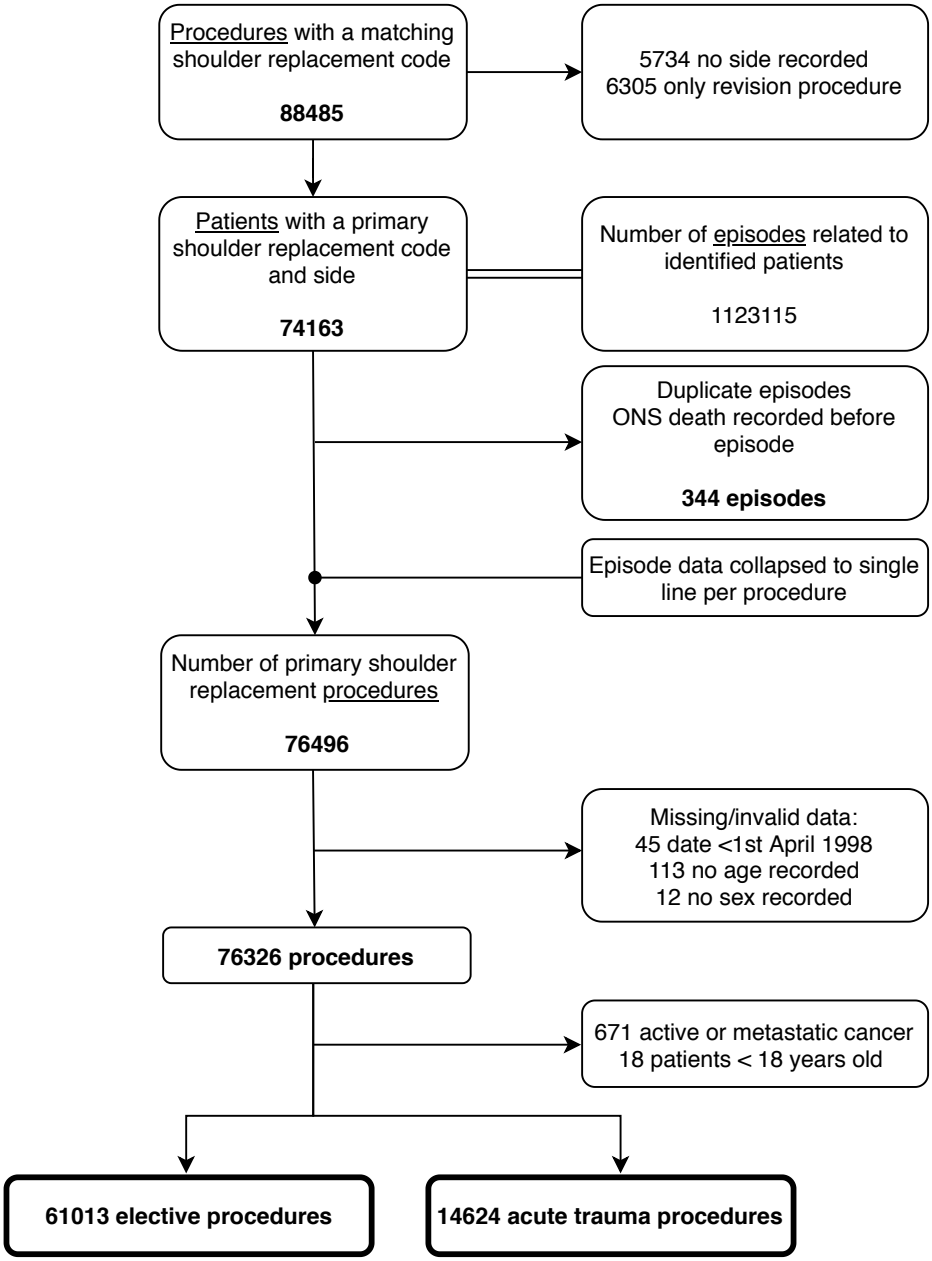


Figure 3.1: Data flow chart for HES studies

3.2.8 Calculation of comorbidity indices

The Charlson Comorbidity Index was devised to classify and adjust for comorbid conditions in longitudinal studies of mortality.[266] It is now widely used for both its original application in mortality studies and more widely to adjust for comorbidities in analyses with other endpoints, including revision risk after joint replacement. For this thesis, Charlson scores were calculated using the Stata **Charlson** package, based on algorithm which has been validated in multiple international administrative databases.[267–269] Separate scores were generated for left and right procedures, correct at the time of surgery.

3.2.9 Handling missing data

The key variables planned for adjusted analyses were age and sex. The total fraction of cases missing this information was less than 0.2%. Regardless of the mechanism of missingness, the likely risk of bias from these missing results is negligible and so they were excluded from the final complete case analyses. Where no comorbidity data is recorded, it is assumed that the individual is simply healthy rather than that the information is missing. No information is available regarding the patient characteristics for those patients exercising a data opt-out. There may be some regional variation in opt-out rates.[270]

3.3 National Joint Registry

The NJR was established in 2002 with the purpose of monitoring the performance of joint replacement implants. Data submission by providers of NHS care is mandatory. Collection of shoulder replacement data began in April 2012. The NJR is commissioned and overseen by the Healthcare Quality Improvement Partnership (HQIP), and the database is administered by Northgate Public Services, an independent provider with technology and data management expertise.

3.3.1 Data quality

The NJR instigated a formal Data Quality Audit programme from the 2014/5 financial year onwards. Local hospital records of joint replacement surgery are cross-referenced to the NJR records for that provider. To date, this process has only covered hip and knee replacements and the data completeness is high, with only 5.82% missing procedures. However, the proportion of missing revision procedures was a little higher (9.47% and 9.87% respectively).[271] The NJR therefore may systematically underestimate the rate of procedure failure when using revision as an endpoint, but this underestimate is likely to be small. Increased automation is available from 2019 and shoulders will soon be incorporated into the audit process.

3.3.2 Data collection procedures and description of the dataset

Minimum Dataset Forms

Data is collected by hospital providers using MDS forms and then uploaded via a secure online portal. MDS data fields include demographic details, the indication for surgery and details of the procedure itself. The forms are subject to a regular formal review process to ensure that they remain fit for purpose. The shoulder database contains records using MDS version 5 (April 2012 to November 2014), version 6 (November 2014 to June 2018) and version 7 (June 2018 onwards). Data for this thesis was all entered using MDS versions 5 and 6. The variables used in this thesis were consistent across the two versions and are summarised in table 3.1. MDS forms are included in appendix A.5.

Component identification numbers

All components used in a joint replacement are uploaded by hospitals into the NJR portal either via barcode scanning or manual entry. The current component catalogue contains manufacturer names and free-text descriptions of the components. Additional database fields classify the humeral and glenoid components. Each procedure is assigned a unique identifier to facilitate

merging of the MDS, component and PROMs data for analyses.

PROMs

A number of different PROMs collection schemes are undertaken for joint replacement surgery. Unlike hip and knee replacements, shoulder replacement surgery is not part of the mandated National PROMs Programme but the NJR does collect Oxford Shoulder Scores (OSSs).[272] Preoperative scores are submitted with the MDS data by the hospital. Postoperative score questionnaires are posted directly to patients by Northgate at approximately six months following surgery. The preoperative Oxford Shoulder Score (Q1 OSS) has been submitted from the outset of the NJR shoulder database. The British Elbow and Shoulder Society (BESS) supported collection of the postoperative Oxford Shoulder Score (Q2 OSS) during two pilot data collection windows between 2012 and 2015. Continuous Q2 OSS collection is now in progress from 2018, together with additional follow-up scores at three and five years. The full OSS questionnaire is provided within the NJR MDS form pack (appendix A.5).

3.3.3 Data linkage and patient consent

Patient-level records in the NJR are linked through unique patient identifiers, usually the NHS number. This allows tracing to other NJR records (i.e. revisions, contralateral procedures, PROMs), and mortality data provided by the ONS. Prior to surgery, all patients undergoing a planned joint replacement are asked to provide voluntary consent for the recording and storage of their personal identifying information. Where patients withhold consent, the operative details only are stored for the purposes of monitoring activity trends, and linkage to internal or external sources cannot be undertaken. Only data from patients providing NJR consent are released for research purposes.

3.3.4 Inclusion criteria

Any patient undergoing their first primary shoulder replacement procedure per side between 1st April 2012 and 13th March 2017 was selected from the NJR database. There were no patients with outlying ages <18 or >100, and no age restrictions were applied. Patients were grouped into those undergoing surgery for acute trauma and those for elective indications.

3.3.5 Exclusion criteria

Primary and revision procedures were linked via a pseudonymised patient ID and the side of surgery. Patients with no valid linkage identifier, an indication of tumour or no indication listed, and revision procedures with no linked primary or revision preceding a primary were excluded. In the absence of a linkage identifier, it is not possible to perform analysis of PROMs, revision risk or mortality. Seven patients had undergone a pyrocarbon interposition arthroplasty within a study group at a single UK centre and they were also excluded. The full NJR study flowchart is shown in figure 3.2. Where the specific type of procedure was important to an analysis, procedures with an incompatible combination between the MDS listed procedure and the component database were excluded, as described in appendix A.5.3. These procedures are described as "unclassified" where relevant.

3.3.6 NJR data flowchart

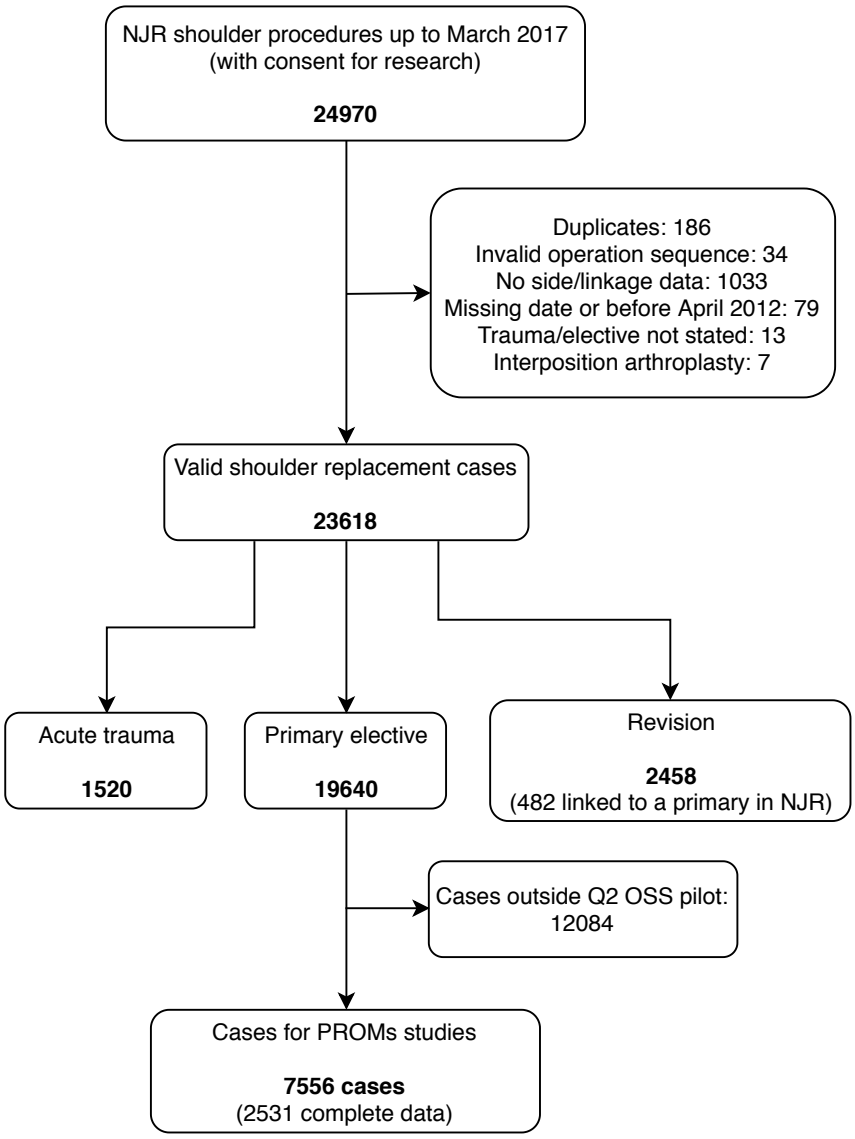


Figure 3.2: Data flow chart for NJR studies

Table 3.1: Description of variables recorded on the NJR minimum dataset forms

Variable name	Description
Age	Age in years
Sex	Male/Female
Operation date	DD/MM/YYYY
Anaesthetic type	General/Regional nerve block
ASA grade	1 to 5
Funding	NHS/Independent
Unit ID	Hospital ID (anonymised in research dataset)
Consultant ID	Consultant in charge ID (anonymised in research dataset)
Surgeon ID	Lead surgeon ID (anonymised in research dataset)
Operating surgeon grade	Consultant/ST3-8/F1-ST2/Speciality Doctor/Other
First assistant grade	Consultant/Other
Side	Left/Right
Indication	8 options
History of previous surgery	None/Excision ACJ/Stabilisation/Subacromial Decompression/Cuff Repair/Other
Procedure type	Five options as per figure 3.3
Fixation type	Cementless/Cemented/Hybrid
Approach	Delto-pectoral/Superior/Deltoid Split/Deltoid Detachment/Posterior
Minimally invasive technique?	Yes/No
Computer guided?	Yes/No
Biological Resurfacing (Glenoid)	None/Microfracture/Reaming/Interposition
Chemical thromboprophylaxis	8 options
Mechanical thromboprophylaxis	5 options
Humeral bone graft	Yes/No
Glenoid bone graft	Yes/No
Long head of biceps tenotomy?	Yes/No
Rotator cuff condition	Normal/Attenuated/Absent or Torn/Repaired
Intraoperative event	Fracture of humerus or glenoid/Nerve or Vascular injury/Shaft penetration

3.3.7 Database reformatting

Main indication for surgery

Several fields in the MDS form allow for more than one option to be selected. In this thesis, a "main" indication has been derived from the MDS options selected using the following hierarchy: 1. Malignancy, 2. Acute trauma, 3. Trauma sequelae (Post trauma), 4. AVN, 5. Inflammatory, 6. CTA, 7. OA, 8. Other.

Timing of Q2 OSS completion

The Q2 OSS is intended for completion at approximately six months following surgery. The form completion date was checked for all submitted scores. If a score was collected before four months or after 12 months, it was deleted from the dataset and treated as missing information in order to reduce any variation in scores resulting from observation during the early phase of recovery. The timing of Q2 OSS completion is shown in figure A.5.

3.3.8 Handling missing data

The dataset of primary procedures was assessed for the presence of missing data. The automated data entry portal prevents most fields from being left blank therefore the rate of missingness for MDS items is generally low or zero. There has been no mandate for the collection of preoperative PROMs by hospitals and submission rates are relatively poor. On average from 2012 to 2017, the Q1 OSS is available for 8902 (45%) of 19640 primary shoulder replacements. If anything, the rates of submission have fallen slightly over time rather than improving (see figure A.1). Postal returns by patients were considerably better. During the pilot collection windows for the Q2 OSS, 5008 (66%) of 7556 postoperative scores were completed within the 4 to 12 month window. Complete pre- and postoperative PROMs data is only available for 2531 (33%) of those 7556 procedures. Simple mean imputation of partial scores was used if only 1 or 2 questions out of the 12 were left blank, as advocated by the OSS designers.[149]

It was assumed that less complete or absent scores were Missing at Random (MAR), which is distinct from similarly named Missing Completely at Random (MCAR). For the Missing Completely at Random (MCAR) assumption, the probability of data being missing does not depend on observed or unobserved data. For the Missing at Random (MAR) assumption, the probability of data being missing is conditional on the observed data but not unobserved data.[273] If data is truly MCAR, deletion of all affected records and complete case analysis should result in unbiased results. However, if the MCAR assumption is falsely applied, a complete case analysis is liable to return biased results, dependent on the features of the responders. There are numerous plausible influences on PROMs completion in joint replacement research including patient age and frailty, and clustering of outcomes by provider. The assumption of MAR is safer in this scenario. Retention of the observed data by using multiple imputation has been shown to be a better approach, and can provide unbiased estimates of regression coefficients and better model performance, irrespective of the proportion of missing data.[274]

For this thesis, multiple imputation was performed using a predictive mean matching algorithm (**aregImpute** function within the **Hmisc** package).[275] Predictive mean matching may be slightly less sensitive to misspecification of the imputation model than standard parametric approaches and does not assume normality of the target variable.[276] All dependent and independent variables for the planned analyses were included in the regression models. The type of humeral implant used, time to revision or death and annual unit caseload were included in the imputation as auxiliary variables. The use of auxiliary variables has been shown to yield more reliable results.[277–279] The number of imputations performed matched the percentage of incomplete records.[280] An additional 100 discarded ("burn-in") imputations were performed first to achieve convergence and improve the imputation stability. Regression analyses were repeated for each imputed dataset and the results were combined by applying Rubin's Rules.[281]

3.3.9 Reclassification of the NJR component database

The existing database fields categorising humeral and glenoid components provided insufficient information to reliably confirm the validity of the component combinations recorded for the type of construct indicated in the MDS forms. The database was not fit for the purposes of procedure-level analysis for either this thesis or the NJR Annual Report. This is, in part, due to the rapid evolution of shoulder implant technology discussed in section 1.5. The wide array of complex, highly modular components at the level of the proximal humerus is one area that is particularly poorly captured by the existing framework.

As part of the data cleaning process of this thesis, a new classification system for shoulder replacement components was devised. A minimum valid set of components was defined for each of the major construct types already illustrated in figure 1.6. 2088 unique components were manually identified by their catalogue numbers from manufacturer product literature. Characteristics of each component were used to assign them within the classification. In the updated MDS version 7, the previous number of different shoulder replacement constructs has been increased from five options to the same nine options shown in figure 1.6 By applying the new classification to the list of recorded components, they can be used to map procedures recorded on the old five-category system to the new nine-category system. The large volume of procedure-level data amassed prior to 2018 can therefore be retained in future analyses.

In 92% of cases, the components recorded could be reconciled to the surgeon-defined procedure on the MDS form. In 4.4% of cases, the component record was incomplete but not in conflict with the MDS form and in 3.6% of cases there were unresolvable conflicts. Specifically, only 38% of patients classed as a resurfacing TSR had the correct combination of components recorded. A simple mapping process is shown in figure 3.3. The reality is a little more disordered and is considered in appendix A.5.3, together with further details of the component definitions and examples of valid constructs.

The reclassification of shoulder implants devised for this thesis was shared with the NJR Component Database Working Group. It was used unaltered to support the analyses for the shoulder chapter of the 15th and 16th Annual Reports.[11,161]

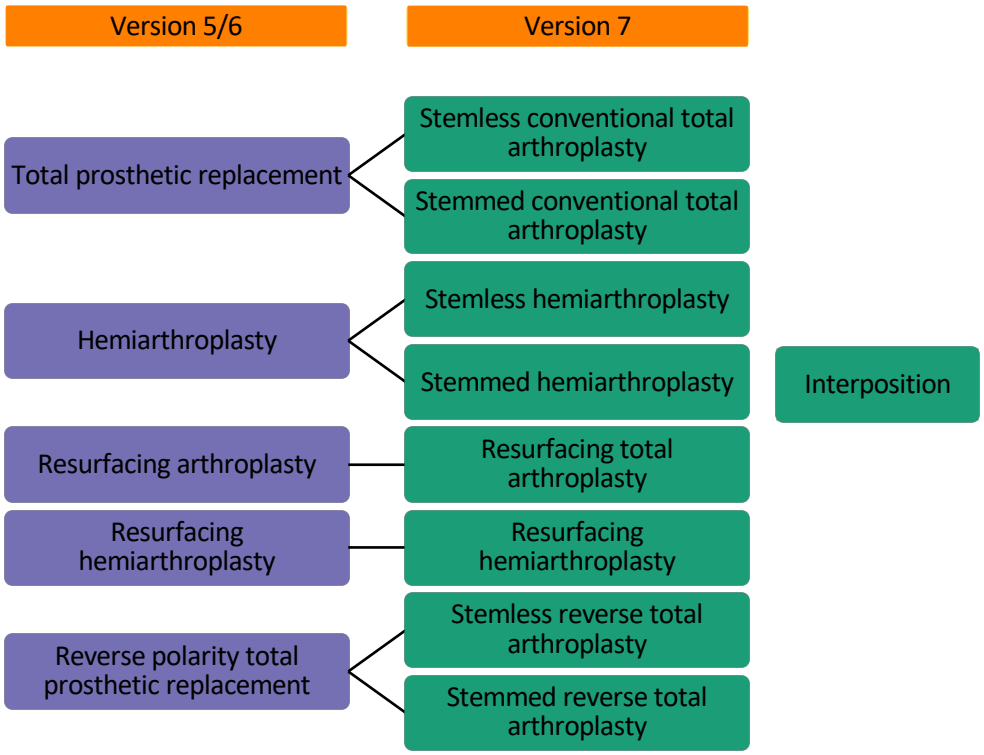


Figure 3.3: Procedure types listed on the NJR minimum dataset forms with intended mapping between versions

3.4 Unit of analysis and patient grouping issues

For all studies in this thesis, the unit of analysis was taken as the procedure, rather than the patient. This is in keeping with the established practice within both the NJR and other established international registries.

Both the NJR and HES datasets included patients undergoing surgery for elective and traumatic indications. These patient groups have been analysed as two distinct cohorts, except where stated.

3.5 Approvals and pre-registration

HES data release with linked Civil Registration of Mortality Data from the ONS was approved and managed under the terms of a data sharing agreement between NHS Digital and the University of Oxford (DARS-NIC-29827-Q8Z7Q). An initial expression of interest was

submitted to NHS digital on 12/11/2016 and the application was submitted on 12/11/2016. A partial data release was available to our research group on 13/12/2017, and then the full release was available on 10/1/2018. Studies were approved and sponsored by the University of Oxford (project ID: 12787). The HES data release was part of a larger data application for a group of research projects by the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS) Upper Limb Epidemiology Group.

NJR data release was approved by the NJR Research Committee and managed under the terms of a data sharing agreement between the University of Oxford and HQIP (RSC2016/34). The NJR research fellowship was awarded on 26/05/2016. A final data application was submitted to the NJR on 05/01/2017. A partial data release was made available on 02/05/2017 followed by PROMs data on 18/10/2017. Studies were approved and sponsored by the University of Oxford (project ID: 12730).

All data extracts were anonymised and no sensitive surgeon/provider/batch fields were received and as such research ethics approvals were not required. Relevant approval and ethics waiver letters are enclosed in appendix A.3.

3.6 Software

All initial data management and cleaning was performed using Stata IC (release version 14.2) for NJR studies and Stata MP (release version 15.0; Statacorp, College Station, TX) for HES studies. All data analyses were carried out using R version 3.4.0 and above (R Core Team; R Foundation for Statistical Computing, Vienna, Austria).

The following R packages were used extensively for the analysis: **tidyverse** (data manipulation and descriptive statistics); **survival**, **flexsurv** (time-to-event survival analyses) **Hmisc**, **rms** (regression analysis and imputation); **ggplot2**, **survminer** (production of plots). Colour-blind friendly palettes were selected for figures using **RColorBrewer**.[275,282–288]

Chapter 4

Temporal trends in the incidence of primary shoulder replacement surgery recorded in Hospital Episode Statistics for NHS England, 1998-2017

4.1 Introduction

Incidence rates of shoulder replacement surgery are rising across a number of international healthcare systems.[13] This comes in the absence of high-quality randomised studies to prove the effectiveness of the various operations performed (chapter 2). However, rates of shoulder replacement are reported to be considerably lower in the UK than those from many other countries. Based on data reported by international registries, the incidence of shoulder replacement in the UK in 2012 was $6/10^5$, compared to $10/10^5$ in Norway, $13/10^5$ in Sweden, $16/10^5$ in both Australia and New Zealand, $19/10^5$ in Denmark, and $34/10^5$ in Germany.[13] Changes in the types of shoulder replacement over time were observed internationally but the duration of data collection by the UK NJR was too short to make such an analysis.

Crude incidence rates are useful for a simple understanding of disease burden and for healthcare commissioners to understand the required resources for a specific condition. However, age structures vary from one population to another. Direct comparison of crude incidence rates between countries for the purpose of comparing intervention rates may be misleading since differences may, in part, be attributable to a difference in the population structure, particularly

where the age distribution is skewed for a given a condition or intervention. Variation based on population structure can be addressed by standardisation to calculate the incidence that would be expected if the study sample matched the structure of a reference population.

This chapter aims to describe the long-term patterns of incidence rates of shoulder replacement surgery in England and to confirm whether the upward trend from around the world is also seen here. In addition, it aims to establish whether the previously calculated rates represent an underestimate of the true incidence of shoulder replacement surgery in England.

4.2 Methods

4.2.1 Data source and participants

Data was obtained from the HES APC database for NHS England. All primary shoulder replacements performed for elective or acute traumatic indications were included. Inclusion criteria and data cleaning processes are described fully in chapter 3. No age restrictions were applied, but procedures performed for primary or metastatic tumours were excluded.

4.2.2 Calculation and standardisation of incidence rates

The number of procedures performed was counted for each complete UK financial year over a 19-year period from 1999 to 2017. Counts were further split according to sex, age (in ten-year bands) and indication (elective or acute trauma). The age groups of less than 25 years, and 85 years or more were pooled.

Mid-year population estimates by year for England from 1999 to 2017 were downloaded from the publicly available sources published by the ONS.[289] The crude incidence of primary shoulder replacement surgery was calculated as a simple rate from the total procedure count divided by the population at risk for each year. Rates for elective and acute trauma subgroups, and age- and sex-specific groups were also calculated. Incidence rates were directly

standardised to the European Population Standard (2013) using the crude age-specific rates to calculate the expected number of procedures.[290] Both the crude and standardised rates are reported.

Patterns in the types of procedure performed are presented simply as the number of cases per year, with no standardisation to the population at risk. Confidence intervals were calculated assuming a normal approximation to the Poisson distribution for count data.[291]

4.3 Results

4.3.1 Standardised incidence rates

The total number of procedures included between 1st April 1998 and 31st March 2017 was 75 655. The yearly crude incidence increased from $3.26/10^5$ (95% CI 2.70 to 3.82) in 1999 to $12.33/10^5$ (95% CI 11.40 to 13.25) in 2017 (figure 4.1). The corresponding standardised incidence rates are slightly higher than the crude incidence rates ranging from $3.89/10^5$ (95% CI 3.24 to 4.52) in 1999 to $13.64/10^5$ (95% CI 12.65 to 14.64) in 2017. The rate of change in incidence approximates to a linear relationship with time. The is no obvious indication of a peak having been reached.

Overall, approximately four elective shoulder replacements were performed for each acute trauma procedure. Whilst the rate of elective surgery has steadily increased from a standardised incidence of $2.72/10^5$ (95% CI 2.18 to 3.26) to $11.79/10^5$ (95% CI 10.87 to 12.70), the incidence of operations performed for acute trauma has only increased by 1.6 times and may have peaked at $2.21/10^5$ (95% CI 1.79 to 2.62) in 2014 (figure 4.2).

4.3.2 Age- and sex-specific incidence rates

Dramatic differential changes in incidence over time are revealed when age- and sex-specific rates are calculated. Figure 4.3 (elective) and figure 4.4 (acute trauma) demonstrate that the

change in the overall population incidence of shoulder replacement surgery is strongly driven by patients aged 65 to 84 years. Rates of change over time in these age categories are considerably higher than in younger patients. For women aged 75 to 84 years the specific incidence of undergoing a shoulder replacement in 2017 was $98.02/10^5$ (95% CI 93.35 to 102.68).

4.3.3 Types of shoulder replacement performed

Further interesting patterns are evident in the types of procedure performed over time. For elective indications (figure 4.5), hemiarthroplasty (including resurfacing) was initially the commonest procedure performed and steadily increased up to a peak of 2578 cases performed in 2010. However, since then the number of hemiarthroplasties has declined rapidly to only 967 cases in 2017. Over the same period of time, conventional TSR use has continued to increase steadily whilst RTSR use has taken off and has become the most commonly performed procedure (2776 cases in 2017). For traumatic indications (figure 4.6), very few conventional TSRs are performed. Hemiarthroplasty has again been the traditional implant of choice but there was a marked downturn seen from 2015 onwards. RTSRs were the most frequently performed replacement for acute trauma in 2017.

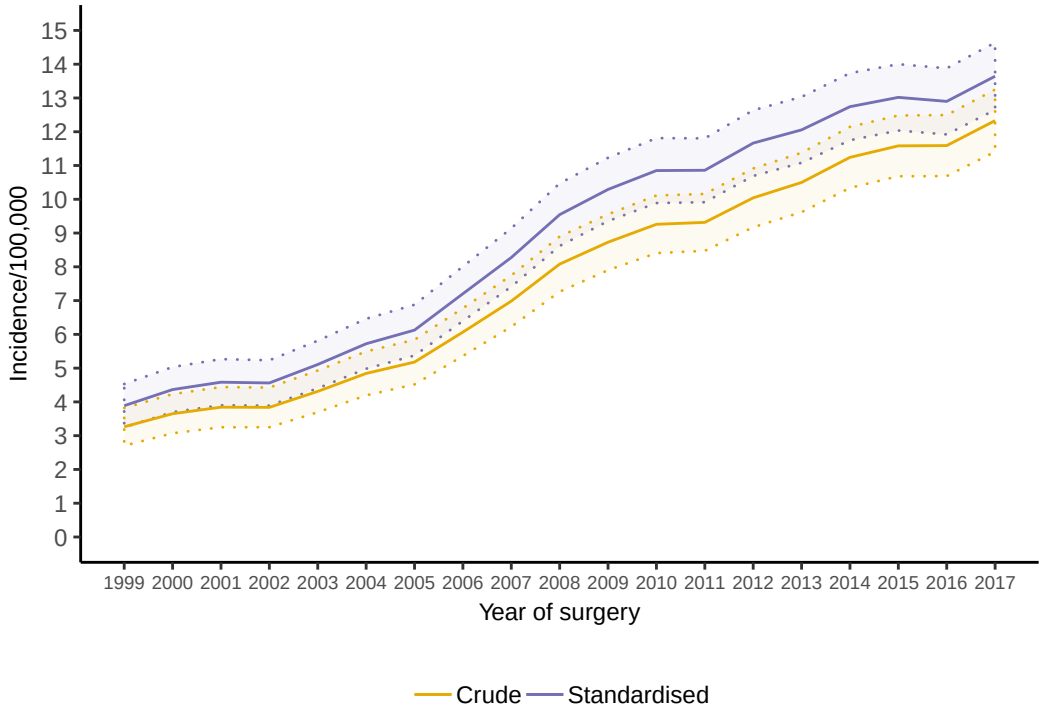


Figure 4.1: Crude and standardised incidence of primary shoulder replacement surgery, England 1998 to 2017

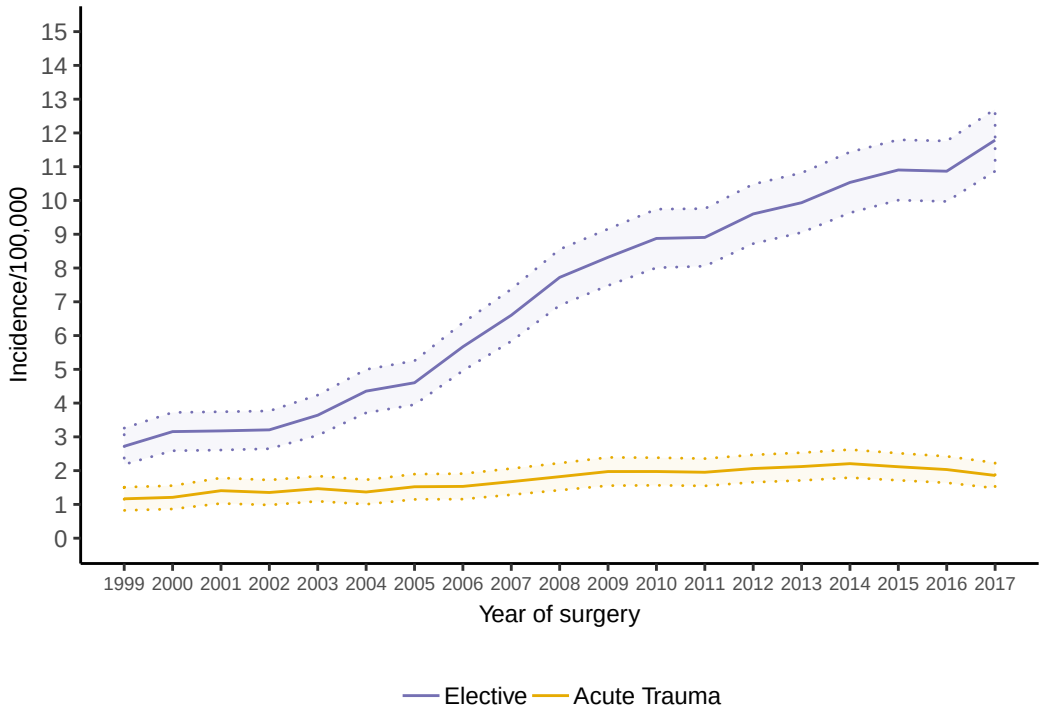


Figure 4.2: Standardised incidence of primary shoulder replacement surgery divided by elective and traumatic indications, England 1998 to 2017

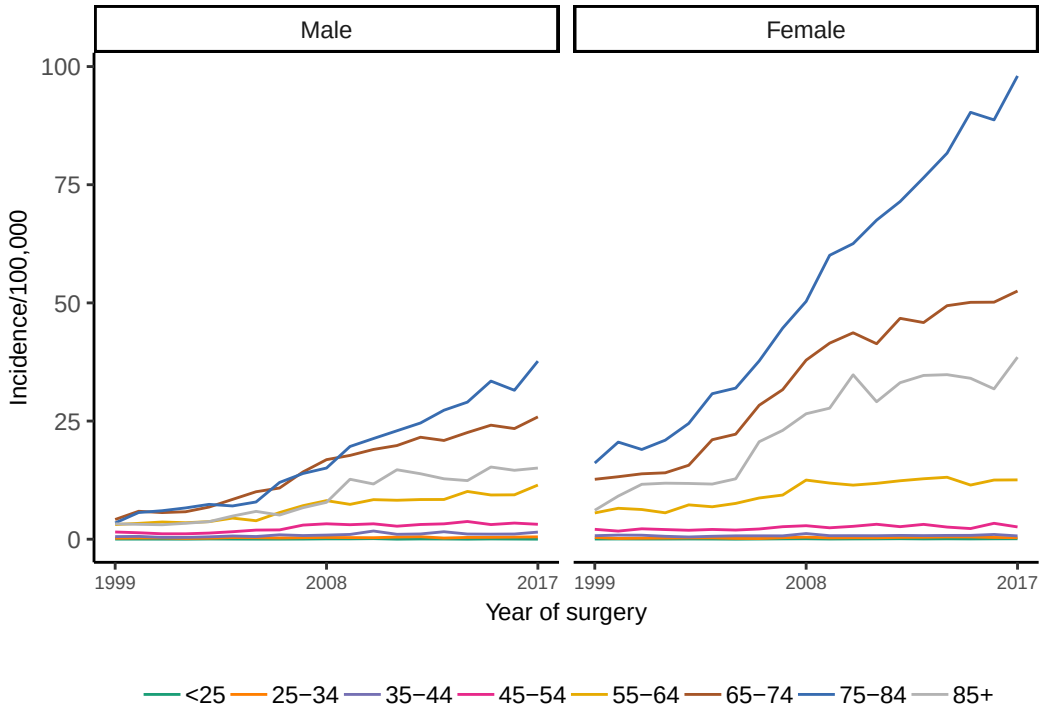


Figure 4.3: Age- and sex-specific incidence of primary elective shoulder replacement surgery, England 1999 to 2017

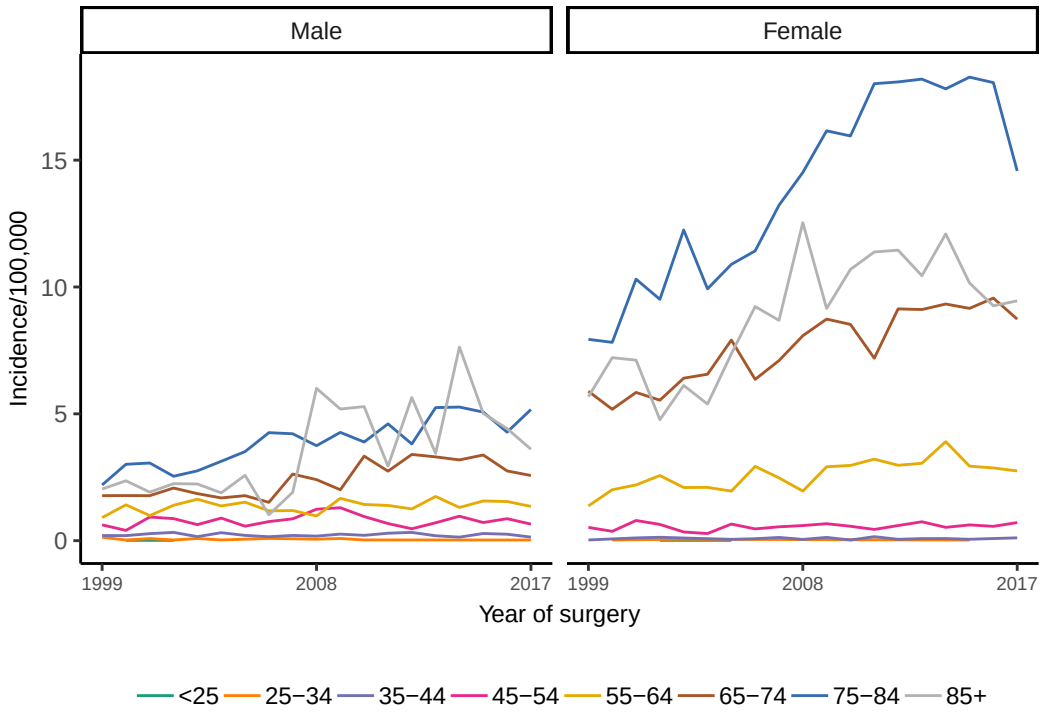


Figure 4.4: Age- and sex-specific incidence of primary shoulder replacement surgery for acute trauma, England 1999 to 2017

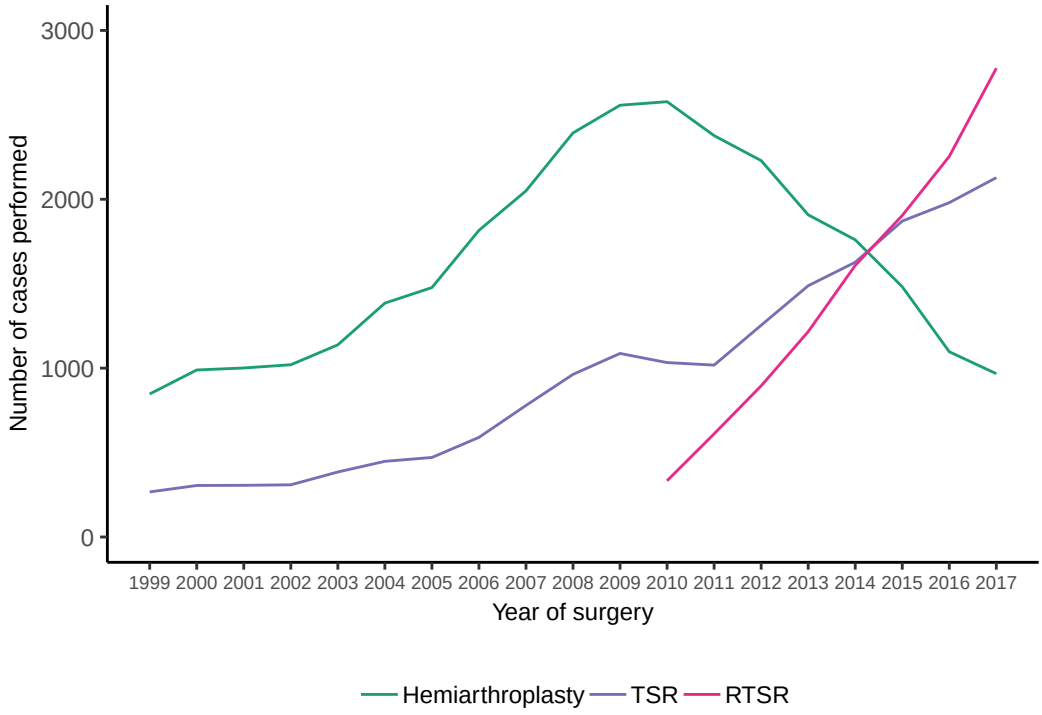


Figure 4.5: Annual volume of primary elective shoulder replacements by procedure type, England 1999 to 2017

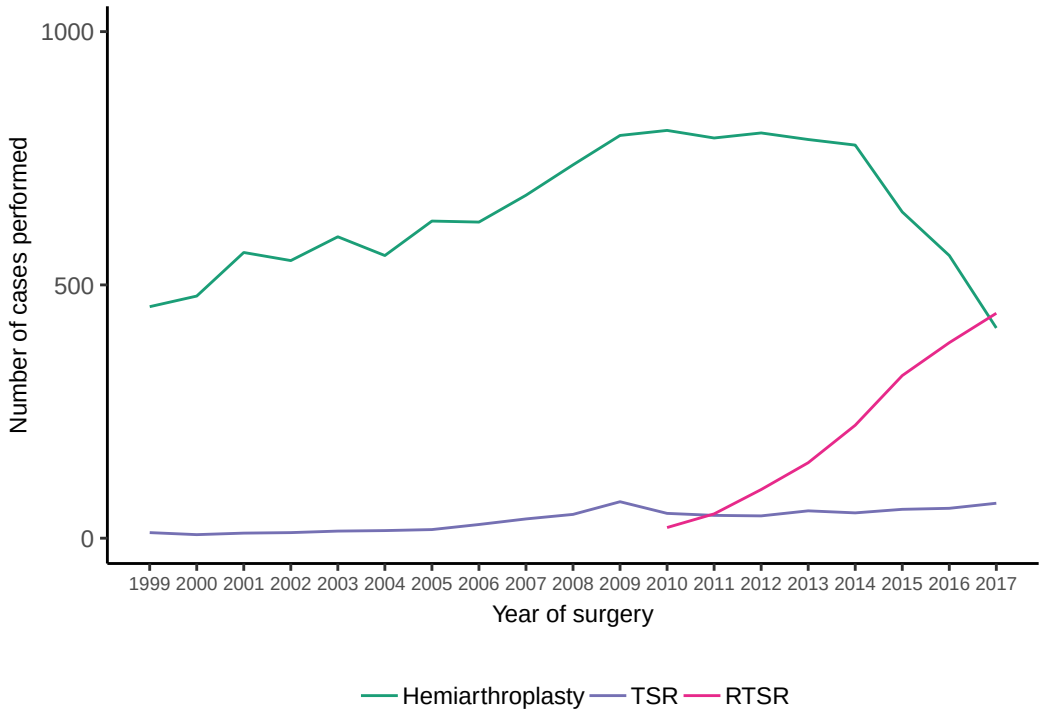


Figure 4.6: Annual volume of primary shoulder replacements for acute trauma by procedure type, England 1999 to 2017

4.4 Discussion

4.4.1 Key findings

The incidence of elective shoulder replacements performed has increased more than fourfold over a 19 year period, with no signs of any slowdown in this trend. In the equivalent time, operations performed for acute trauma have only increased by 1.59 times and their use may have peaked. Underlying these trends, there has been a considerable shift in the types of surgery performed and RTSRs are now the most commonly performed procedure in both settings. The rapid rate of uptake of RTSR is quite clearly shown by figures 4.5 and 4.6. The corresponding drop-off in the usage of hemiarthroplasty over the second half of the study period may be blunting the increase in rates of surgery overall. For acute trauma, the currently observed plateau may yet prove to be a temporary effect if RTSR use continues to increase at the current rate. The most profound expansion in elective surgery has come within patients aged 75 to 84 years, independent of the number of older people at risk.

4.4.2 Strengths and limitations

This study provides the largest and longest time series of shoulder replacements from UK data to date and provides reliable estimates through use of a population-level dataset with universal coverage of NHS England, including procedures funded by the NHS but performed in the independent sector. A small additional fraction of operations performed annually with purely independent funding will not be captured by this study.

As described in chapter 3.2.1, the overall completeness of shoulder replacement records is good. However there is some variability in the coding of the type of shoulder replacement performed, which is correct in approximately 87% of cases.[261] The procedure-specific estimates should be interpreted in this context. In particular, the specific code for RTSRs was only introduced for use in the 2010 financial year, despite the approved use of RTSR preceding this. Prior to the introduction of the new code, the number of RTSRs performed is most likely hidden in the TSR

group. Minor humps in the TSR plots in figures 4.5 and 4.6 from around 2007 to 2010 may include some of these miscoded procedures.

Given the rapid increase in RTSR in older patients, it would be interesting to understand this in the context of associated tears of the rotator cuff or even CTA. Whilst there are ICD-10 codes for rotator cuff tears in HES, there is no code for CTA and no way of differentiating between complete and partial tears of the cuff. Analysis by rotator cuff status was not feasible. A more accurate picture of indications for surgery and status of the rotator cuff could potentially be gained from NJR data, but since that dataset was only conceived in 2012 it is of insufficient duration to robustly examine trends over time.

4.4.3 Comparison to other studies

The estimated incidence of shoulder replacement surgery from this study is considerably higher than that reported between 2012 to 2014 by Lubbeke et al, using publicly available data from the UK NJR.[13] During the first few years of the NJR shoulder database, reporting rates may have been incomplete. In addition, the NJR also includes patients from Wales, Northern Ireland and the Isle of Man. We do not know if the incidence rates for these areas are comparable to England. The revised estimates from this study are similar to those from Norway and Sweden, but still some way lower than those of Australia, New Zealand, Denmark, and Germany.[13] Some of these higher crude incidence rates may be in part explained by increased demand from older populations. For example, Germany is known to have a higher Old Age Dependency ratio than the UK.[292] However, variation in healthcare is usually a function of more than just variation in demographic determinants of illness. The 2011 King's Fund Report, "Variations in Health Care" outlined supply factors which affect rates of service use including local and national guidelines, government policy, prevailing customs and clinical decision making, service configuration, resource availability, and private provision. Demand-factors include GP and patient decision making, and commissioning priorities.[293] Shoulder replacement surgery falls into the category of "preference-sensitive" or arguably "supply-sensitive" care, and thus local, regional and international variation is expected.[294,295]

The age-specific patterns observed here were also found in studies using data from the US Nationwide Inpatient Sample and from the Kaiser Permanente healthcare system in California.[12,14,296] However, the absolute rates in the USA were considerably higher. Procedure-specific trends from the Kaiser Permanente study also demonstrate marked increases in TSRs and RTSRs, alongside a decline in humeral hemiarthroplasty.[14] Kim et al found that the increase in overall rates of TSR and RTSR occurred abruptly with the approval of RTSR for use in the USA by the FDA in 2003.[12]

4.4.4 Meaning of the study

The overall increase in the number of shoulder replacements performed reflects growing confidence of patients and clinicians in their use for treating glenohumeral arthropathies. There are important implications for those responsible for commissioning care and managing healthcare budgets. The huge expansion in incidence rates within older age groups may have a magnified impact in the context of an increasingly top-heavy population.[292] Whilst revision surgery was not the focus of this study, a larger number of primary shoulder replacements at risk will also inevitably result in more patients requiring revision surgery for failed implants if revision rates remain constant.

The growing confidence in elective shoulder replacement is not supported by evidence from high-quality studies, since the systematic review in chapter 2 showed that these are lacking. Such diffusion of technological advances ahead of high-quality evidence is a widespread phenomenon.[295] The substantial increase in the use of RTSR is an example of how a new procedure has the potential to change the traditional indications for surgery. The traditional use of humeral hemiarthroplasty was unsatisfactory for the management of the profound functional deficits seen in some patients with a rotator cuff deficient shoulder.[297] RTSR seems to provide a better solution for these patients.[298,299]

Early authors of RTSR series were cautious to recommend against its use outside of a well defined group of older patients with CTA.[117–119] A survey of the BESS membership found that UK surgeons almost universally offer an RTSR to patients aged 65 years or older with CTA.

Over half of surgeons also offer RTSR to patients with an attenuated rotator cuff aged 75 years or older.[300] The increased rates of surgery in older patients probably reflects the considerable burden of prevalent rotator cuff disease combined with progressively lower thresholds for using RTSR. We know from the NJR that the median age of patients undergoing primary shoulder replacement is 73 years, four years older than the corresponding median ages for hip and knee replacements.[11] It is probable that the median age for shoulder replacement may fall over time as the surgical community expands the use of RTSR into younger patients and broader indications.

Whilst the diffusion of surgical technology will occur, RCTs may still have the potential to impact the practice of UK shoulder surgeons. The PROFHER trial, which included hemiarthroplasties in the surgical arm, found no difference when comparing surgery to non-operative treatment for acute proximal humeral fractures.[128] Its conduct and reporting in 2015 corresponds to the start of a notable downturn in the rates of hemiarthroplasty performed for acute trauma found in this current study. The influence of this evidence may have been enhanced since it supported a prevailing pessimistic view of the surgical community towards the outcomes of hemiarthroplasties.[301] We can see that hemiarthroplasties are rapidly being replaced in the acute trauma setting by RTSRs, which are surrounded by greater optimism but no high-quality evidence as yet.

4.4.5 Unanswered questions and future research

This study has revealed striking upward trends in the use of shoulder replacements and considered some plausible explanations and underlying factors. These trends in the use of shoulder replacement have not been driven by high-quality evidence. Subsequent chapters of this thesis will examine whether the observed outcomes achieved within the same population support the expanded practice of shoulder replacement surgery.

Chapter 5

What are the long-term and lifetime risks of revision surgery following a primary shoulder replacement recorded in Hospital Episode Statistics for NHS England?

5.1 Introduction

The risk of revision surgery is one of the commonest metrics used to assess the outcomes of joint replacement surgery and is the mainstay of national joint registry reports. It is most often calculated and presented as time-to-event survival analyses at fixed time points, typically five and ten years.[140] The most recent annual report of the NJR only contains shoulder data up to six years.[11] How long a joint replacement will last is also a common question posed by patients. A study of patient preferences found that patients may favour risk description in terms of lifetime risk, and prefer longer over shorter follow-up for fixed time estimates.[207] Lifetime risk statistics are commonly used in cancer epidemiology, but have more recently been used for the first time to describe revision risk following lower limb joint replacement surgery, where the increased

Elements of this chapter have been published in the BMJ and as a conference abstract from the BESS Annual Congress.[302,303]

risks for younger patients have been starkly exposed.[304]

Increasingly there are reports from single institution studies, case series and opinion pieces that report encouraging results for shoulder replacement surgery performed in younger age groups.[305–310] The availability of new bone-preserving options (stemless humeral components) and new bearing materials (pyrocarbon) combined with an absence of reliable alternatives may drive higher intervention rates in these patients. The outcomes of surgery reported in small series by early adopting surgeons may not be generalisable to the outcomes achieved in everyday practice. More reliable information regarding the true effects of patient factors may be obtained from datasets covering a wider range of practice and patients. A Californian registry study including 504 patients under 59 years found that found revision rates at two years after primary shoulder replacement were much higher in this group than the 2981 patients aged 60 years or older.[188] This finding warrants further investigation in a series with longer follow-up.

Previous reports are limited to short follow-ups. This study uses national data with up to 19 years of follow-up to estimate the long-term risks of revision framed in the context of an individual's life expectancy. Age and sex are non-modifiable factors influencing outcome, and their effect on revision risk is considered.

5.2 Methods

5.2.1 Data source and participants

Data was obtained from the HES APC database for NHS England. All primary shoulder replacements performed for elective or acute traumatic indications were included. Inclusion criteria and data cleaning processes are described fully in chapter 3. Patients aged less than 18 years were excluded, along with procedures performed for primary or metastatic tumours. For the calculation of long-term time-to-event revision risks, the whole study cohort was used. Lifetime risk of revision was only calculated for patients aged 50 years or more undergoing elective surgery. In younger patients, longer follow-up is required to more reliably inform estimates in the context of life-expectancy. In addition, fewer patients underwent surgery below

50 years of age and so too few patients remain alive and at risk to produce reliable estimates.

5.2.2 Data processing

Data processing steps to establish a clean dataset of primary procedures together with dates of linked revision surgery episodes were completed as described in chapter 3. The resulting time-to-event dataset was in a format suitable for survival analysis. The following additional categories of further surgery were also added in as time-to-event variables and described as reoperations: subacromial decompression and/or acromioclavicular joint excision, rotator cuff repair, manipulation under anaesthesia and/or release, washout/debridement, surgery for instability, fixation or periprosthetic fracture. The OPCS-4 codes used to identify these procedures are listed in appendix A.4.

5.2.3 Statistical analysis

Patients were grouped into five-year bands based on their age at the time of primary surgery. Those aged 18 to 49, and those aged 85 years and older were pooled. For the primary analysis, age- and sex-specific risks of revision and reoperation were calculated for combined elective indications and procedure types. Secondary analyses were stratified by procedure type and indication. For patients who had undergone both a left and a right shoulder replacement the implant survival and censoring were assumed to be independent, in line with routine joint replacement registry practices.[311]

Lifetime revision risk calculations

The actuarial life table method was used to calculate revision-free and reoperation-free implant survival rates over the 19-year follow-up period available. Revision, reoperation, and mortality events were applied as multiple decrements at the end of each single-year period. Late events occurring in small residual populations at risk have the potential to generate unusual leverage effects on the overall estimates, therefore years of follow-up with fewer than 100 participants

remaining at risk were arbitrarily excluded. Patients who were censored at the end of the study period were assumed to contribute to half of their final year at risk.[291]

The methodology to calculate lifetime risk was adapted from standard cancer research methods, recently applied in a study calculating risks of revision after hip and knee replacement surgery.[304,312–314] Person-time incidence rates for each year of follow-up were calculated as the quotient of the number of revisions (or reoperations) performed divided by the total time at risk. A standardised hypothetical population was created for each of the five-year age bands from 50 to 54 years through to 85 years and older. Age- and sex-specific person-time incidence rates and ONS mortality rates were applied to these hypothetical groups to calculate their "current probability" of revision for that year of follow-up. Each of the current probabilities for sequential years were summed to produce the overall estimated lifetime risk. The same method was applied for implant revisions only, as well as reoperations including revisions. The formulae for calculations are listed in full in appendix A.6.

Long-term revision risk calculations

The Kaplan-Meier method was used for time-to-event implant survival analyses, stratified where indicated by age and sex. In Kaplan-Meier estimates, censoring is non-informative i.e. those censored through loss to follow-up or death are presumed to have the same risk of experiencing an event as those remaining at risk for analysis.[315] This method is commonly used in studies of joint replacement outcomes and allows direct comparison of the estimates in this study to those already published in the literature. Revision risk was calculated as 1 minus the Kaplan-Meier survival estimate. Confidence intervals were calculated by the Greenwood method on a log-log scale.[316] Revision estimates calculated in this manner are sometimes referred to as "net" revision risks.[317]

Sensitivity analyses

Calculation of long-term and lifetime revision rates are, by definition, dependent on the results of procedures performed in the past with adequate periods of follow-up. In chapter 1, the

evolution of shoulder replacement surgery was discussed. To take account of changes over time, Kaplan-Meier estimates were calculated separately by the year when the primary surgery was undertaken and presented as trends. Revision risks were also calculated by the cumulative incidence method, with death as a competing risk, and compared to the estimates obtained from Kaplan-Meier analysis.[316,318] These estimates may be termed "crude" revision risk.[317]

5.3 Results

5.3.1 Demographics

After data cleaning, 75 637 patients were included in time-to-event revision or reoperation analyses (as per the study flow chart in chapter 3, figure 3.1). Lifetime risks were calculated in a subset of 58 054 patients aged 50 years or older at the time of elective primary shoulder replacement. The overall demographics of the cohort are shown in table 5.1. The median age for elective patients was 72 years (range 18 to 109) and for acute trauma patients was 74 years (range 21 to 104). The mean Charlson comorbidity index was 1.04 (95% CI 1.03 to 1.05) for elective patients compared to 0.95 (95% CI 0.92 to 0.97) for acute trauma patients. This difference was mainly accounted for by more patients scoring points for rheumatoid diseases in the elective group. The commonest indications for surgery were OA and CTA. For elective indications, 51% of patients received hemiarthroplasties, 30% TSRs and 19% RTSRs. For acute trauma the procedure split was 84%/5%/11%. In chapter 4, a recent shift towards more TSRs and RTSRs was demonstrated.

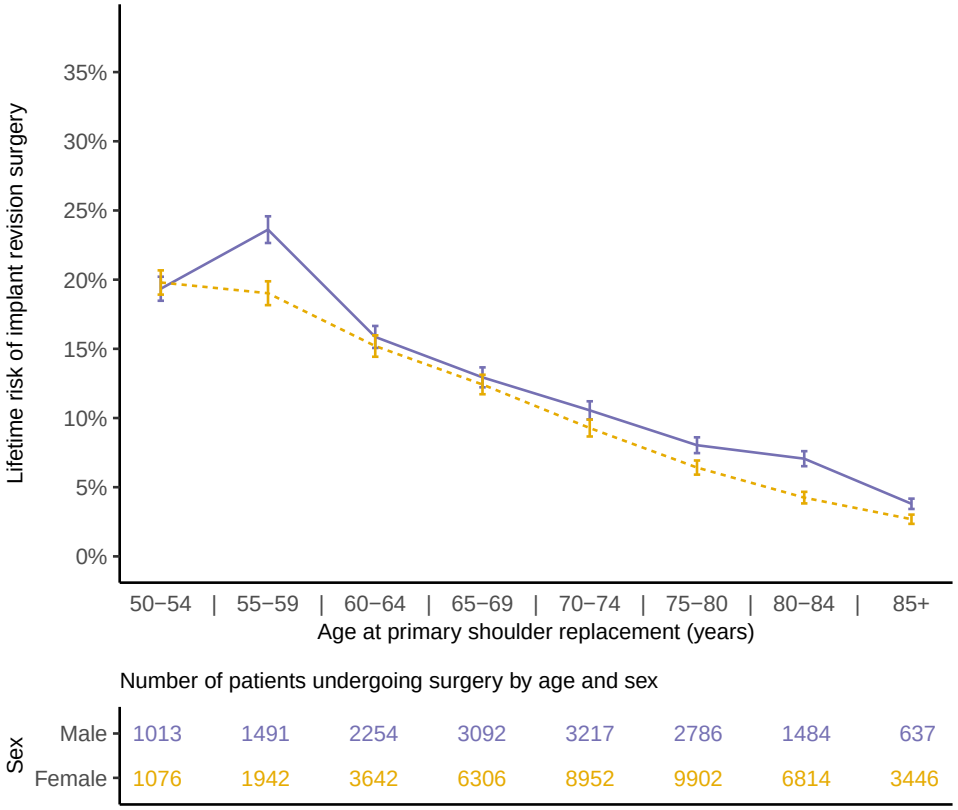
Table 5.1: Demographics of the HES cohort used in analyses of revision risk and serious adverse events

		Elective		Acute trauma	
		Male	Female	Male	Female
Sex		17532 (28.7%)	43481 (71.3%)	3610 (24.7%)	11014 (75.3%)
Age (years)	<25	52 (0.3%)	80 (0.2%)	*	*
	25-29	87 (0.5%)	75 (0.2%)	*	*
	30-34	143 (0.8%)	143 (0.3%)	*	*
	35-39	229 (1.3%)	207 (0.5%)	79 (2.2%)*	27 (0.2%)*
	40-44	417 (2.4%)	340 (0.8%)	112 (3.1%)	35 (0.3%)
	45-49	630 (3.6%)	556 (1.3%)	210 (5.8%)	109 (1%)
	50-54	1013 (5.8%)	1076 (2.5%)	305 (8.4%)	264 (2.4%)
	55-59	1491 (8.6%)	1942 (4.5%)	347 (9.6%)	552 (5.0%)
	60-64	2254 (13.0%)	3642 (8.4%)	395 (10.9%)	915 (8.3%)
	65-69	3092 (17.8%)	6306 (14.5%)	498 (13.8%)	1459 (13.2%)
	70-74	3217 (18.5%)	8952 (20.6%)	499 (13.8%)	1965 (17.8%)
	75-79	2786 (16.1%)	9902 (22.8%)	536 (14.8%)	2468 (22.4%)
	80-84	1484 (8.6%)	6814 (15.7%)	385 (10.7%)	1940 (17.6%)
	85+	637 (3.7%)	3446 (7.9%)	244 (6.8%)	1280 (11.6%)
Charlson Index (mean + 95% CI)		1.07 (1.04 to 1.09)	1.03 (1.02 to 1.05)	1 (0.95 to 1.05)	0.93 (0.90 to 0.96)
Indication	OA/CTA	11703 (67.4%)	28935 (66.5%)		
	Previous trauma	1608 (9.3%)	3788 (8.7%)		
	AVN	321 (1.8%)	728 (1.7%)		
	Inflammatory	1685 (9.7%)	5328 (12.3%)		
	Other	2215 (12.8%)	4702 (10.8%)		
Procedure type	Hemiarthroplasty	8951 (51.1%)	22125 (50.9%)	3129 (86.7%)	9110 (82.7%)
	TSR	5115 (29.2%)	13191 (30.3%)	149 (4.1%)	547 (5.0%)
	RTSR	3452 (19.7%)	8145 (18.7%)	331 (9.2%)	1357 (12.3%)
	Unspecified	14 (0.1%)	20 (0%)	1 (0%)	0

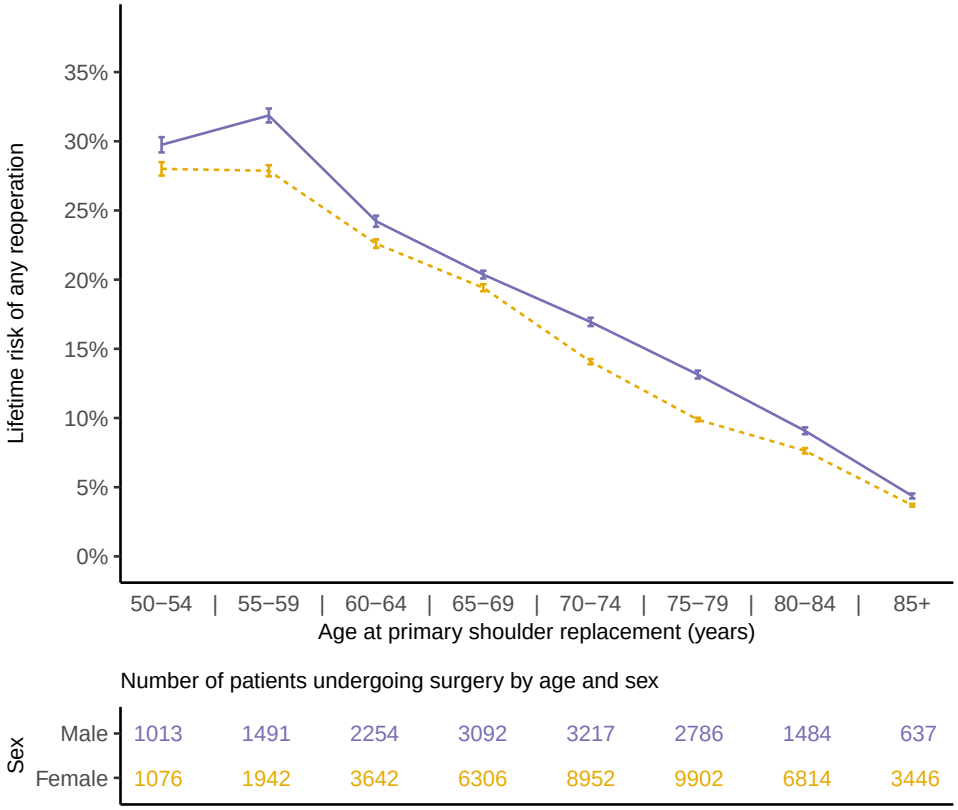
All values presented as absolute counts except Charlson Index, which is the mean + 95% CI.
*Small numbers in acute trauma category aggregated to satisfy disclosure rules - lowest age group is <40 years

5.3.2 Lifetime risk of revision or reoperation following elective surgery

The lifetime risk of revision surgery varied considerably with age. It was much higher in younger patients (aged 50 to 59) and decreased linearly with increasing age (figure 5.1a). Across all age groups, the risk in men was slightly higher than in women, however the difference was small and may not be clinically important. In men aged 55 to 59 years the lifetime risk of implant revision was 23.6% (95% CI 22.7 to 24.6) compared to 2.7% (95% CI 2.4 to 3.0) in women aged 85 years and older. When other types of additional shoulder surgeries were included, the lifetime risk of reoperation was approximately 1.5 times higher for all age groups (figure 5.1b), and may be as high as one in three for younger men.



(a) Lifetime risk of implant revision



(b) Lifetime risk of any reoperation, including implant revision

Figure 5.1: Age- and sex-specific lifetime risks of revision or reoperation following primary elective shoulder replacement surgery. 95% CIs shown. Figure appears in the published version of this work.[302] Reformatted here under CC BY-NC licence.

5.3.3 Unadjusted estimates of revision and reoperation risk

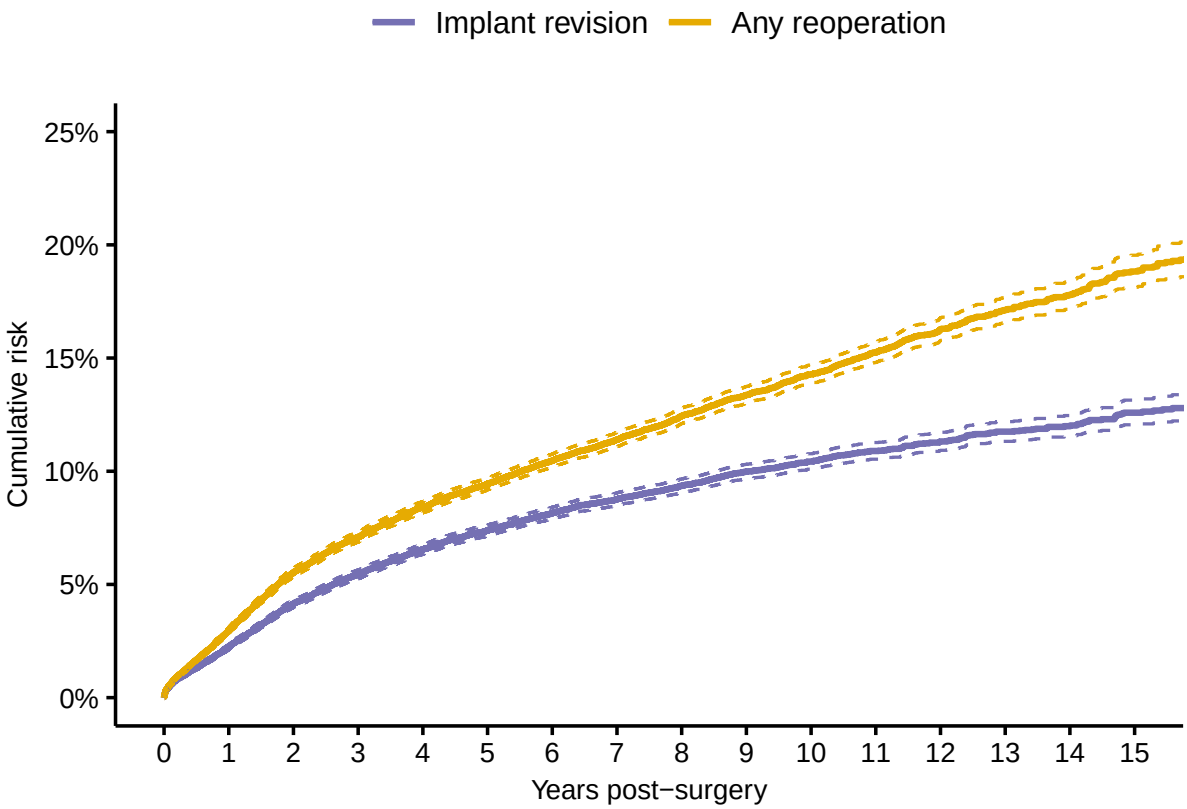
A total of 4270 elective and 701 acute trauma procedures required implant revision. Of these implant revision operations, 516 and 60 cases respectively were preceded by another reoperation. An additional 1523 elective and 235 acute trauma procedures required other operations without revision.

The risk of revision surgery estimated at 19 years of follow-up was 13.2% (95% CI 12.5 to 13.9) for elective procedures and 6.8% (95% CI 6.2 to 7.5) for trauma. Table 5.2 summarises the undjusted revision rates at up to 19 years post surgery. The patterns of revision and reoperation over time are illustrated by figure 5.2. In the first year after surgery, the revision rate was the same or slightly higher for acute trauma compared to elective procedures. Thereafter, revisions in the acute trauma group became less common and plateaued at approximately five years. In the elective group, the instantaneous hazard was also highest in the first year but an important incremental risk of revision persisted beyond five years.

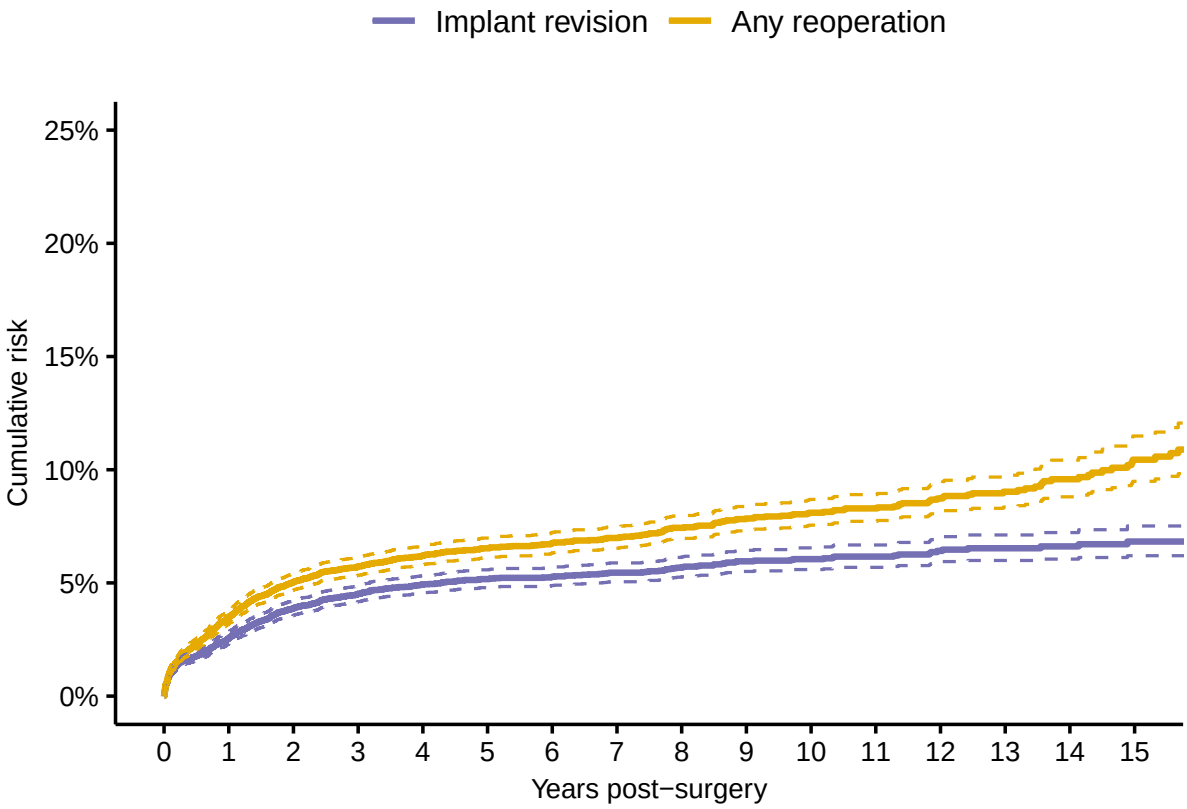
For patients whose first reoperation was not an implant revision, the commonest types of further surgery following elective shoulder replacement were subacromial decompression/acromioclavicular joint (ACJ) excision (n=968), manipulation under anaesthesia (n=583), washout/debridement (n=349) and rotator cuff repair (n=179). In the acute trauma setting, the common reoperations were manipulation under anaesthesia (n=153), subacromial decompression/ACJ excision (n=58) and washout/debridement (n=53). These procedures were not mutually exclusive.

Table 5.2: Risk of revision surgery up to 19 years following primary shoulder replacement surgery. Outcomes of elective and acute trauma procedures estimated separately.

Time (Years)	Number at risk	Number revised	Censored (Death)	Censored (Study end)	Implant revision (95% CI)
Elective					
0	61013				
1	52793	1306	1217	5697	2.25 (2.13 to 2.37)
2	45576	967	1261	4989	4.15 (3.99 to 4.32)
3	38951	575	1335	4715	5.44 (5.25 to 5.64)
4	32847	420	1435	4253	6.54 (6.33 to 6.77)
5	27593	274	1279	3697	7.38 (7.14 to 7.62)
6	22869	211	1229	3284	8.15 (7.90 to 8.42)
7	18723	140	1190	2816	8.76 (8.49 to 9.05)
8	15004	109	1053	2567	9.36 (9.06 to 9.66)
9	11738	92	892	2272	9.98 (9.66 to 10.3)
10	8974	51	753	1960	10.43 (10.09 to 10.77)
11	6821	42	576	1535	10.89 (10.53 to 11.27)
12	5118	28	475	1206	11.32 (10.92 to 11.72)
13	3816	22	379	895	11.75 (11.32 to 12.19)
14	2785	9	284	738	11.99 (11.54 to 12.46)
15	1954	16	206	609	12.59 (12.05 to 13.14)
16	1335	5	150	464	12.85 (12.27 to 13.45)
17	795	2	101	437	13.00 (12.39 to 13.63)
18	349	1	35	410	13.18 (12.48 to 13.92)
19	0	0	8	341	13.18 (12.48 to 13.92)
Trauma					
0	14624				
1	12641	357	774	852	2.55 (2.30 to 2.82)
2	11029	162	564	886	3.87 (3.55 to 4.21)
3	9575	70	537	847	4.51 (4.17 to 4.89)
4	8150	39	565	822	4.93 (4.56 to 5.32)
5	6897	19	523	710	5.17 (4.79 to 5.58)
6	5792	7	464	634	5.27 (4.89 to 5.69)
7	4853	10	375	554	5.45 (5.05 to 5.88)
8	4039	11	308	497	5.70 (5.27 to 6.15)
9	3312	10	272	443	5.96 (5.51 to 6.44)
10	2685	3	243	381	6.05 (5.59 to 6.55)
11	2147	3	205	330	6.16 (5.69 to 6.68)
12	1717	5	144	283	6.42 (5.90 to 6.98)
13	1331	2	114	268	6.53 (5.99 to 7.12)
14	1020	1	89	221	6.61 (6.05 to 7.22)
15	745	2	74	199	6.83 (6.20 to 7.52)
16	519	0	49	177	6.83 (6.20 to 7.52)
17	314	0	37	168	6.83 (6.20 to 7.52)
18	143	0	21	150	6.83 (6.20 to 7.52)
19	0	0	3	140	6.83 (6.20 to 7.52)



(a) Elective



(b) Acute Trauma

Figure 5.2: Unadjusted risk of revision or reoperation following primary shoulder replacement. 95% CIs shown.

5.3.4 Age- and sex-specific estimates of revision risk

Estimates of revision risk are shown stratified by age and sex in figures 5.3 (elective) and 5.4 (acute trauma). At all ages, the likelihood of revision following elective procedures was higher than the risk for patients undergoing primary shoulder replacement procedures for acute trauma. In older patients, the additional risk of revision surgery beyond five years of follow-up was relatively low. In younger patients, particularly following elective surgery, the risk of revision surgery seems to persist beyond 10 years of follow-up. The influence of age on the likelihood of revision is striking, particularly in the elective setting.

When death was accounted for as a competing risk, the risks of revision surgery were slightly lower at all time points. The difference was more apparent at longer-term follow-up. In the elective cohort the 15-year Kaplan-Meier estimate for revision risk was 12.6% compared to 10.5% by the cumulative incidence method. The corresponding estimates for the acute trauma cohort were 6.8% and 5.7%. The differences are presented in full in table A.11 (appendix A.6).

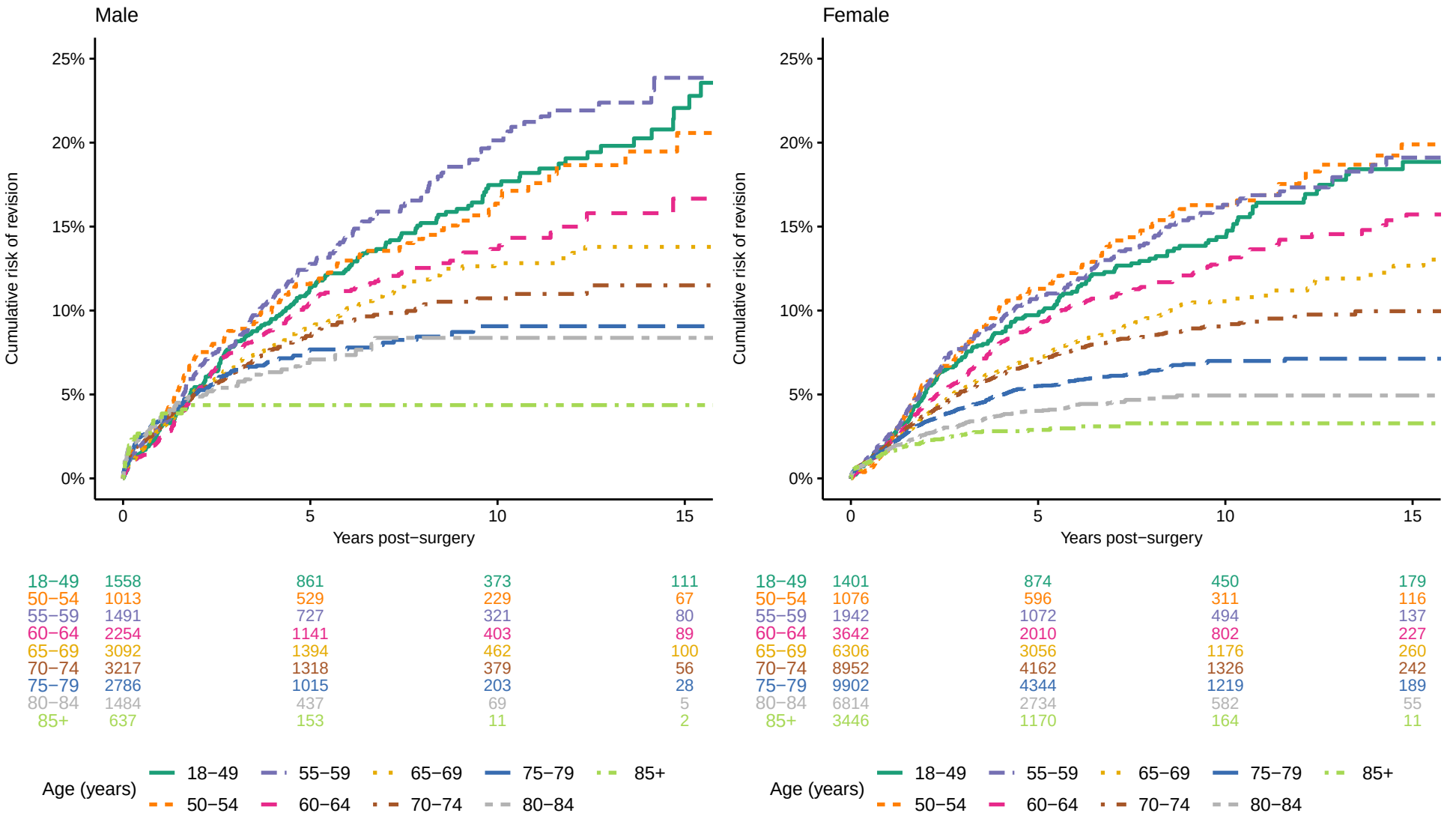


Figure 5.3: Age- and sex-specific risk of revision following elective shoulder replacement. Numbers remaining at risk shown in tables.

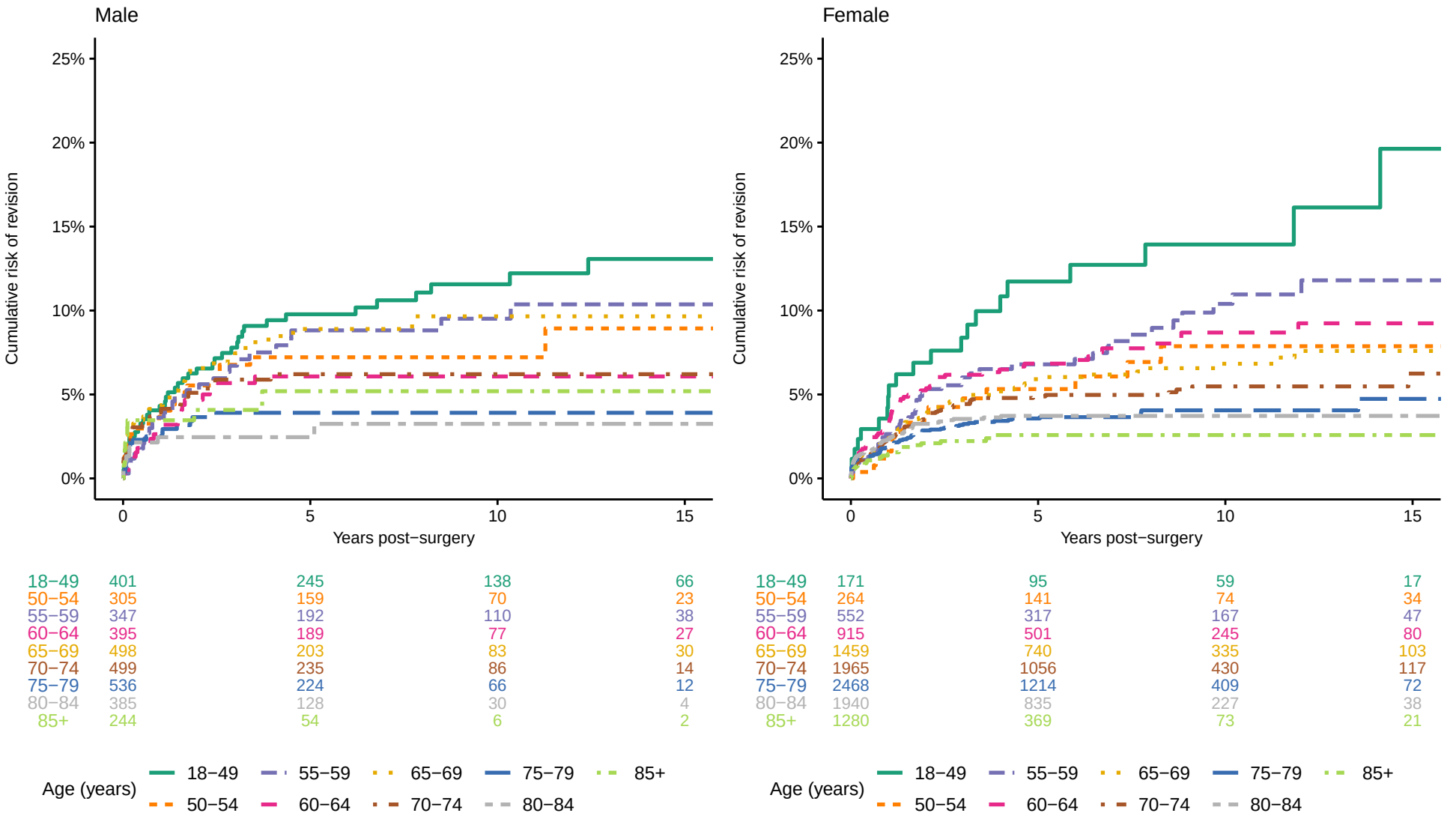


Figure 5.4: Age- and sex-specific risk of revision following shoulder replacement for acute trauma. Numbers remaining at risk shown in tables.

5.3.5 Revision risk by indication for surgery

The long-term risk of revision was lowest for patients undergoing a shoulder replacement for acute trauma (figure 5.5). Risk for patients with OA/CTA, inflammatory and other pooled indications were all similar to each other. The risk of revision for patients who had a shoulder replacement sequelae of previous fractures was much higher than for all other indications. At one year this risk was 5.1% (95% CI 4.6 to 5.8), at two years it was 8.6% (95% CI 7.9 to 9.5) and at three years it was 10.4% (95% CI 9.6 to 11.4).

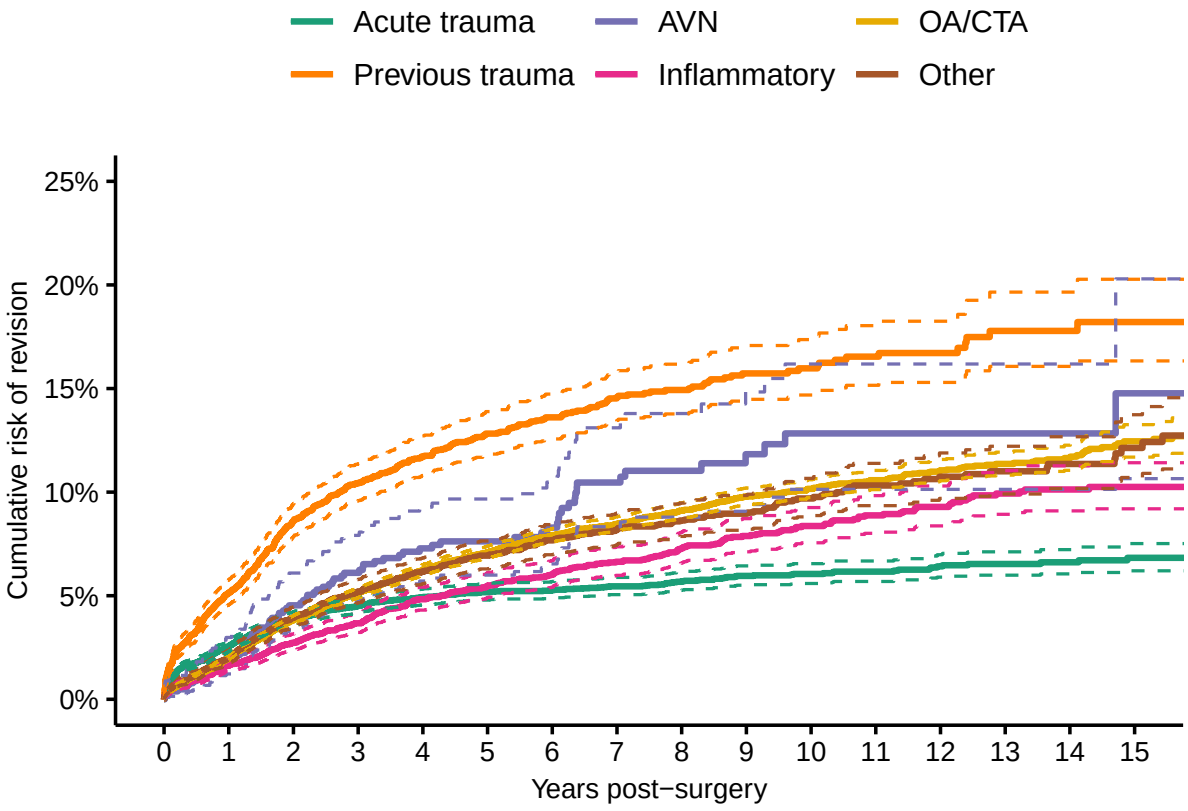


Figure 5.5: Risk of revision by indication for primary shoulder replacement

Cause-specific revision rates cannot be reliably obtained from HES data, however there are codes for septic arthritis and periprosthetic infections. Table 5.3 shows the number and proportion of cases where infection was coded at the time of the first revision or reoperation procedure. It shows that the absolute rate of infection seems to be higher in cases with previous trauma sequelae (1.9%) compared to other indications (0.8 to 1.2%).

Table 5.3: Crude risk of reoperation for infection (reported by indication for the primary shoulder replacement).*

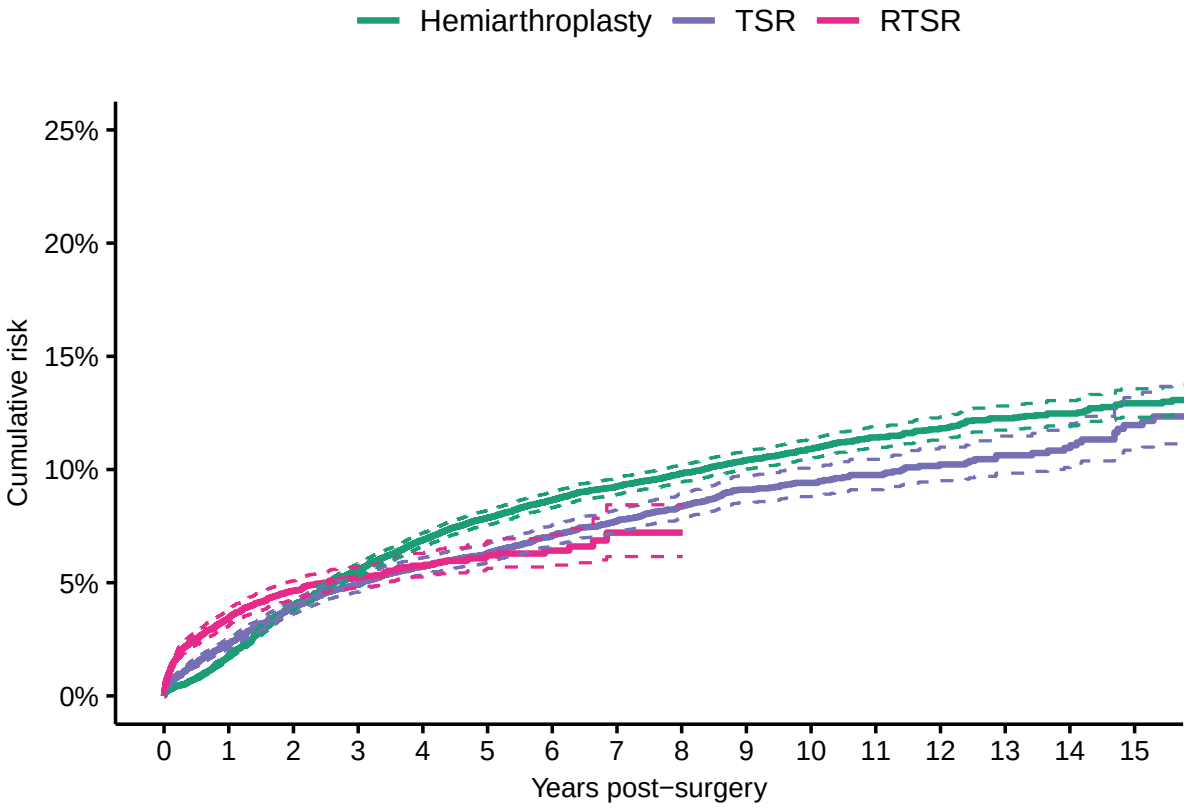
Indication	N	Any reoperation	Infections		
			n	% of total	% of reoperations
OA/CTA	40638	3697	358	0.9%	9.7%
AVN	1049	115	10	1.0%	8.7%
Inflammatory	7013	570	69	1.0%	12.1%
Other	6917	630	82	1.2%	13%
Previous trauma	5396	781	103	1.9%	13.2%
Acute trauma	14624	936	119	0.8%	12.7%

*Includes just the first reoperation or revision, and crude rates do not take censoring into account.

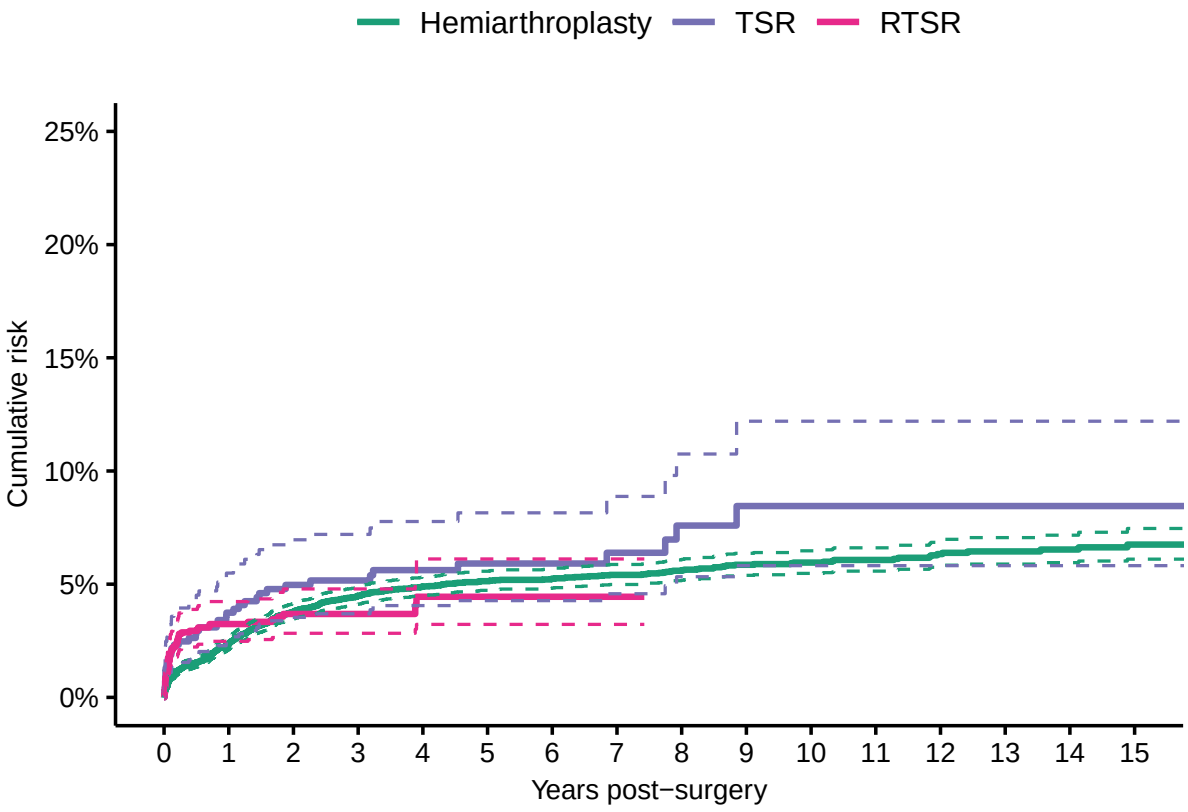
5.3.6 Revision risk by procedure type

Hemiarthroplasties, TSRs and RTSRs all exhibit different failure patterns over time. For elective indications (figure 5.6a) , hemiarthroplasties have the lowest revision rate in the first year (1.8%, 95% CI 1.6 to 1.9), strikingly lower than RTSRs (3.4% 95% CI 3.1 to 3.8). From around five years onwards, the revision risk was around 1.5% higher for hemiarthroplasties compared to TSRs. At 10 years, the revision rates were 10.9% (95% CI 10.5 to 11.3) and 9.4% (95% CI 8.8 to 10.1) respectively. Data for RTSRs beyond five years was limited, but it appears that the risk of revision may begin to plateau.

In the acute trauma setting (figure 5.6b), the high risk of one-year revision following RTSR and subsequent sharp levelling off was more pronounced. Cases recorded as TSRs seem to exhibit a similar pattern. This was a small group of patients and some of the procedures may actually represent RTSRs that were incorrectly coded prior to the introduction of specific OPCS-4 codes.



(a) Elective



(b) Acute Trauma

Figure 5.6: Risk of revision by type of primary shoulder replacement. 95% CIs shown.

5.3.7 Temporal trends in revision risk by year of surgery

Trends in the rate of revision surgery were estimated by year of the primary procedure for patients having an elective shoulder replacement (figure 5.7). No obvious patterns were observed for revisions within one or three years of surgery respectively. Five-year revision risk increased from 5.3% (95% CI 4.1 to 6.8) for primary procedures performed in 1999 to 8.4% in 2012 (95% CI 7.6 to 9.3), seven-year revision risk increased from 6.3% in 1999 (95% CI 5.0 to 7.9) to 11.0 in 2010 (95% CI 10.0 to 12.0), and 10-year revision risk increased from 7.9% in 1999 (95% CI 6.3 to 9.8) to 10.7 in 2007 (95% CI 9.6 to 12.0). No obvious patterns were seen in 15-year revision risks, but the data for this was limited to four financial years.

Based on the same principle, figure 5.8 illustrates the trends over time in revision risk according to the type of procedure performed. Splitting the data into these smaller groups sacrifices statistical power but a visual inspection of the five- and 10-year trends does not find any substantially different patterns to those overall. Revision risk following RTSR at one and three years may have fallen over time, but the confidence intervals overlap and this apparent difference may not be important.

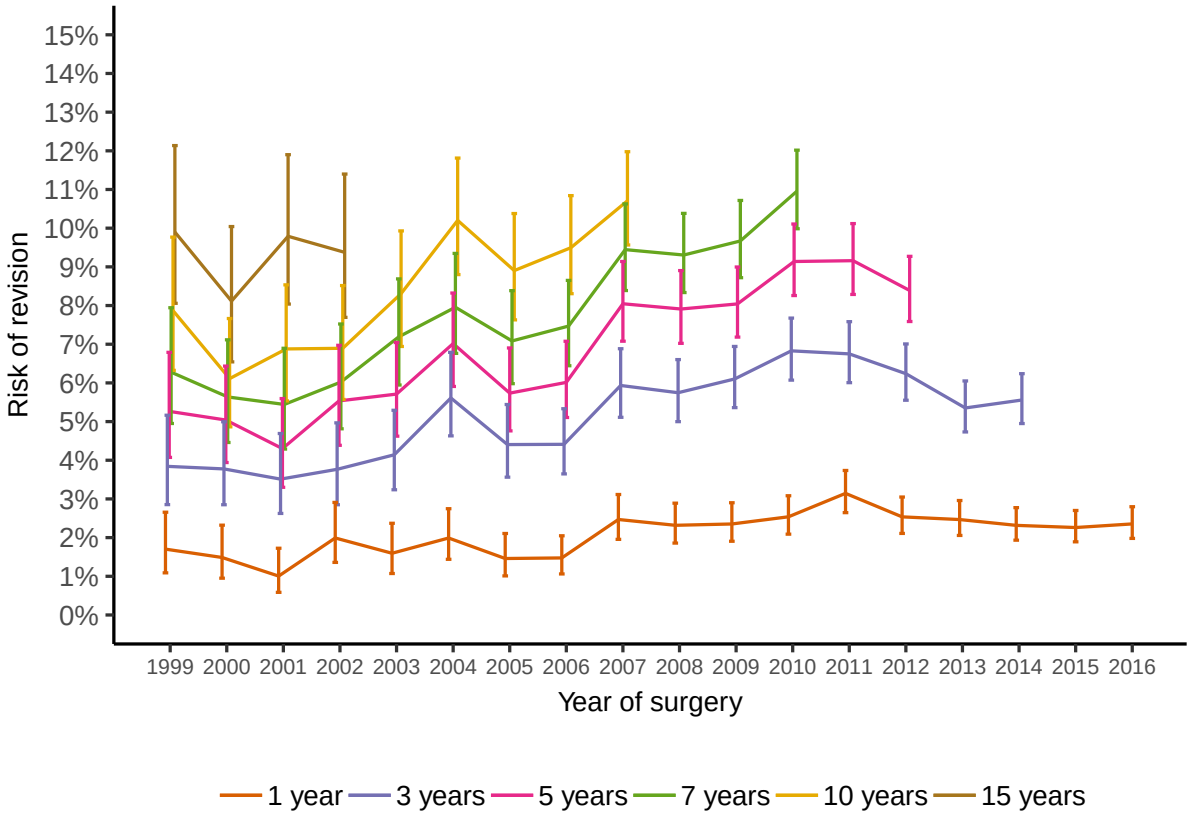


Figure 5.7: Temporal trends in revision risk at 1, 3, 5, 7, 10 and 15 years post surgery presented by year of primary surgery. 95% CIs shown.

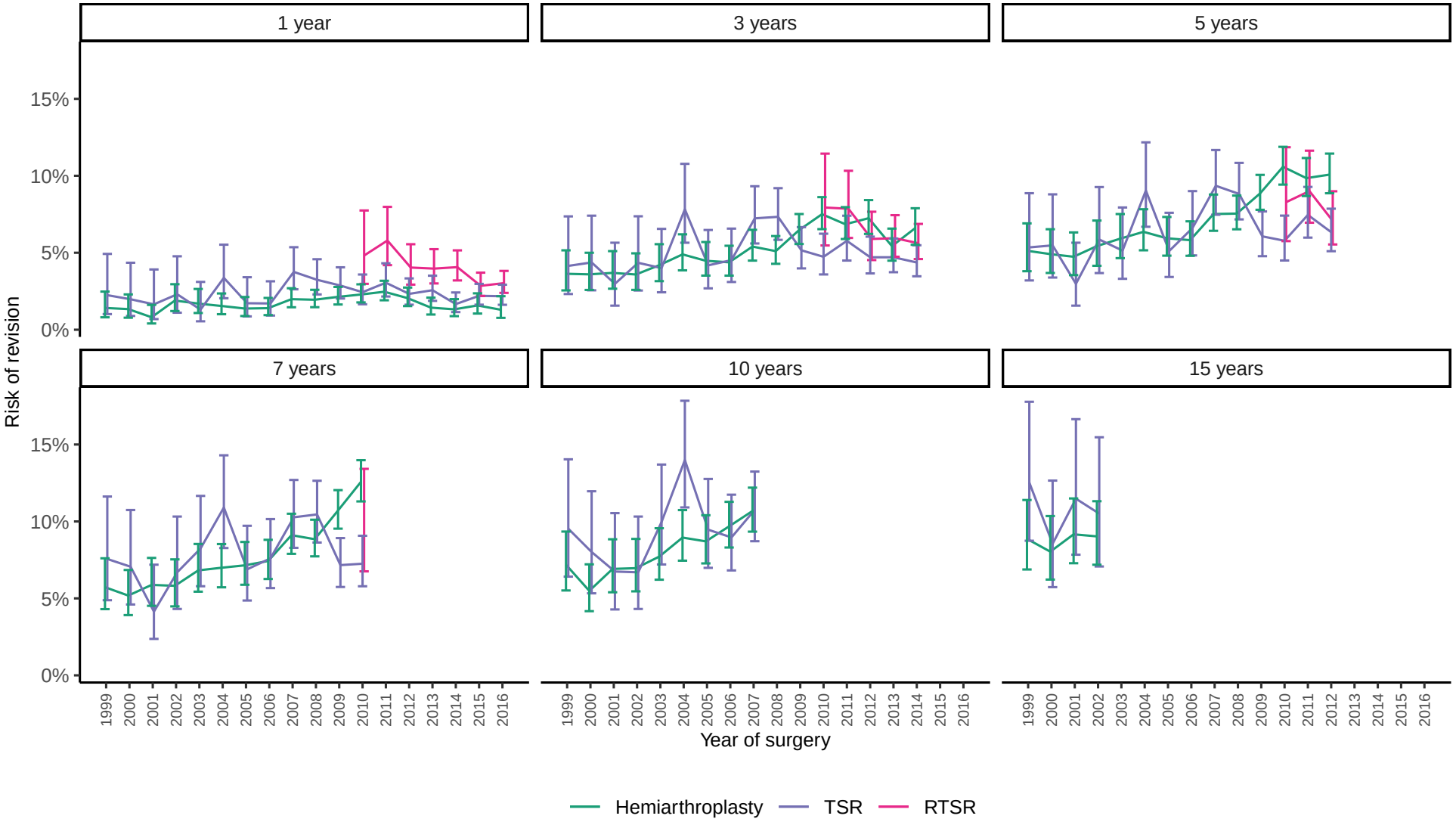


Figure 5.8: Temporal trends in revision risk at 1, 3, 5, 7, 10 and 15 years post surgery presented by year of primary surgery and procedure type. 95% CIs shown.

5.4 Discussion

5.4.1 Key findings

This study is the first to estimate the overall risk of revision surgery in the context of a patient's life-expectancy following elective primary shoulder replacement. The lifetime risk of revision surgery was strongly linked to age at the time of surgery. In particular, younger men aged 55-59 years have a lifetime revision risk of at least 23.6% (95% CI 22.7 to 24.6) compared to 2.7% (95% CI 2.4 to 3.0) in women aged 85 years and older. Standard time-to-event analyses demonstrated that more than half of revision risk occurs in the first five years following primary surgery across all age groups. When additional types of reoperation were included, the risk of further surgery was approximately 1.5 times higher than estimates for implant revision alone. Despite similar age and comorbidity profiles, the risks of revision surgery beyond five years were considerably lower following procedures performed for acute trauma. However, patients undergoing a shoulder replacement for the sequelae of previous fractures experienced an alarmingly high risk of subsequent revision surgery.

The patterns of failure observed were different for hemiarthroplasties, TSRs and RTSRs. RTSRs required more early revisions but required fewer later revisions beyond five years, whereas hemiarthroplasties and TSRs failed more consistently over time. Indications for revision were not examined in this study but the higher early revision rate for RTSRs is likely due to higher rates of instability, particularly in the trauma setting.[319–321] Patients receiving RTSRs were, on average, older than patients receiving other types of shoulder replacement and this may influence the apparently lower revision rates seen after the first three years. Medium-term revision rates (aggregated for all procedure types) at five, seven and ten years following surgery increased over the study period.

5.4.2 Strengths and limitations

This study used large-scale population-level data to estimate the risk of revision surgery in the context of a patient's life-expectancy, information which has not previously been available for patients considering a shoulder replacement. The large sample size available has facilitated a detailed analysis of the effects of age, sex, indication and procedure type on the long-term risk of revision surgery. In addition to the lifetime risk estimates, the standard time-to-event calculations allow for direct comparison to the existing literature to set the findings in context.

Lifetime risk estimates were based on non-parametric analysis of the available data. Their accuracy is influenced by the overall length of available follow-up and the censoring profile of included patients. Estimates for older patients, with a high proportion of censoring for death, are likely to be most accurate. In younger patients, with a higher proportion of censoring at the end of follow-up, lifetime risk may be underestimated. Despite this, the absolute number of revisions (n=8) and incremental risk observed beyond 15 years of follow-up were both very small, even though 1954 patients remained alive and at risk of revision. Therefore, any underestimate is likely to be unimportant. Parametric modelling could be used to predict outcomes at longer follow-ups but this would require assumptions to be made about the survival distribution.

The use of Kaplan-Meier methodology to estimate revision risk is the standard in the field of joint replacement surgery.[140,322] This may be partly due to its ease of use as well as a familiarity with both Kaplan-Meier methodology and the commonly used Cox Proportional Hazards models, which apply the same logic to censoring to estimate relative risks as hazard ratios. One of the main functions of joint replacement registries is the comparison of different implants intended for the same indications. Calculation of "net" failure by the Kaplan-Meier method is the most appropriate choice in this setting. Competing risks estimates of "crude" failure may be more useful when considering future resource use.[323,324] Regardless of the method used in this study, the differences observed at short-term follow-up were small and not clinically important. At longer-term follow-up in younger patients, the difference between Kaplan-Meier and competing risks estimates was larger (absolute difference >2%). Clinicians

need to be aware that results from studies with these different methodologies may not be directly comparable.

Submission of data to HES for each episode of patient care is mandatory and so the risk of bias from under-reporting of revision procedures is low. The capture of other types of reoperation, beyond implant revision, provides information that is rarely available from international registry reports. The coverage of the dataset is universal for NHS-funded care in England and as such the potential for confounding from local geographical, socioeconomic, and commissioning factors is lower than that for smaller studies from single institutions or regional registries. The findings are broadly generalisable to patients undergoing treatment in an NHS setting and may be useful to other healthcare systems with similar population structures.

Data quality for the HES dataset has been discussed in chapter 3. Coding for the occurrence and side of shoulder replacement surgery is highly accurate (96%) but accuracy of coding for the primary diagnosis and type of shoulder replacement surgery may be less accurate (80 to 87%).[261] The aggregate estimates by setting (elective vs acute trauma) and analyses by patient factors (age, sex) are therefore most accurate. Analyses at a diagnosis- and procedure-level should be interpreted with more caution. No codes to specifically identify CTA are available in ICD-10 and so no attempt was made to differentiate these from OA. In 11% of elective cases, the primary diagnosis was unclear from the ICD-10 codes available. It is likely that a considerable number of these patients also had OA/CTA and the patterns of failure were the same (figure 5.5). For procedure types, a specific OPCS-4 code for RTSR was only introduced in the 2010 financial year. Prior to this a small number of cases should be assumed to be misclassified within the hemiarthroplasty and TSR groups.

5.4.3 Comparison to other studies

The Mayo clinic (Rochester, Minnesota) have published long-term follow-up results from their large institutional database that are widely cited. At 20 years, they estimated a 15% revision risk for 1431 humeral hemiarthroplasties and 19% revision risk for 2588 TSRs. These estimates are higher than those from the present study but the mean age of patients in their studies was 63

and 65 years respectively, also much lower than this study. The strong association between age and revision risk may help to explain the difference between the study outcomes.

The most recent annual report of the NJR only provides age-adjusted estimates in coarse age-bands.[161] The five-year revision rates in the 2019 report were 4.8% (95% CI 4.5 to 5.1) for elective procedures and 3.6% for acute trauma procedures (95% CI 2.8 to 4.5), both considerably lower than those found in this study based on patients drawn from the same population. Definitions of what constitutes revision surgery are not uniform between registries but the NJR clearly considers a revision as: "any further operation to remove, change, add or reduce any of the components".[159,160] This guidance is not applicable to HES data and there is some potential for variation in how hospitals report these cases. However, previous audits of the NJR hip and knee datasets have found that there may be systematic under-reporting of revision surgery.[271] The shoulder dataset is less well established and may suffer from the same risk of bias. HES data does not suffer from this bias and therefore the higher revision risks found here may potentially be a closer reflection of true outcomes.

The NZJR (New Zealand) has data at up to 18 years on a cohort of 9250 primary shoulder procedures. Their estimated 15-year revision risk of 10.8% was very close to the estimate in this thesis, in a population with similar reported demographics. The Norwegian registry has been collecting shoulder data since 1994 and report survival curves as up to 20 years of follow-up. The patterns of failure are similar to this study but the underlying data from the Norwegian curves are not published therefore direct comparisons are not easily made.[325] Revision risks at 10 years in the 2018 Swedish report were somewhat lower but their report did not provide any demographic data to help put these findings into context.[326] The AOANJRR (Australia) has data up to 11 years and their reports are highly detailed, breaking down results by implant type, diagnosis, age and sex.[319] Failure patterns are consistent with the trends seen here, but few patients remain at risk beyond 10 years for each individual analysis and so implant survival estimates are accompanied by wide confidence intervals. None of these national databases have a large enough dataset with sufficiently long follow-up to provide accurate age- and sex-specific estimates of long-term revision risk to compare with those presented in this study.

5.4.4 Meaning of the study

The information regarding both the lifetime risk of further surgery and the time-specific risks of revision have been calculated and presented in sufficient detail to support meaningful preoperative conversations with patients regarding the longevity of their proposed shoulder replacement surgery. Age has been shown to be a powerful predictor of revision risk and younger patients should be made aware that their risks are much higher. Unadjusted estimates of revision risk are strongly weighted by the demographic composition of the patients included in the study and should not be applied universally across different patient groups. Discussion of the very much higher rates of revision surgery with younger patients may help to guide their expectations, and support the comprehensive use of other treatment options prior to considering joint replacement surgery. Revision risks in older patients are very much lower, although still of sufficient importance to be carefully discussed with them.

5.4.5 Unanswered questions and future research

Patient frailty, personal circumstances, quality of bone stock and soft tissues, and surgeon skill in revision surgery may all influence the decision to undergo or undertake revision surgery. Not all patients with a poor outcome or failing shoulder replacement will have surgery. Further research is therefore needed to understand what the long-term outcomes are in terms of function, quality of life and satisfaction, which this study does not address. More precise prediction of outcomes, taking into account additional factors, may be desirable. The non-parametric methodology applied for lifetime risk calculations in this study does not lend itself to modelling of outcomes at an individual patient-level. An alternative parametric modelling approach could allow the study of multiple covariates.

Rates of failure in patients treated for the sequelae of previous trauma are alarmingly high. In this study the previous trauma group did include patients who had undergone a previous osteosynthesis. Cause-specific reasons for revision were not investigated independently but this challenging patient group may experience more early failures due to infections. Future studies may test this theory, and identify modifiable factors and strategies to improve outcomes. They

should also investigate whether earlier intervention in the acute trauma phase would improve outcomes and reduce the frequency of complications.

Any long-term outcomes study, this one included, is unavoidably weighted by the inclusion of older designs of shoulder joint replacement. It could be expected that innovation in implant technology would result in better-performing, longer-lasting shoulder replacements. Unexpectedly, the analysis of trends over time suggested that medium-term revision risk actually may be increasing over time. A similar pattern of worsening outcomes was observed in NJR data for hip replacements, where the seven-year revision risk doubled for primary procedures done between 2003 and 2008.[161] It was thought to be largely driven by metal-on-metal bearing failures, which were prevalent at the time. Less-marked patterns were seen for knee replacement surgery. The NJR clinician-feedback process is one suggested mechanism for the timing of improvements in hip and knee outcomes. Wider availability of surgeons skilled to perform reverse total shoulder replacement as a revision option may be contributing to the higher contemporary revision rate seen in this shoulder replacement study. It is not clear whether patient selection and technical aspects of the primary procedure are also contributing to this trend. Whilst the findings in this shoulder dataset are not as dramatic as those for hip replacements, it identifies an area requiring close future scrutiny.

A systematic review and meta-analysis drawing in estimates from all published cohort studies and registry sources is in progress (PROSPERO registration: CRD42019140221). This will formally appraise the quality of the available studies, allow direct comparison of different sources, and provide pooled estimates of long-term revision risk.

Chapter 6

What are the risks of serious adverse events following shoulder replacement surgery recorded in Hospital Episode Statistics for NHS England?

6.1 Introduction

The OMERACT group advised that adverse events, including death, should be reported as a core outcome in studies of shoulder disorders.[143] The systematic review in chapter 2 found that randomised controlled trials in this area are inadequately powered and not designed to identify serious adverse events. Chapter 1 identified an absence of consensus over what conditions should constitute a serious adverse event after shoulder replacement surgery. There has been a substantial focus across the NHS on the rates of potentially preventable venous thromboemboli following orthopaedic surgery. One previous study using HES data by Jameson et al estimated the risks of deep vein thrombosis, pulmonary embolism and death occurring within 90 days of shoulder surgery, including shoulder replacement.[172] That study did not consider any other adverse events and there are no other large studies representative of NHS shoulder surgery practice.

Elements of this chapter have been published in the BMJ and as a conference abstract from the BESS Annual Congress.[302,303]

Too narrow a focus on thromboembolic complications may overlook other important causes of morbidity and service use, such as cardiorespiratory problems, renal problems and common infections. Partridge et al studied medical complications and mortality rates in patients undergoing primary total hip replacement surgery.[327] They found that whilst mortality has fallen, the rates of lower respiratory tract infections and acute kidney injuries have risen over a 10-year period. These complications carried an increased risk of 90-day inpatient mortality. Our understanding of these events in patients undergoing shoulder replacement surgery is more limited. There are some data from large North American studies but their population demographics, health service structures, and methods of recording events may limit the transferability to UK practice. Shoulder replacement surgery is being more frequently performed in older patient groups (chapter 4: figures 4.3 and 4.4). An improved understanding of the risks of medical complications following these procedures is important to ensure that surgeons and hospitals can provide high-quality, effective, and safe care.

The main objective of this chapter was to provide estimates of the absolute rates of common serious systemic adverse events following shoulder replacement surgery in routine NHS care. Secondary objectives were to measure the association of these adverse events with age, sex and comorbidities, to examine how risks vary with time following surgery, and to examine temporal trends in the risks.

6.2 Methods

6.2.1 Data source and participants

Data was obtained from the HES APC database for NHS England for the financial years ending 1998 through to 2017. Patients undergoing primary shoulder replacements performed for elective or acute traumatic indications were included. Inclusion criteria and data cleaning processes are described fully in chapter 3. Patients aged less than 18 years were excluded, along with procedures performed for primary or metastatic tumours. In addition, patients with less than 90 days of possible follow-up were excluded (i.e. those whose procedure was performed

after 31st December 2016). Cause-specific mortality was obtained from the linked ONS record.

6.2.2 Adverse event outcome definitions

Adverse events identifiable within HES APC data are inherently severe enough to require inpatient treatment. This aligns well with the usual NIHR definition that requires a serious adverse event to result in prolonged admission, readmission, or a threat to life.[165] The following conditions were identified within episodes of care following the index shoulder replacement surgery:

- Death
- Pulmonary embolism
- Myocardial infarction
- Lower respiratory tract infection
- Acute kidney injury
- Urinary tract infection
- Cerebrovascular event
- Any one of the above events +- death

The list of ICD-10 codes used to identify relevant conditions is provided in appendix A.4. All relevant events occurring within the index episode or subsequent episodes were considered to be incident events. No specific washout period was applied. Procedure-specific complications severe enough to require further surgery have been considered within the analyses of revision and reoperation in chapter 5 and are not included this analysis of serious adverse events, which focuses on systemic complications. The primary cause of death was extracted from the registered ICD-9 or ICD-10 codes.

6.2.3 Data processing

Data processing steps were completed as described in chapter 3 to establish a clean dataset of primary procedures. For each primary procedure in the dataset, other episodes were identified as either past medical history or follow-up in relation to the index episode. Exact dates are recorded in HES for surgical procedures but not for diagnosis codes. The occurrence of each adverse event was grouped into divisions of 30 days following surgery, based on the admission date recorded for the relevant episode.

6.2.4 Statistical analysis

The incidence of total and individual types of serious adverse event were calculated as simple rates with 95% confidence intervals, assuming a normal approximation to the Poisson distribution for count data. Rates were stratified by age, sex, and elective vs acute trauma indications. Initially it was planned to group age into five-year bands, but there were too few patients in these age bands to reliably estimate the rates of low frequency events, therefore the following broader age groupings were chosen: 18 to 49, 50 to 64, 65 to 74, 75 to 84, and 85 years or older.

Rates are primarily reported at 30 and 90 days following surgery. Cumulative events (total and by type) were also counted in 30-day intervals, up to 360 days post surgery, to evaluate the temporal association between complications and the index surgical procedure. Incidence rate ratios were calculated to compare the relative risk of death observed at 30 days in the study cohort to the expected number of deaths in an age- and sex-matched cohort, calculated from publicly available ONS data for the 2013 to 2017 financial years.[328] The probability of death over time following surgery was also calculated using Kaplan-Meier survival analysis, based on the exact dates of death available from linked ONS data.

The relative effects of age, sex and comorbidities were estimated using multiple logistic regression. Comorbidities were measured using the Charlson Index as described in chapter 3 and grouped into scores of 0, 1 or 2+. Age was modelled as a continuous variable using

restricted cubic splines with three knots to account for non-linearity. Knots were placed in the default positions at the 10th, 50th and 90th centiles (56, 73 and 83 years). Effect estimates for each factor are reported as odds ratios. The coefficients obtained from restricted cubic spline models are not intuitive, therefore odds ratios for age are reported for the contrast between the 1st and 3rd quartiles. Predicted probabilities from the model are summarised as a figure. A sensitivity analysis was performed using data recorded just from the 2010 financial year onwards to evaluate the effect of adding type of surgery as an interaction term with the acute trauma/elective variable. A second logistic regression model was used to test the association between each type of adverse event and the risk of death within 90 days, including adjustment for age, sex and grouped Charlson Index.

6.3 Results

6.3.1 Demographics

The overall population for this study was the same as that used in chapters 4 and 5. The patient characteristics have been described in table 5.1 and are not repeated here. 2523 patients with less than 90 days of possible follow-up were excluded from this analyses of serious adverse events, leaving 59 605 patients in the elective group and 14 368 in the acute trauma group. Frequency density plots for the age and comorbidity profiles of acute trauma and elective patients are available in appendix A.7. These demonstrate little difference between the two cohorts.

6.3.2 Mortality after shoulder replacement surgery

For patients undergoing an elective shoulder replacement, the mortality rate at 30 days was 0.34% (95% CI 0.29 to 0.39) and at 90 days was 0.66% (95% CI 0.60 to 0.73). The corresponding rates following surgery for acute trauma were three-times higher, at 1.16% (95% CI 1.00 to 1.35) and 2.28% (2.04% to 2.54%). Figure 6.1 demonstrates that the acute trauma cohort has a consistently higher mortality rate than the elective cohort, and 47% of the acute

trauma patients had died within 10 years of surgery. Figure 6.2 shows a stark difference in the patterns of all-cause mortality over the first year following surgery.

Compared to age- and sex-matched estimates from ONS data, the relative risk of death following elective surgery was 0.95 (95% CI 0.66 to 1.37) at 30 days and 0.64 (95% CI 0.50 to 0.84) at 90 days. The corresponding relative risks in the acute trauma cohort were 2.97 (95% CI 1.63 to 5.81) at 30 days and 2.20 (95% CI 1.52 to 3.24) at 90 days. The commonest causes of death were chronic ischaemic heart disease, acute myocardial infarction, trauma (including accidents and falls), and pneumonia, accounting for a total of 23% of all deaths.

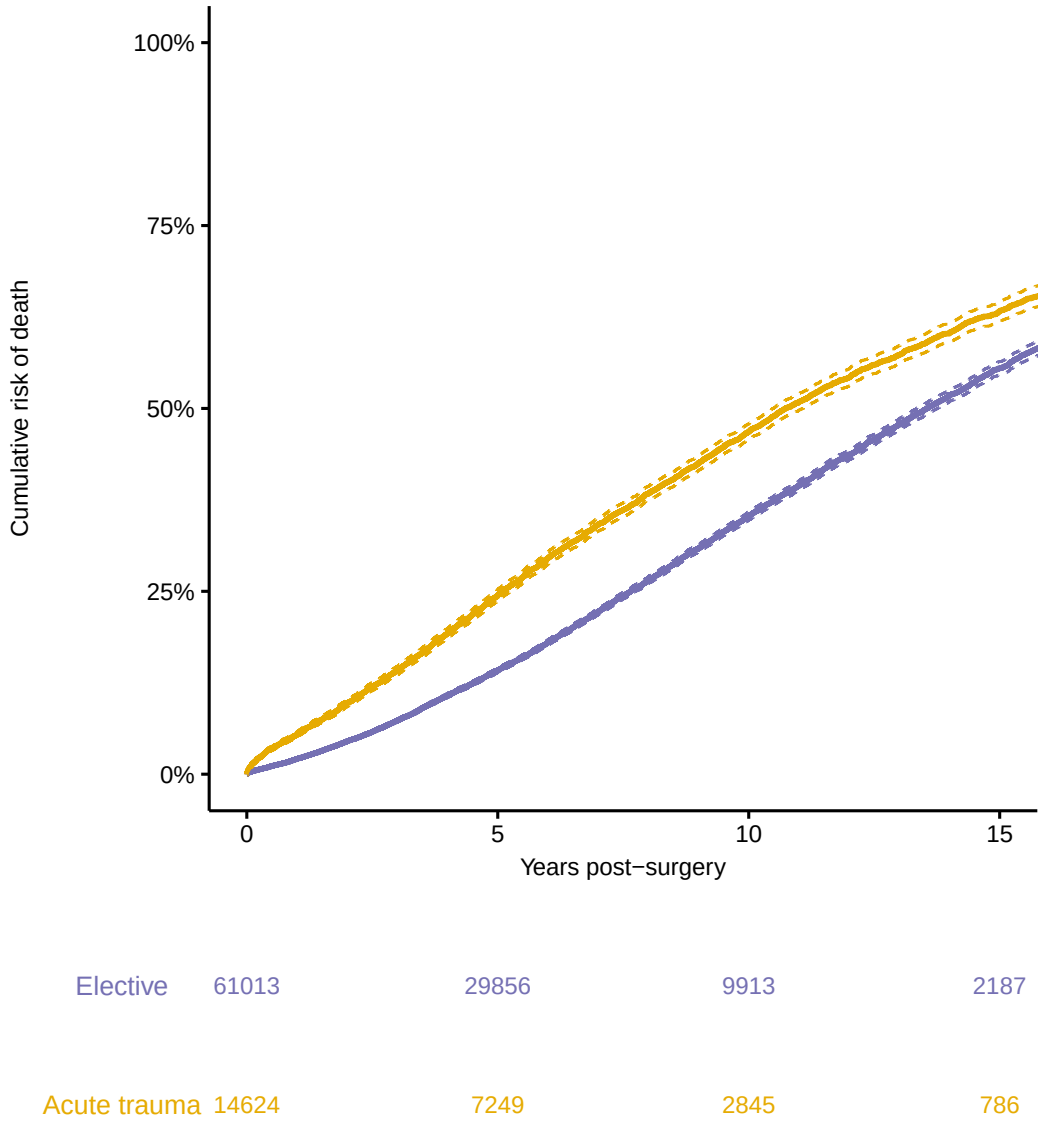


Figure 6.1: Cumulative risk of death up to 15 years following primary shoulder replacement. Table shows numbers at risk. 95% CIs shown.

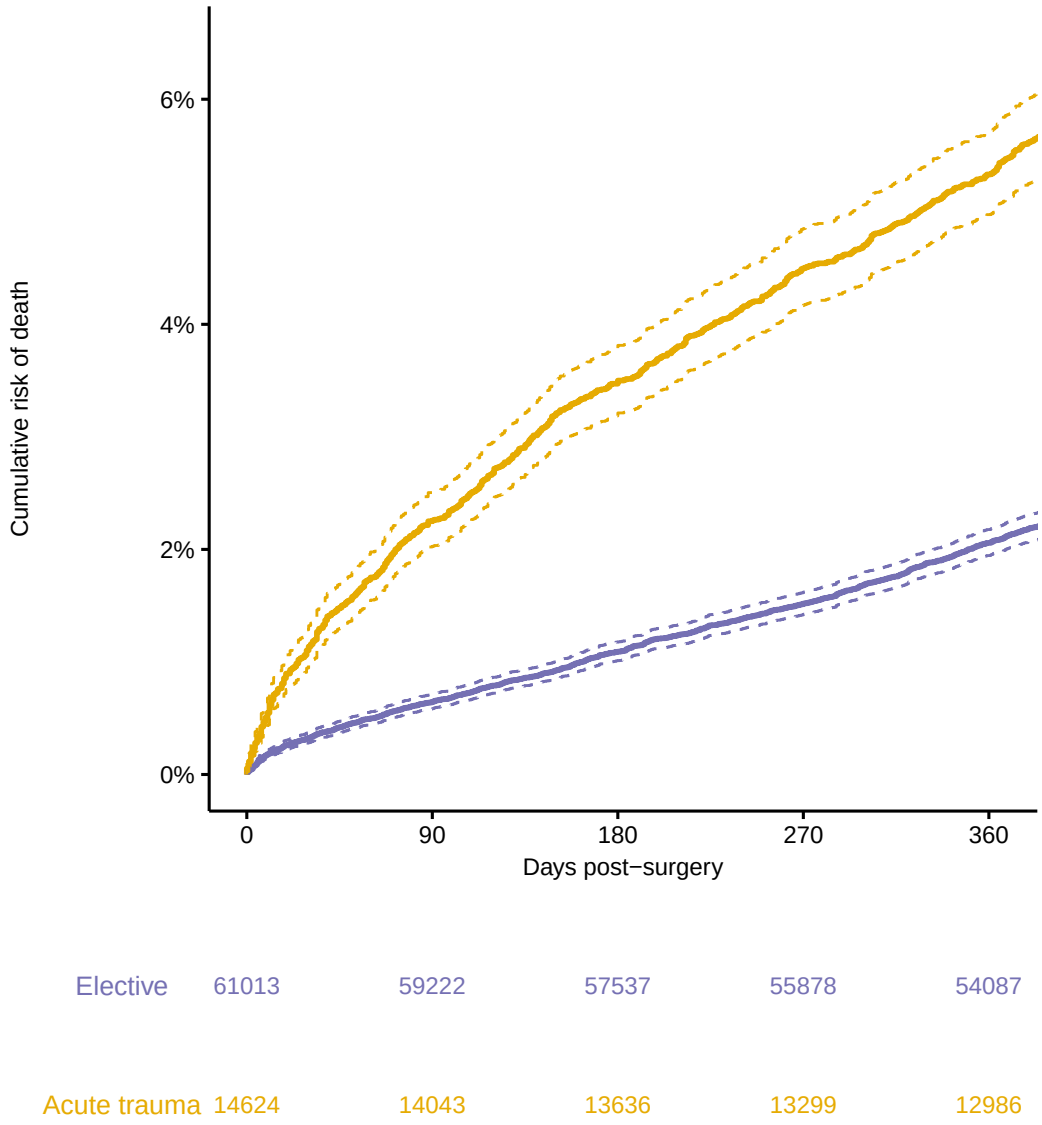


Figure 6.2: Cumulative risk of death up to one year following primary shoulder replacement. Table shows numbers at risk. 95% CIs shown.

All of the serious adverse events studied, except for urinary tract infection, were associated with increased risk of 90-day mortality. Odds ratios from the logistic regression model are shown in table 6.1. Increased age, higher Charlson Index and, to a lesser extent, male sex were also associated with an increased risk of 90-day mortality. At 90 days following surgery for acute trauma, 1 in 8 male patients aged 85 years or older had already died (n=239). This compares to no deaths occurring in women aged 18 to 49 years undergoing elective surgery (n=1391).

Table 6.1: Association of serious adverse events with 90-day mortality: reported as odds ratios from multiple logistic regression.

Variable	Odds Ratio (95% CI)
Pulmonary embolism	6.93 (4.22 to 11.38)
Myocardial infarction	7.75 (5.25 to 11.43)
Lower respiratory tract infection	7.08 (5.82 to 8.60)
Urinary tract infection	1.30 (0.98 to 1.73)
Acute kidney injury	4.30 (3.28 to 5.62)
Cerebrovascular event	10.60 (7.11 to 15.80)
Age (65 vs 79 years)	3.21 (2.84 to 3.62)*
Sex - Male:Female	1.42 (1.19 to 1.69)
Charlson index (1 vs 0)	1.78 (1.44 to 2.20)
Charlson index (2+ vs 0)	2.23 (1.83 to 2.71)

*Older=higher risk. The effect of age is presented as a contrast between the 1st and 3rd quartiles to simplify the underlying restricted cubic spline function

6.3.3 Total serious adverse events at 30 and 90 days following surgery

Total counts and rates at 30 and 90 days for death, pulmonary embolism, myocardial infarction, lower respiratory tract infection acute kidney injury, urinary tract infection, cerebrovascular event, and any serious adverse event +- death are detailed in table 6.2. As was observed in the analysis of mortality rates, serious adverse events were more commonly observed following acute trauma. At 30 days, the risk of any serious adverse event including death was 1 in 30 following elective surgery and 1 in 11 following surgery for acute trauma.

The commonest adverse events were lower respiratory tract infections, urinary tract infections and acute kidney injury. Whilst rates of pulmonary embolism and myocardial infarction were comparatively lower, they were not negligible. The incidence of pulmonary embolism in the acute trauma cohort was 1 in 200 within 90 days. Full age- and sex-stratified estimates for each type of adverse event, including death, are provided in appendix A.7 (tables A.12 and A.13).

Table 6.2: Rates of serious adverse events within 30 and 90 days following surgery

Time	Elective n	Incidence	Trauma n	Incidence
Number at risk	59605		14368	
All cause death				
30 days	201	0.34 (0.29 to 0.39)	167	1.16 (1.00 to 1.35)
90 days	392	0.66 (0.60 to 0.73)	327	2.28 (2.04 to 2.54)
Pulmonary embolism				
30 days	106	0.18 (0.15 to 0.22)	64	0.45 (0.35 to 0.57)
90 days	156	0.26 (0.22 to 0.31)	73	0.51 (0.40 to 0.64)
Myocardial infarction				
30 days	121	0.20 (0.17 to 0.24)	78	0.54 (0.43 to 0.68)
90 days	163	0.27 (0.23 to 0.32)	93	0.65 (0.53 to 0.79)
Lower respiratory tract infection				
30 days	861	1.44 (1.35 to 1.54)	562	3.91 (3.60 to 4.25)
90 days	1110	1.86 (1.76 to 1.98)	669	4.66 (4.32 to 5.02)
Acute Kidney Injury				
30 days	442	0.74 (0.68 to 0.81)	246	1.71 (1.51 to 1.94)
90 days	531	0.89 (0.82 to 0.97)	285	1.98 (1.77 to 2.23)
Urinary Tract Infection				
30 days	582	0.98 (0.90 to 1.06)	508	3.54 (3.24 to 3.86)
90 days	803	1.35 (1.26 to 1.44)	611	4.25 (3.93 to 4.60)
Cerebrovascular event				
30 days	82	0.14 (0.11 to 0.17)	67	0.47 (0.37 to 0.59)
90 days	138	0.23 (0.20 to 0.27)	97	0.68 (0.55 to 0.82)
Any serious adverse event				
30 days	1921	3.22 (3.08 to 3.37)	1267	8.82 (8.35 to 9.32)
90 days	2464	4.13 (3.97 to 4.30)	1474	10.26 (9.75 to 10.80)
Any serious adverse event or death				
30 days	2008	3.37 (3.22 to 3.52)	1350	9.40 (8.91 to 9.91)
90 days	2652	4.45 (4.28 to 4.62)	1628	11.33 (10.79 to 11.89)

Rates reported as percentage with 95% CIs

6.3.4 Timing of serious adverse events over first year following surgery

Patients included in this analysis required a minimum of 360 days possible follow-up and so an additional 5004 patients were excluded, leaving 55244 in the elective cohort and 13725 in the acute trauma cohort for this analysis. The cumulative incidences of different adverse events were calculated for each 30-day period following surgery up to 360 days. Figures 6.3 and 6.4 show the rates of additional events for each of these 30-day periods (note the different y-axis scales). The period of highest risk is clearly in the first 30 days following surgery. By 90 days the incidence of new events falls to a baseline level. For patients experiencing an event within the first 360 days, 50% and 66% of the cumulative incidence was reached within the first 90 days following elective and acute trauma surgery respectively.

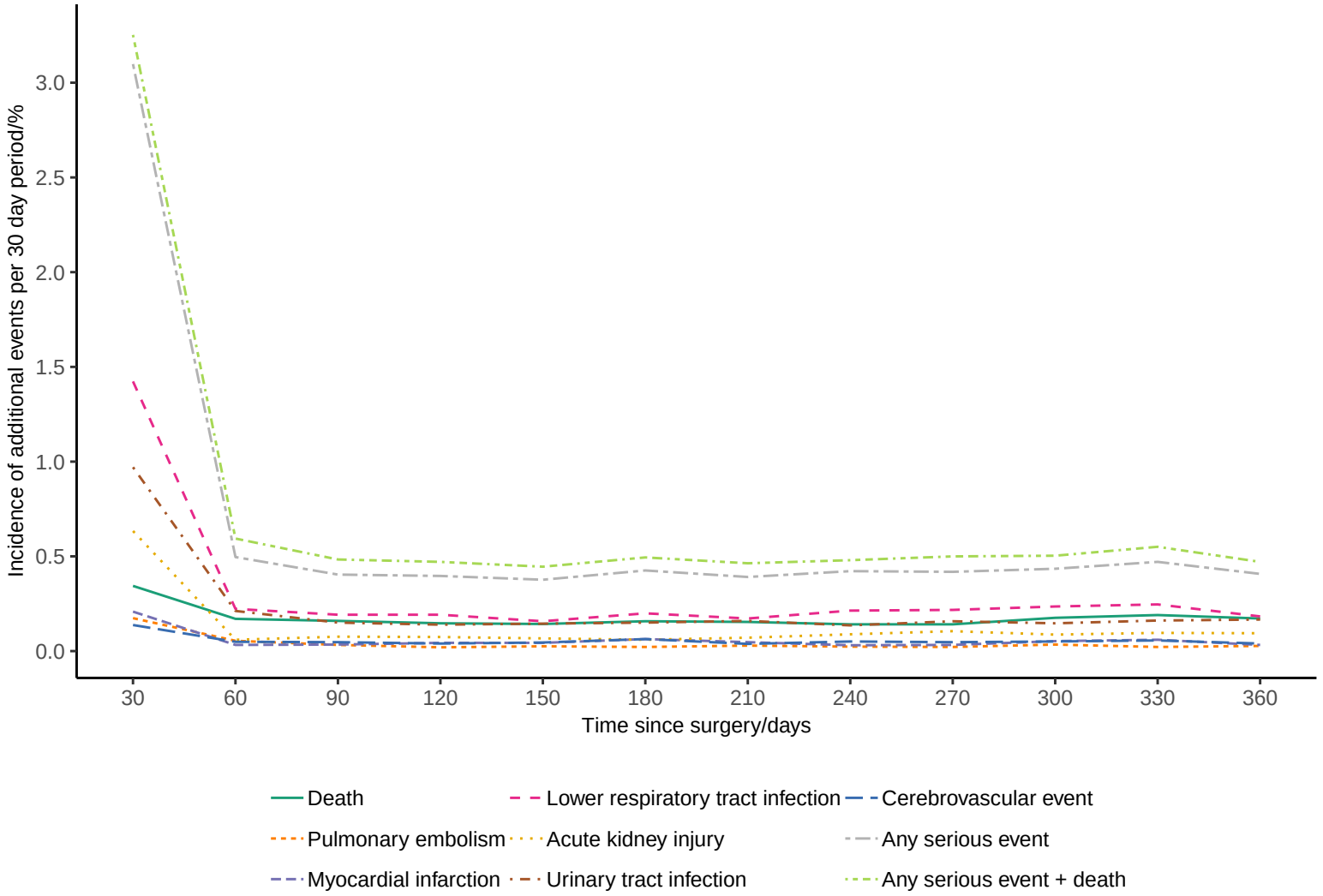


Figure 6.3: Rates of serious adverse events within one year following elective shoulder replacement surgery

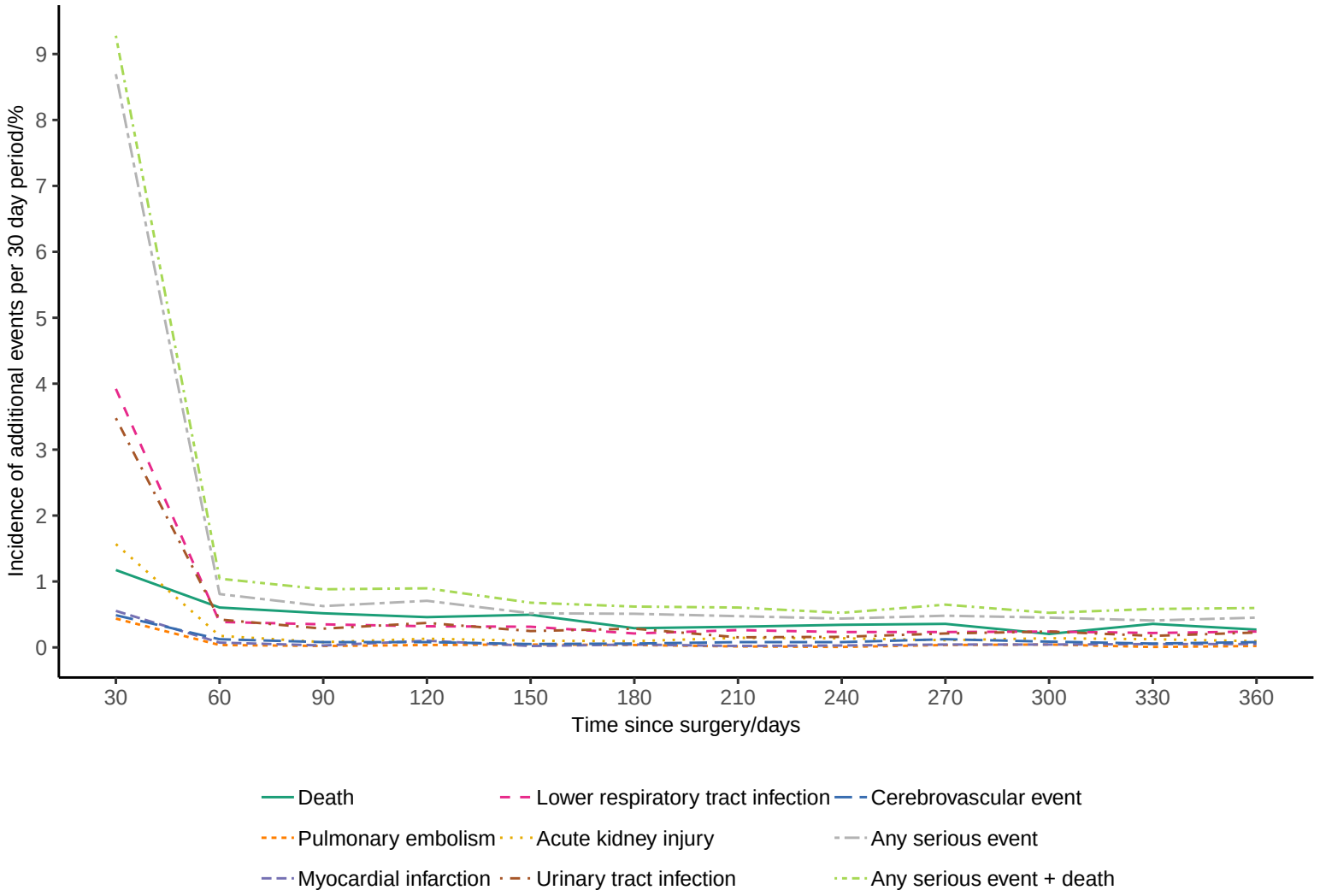


Figure 6.4: Rates of serious adverse events within one year following shoulder replacement surgery for acute trauma

6.3.5 Risk of serious adverse events within 30 days adjusted for age, sex and comorbidities

The predicted probability of experiencing any serious adverse event including death within 30 days is shown in figure 6.5, adjusted for age, sex, comorbidities, and elective vs acute trauma indications. This confirms that the likelihood of an event is much higher in the acute trauma cohort (OR 3.02, 95% CI 2.81 to 3.26), higher in men (OR 1.24, 95% CI 1.14 to 1.35), higher in patients with a Charlson Index $\geq$2 (OR 3.64, 95% CI 3.34 to 3.97, reference level = 0), and higher with increasing age (OR 2.35, 95% CI 2.23 to 2.47 - contrast between 65 years and 79 years). Increasing age had important effects beyond 85 years. This is best appreciated in figure 6.5. In men aged 85 years or older, the likelihood of experiencing any event including death within 30 days of elective surgery was 17% (95% CI 14 to 20) and after acute trauma surgery was 21% (95% CI 16 to 28). The model equation is provided in appendix A.7.

In the sensitivity analysis, the effects of the main set of independent variables were not meaningfully changed by the addition of the recorded operation type in the regression model. Different types of shoulder replacement did not result in any important differences in the likelihood of a serious adverse event at 30 days. The odds ratios from the updated model were as follows: elective TSR vs hemiarthroplasty 0.85 (95% CI 0.74 to 0.98), elective RTSR vs hemiarthroplasty 1.08 (0.95 to 1.23), and acute trauma RTSR vs hemiarthroplasty 0.96 (0.81 to 1.15).

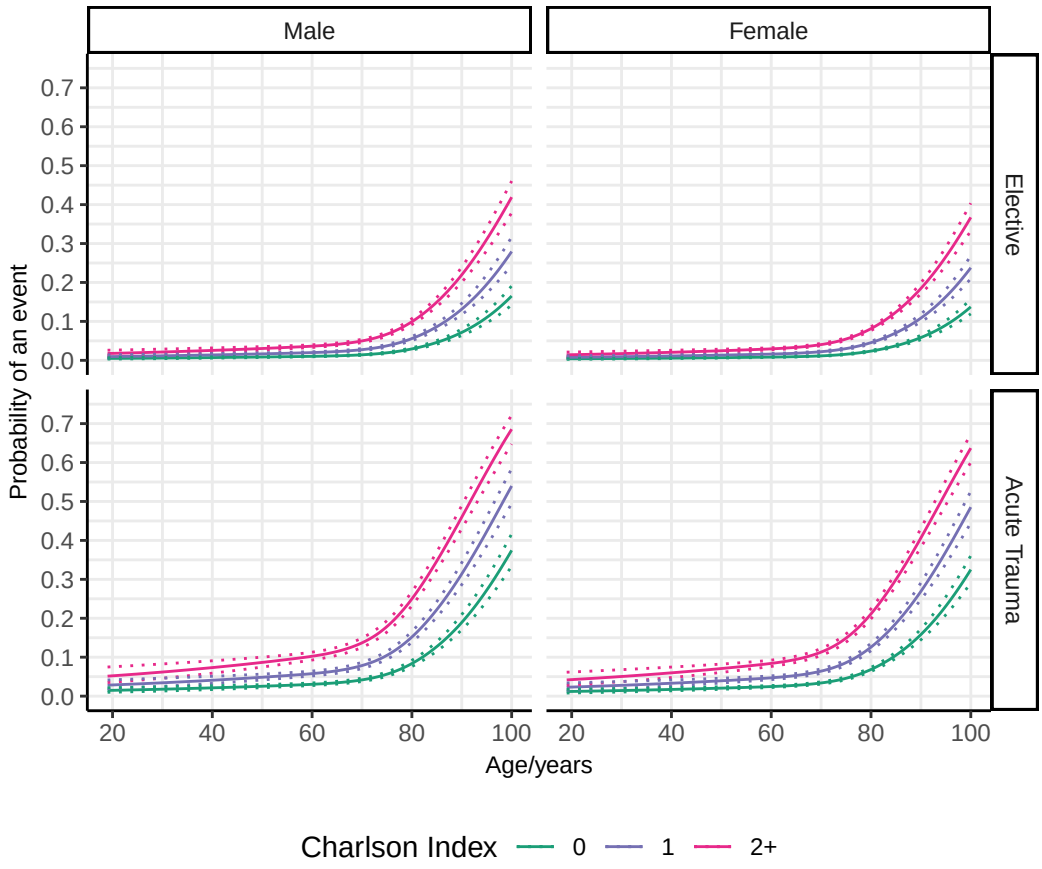


Figure 6.5: Predicted probability of any adverse event. Multiple logistic regression adjusted for age, sex and Charlson Index (95% CIs shown)

6.3.6 Trends in serious adverse events and patient demographics over time

The 30-day mortality rate following elective surgery nearly halved from 0.44% (95% CI 0.30% to 0.66%) in 1998, to 0.25% (95% CI 0.21% to 0.29%) in 2017. There was a noticeable improvement between 2007 and 2008 (figure 6.6). Mortality rates for acute trauma also seemed slightly lower at the end of the study period but the yearly sample sizes are small and this observation cannot be confirmed. The risk of any serious adverse event increased from 2.21% (95% CI 2.05 to 2.39) to 4.86% (95% CI 4.82 to 4.90) for elective surgery, and from 3.59% (95% CI 3.20 to 4.02) to 12.80% (95% CI 12.51 to 13.09) for acute trauma surgery (figure 6.7).

Over the same time period, the mean age increased by 5.1 years for patients undergoing elective surgery and by 1.6 years for those with acute trauma (figure 6.8). The mean Charlson Index increased annually, approximating to a linear relationship over time for both elective and acute

trauma patients (figure 6.9). In 2017, the mean scores were 1.40 (95% CI 1.26 to 1.54) and 1.42 (95% CI 1.36 to 1.47) respectively.

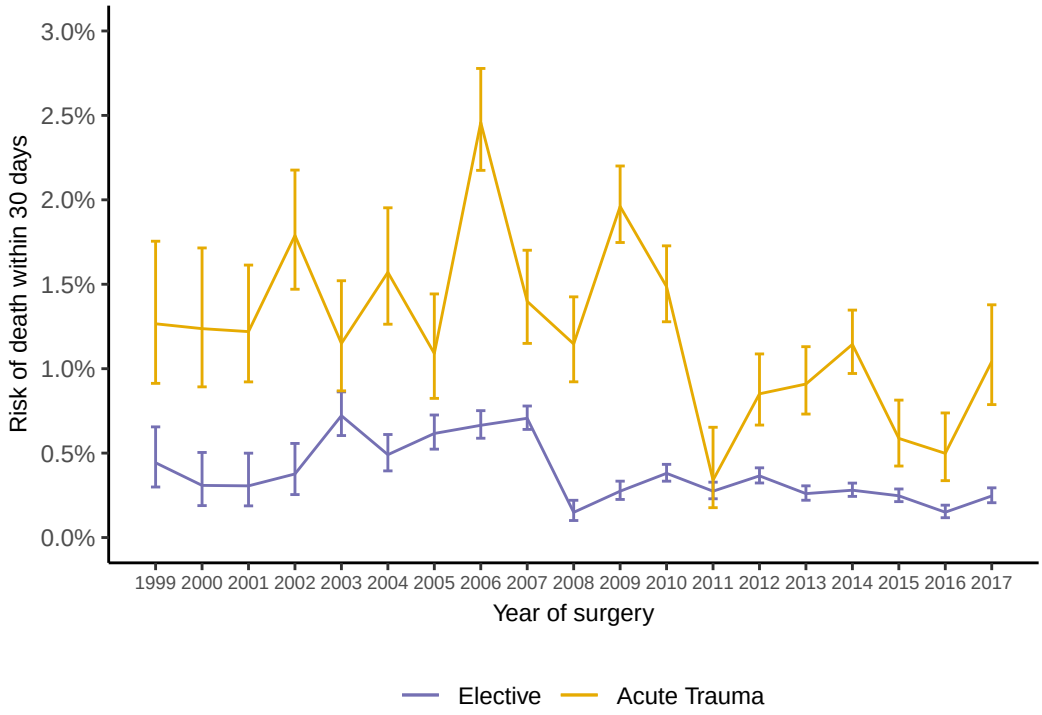


Figure 6.6: Temporal trends in 30-day mortality (95% CIs shown)

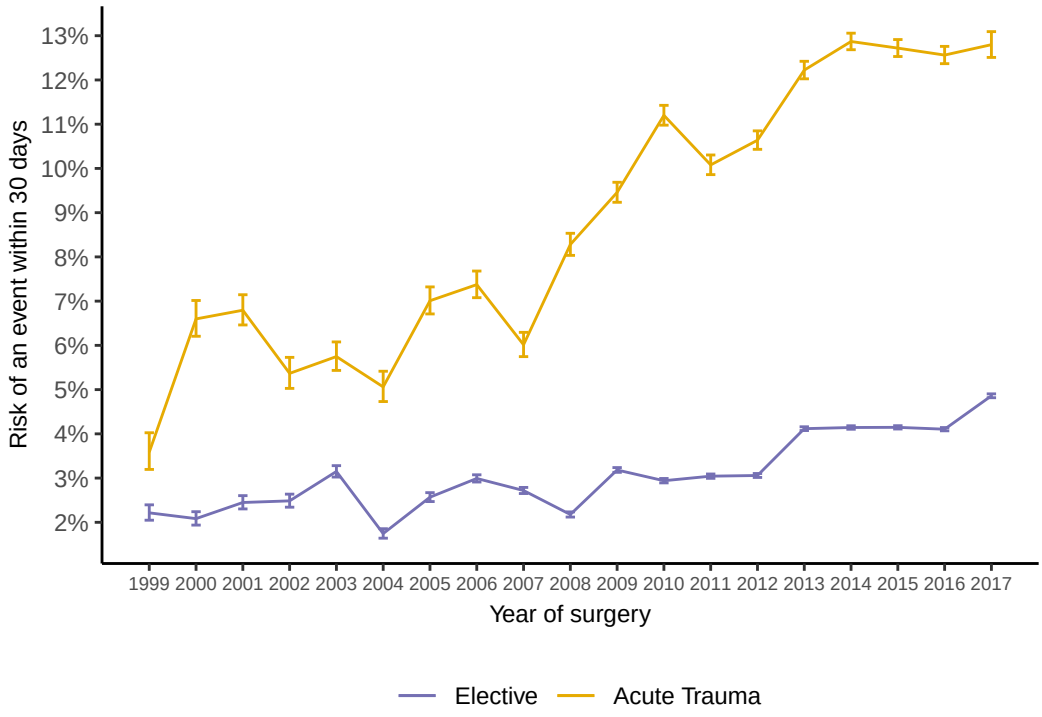


Figure 7: Temporal trends in serious adverse events within 30 days (95% CIs shown)

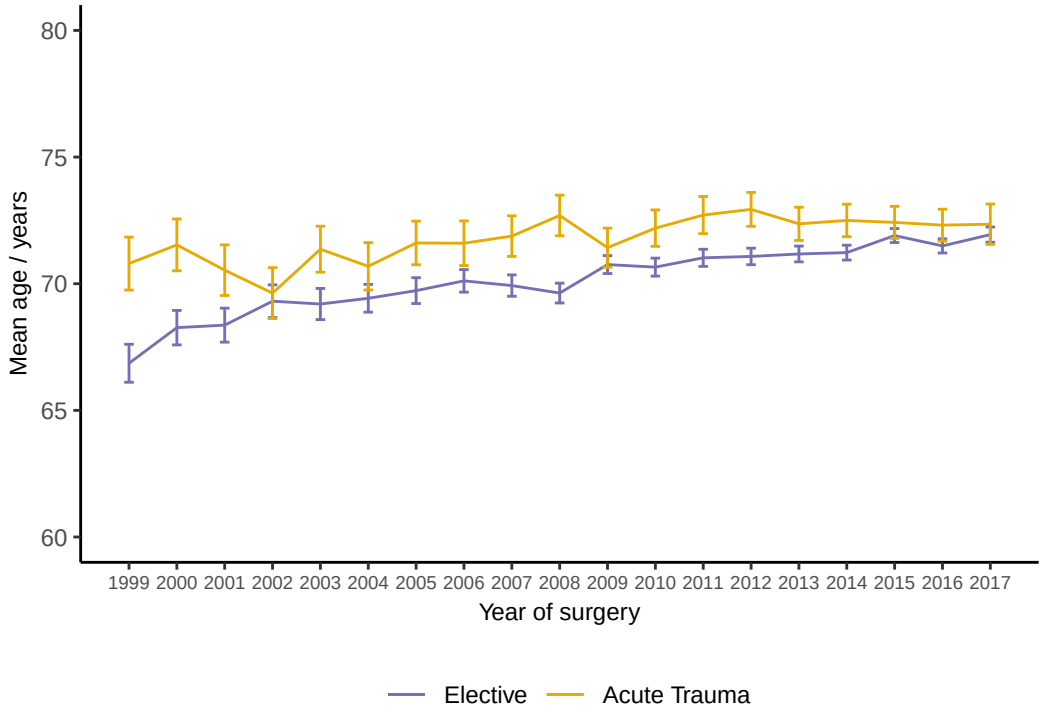


Figure 6.8: Temporal trends in the mean age of patients undergoing shoulder replacement surgery (95% CIs shown)

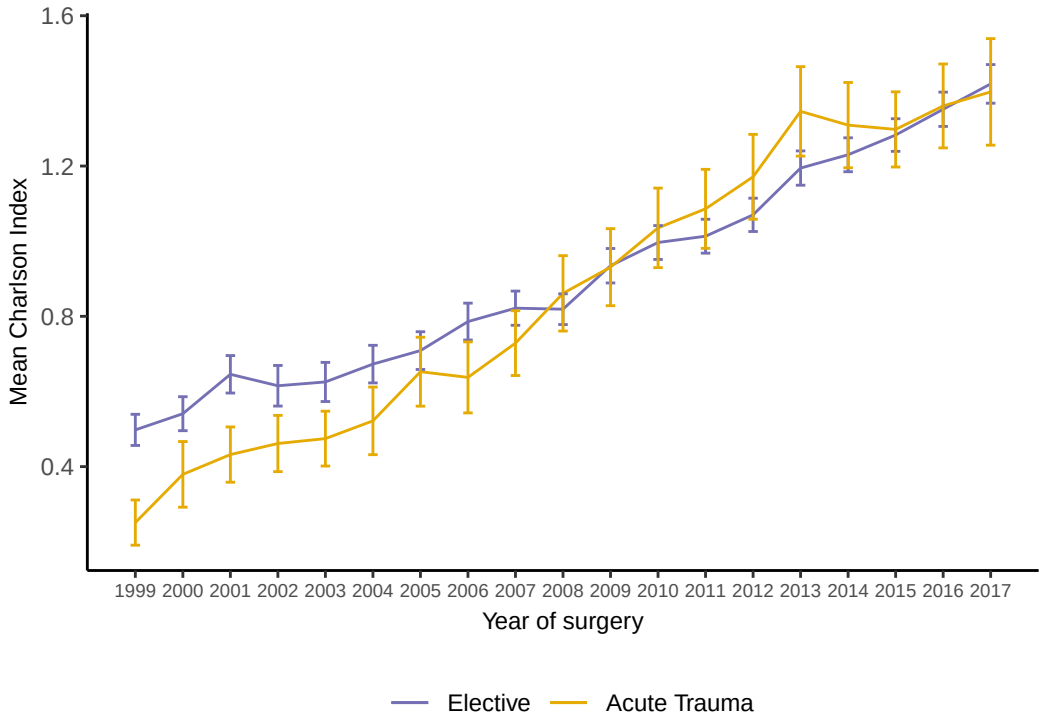


Figure 6.9: Temporal trends in the mean Charlson Comorbidity Index of patients undergoing shoulder replacement surgery (95% CIs shown)

6.4 Discussion

6.4.1 Key findings

The rates of serious systemic adverse events were much higher than expected, and higher than previously estimated. At 30 days, the unadjusted risk of any serious systemic adverse event was 1 in 30 following elective surgery and 1 in 11 following surgery for acute trauma. The commonest adverse events were lower respiratory tract infections, urinary tract infections and acute kidney injuries. A clear temporal association between surgery and the occurrence of a serious adverse event or death was observed.

The risk of death following elective shoulder replacement is at, or lower than, the expected rate for an age- and sex-matched population estimate. For acute trauma, mortality rates were considerably higher than those for elective surgery, and two to three times higher than matched population estimates. Whilst early postoperative mortality has fallen between 1998 and 2017 for elective surgery, the presence of improvement in the acute trauma setting is less certain. This comes against a backdrop of increasing rates of serious adverse events in a patient group that is older and has a greater number of reported comorbidities in both settings.

With the exception of urinary tract infections, each of the recorded adverse events were in themselves strongly associated with subsequent risk of death within 90 days. Logistic regression analysis found that acute trauma, comorbidities, increasing age and male sex were all associated with higher risks of experiencing an adverse event. In men aged 85 years or more, the risk of any adverse event may be as high as 1 in 5, including up to a 1 in 8 risk of death within 90 days of surgery for acute trauma.

6.4.2 Strengths and limitations

This study has provided a detailed analysis of the rates of common serious adverse events and death for patients undergoing shoulder replacement surgery in the NHS. The large scale of the dataset has allowed for estimates to be made with reasonable precision and to mitigate for some

of the variation reported by studies from smaller samples. Analyses are sufficiently powered to evaluate the effects of age, sex, comorbidity, and the setting of surgery. Analysis of a single secondary care system, with universal coverage, provides results that should be generalisable and representative of expected NHS outcomes.

The definition of serious adverse events for this study was pragmatic, based on a requirement for inpatient hospital-level care. The relative severity of medical conditions is not objectively recorded in HES data. However, the need for additional or prolonged inpatient hospital care is conceptually straightforward to explain to patients. Urinary tract infections were not associated with subsequent mortality and they may be of lesser importance than the other conditions studied. Urinary tract infections still satisfied the predetermined criteria of inpatient treatment, and they may be associated with higher rates of periprosthetic joint infection.[329] The ICD-10 coding system does provide a code for hospitals to indicate potential causality between subsequent events and a previous surgical procedure (Y83). This has been shown to be poorly implemented in the HES database and was not used in this study.[330] More precise timings for adverse events could increase the strength of causal inference between surgery and the events, but diagnostic codes are only provided in HES for an episode as a whole, rather than a specified date. Hence, it was necessary to dichotomise the time scale.

The estimated odds ratios for the association between adverse events and subsequent mortality may represent an underestimate of the true strength of association. These estimates suffer from an immortal time bias since the precise event time is not known.[331] Hence, some patients may not have sufficient follow-up from the time of their adverse event to develop the endpoint (death).

The accuracy of coding for the conditions of interest in this study is generally high. Wright et al compared vascular disorders recorded in HES data to GP records.[332] They found agreement between the records for pulmonary embolism of 91.2% to 92.6%, for stroke 72.2% to 94.1%, and for myocardial infarction 89.2% to 97.7%. In their 2014/5 national audit of community acquired pneumonia, the British Thoracic Society found that cases of pneumonia identified by discharge codes met their established diagnostic criteria in 64% of cases.[333] However, most of these cases only missed the criteria due to the absence of required chest x-ray signs and 92% had

appropriate signs, symptoms, and treatment of a lower respiratory tract infection. The accuracy of acute kidney injury recorded in HES data compared to national criteria is approximately 95%.[334] Deep vein thromboses were intentionally not included in the study since patients are frequently treated in emergency departments or ambulatory units. They are therefore not reliably captured in HES inpatient data.[335]

This study has reported absolute risks for adverse events following surgery rather than relative risks compared to a matched cohort. The absence of adequate data on adverse events from RCTs has already been demonstrated in chapter 2, and there are no natural cohorts from which to draw an matched comparator group that is substantially free from unmeasured confounding. Kuan et al have recently published a comprehensive analysis of 308 conditions identified using linked Clinical Practice Research Datalink (CPRD) and HES data.[336] The conditions reported in this thesis study are all included in their work but their estimates are presented as the period prevalence and cumulative incidence over five-year periods, which cannot be directly compared.

The absence of a washout period may lead to inflated estimates for the frequency of adverse events since some comorbid conditions may be have been captured. One possible method to account for this and to estimate relative risks associated with shoulder replacement surgery would be to use a self-controlled case series (SCCS) model, where a period of presurgical history for each patient is used as a within-patient control.[337] This method has not yet been widely used in surgical studies. Patients undergoing planned surgery may be fitter than the general population, the so-called "healthy user bias". The comparisons made here to ONS mortality data have a useful face-value, but they may not be an accurate reflection of the true risks in this cohort.

6.4.3 Comparison to other studies

Between 2012 and 2018, the NJR reported mortality rates after elective surgery to be 0.11% (95% CI 0.08 to 0.15) and after surgery for acute trauma to be 0.67% (95% CI 0.45 to 1.01).[161] In the other HES study previously referred to, the mortality rates at 90 days were 0.37% for elective surgery and 3.76% following hemiarthroplasty for trauma.[172] The Jameson et al study

did not have access to reliable ONS mortality data, relying only on inpatient mortality recording that is known to be less accurate. Their data collection period was 2005 to 2008, a time that corresponds to noticeable fluctuation in mortality rates observed in the long-term trends of the current study (figure 6.6). Jameson et al also observed much lower rates of pulmonary embolism in the elective setting (0.11%), which may reflect a sampling error. The large size and ability of this current study to account for year-to-year fluctuations is an important strength.

The study by Partridge et al of outcomes following hip replacement provides useful parallels since they employed a similar methodology in a HES data sample.[327] They found similar temporal trends for mortality, complications and comorbidities, although the mean age was very slightly lower in 2014 compared to 2005. The most striking difference was that the rate of lower respiratory tract infection was only 0.84% in 2014, less than half the rate found here following elective shoulder replacement surgery. On careful inspection of the ICD-10 codes, Partridge et al failed to include all of the most common codes for pneumonia captured by British Thoracic Society Audits (J12-J17) and so their estimate is likely to be somewhat lower than the true value.[333]

A review of the North American literature identified five studies with considerable overlap between cohorts using the Nationwide Inpatient Sample (NIS).[166–168,338,339] The NIS includes a stratified sample of 20% of US hospital discharges and is a similar administrative database to HES APC, but no identifiers are held in the database therefore no linkage to identify events beyond the index hospital stay can be performed.[340] Therefore direct comparisons to the 30- and 90-day rates in this study are not particularly useful. Six studies using the American College of Surgeons National Surgical Quality Improvement Program (ACSNSQIP) database reported complications and reasons for readmission within 30 days of surgery.[171,341–345] There is again considerable overlap between the cohorts and results are reported in some studies as absolute rates and in others simply as regression results for risk factor analyses. All of these studies seem to pool the outcomes of surgery for elective and trauma procedures. Their results are summarised in table A.14 (appendix A.7). The reported event frequencies from Waterman et al for a 2006 to 2011 cohort were: death 0.25%, pulmonary embolism 0.25%, myocardial infarction 0.20%, stroke 0.20%, AKI 0% (no events), urinary tract infection 1.50% and lower

respiratory tract infection 0.50%.[342] Shields et al found that the risks of any major complication were higher for emergency versus planned procedures (odds ratio 3.7), and that there was no difference in risks between hemiarthroplasty and TSR/RTSR.[341] Jiang et al found no association between obesity and the likelihood of experiencing an adverse event, however their sample size (n=4267) was probably too small to be sure that no difference exists.[343] Much higher reported complication rates were reported by Cancienne et al using the PearlDiver Humana Insurance database.[170] However, the level of comorbidity in their study was extraordinarily high (mean Charlson Index = 4.9), and over 40% of their cohort was diabetic and/or obese (BMI$\geq$30 kg/m^2).

6.4.4 Meaning of the study

This study reinforces that shoulder replacement surgery is a major procedure carrying considerable risks. It is increasingly performed in older patients with more comorbid conditions. These patients have a high risk of needing additional hospital care for serious adverse events, and the overall frequency of adverse events has increased considerably. Event frequencies are low enough that some surgeons may not have appreciated these risks first-hand. However, there are important implications for both individual patients considering surgery and at a wider health service level. As part of delivering safe, effective and efficient care, hospitals and surgeons should aim to achieve low rates of adverse events. Measures targeted at reducing respiratory tract infections, urinary tract infections and acute kidney injury should perhaps be considered first since due to their relatively high incidence. The National Institute for Health and Care Excellence (NICE) guidelines suggest that thromboprophylaxis should be considered for patients undergoing upper limb surgery who will have poor mobility after surgery or where the duration of general anaesthesia is over 90 minutes.[346] These NICE guidelines do not specify the type or duration of thromboprophylaxis. Current BESS venous thromboembolism guidelines do not proscribe the use of extended pharmacological prophylaxis after elective shoulder replacement surgery but do suggest that it is considered following acute trauma.[347] Both of these recommendations remain reasonable in light of the risk levels identified by this study.

6.4.5 Unanswered questions and future research

This study identified a specified list of complications that can be identified within HES data with reasonable confidence. The list is not exhaustive and other important complications not routinely requiring hospital admission are overlooked, such as nerve damage and deep vein thrombosis. These conditions could perhaps be studied using linked primary care data.

All types of adverse events were much more frequent in patients undergoing surgery for acute trauma. This was not explained by differences in the overall comorbidity indices and demographic profiles. A more detailed investigation of specific comorbidities may identify specific risk factors, however these patients may have markers of frailty (e.g. falls/fractures, under-nutrition, poor mobility) that are not well represented by traditional comorbidity indices. The Dr Foster group have shown that it is feasible to calculate frailty scores using secondary care data.[348,349] Further research should also examine the outcomes of patients with proximal humeral fractures who do not undergo surgery, to understand to what extent adverse events are related to pain, disability and associated injuries, rather than the surgical intervention.

The factors driving high rates of lower respiratory tract infections, urinary tract infections and acute kidney injuries in this cohort need to be better understood. HES data does not allow for the analysis of many important treatment-specific factors such as anaesthesia type, urinary catheterisation, fluid balance status and prescribing of nephrotoxic drugs. Whilst it is well-known that general anaesthesia is a risk factor for atelectasis and lower respiratory tract infections, it is plausible that they maybe an association with increased use of regional anaesthesia techniques, since interscalene blocks are known to induce temporary diaphragmatic paralysis.[350] Although a systematic review of 15 RCTs comparing single-injection interscalene blocks to continuous infusions recommended the use of continuous blocks and found limited evidence of important respiratory function deficits on spirometry, the studies did not count clinically respiratory infections.[351]

National publication of surgeon- and unit-level outcomes data for lower limb arthroplasty and hip fractures, together with financial incentivisation may have contributed to improved outcomes[352] National outcomes reporting and tariff incentives are not yet in place for shoulder

replacement surgery. Adverse events should be included as a quality marker for future service improvement exercises.

Chapter 7

Development and validation of models to predict the outcomes following shoulder replacement surgery using National Joint Registry data

7.1 Introduction

Predicting which patients will do well with surgery for different shoulder conditions was the highest ranked research priority from the James Lind Alliance Priority Setting Partnership exercise run in 2015.[10] This comes against a backdrop of an increased desire from patients, clinicians and commissioners for precision and tailored treatment solutions in many other aspects of healthcare.

Prediction models can be employed for a number of different purposes, such as individualised preoperative decision making, risk stratification to target perioperative resources, and casemix adjustment for surgeon, unit and implant comparisons. The established NJR outlier analyses and unit/surgeon dashboards for hip and knee replacements are a good example where casemix adjustment uses regression models.[353]

The statistical models used for prediction are the same as those used more commonly in explanatory studies of association or causal inference. The primary purposes of prediction, explanatory, and descriptive models are quite different. In prediction, the ability of the model

to accurately predict outcomes in new data is the most important goal, with minimisation of the bias and variance of predictions. The magnitude and plausibility of individual covariate effects in the context of prediction is not important.[354]

There are no widely used prediction models for shoulder replacement surgery. The designs of existing prediction models for the outcomes of shoulder replacement surgery described previously in chapter 1.9 are now summarised here in table 7.1. These three models all demonstrate reasonably good performance measured by the AUC. However, the models are all based on small samples from single centres. Despite counter claims in one of the studies, there is likely to be considerable overfitting of these models and a lack of wider generalisability. The models from Matsen et al and from Chen et al both used baseline adjustment whilst predicting change scores.[177,180] This is not a widely recommended approach. Analysis of change scores does not in itself account for baseline imbalance because of regression to the mean, but inclusion of change score as the outcome variable with baseline adjustment is likely to induce bias in the effects of other predictors.[355–357]

Table 7.1: Summary of previously published prediction models

Study reference	Simmen 2008	Matsen 2016	Chen 2019
Setting	Single centre; Switzerland	Single centre; USA	Single centre; USA
Outcome predicted	Constant-Murley score $\geq$80 points	$\geq$30% improvement in SST	$\geq$MCID for PROMIS
Timepoint	1 year	2 years	3 months
Procedures	Promos^{TM} TSR	Mixed	TSR
Indications	Mixed	Mixed elective	Not stated
Variables included in final model	Age>75/ prior shoulder surgery/ SF-36 mental component/ DASH score	Age/ gender/ work related/ ASA/ baseline SST/ anxiety/ depression/ smoking/ prior shoulder surgery/ elevated head on AP radiograph/ Walch classification of glenoid	Age/ sex/ BMI/ ASA grade/ PROMIS (Physical Functioning/ Pain Interference/ Depression domain
Model type	Logistic	Logistic	Logistic
Variable selection	Backwards stepwise ($p<0.05$)	Backwards stepwise (AIC)	
N for development	100	337 (275 complete data)	62
N for validation	40	10 equal groups	No validation
Validation	Split-sample	Cross-validation	
Proportion of patients with outcome	31% development / 40% validation	81%	
AUC	0.79 (95% CI 0.70 to 0.88)	0.79	0.70 for PF/ 0.87 for PI
Measures for overfitting	No overfitting detected	No	No
Notes	Brier score planned but not reported		137 excluded for incomplete scores

A number of different outcomes are suitable for prediction and may provide useful information to patients considering surgery. The NJR-hosted Patient Decision Support Tool for Joint Replacement (JointCalc, https://jointcalc.shef.ac.uk) provides estimates for the outcomes of hip or knee replacement surgery.[358] It predicts the level of pain and function following surgery (on the Oxford Hip or Knee Score), the likelihood of revision surgery within 10 years, and the likelihood of death within a year of surgery. Whilst these outcomes are all useful, none addresses the question of how likely a patient is to get better, or indeed worse, with surgery.

Chapter 6 demonstrated that the risk of death within 90 days of shoulder replacement surgery was less than the age- and sex-matched mortality rates for the population at large. The likelihood of experiencing a serious adverse event within 90 days was predicted by the patient age, sex and comorbidities with modest accuracy. Death and adverse events are not considered further in this chapter.

The NJR has one of the largest datasets of pre- and postoperative PROMs and has released limited results suggesting that 8% of patients have a worse OSS at six months following surgery.[359] Prediction of patients' likely level of postoperative pain and function, and identification of those least likely to improve with surgery may help to better inform patient and surgeon expectations. In some cases, resources may be better used on alternative treatment strategies.

In this chapter, three prediction models are developed and internally validated with the objective of informing preoperative shared decision making for patients considering an elective shoulder replacement. Model 1 predicts the absolute Q2 OSS, model 2 predicts the risk of early revision surgery (within two years), and model 3 predicts the likelihood of a poor outcome (defined as a failure to achieve an improved OSS and/or requiring revision surgery with one year. The models are designed based on preoperative patient factors that could reasonably be known at the point of referral to secondary care.

7.2 Methods

The underlying design and validation principles for prediction modelling are laid out by the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.[360,361] These principles have been followed in the design, conduct and reporting of this chapter.

7.2.1 Data source and participants

All of the models described in this chapter were developed and validated using NJR data. The dataset was cleaned and prepared for analysis as described in chapter 3. Patients with a primary shoulder replacement recorded for any elective indication were included. For models 1 and 3, 7556 procedures were included only if they took place during the Q2 OSS pilot window between 2012 and 2015. Model 2 did not rely on postoperative scores for the analysis and therefore all 19640 primary procedures were eligible for inclusion. Sensitivity analyses for model 2 relied on only the 8792 patients with complete Q1 OSS data.

7.2.2 Definition of outcome variables

Model 1: Postoperative Oxford Shoulder Score (Q2 OSS) recorded between 4 and 12 months post surgery

Model 2: Risk of revision surgery (exchange, addition, or removal of any prosthetic component)

Model 3: Poor outcome = Change in OSS $\leq$ 0 points and/or revision surgery within one year

Whilst the outcome variable for the regression in model 1 is the postoperative score, knowledge of the preoperative score allows the calculation of predicted change score from the observed Q1 OSS and predicted Q2 OSS.

7.2.3 Independent variables

The factors known to influence the outcomes of shoulder replacement were identified in chapter 1.10 and matched to those available in the routinely collected NJR MDS dataset to define the variable lists *a priori*. The following variables were used in fitting all of the models: age, sex, indication for surgery, status of the rotator cuff, history of previous surgery, American Society of Anaesthesiologists Grade (ASA), and dominance of the operated arm. Baseline Q1 OSS was also included in models 1 and 2, but not in model 3. The use of baseline adjustment within a model predicting change scores will result in inflated regression coefficients since it is used as both a predictor and in the definition of the outcome.[355–357] Sensitivity analyses was performed adding in the procedure type and by removing baseline Q1 OSS adjustment from model 1.

Variables and their coding are described in table 7.2. The continuous predictors (age, Q1 OSS) were modelled using restricted cubic splines with four knots to account for any non-linear relationships. Spline functions allow for a model to represent a relationship between a variable and the outcome with a smooth flexible curve. This is a useful approach for biological modelling, where true relationships are rarely linear. Restricted cubic splines apply a cubic polynomial transformation between each knot and constrain the effects beyond the outer knots to be linear to avoid overfitting in the tails.[362] Default knots were placed at the 5th, 35th, 65th and 95th centiles for the distribution of each continuous variable.

The predictor list was assessed for variable redundancy using the **redun** command of the **Hmisc** package, which applies a flexible additive model to determine how well any predictor can be predicted by the other variables.[275] A threshold of 0.7 for the adjusted R^2 was used to identify redundant variables. None of the candidate variables were excluded on this basis. There was evidence of collinearity of procedure type with the other factors, largely driven by the general preference to use RTSR in older patients with more advanced rotator cuff disease. This supported the prior decision not to include it in the list of candidate variables.

Table 7.2: Description of the variables used to fit prediction models

Variable	Description
Age	Whole years
Sex	Female Male
Main indication	OA or CTA Inflammatory Post-trauma sequelae AVN Other
Rotator cuff status	Normal Attenuated Repaired Absent / torn CTA
Previous ipsilateral shoulder surgery	No Yes
ASA	1 2 3 4+
Operated side	Dominant Non-dominant
Q1 and Q2 OSS	Points (0-48 scale)
Procedure type	Hemiarthroplasty TSR RTSR

7.2.4 Model fitting procedures

Models were fitted using maximum likelihood estimation. All variables were prespecified and entered into the models, such that no variable selection procedures were undertaken. Missing OSS data was imputed by the procedures described in chapter 3, model development processes were repeated for each imputed dataset and the results were combined by applying Rubin's Rules.[281] The specific modelling methods depend on the type of outcome being predicted and these are summarised in table 7.3.

Model assumptions were checked by visual inspection of residual plots. Where significant non-proportionality was observed in the Cox models, the use of flexible parametric methods

was explored. Royston and Parmar described a method that models the baseline hazard function using restricted cubic splines and does not require hazards to remain proportional.[363] It has been implemented in R using the **flexsurv** package.[284] For both Cox and flexible parametric models, censoring was assumed to be non-informative. Combining flexible parametric modelling, multiple imputation and robust internal validation is computationally intensive. Therefore the flexible parametric model was developed using complete case data only and the Cox model development was repeated on the same complete case data to facilitate a meaningful comparison of performance.

Table 7.3: Model fitting procedures and performance measures

Model	Model 1	Model 2	Model 3
Outcome	Q2 OSS	Revision risk	Poor outcome
Model type	Ordinary least squares	Cox Proportional Hazards Flexible parametric	Logistic
Performance measures	R-squared Root mean squared error	c-index Scaled Brier score	c-index Scaled Brier score

7.2.5 Sample size considerations

The models were designed based on the long-established rules of thumb requiring a minimum of 10 events per variable. More recently, and after the models were developed, it has been shown that there is a limited rationale for this and similar rules of thumb.[364] These rules have been superseded by more robust methods designed to minimise model overfitting, minimise model optimism, and produce precise estimates of key model parameters.[365,366] The adequacy of the sample sizes in light of these new recommendations will be considered in the discussion.

7.2.6 Model performance and internal validation

The specific performance measures used for each model are shown in table 7.3. Overall goodness of fit for different modelling approaches can be measured using the log likelihood function, where a score closer to zero is better. The Akaike Information Criterion (AIC) also measures goodness of fit but is penalised by increasing model complexity (i.e. more degrees

of freedom/parameters in the model). Both values are dependent on sample size and are therefore mainly suitable for comparing nested models, or different modelling approaches on the same dataset. The improvement of model fit was tested by calculating the likelihood ratio (Chi-squared) statistic.

R-squared (R^2) is sometimes described as the proportion of the variance in outcome that is explained by the model. It is most frequently used for models predicting a continuous outcome. It can be generalised to models with a binary outcome but the interpretation is less intuitive.[367]

The c-index, described by Harrell, is a measure of model discrimination and is equivalent to the AUC. It is primarily used for binary outcomes but can be generalised to survival models. The c-index represents the probability that a randomly sampled patient experiencing an event has a higher predicted risk than a patient who did not experience an event.[368] A score of 1 represents perfect discrimination and a score of 0.5 indicates that discrimination is no better than a coin toss.

The Brier score was originally developed to assess the accuracy of weather forecasts.[369] Used for binary outcomes, it is calculated as the mean squared difference between the observed outcome (0/1) minus the predicted probability. A perfect model has a Brier score of 0% but the maximum score depends on the incidence of the outcome. The scaled version of the Brier score is used in this study. This takes into account the incidence of the outcome and a score of 100% is best. The integrated Brier Score is a special case which has been modified for use in survival modelling with censored data.[370]

These performance measures do not directly relate to the subsequent clinical utility of the models. A number of so-called "decision analytic" measures have been proposed, including classification tables, net reclassification improvement, and integrated discrimination improvement.[371] The plotted distributions of observed versus predicted values may be more intuitive and will be used here.[372]

A number of methods are available to quantify or reduce the risk of model overfitting. A well-established approach is to estimate a uniform or global shrinkage factor. This takes a value between 0 and 1, where a value of 1 indicates that no overfitting is present. A shrinkage factor of

$\geq$0.9 is considered acceptable evidence of minimal overfitting.[368] The fitted model coefficients can be revised by multiplication with the shrinkage factor, and then the model intercept can be recalculated.[365,366] The shrinkage factor was calculated for this thesis using the heuristic formulas described by Copas for linear models, and by Van Houwelingen and Le Cessie for binary or time-to-event models.[373,374]

Internal validation of model performance was undertaken using the bootstrap method. From a development dataset (size=*n*), *n* individuals were drawn with replacement after each draw to create a bootstrap sample of size *n*. A model was developed on this bootstrap sample and the performance was assessed using the standard measures (R^2, root mean squared error, c-index, Brier score). The process was repeated 200 times and the mean of each performance measure was calculated together with confidence intervals. Only cases with complete data were used as the validation dataset. Bootstrapping is a robust method of internal validation and provides an estimate of the likely model performance that would be seen if the model was applied to a new external dataset.[375]

Model fit was assessed visually by plotting calibration in tenths, together with a smoothed line of best fit using splines (3 knots). Best fit lines were plotted for both original and shrinkage adjusted model estimates.

7.3 Results

7.3.1 Demographics

The patient demographics for the different samples used in the model development process are described in table 7.4. The main difference was that only patients undergoing surgery at the time of the Q2 OSS pilot data collection were eligible for inclusion in models 1 and 3.

Table 7.4: Modelling sample demographics

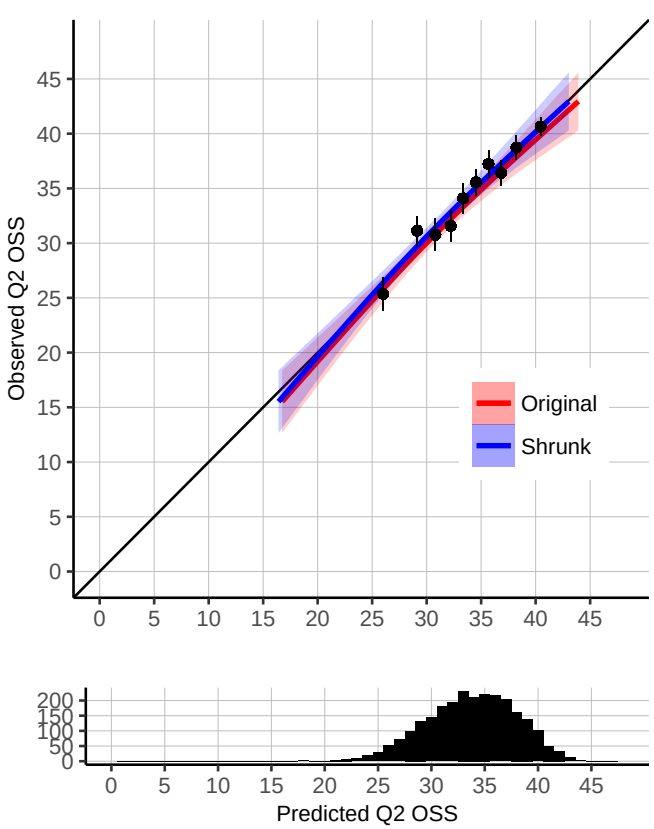
Variable	Level	Models 1 + 3	Model 2	Model 2 (complete cases)
N		7556	19640	8792
Age (years)	Median (IQR)	73 (67 to 79)	74 (67 to 79)	74 (67 to 79)
Sex	Female	5323 (70.4%)	13789 (70.2%)	6209 (70.6%)
	Male	2233 (29.6%)	5851 (29.8%)	2583 (29.4%)
Main indication	OA / CTA	6427 (85.1%)	16765 (85.4%)	7577 (86.2%)
	Post-trauma	450 (6.0%)	1204 (6.1%)	483 (5.5%)
	Inflammatory	351 (4.6%)	902 (4.6%)	430 (4.9%)
	AVN	191 (2.5%)	489 (2.5%)	206 (2.3%)
	Other	137 (1.8%)	280 (1.4%)	96 (1.1%)
Rotator cuff status	Normal	3117 (41.3%)	7660 (39.0%)	3468 (39.4%)
	Attenuated	1685 (22.3%)	4479 (22.8%)	1973 (22.4%)
	Repaired	132 (1.7%)	299 (1.5%)	87 (1.0%)
	Absent / torn	673 (8.9%)	1948 (9.9%)	776 (8.8%)
	CTA	1949 (25.8%)	5254 (26.8%)	2488 (28.3%)
Previous surgery	No	6135 (81.2%)	15960 (81.3%)	7175 (81.6%)
	Yes	1421 (18.8%)	3680 (18.7%)	1617 (18.4%)
ASA	1	567 (7.5%)	1473 (7.5%)	643 (7.3%)
	2	4993 (66.1%)	12874 (65.5%)	5748 (65.4%)
	3	1921 (25.4%)	5141 (26.2%)	2336 (26.6%)
	4+	75 (1.0%)	152 (0.8%)	65 (0.7%)
Operated side	Dominant	4939 (65.4%)	12419 (63.2%)	5418 (61.6%)
	Non-dominant	2817 (37.3%)	7221 (36.8%)	3374 (38.4%)
Annual unit caseload	Median (IQR)	9 (5 to 18)	9 (4 to 18)	10 (5 to 20)
Procedure type	Hemiarthroplasty	1851 (24.5%)	3919 (20.0%)	1629 (18.5%)
	TSR	2556 (33.8%)	6734 (34.3%)	3658 (41.6%)
	RTSR	2797 (37.0%)	8240 (42.0%)	3210 (36.5%)
	Unclear	352 (4.7%)	747 (3.8%)	295 (3.4%)
Q1 OSS	Median (IQR)	16 (11 to 22)	16 (11 to 23)	16 (11 to 23)
	Proportion missing	53%	55%	0%
Q2 OSS	Median (IQR)	37 (26 to 43)	37 (26 to 43)	37 (27 to 43)
	Proportion missing	34%	75%	71%
Change score (Q2-Q1)	Median (IQR)	18 (9 to 26)	18 (9 to 26)	17 (9 to 26)
	Proportion missing	67%	87%	71%

7.3.2 Model 1: Prediction of postoperative Oxford Shoulder Scores

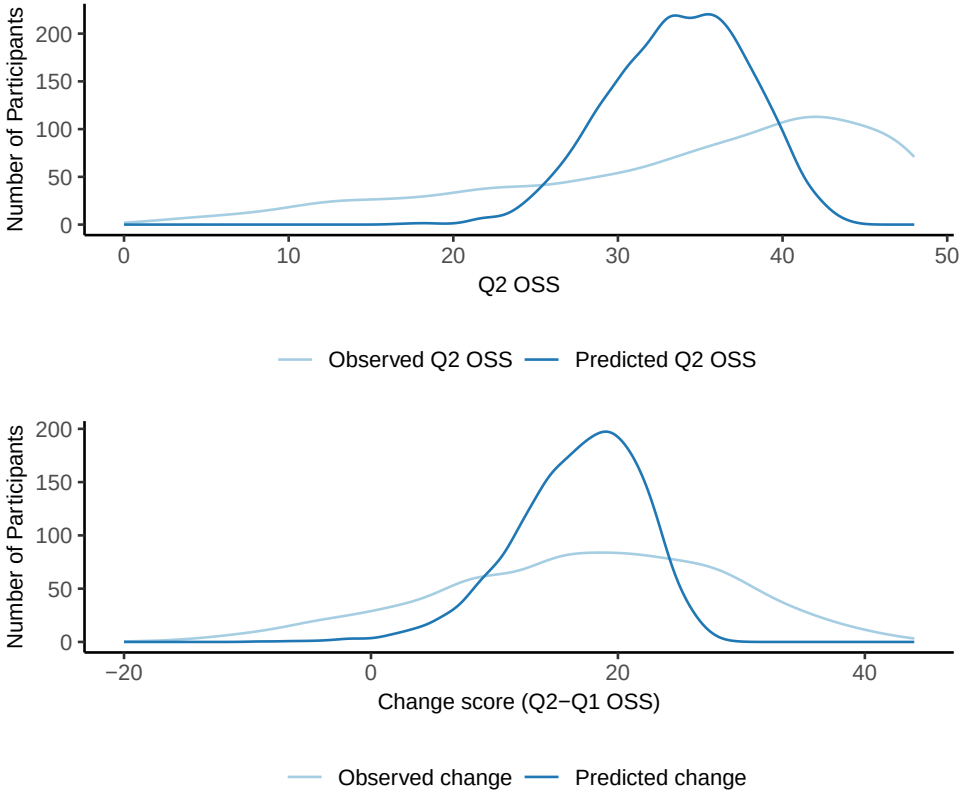
7556 records were used to develop an ordinary least squares (linear) regression model. 67 imputations were performed to account for missing Q1 OSS and Q2 OSS data. The median Q2 OSS was 37 (IQR 26 to 43). The observed distribution of the Q2 OSS shows a slight left skew with a ceiling effect, compared to the normally distributed model predictions. However, the calculated observed and predicted change scores were both normally distributed (figure 7.1b) and the model residuals approximated to a normal distribution. The global shrinkage estimate was 0.98, indicating that there was negligible overfitting present in the model. The corrected R^2 was 0.16 (95% CI 0.16 to 0.16) and the root mean squared error was 10.45 (95% CI 10.45 to 10.46). The calibration plot in figure 7.1a demonstrates that there was minimal evidence of systematic miscalibration, despite the Q2 OSS ceiling effect. If the validation and calibration processes were repeated using the calculated change score as the outcome (predicted Q2 OSS minus observed Q1 OSS) then the results would be the same since the standard error of the model depends only on variation in the predicted Q2 OSS.

With the exception of sex and hand-dominance, an analysis of covariance found that all of the variables included had significant effects ($p<.05$) including the non-linear effects of age and Q1 OSS. The partial-effects plot of predicted outcomes in figure 7.2 demonstrates that the most influential variables were Q1 OSS, health status (measured by the ASA), shoulder replacement performed for the sequelae of trauma or CTA, and a history of previous surgery. Age was also important and the highest Q2 OSS outcomes were observed in patients aged 70 to 80 years.

Compared to a null model, the likelihood ratio (Chi-squared) statistic was 1228 for 20 degrees of freedom ($p<.0001$). Two sensitivity analyses were performed. A model without the inclusion of the baseline Q1 OSS performed considerably worse (Chi-squared 598, 3 degrees of freedom, $p<.0001$), with a considerable drop off in the proportion of explained variance (corrected R^2 0.08). A model including the type of procedure performed better (Chi-squared 288, 2 degrees of freedom, $p<.0001$), with a small increment in the proportion of explained variance (R^2 0.19).



(a) Calibration plot plus histogram of predicted probabilities



(b) Distribution of observed vs predicted scores.
Upper panel - Q2 OSS. Lower panel - Change score

Figure 7.1: Performance of model 1.
Calibration plot of predicted vs observed Q2 OSS in tenths with overlying smoothed curves.

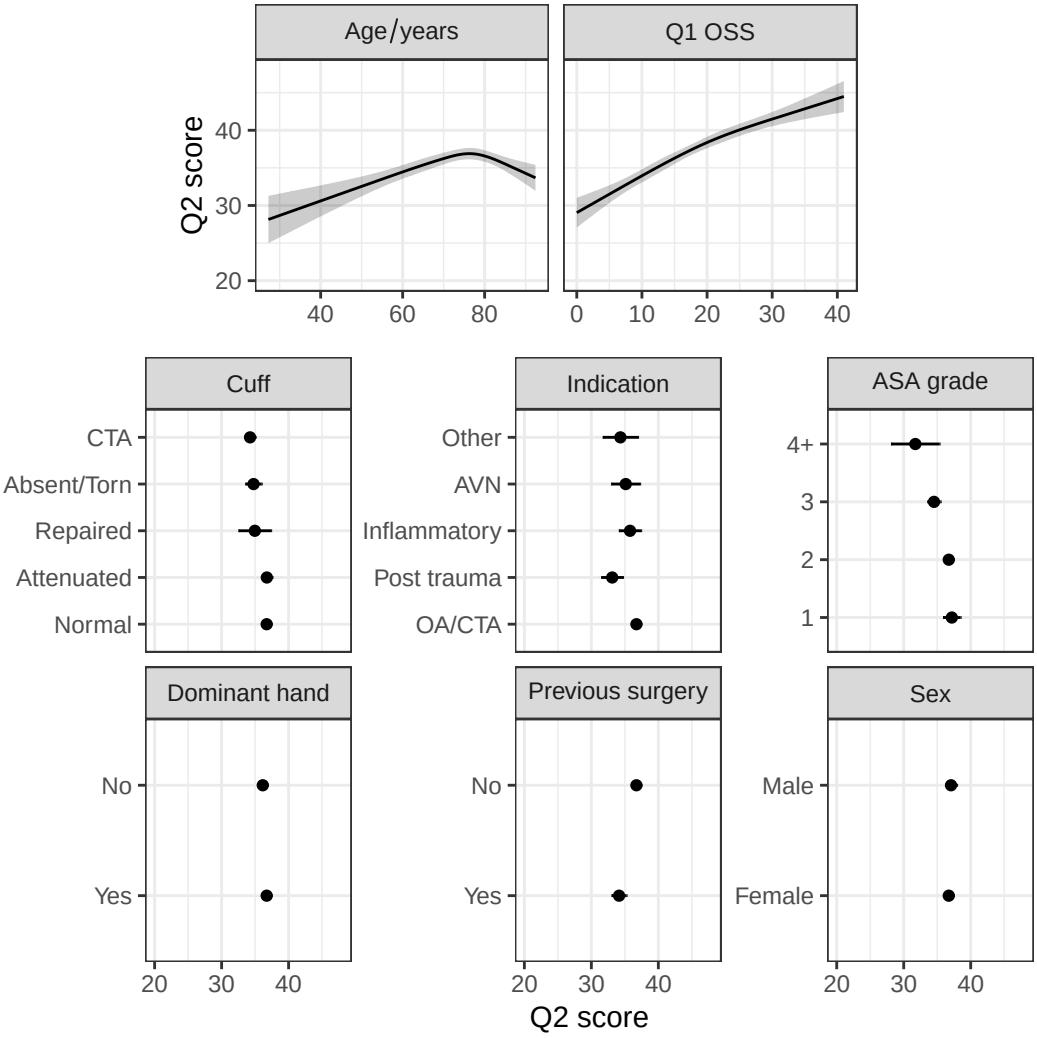


Figure 7.2: Model 1: Partial-effects plot for predicted Q2 OSS

7.3.3 Model 2: Prediction of early revision risk (within two years)

A: Proportional hazards approach - imputed dataset

The full dataset of 19640 records was used to develop this model, including 438 patients who had been revised ("events") prior to censoring or death. 55 imputations were performed to account for missing Q1 OSS data. The proportional hazards assumption was breached and Schoenfeld residual plots found that hazard rates varied over time for sex, rotator cuff status and indication for surgery (not shown). The global shrinkage estimate was 0.82, indicating that some model overfitting was present. At predicted probabilities of revision greater than 5%, the model underestimated the risk of revision surgery, shown by the calibration plot for revision risk

at two years in figure 7.3a. This was not improved by model shrinkage. The overall predictive performance was poor. The raw integrated Brier score was 0.037 and the scaled score was 0% (95% CI 0 to 1). The risk distribution plots demonstrated limited separation between the risks predicted for those revised and not revised (figure 7.3b). This is consistent with the c-index of 0.64 (95% CI 0.62 to 0.65), which indicates modest discrimination at best. Compared to a null model the likelihood ratio (Chi-squared statistic was 108 on 20 degrees of freedom, p<.0001).

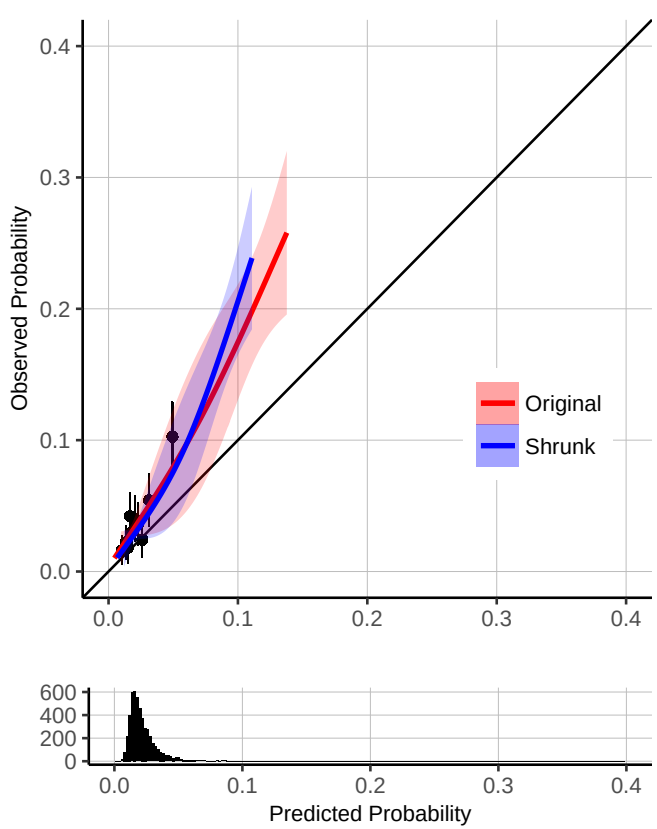
B: Proportional hazards approach - complete cases dataset

This analysis was performed to facilitate direct comparison with the subsequent flexible parametric survival model. A subset of 8792 records, including 192 revision events, were used to develop the model. The same problem was observed with a breach of the proportional hazards assumption. The global shrinkage estimate was 0.71, indicating that considerable model overfitting was present. At a predicted probabilities of revision greater than 5%, this model also underestimated the risk of revision surgery, shown by the calibration plot for revision risk at two years in figure 7.4a. The risk distribution plots were not substantially different from those for model 2A or 2C (figure 7.4b. The raw integrated Brier score was 0.037 and the scaled score was 1% (95% CI 0 to 1) and the c-index was 0.64 (95% CI 0.63 to 0.66). Compared to a null model the likelihood ratio (Chi-squared statistic was 71 on 20 degrees of freedom, p<.0001).

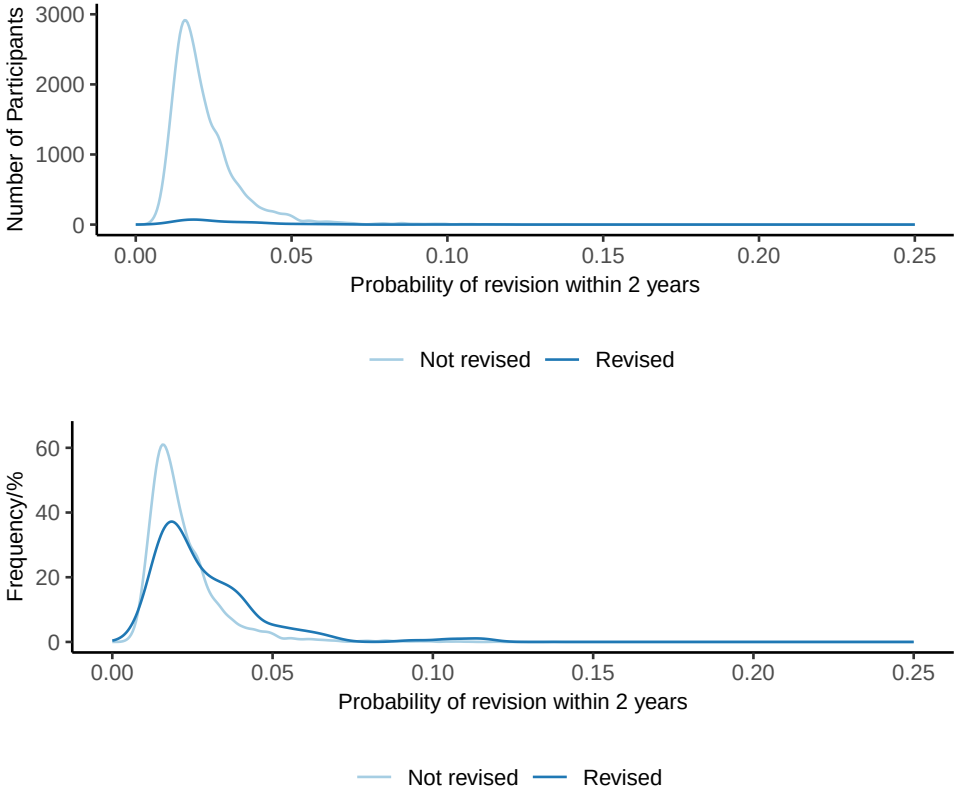
C: Flexible parametric approach - complete cases dataset

The same subset of 8792 records was used to develop a flexible parametric model, intended to mitigate for the breach of the proportional hazards assumption. The baseline hazard was fitted using splines with one knot, placed at one year based on the observed time-varying hazards seen in residual plots for the Cox models. Sex, rotator cuff status and indication for surgery were included as ancillary variables. The global shrinkage estimate was 0.86, indicating that some model overfitting was present but considerably less than that seen for the equivalent Cox modelling approach. The risk distribution plot in figure 7.5b shows that this model represents the spread of observed revision risk marginally better, supported by the calibration plot in figure 7.5a. However, it should be noted that the tails of the calibration plots are highly sensitive to

small changes in the data used for model development and validation, and a shrinkage adjusted calibration has not been performed. The model discrimination was also limited with the flexible parametric approach. The raw integrated Brier score was 0.037 and the scaled score was 1% (95% CI 0 to 14), the c-index was 0.66 (95% CI 0.61 to 0.70). The addition of time varying covariates increased the number of degrees of freedom to 32. This model is therefore likely considerably underpowered, based on the 10 events per variable rule of thumb. Compared to a null model the likelihood ratio (Chi-squared) statistic was 230 on 32 degrees of freedom ($p<.0001$).

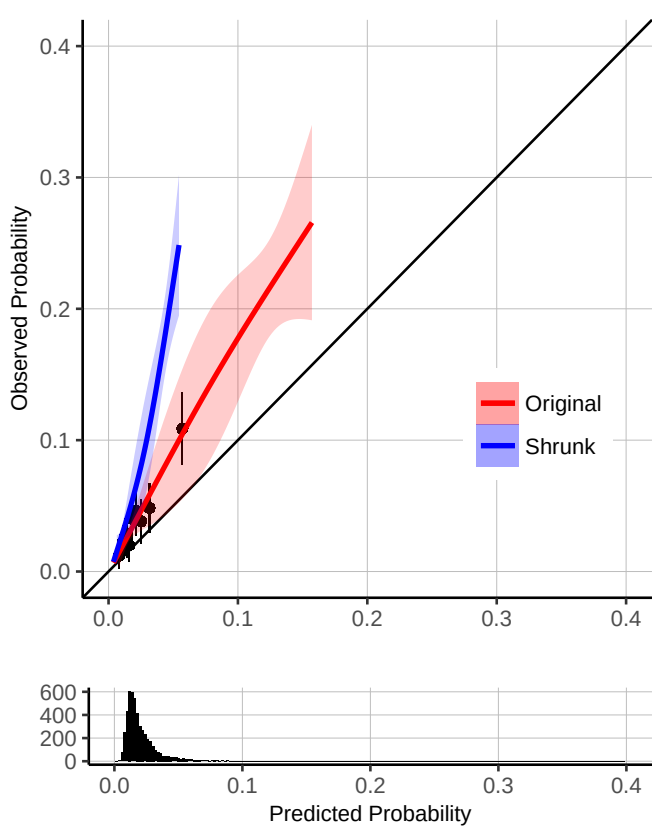


(a) Calibration plot plus histogram of predicted probabilities

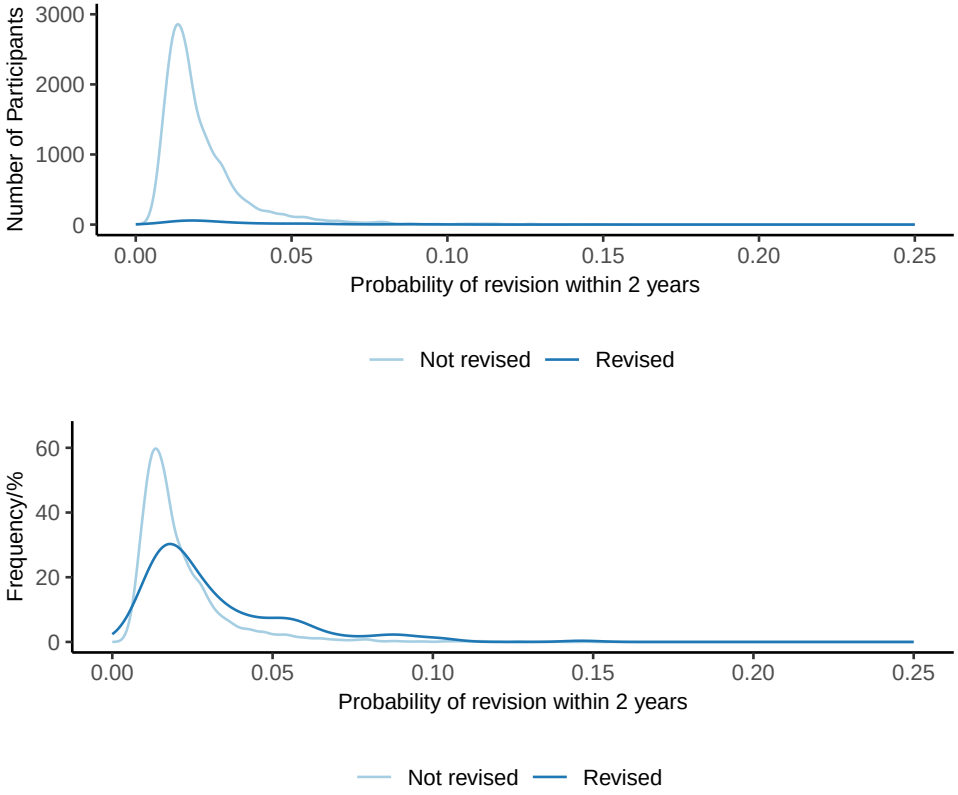


(b) Distribution of predicted probabilities by observed outcome.
Upper panel - absolute numbers. Lower panel - within-group proportion.

Figure 7.3: Performance of model 2A (Cox proportional hazards with imputation).
The calibration plot shows predicted vs observed probabilities in tenths with overlying smoothed curves.

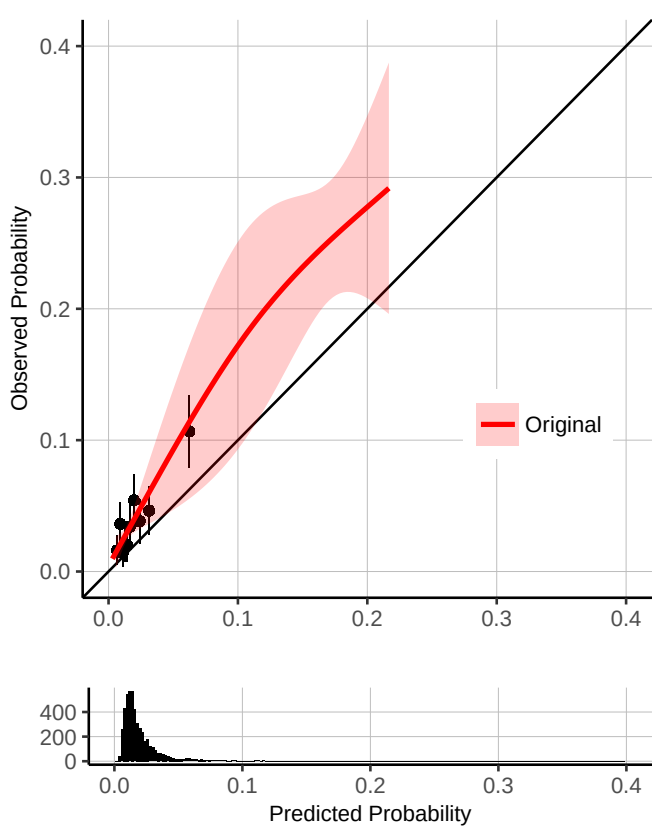


(a) Calibration plot plus histogram of predicted probabilities

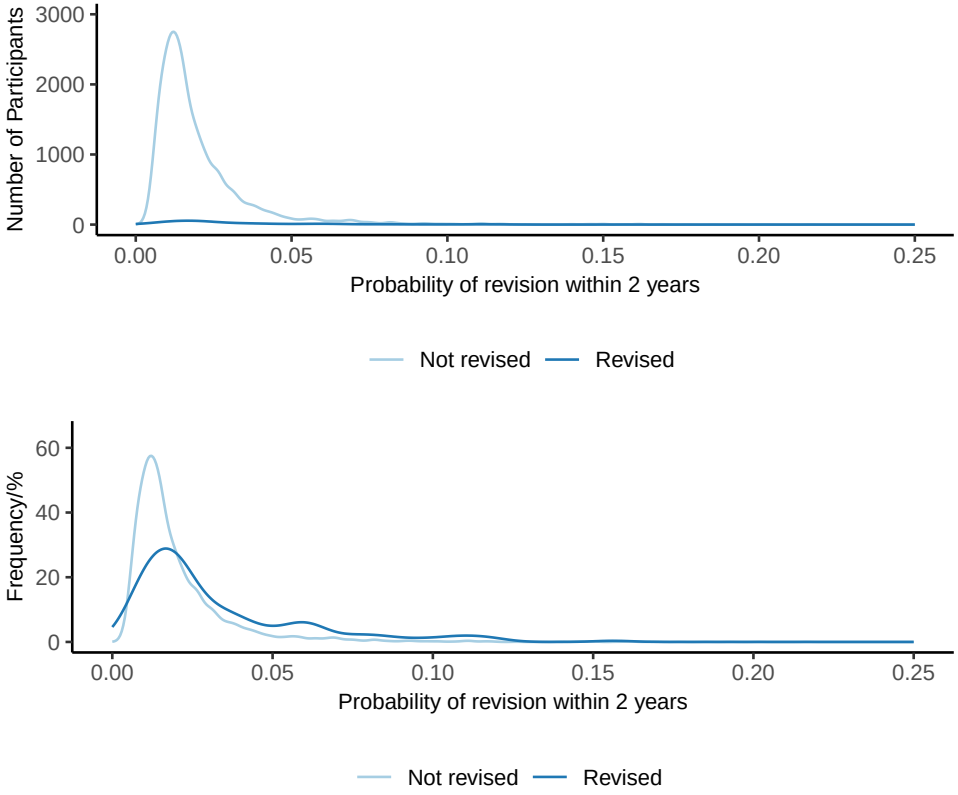


(b) Distribution of predicted probabilities by observed outcome.
Upper panel - absolute numbers. Lower panel - within-group proportion.

Figure 7.4: Performance of model 2B (Cox proportional hazards on complete cases).
The calibration plot shows predicted vs observed probabilities in tenths with overlying smoothed curves.



(a) Calibration plot plus histogram of predicted probabilities



(b) Distribution of predicted probabilities by observed outcome.
Upper panel - absolute numbers. Lower panel - within-group proportion.

Figure 7.5: Performance of model 2C (Flexible parametric on complete cases).
The calibration plot shows predicted vs observed probabilities in tenths with overlying smoothed curves.

Comparison between model fits

The AIC for the complete case Cox model was 8194 compared to 2236 for the flexible parametric model. The likelihood ratio (Chi-squared) statistic comparing the two was 1086 on 12 degrees of freedom. Both indicate that the fit of the flexible parametric model was considerably better than the fit of the Cox model (p<.0001). Model partial effects are not reported due to the considerable potential for overfitting, lack of model power and likely unreliability of estimates.

7.3.4 Model 3: Prediction of patients with a poor outcome

This final model was based on 7556 records, including 793 records of patients with a poor outcome. 67 imputations were performed to account for missing Q1 OSS and Q2 OSS data. The global shrinkage estimate was 0.90, indicating that there was limited overfitting present in the model. The raw Brier score was 0.098 and the scaled score was 4% (95% CI 0 to 13), the c-index was 0.64 (95% CI 0.61 to 0.68). The calibration plot in figure 7.7a demonstrates that there is minimal evidence of systematic miscalibration. The risk distribution plot in figure 7.7b shows a better separation of risks between those observed to have a poor or satisfactory outcome than was found for any of the pure revision-risk models.

An analysis of covariance showed that age, sex, and history of previous surgery had significant effects (p<.05). Accounting for non-linear age effects was not significant (p=0.10). The partial-effects plot of predicted probabilities in figure 7.6 demonstrates that the most influential variables were shoulder replacement performed for the sequelae of trauma, a history of previous surgery and age. Mirroring the observations from model 1, patients aged 70 to 80 were least likely to have a poor outcome.

Compared to a null model, the likelihood ratio (Chi-squared) statistic was 176 for 17 degrees of freedom (p<.0001). Two sensitivity analyses were performed. A model with the inclusion of the baseline Q1 OSS performed considerably better (Chi-squared 167, 3 degrees of freedom, p<.0001), produced near perfect calibration (not shown) but resulted in no meaningful

improvement in discrimination indices. For the reasons already stated regarding bias in baseline adjusted analyses of change scores, this approach was not considered further. A model including the type of procedure also performed better (Chi-squared 76, 2 degrees of freedom, p<.0001) but with similarly no meaningful improvement in model discrimination.

Full equations for each of the models are reported in appendix A.8.

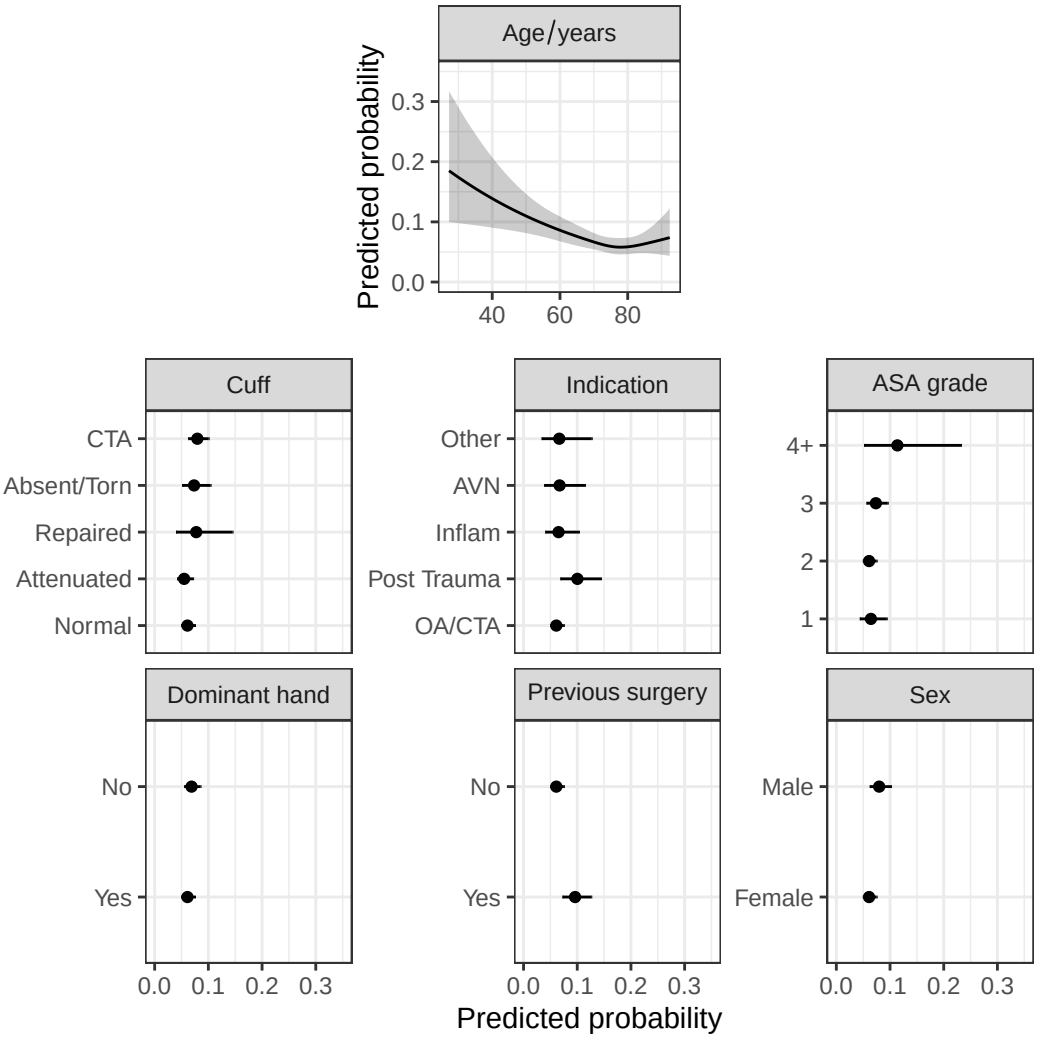
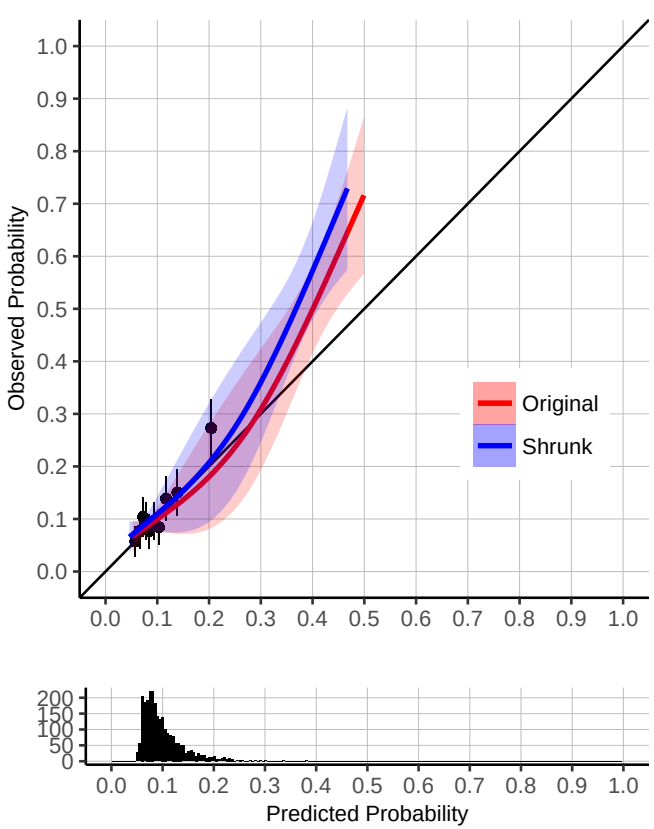
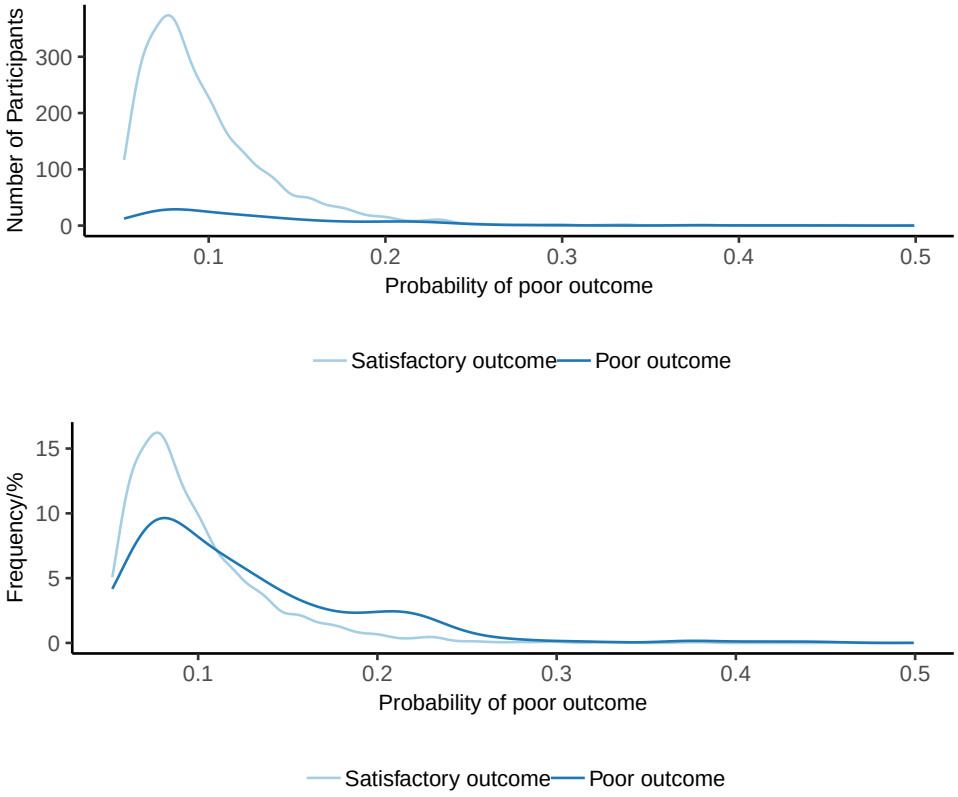


Figure 7.6: Model 3: Partial-effects plot for predicted probability of a poor outcome



(a) Calibration plot plus histogram of predicted probabilities



(b) Distribution of predicted probabilities by observed outcome.
Upper panel - absolute numbers. Lower panel - within-group proportion.

Figure 7.7: Performance of model 3.
Calibration plot shows predicted vs observed probabilities in tenths with overlying smoothed curves.

7.3.5 Example scenarios

The following contrasting scenarios may be useful to put the model predictions into context. A 75 year-old woman with osteoarthritis, a normal rotator cuff, no history of prior surgery, and a starting Q1 OSS of 16 points has a predicted Q2 OSS of 37 points (95% CI 36.1 to 37.6), predicted improvement of 21 points (95% CI 10.1 to 11.6), and a 5.9% (95% CI 4.8 to 7.4) likelihood of experiencing no improvement in function and/or requiring revision surgery within one year. A 50 year-old man with osteoarthritis, a torn rotator cuff, a history of previous surgery and a starting Q1 OSS of 16 points has a predicted Q2 OSS of 28 points (95% CI 26.3 to 30.1), predicted improvement of 12 points (95% CI 10.3 to 14.1), and a 24.9% (95% CI 16.8 to 34.6) likelihood of experiencing no improvement in function and/or requiring revision surgery within one year.

7.4 Discussion

7.4.1 Key findings

On average, patients report considerable improvement in pain and function measured on the Oxford Shoulder Score at 4 to 12 months (median 18 points, IQR 9 to 26), well in excess of the suggested value for the minimal clinically important difference of 4.5 points.[130] Up to 10.5% of patients will report no improvement in their postoperative scores and/or require revision surgery within one year.

Regression models were developed to predict three different outcomes following elective shoulder replacement surgery based on eight preoperative variables collected prospectively by the NJR. The first model, predicting postoperative outcome scores (Q2 OSS), performed modestly (R^2 0.16). The model calibration was good with no evidence of systematic bias and no evidence of overfitting (optimism). The second group of models used survival modelling methods to predict early revision risk. Regardless of whether Cox proportional hazards or flexible parametric methods were used, the survival models performed poorly with considerable

evidence of model overfitting. The final model, predicting the likelihood of a poor outcome following surgery also demonstrated modest performance (c-index 0.64), reasonable calibration and minimal overfitting.

All candidate variables were included in the models without further selection processes. So, whilst the individual effects of the independent variables were not the key focus of the study, their relative influence on outcome may be of interest. In the presence of the likely considerable unmeasured confounding, no causal association can be inferred. Future prediction modelling exercises should consider these factors when selecting their candidate variables. In model 1, the lowest postoperative scores (Q2 OSS) were observed in younger patients, those with a lower baseline Q1 OSS, an ASA greater than 2, a history of previous trauma and/or previous surgery, and full-thickness tear of the rotator cuff or CTA. If change scores were the outcome of interest, the direction of effect of the Q1 OSS was reversed (i.e. lower Q1 OSS larger change score) but the effects of the other covariates is unchanged. The Q1 OSS accounted for half of the explanatory value of the model. In model 3, the patients least likely to improve after surgery were those who were younger, with a history of previous trauma and/or previous surgery.

7.4.2 Strengths and limitations

The main strength of this study is the robust approach taken to the steps of model development and validation, in line with TRIPOD guidance.[360,361] Steps were taken throughout to avoid inducing bias into the model predictions or model overfitting. These steps included *a priori* definition of independent variables, avoiding the use of stepwise model selection processes or dichotomisation of variables, and using multiple imputation where possible to increase statistical efficiency and to avoid the potential for bias from complete case analysis.[274,376] The robust methods of internal validation provide a fair estimation of how the models may perform on unseen data (i.e. new patients).

The dataset used for this study is the largest known resource of pre- and postoperative PROMs for shoulder replacement surgery. The dataset represents unselected patients undergoing shoulder replacement across NHS and independent providers in the UK (excluding Scotland),

and as such the observed effects of patient factors on outcomes and model predictions are likely to be more generalisable and reliable than those from smaller models developed on single-centre datasets.

An important limitation of the study is the incompleteness of PROMs data in the NJR dataset. Robust imputation methods were employed to reduce the potential for this missing data to bias results, but any misspecification of the imputation model may in itself introduce bias.[276] Clustering of missing data by centre or surgeon is one potential source of misspecification.

Postoperative scores were recorded at a median of 7.4 months (range 5.1 to 12, IQR 6.1 to 9.4). This may be earlier than the point at which maximal improvement is achieved following surgery. Simovitch et al have shown that approximately 80% of maximum improvement in the ASES Score occurs in the first 6 months, and they found no important differences in the rate of improvement between patients undergoing TSRs or RTSRs.[377] So the overall benefits of surgery may be slightly underestimated by this current study.

For model 3, the endpoint chosen was worsening of the OSS (change score $\leq$0 points), incorporating cases requiring early revision to ensure that outright failures were not censored from the outcomes. There is no clearly defined anchor-based MCID for the OSS in relation to shoulder replacement, although a value of 4.5 points has been suggested in the literature.[130] This corresponds closely to a value that can be obtained from a distribution based method using half of the standard deviation of the Q1 OSS in this study (8.2/2=4.1).[378] Arguably, an improvement of 4.5 points on the 48-point OSS scale represents a small potential gain after major shoulder surgery. The value may prove to be higher in the future once research has been conducted to anchor this commonly used tool to patient-reported satisfaction after this type of major surgery.[379] The 10.5% of patients with a poor outcome may represent an underestimate of the number who fail to achieve meaningful benefit or satisfaction from their surgery. Further work is needed to more accurately identify this group and to optimise their outcomes.

The revision risk models (2A-C) provided no useful predictions to inform clinical practice. The absolute number and relative frequency of revision events in the first two years is too small to support reliable predictions. Any discrepancies in revision risk at this early stage may have more to do with choice and conduct of surgery than the patient factors available for this study.

Factors such as patient age are likely to have more influence on longer-term risk of revision surgery, as discussed in chapter 5.

Sample size

Rules of thumb, such as 10 events-per-variable, are widely used but are now being replaced by more statistically sound methods. The series of criterion based rules were retrospectively applied, as described by Riley et al, to put the study findings into the context of current recommendations.[365,366] For model 1, the shrinkage factor was close to 1, a small absolute difference was observed in the apparent and adjusted R^2, the sample was considerably larger than a threshold of 254 patients (234 + number of parameters). Based on the 7556 available patients and the known sample Q2 OSS variance of 125 points, the precision of the intercept could be estimated with a 95% CI of +- 0.36 points.

For model 3, the estimated shrinkage factor was at an acceptable threshold of 0.9. Based on an outcome frequency of 11%, the sample size was considerably larger than a calculated sample size threshold of 3761 to achieve an absolute margin of error in the model intercept of 1%. The difference between the apparent and adjusted values for Nagelkerke's R^2 was 0.005, less than Riley et al's recommended threshold of 0.05.

7.4.3 Comparison to other studies

The influences of age, health/comorbidities, indication for surgery, history of prior surgery, and the state of the rotator cuff are consistent with the factors identified in the literature search informing the model development process (section 1.10). None of the outcomes modelled by Simmen, Matsen or Chen are directly comparable to those in this chapter and hence comparison of the magnitude of individual variable effects is not relevant.[176,177,180] The reported performance of their models, based on the AUC, was better than that achieved in this study. Numerous methodological flaws (lack of power, stepwise model selection, arbitrary outcome thresholds) mean that their results would be extremely unlikely to be reliably reproduced in practice. Simmen and Chen included validated scores for preoperative

psychosocial factors and in both cases these factors had important effects in the final models. This information may have contributed to improved model performance.

There are some parallel models available in for NJR hip and knee outcomes. Whilst the R^2 value was only 0.16 for model 1 here, the reported values for the hip and knee models in established use for outlier analysis are only 0.24 and 0.26 respectively.[380] These lower limb models include some additional comorbidity data, self-reported disability and depression scores, and some postoperative factors. These variables may well explain the better performance. However, there may also be greater heterogeneity in outcomes following shoulder replacement, suggested by greater observed variation in the PROMs following shoulder replacement than those reported after hip and knee replacement.[381] The models and performance statistics behind the public-facing NJR JointCalc are not published.

7.4.4 Meaning of the study

On average, patients can expect to experience considerable improvement in pain and function following elective shoulder replacement surgery. The level achieved postoperatively and the change from baseline can be predicted with modest accuracy based on routinely collected preoperative information. The likelihood of failing to improve with surgery can also be predicted with modest accuracy. These predictions could be used to inform preoperative discussions between patients and surgeons. However, it should be acknowledged that there is considerable variation in outcomes that is not captured within NJR data and not predicted by these models. The model predicting poor outcomes does not produce extreme predictions (model 3). The 90th, 99th and 100th centiles of predicted probability were 16%, 25% and 50% respectively. Individuals have different attitudes to risk and may reach different decisions based on this information. The model estimates provide guidance rather than a standalone tool for treatment decision making. They should be interpreted in the context of other factors such as long-term revision risk and the likelihood of adverse events discussed in chapters 5 and 6.

7.4.5 Unanswered questions and future research

Increasingly it is becoming apparent that patient reported function and disability after upper limb injury may be strongly correlated with factors such as depression, anxiety and self-efficacy.[382,383] It is highly likely that the same is true of outcomes following elective surgery, and this was indeed one of the findings of the Chen et al study.[180] Other studies have also related a diagnosis of depression to smaller, but still significant improvements in patient reported outcome scores.[182,384] Future research should assess the incremental value of adding a validated psychosocial measure to the model predictors used here. The trade-off between any improved performance and the burden of additional data collection should be considered as part of any future analyses. A model producing imperfect predictions may be preferable to one with too much complexity.

During the Q2 OSS pilot data collection period, the proportion of scores returned by patients postoperatively was much higher than the proportion of preoperative scores returned by hospitals. Reasons for poor compliance with Q1 OSS reporting should be urgently investigated. The number of replacements performed and at risk of subsequent revision surgery is much lower for shoulders than for hips or knees. As such, the statistical power for outlier analyses is limited. PROMs have the potential to be useful for outlier analyses in shoulders, but compliance rates need to drastically improve if they are to support valid comparisons between providers, surgeons and implants. Further research should also establish anchor-based values for acceptable symptom states and minimally important improvements in the OSS for patients undergoing shoulder replacement surgery.

The model development process for revision risk has identified that Cox proportional hazards methods may not adequately represent the patterns of failure following shoulder replacement. This is likely to be more important for analyses by procedure type where it is known that hemiarthroplasties, TSRs and RTSRs all produce different (and crossing) survival curves (figure 5.6a in chapter 5). Further research should explore the value of flexible parametric modelling for these situations in suitably large samples.

The effect of different types of shoulder replacement on outcome was specifically not the

focus of this chapter, but the sensitivity analyses suggested that procedure type may influence outcomes. The effect of implant type in different patient groups will be considered in detail in the final study in chapter 8.

Chapter 8

Variation in patient reported outcomes by implant type, age and rotator cuff status recorded in the National Joint Registry

8.1 Introduction

In chapter 1, three key uncertainties for shoulder replacement surgery from the James Lind Alliance (JLA) Priority Setting Partnership were outlined. This chapter addresses the last of these uncertainties: "In patients with shoulder arthritis, is a hemiarthroplasty or a total shoulder replacement or a reverse replacement most effective?".[10] The systematic review in chapter 2 found low-quality evidence, based on two small RCTs, to suggest that there may be a clinically important improvement in function/disability favouring TSR over hemiarthroplasty for the treatment of glenohumeral OA. No studies comparing RTSR to either TSR or hemiarthroplasty were eligible for inclusion.

The RTSR concept was designed with the primary intention of managing patients with severe functional deficits (pseudoparalysis) in rotator cuff arthropathy[116] The indications expanded into patients with irreparable massive cuff tears without OA, rheumatoid arthritis with cuff tears, acute fractures, fracture sequelae and revision.[385] There is little written on more recent trends towards the use of RTSR in older patients for OA, regardless of the condition of the rotator cuff.

Variation in outcomes by the state of the rotator cuff has not been reported by any RCTs and many of the decision-making scenarios in routine clinical practice are poorly represented by the available trials data. In the prediction of postoperative outcomes in chapter 7, the modelled Q2 OSS was approximately two points lower for patients with an absent/torn rotator cuff (+- CTA) than those with a normal or attenuated cuff. Outside of the carefully controlled conditions of a trial, comparison of outcomes between different procedure types should account for patient age and the condition of the rotator cuff.

This study uses routinely collected data from the NJR to address two different questions. Firstly, what type of procedure do surgeons and patients currently choose depending on age and the status of the rotator cuff tendons? Secondly, what are the patient reported outcomes in everyday practice for hemiarthroplasties, TSRs and RTSRs in those different patient groups? The NJR is a good potential source of data to study these stratified patient reported outcomes, since it collects details of the condition of the rotator cuff, and the record of implanted components allows for validation of the type of procedure performed, as described in chapter 3.

8.2 Methods

8.2.1 Data source and participants

The data for this chapter was obtained from procedures recorded in the NJR during the Q2 OSS pilot window between 2012 and 2015. The dataset was cleaned and prepared for analysis as described in chapter 3. Cases were included if they were performed for an indication of OA or CTA. Rotator cuff status was classified according to the details recorded by surgeons on the MDS forms into three groups: normal, attenuated, and absent/torn (including CTA). Cases were excluded if the type of procedure could not be confirmed or if the rotator cuff status was recorded as "repaired". This group of patients with repaired tendons may achieve different outcomes to those with an unrepaired or unrepairable tendon, but the group is too small to undertake a separate analysis.

8.2.2 Interventions/comparisons

Outcomes were compared between three procedure types: hemiarthroplasty, TSR and RTSR. Procedures were grouped into the three major categories according to the description in chapter 1 (figure 1.6), without further subdivision into resurfacing, stemless and stemmed implants.

8.2.3 Outcome measures

The primary outcome measure was the postoperative Oxford Shoulder Score (Q2 OSS) recorded up to one year following surgery (range 4 to 12 months). The established threshold of 4.5 points was used to defined the minimum clinically important difference.[130] The secondary outcome measure was revision risk at two years post surgery. The distribution of implant choice by rotator cuff status, age and Q1 OSS was evaluated graphically.

8.2.4 Statistical analysis

The Q2 OSS by procedure type was estimated using ordinary least squares (linear) regression, adjusted for age, sex, ASA, any history of previous surgery and baseline Q1 OSS. Separate models were constructed for each of the three rotator cuff status groups. Age was modelled using restricted cubic splines to account for non-linearity, with three knots placed at the 10th, 50th and 90th centiles (60, 75 and 85 years). Missing data was replaced using multiple imputation to generate 67 complete datasets. The modelling process was repeated 67 times before pooling effect estimates and standard errors according to Rubin's Rules.[281] Treatment effects were calculated using the **contrast** function of the **rms** package.[285] Predictions were plotted across the range of observed ages from the 10th smallest value to the 10th largest value (43 and 92 years respectively).

A sensitivity analysis was performed by repeating the modelling process with the final rotator cuff category split into a) Absent/torn, and b) CTA.

Risk of revision surgery was estimated using the 1 minus Kaplan-Meier method employed

elsewhere in this thesis, calculating 95% CIs using the Greenwood method on a log-log scale.[316]

8.2.5 Sample size

The minimum sample size for the primary outcome was calculated using the **pwr** package in R, based on the description in Kutner et al.[386] 546 cases were required to detect any difference between procedure types based on a significance level of 0.05, a power of 80%, and a model with 10 degrees of freedom and an arbitrarily small R^2 of 0.03.

8.3 Results

8.3.1 Demographics

A total of 6037 out of a possible 7556 shoulders met the inclusion criteria within the Q2 OSS pilot window, of whom 2488 had a normal rotator cuff, 1196 had an attenuated rotator cuff and 2222 had a full-thickness tear of the rotator cuff or established CTA. The reasons for exclusion were: 352 cases with an unclear procedure type, 1082 with indications for surgery other than OA or CTA, and 85 patients undergoing simultaneous repair of the cuff at the time of shoulder replacement surgery.

Demographics and unadjusted PROMs are reported by rotator cuff status in table 8.1. Patients with a full-thickness rotator cuff tear or CTA were older and had a higher frequency of previous surgery (27% compared to 10% in the normal cuff group). They had a slightly lower Q1 OSS, Q2 OSS and change score.

Table 8.1: Demographics by rotator cuff status

	Normal	Attenuated	Absent/torn +- CTA	Absent/torn	CTA
				Sensitivity*	Sensitivity*
N	2521	1221	2295	365	1930
Age (median/IQR)	72 (66 to 77)	75 (69 to 80)	76 (71 to 81)	76 (72 to 82)	76 (71 to 81)
Male	790 (31.3%)	302 (24.7%)	715 (31.2%)	102 (27.9%)	613 (31.8%)
Female	1731 (68.7%)	919 (75.3%)	1580 (68.8%)	263 (72.1%)	1317 (68.2%)
ASA 1	258 (10.2%)	71 (5.8%)	108 (4.7%)	18 (4.9%)	90 (4.7%)
ASA 2	1735 (68.8%)	843 (69.0%)	1470 (64.1%)	238 (65.2%)	1232 (63.8%)
ASA 3	511 (20.3%)	290 (23.8%)	693 (30.2%)	104 (28.5%)	589 (30.5%)
ASA 4+	17 (0.7%)	17 (1.4%)	24 (1.0%)	5 (1.4%)	19 (1.0%)
Previous surgery	257 (10.2%)	142 (11.6%)	611 (26.6%)	60 (16.4%)	551 (28.5%)
Q1 OSS (median/IQR)	18 (12 to 24)	16 (10 to 21)	15 (10 to 21)	15 (10 to 20)	16 (10 to 21)
Q1 OSS (mean/95%CI)	18.1 (17.8 to 18.4)	16.1 (15.7 to 16.5)	16.2 (15.9 to 16.5)	15.4 (14.6 to 16.1)	16.3 (16 to 16.7)
Completeness	48%	48%	49%	42%	51%
Q2 OSS (median/IQR)	39 (30 to 44)	38 (29 to 43)	35 (25 to 42)	35 (25 to 42)	35 (25 to 41)
Q2 OSS (mean/95%CI)	35.8 (17.7 to 18.5)	34.8 (15.5 to 16.7)	32.5 (15.7 to 16.7)	33.2 (14.2 to 16.5)	32.4 (15.8 to 16.8)
Q2 complete	67%	68%	69%	62%	70%
Change (median/IQR)	19 (11 to 26)	20 (12 to 27)	17 (8 to 24)	17 (9 to 27)	17 (8 to 24)
Change (mean/95%CI)	18.1 (17.7 to 18.6)	19.3 (18.7 to 19.9)	16.1 (15.6 to 16.5)	16.9 (15.7 to 18.1)	15.9 (15.4 to 16.5)
Change complete	34%	35%	36%	28%	36%
Hemiarthroplasty	784 (31.1%)	474 (38.8%)	196 (8.5%)	67 (18.4%)	129 (6.7%)
TSR	1639 (65.0%)	543 (44.5%)	60 (2.6%)	36 (9.9%)	24 (1.2%)
RTSR	98 (3.9%)	204 (16.7%)	2039 (88.8%)	262 (71.8%)	1777 (92.1%)

*The absent/torn +- CTA group was split for sensitivity analysis

8.3.2 Implant choice by age and rotator cuff status

The choice of procedure by age and rotator cuff status is shown in the frequency density plots in figure 8.1. The distribution of Q1 OSS by procedure type and rotator cuff status is shown in figure A.9 in appendix A.9.

In patients with a normal rotator cuff, 31% underwent hemiarthroplasty, 65% underwent TSR and 4% underwent RTSR. No obvious important differences were observed between the ages of patients having hemiarthroplasties or TSRs. The median age for both was 71 years (IQR 65 to 77). Patients undergoing an RTSR were older (median 76 years, IQR 70 to 81). The distribution of baseline Q1 OSS was similar for the different procedure types.

In patients with an attenuated rotator cuff, 39% underwent hemiarthroplasty, 44% underwent TSR and 17% underwent RTSR. The patients undergoing RTSR were slightly older than those having a hemiarthroplasty or TSR, and they had a slightly lower Q1 OSS.

In patients with an absent/torn rotator cuff or CTA, 8% underwent hemiarthroplasty, 3% underwent TSR and 89% underwent RTSR. There were no differences in the age or Q1 OSS profiles between patients treated with hemiarthroplasty compared with RTSR or TSR. Fewer than six patients aged under 60 years underwent RTSR with a normal or attenuated rotator cuff, compared to 51 with a torn rotator cuff.

8.3.3 PROMs

All mean treatment effects reported in the text are adjusted to a patient aged 74 years with a Q1 OSS of 16 points, an ASA grade of 2 and no history of previous surgery. These represent the median values in the sample. The reference procedure type was set as hemiarthroplasty for each of the analyses. In figure 8.2, predicted Q2 OSS is plotted against age. Since the relationships were modelled without interactions, the mean treatment effect of procedure type is constant with age. However, the uncertainty (width of confidence intervals) of the estimated effect size varies according to the number of patients in the models at different ages.

For patients with a normal rotator cuff, the mean predicted Q2 OSS following hemiarthroplasty was 32.59 points (95% CI 31.55 to 33.63). Following TSR the Q2 OSS was 5.49 points higher (95% CI 4.43 to 6.54), and following RTSR there may be little or no difference (2.57 points, 95% CI -0.08 to 5.22). The difference between hemiarthroplasty and TSR was present throughout the age range, with no overlap between confidence intervals. Few patients underwent RTSR and there was consistent overlap with both TSR and hemiarthroplasty estimates.

For patients with an attenuated rotator cuff, the mean predicted Q2 OSS following hemiarthroplasty was 33.41 points (95% CI 32.08 to 34.74). Following TSR the Q2 OSS was 5.54 points higher (95% CI 4.06 to 7.03), and following RTSR it was 3.79 points higher (95% CI 1.72 to 5.86). There was no significant difference between TSR and RTSR. Below the age of 55 years for TSR, below 65 years for RTSR, and above 85 years for RTSR, the inferiority of hemiarthroplasty is uncertain since there were fewer patients in the model and the resulting wide confidence intervals overlap.

For patients with an absent/torn cuff or CTA, the mean predicted Q2 OSS following hemiarthroplasty was 29.78 points (95% CI 27.74 to 31.81) following hemiarthroplasty. Following RTSR the Q2 OSS was 4.58 points higher (95% CI 2.59 to 6.57), and following TSR the Q2 OSS was 5.62 points higher (95% CI 1.85 to 9.39). Below the age of 60 years very few patients underwent any procedure other than RTSR. No meaningful comparisons can be made for this patient group.

Since very few patients underwent RTSR with a normal cuff or TSR with an absent/torn cuff, the effect estimates are extremely unreliable with wide confidence intervals. Little should be concluded in those scenarios.

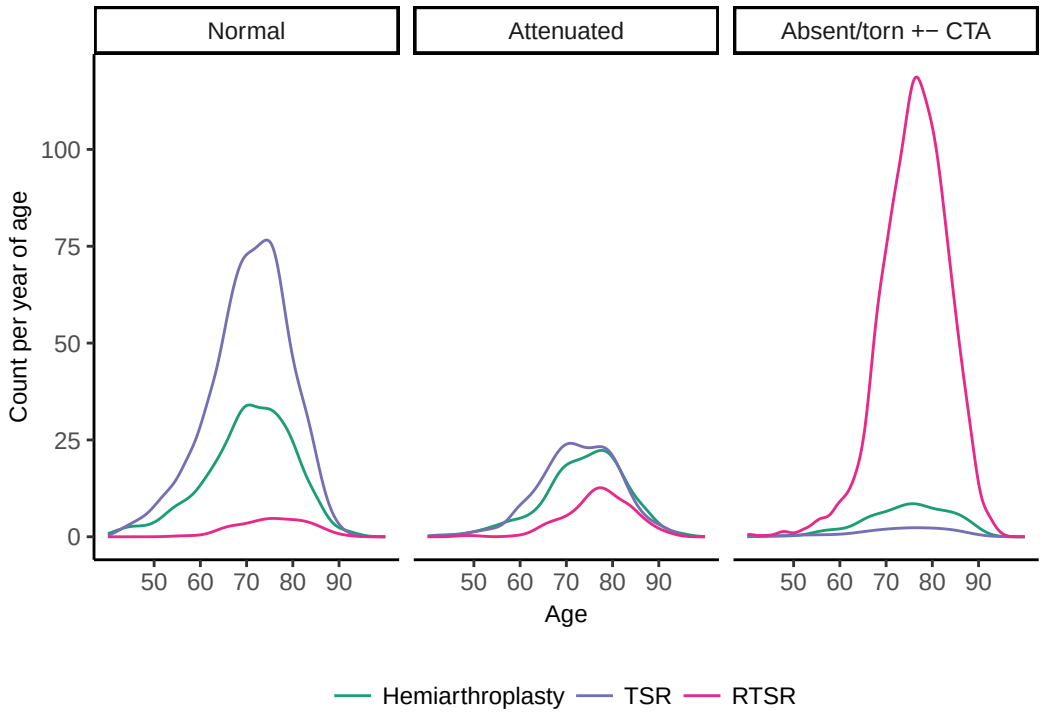


Figure 8.1: Age distributions by state of rotator cuff and type of surgery performed

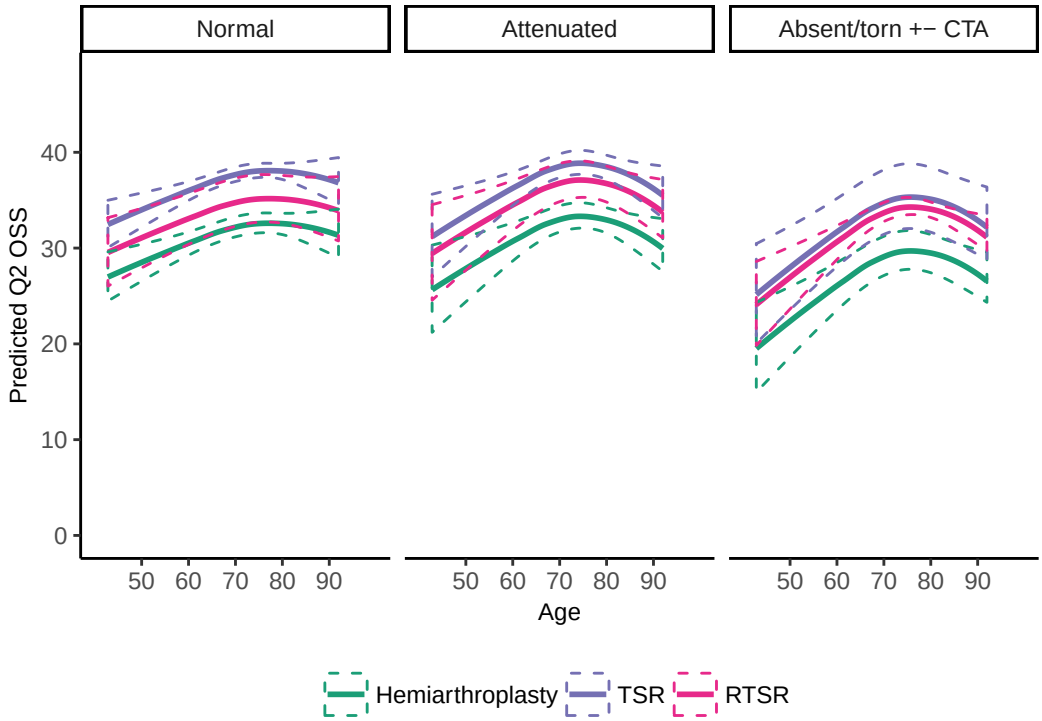


Figure 8.2: Partial effects plots of predicted Q2 OSS by rotator cuff status, age and procedure type. Predictions adjusted to Q1 OSS=16, ASA=2, no previous surgery. 95% CIs shown.

Sensitivity analysis

The adjusted point estimates for differences between procedure types were not changed by splitting the final rotator cuff status group. However, each of the new independent analyses was underpowered and the confidence intervals were very wide. For patients with an absent/torn rotator cuff, the predicted Q2 OSS was 30.7 points after hemiarthroplasty (95% CI 27.18 and 34.31). After RTSR it was 4.8 points higher (95% CI 1.20 to 8.37) and after TSR it was 5.1 points higher (95% CI -0.27 to 10.55). For patients with CTA the corresponding estimates were: hemiarthroplasty 29.2 points (95% CI 26.81 to 31.54), RTSR 5.0 points higher (95% CI 2.69 to 7.32) and TSR 4.9 points higher (-0.61 to 10.40).

8.3.4 Revision risk at two years

A total of 63 patients required revision surgery within two years of their primary procedure. For patients with a normal rotator cuff, the unadjusted revision risk was higher following hemiarthroplasty (2.21%, 95% CI 1.17 to 3.24) compared to TSR (0.75%, 95% CI 0.33 to 1.17). No other differences between procedure types were observed (table 8.2).

Table 8.2: Risk of revision at two years by implant type and rotator cuff status

Status of rotator cuff	Procedure type	Number at start	At risk at 2 years	Censored	Revised	Revision risk at 2 years % (95% CI)
Normal	Hemiarthroplasty	775	731	27	17	2.21 (1.17 to 3.24)
	TSR	1617	1552	53	12	0.75 (0.33 to 1.17)
	RTSR	96	91	5	0	-
Attenuated	Hemiarthroplasty	466	442	18	6	1.30 (0.26 to 2.33)
	TSR	532	493	33	6	1.14 (0.23 to 2.05)
	RTSR	198	176	20	2	1.02 (0 to 2.42)
Absent/torn +- CTA	Hemiarthroplasty	190	174	11	5	2.69 (0.34 to 5.00)
	TSR	58	52	5	1	1.82 (0 to 5.29)
	RTSR	1974	1862	100	14	0.72 (0.34 to 1.09)

8.4 Discussion

8.4.1 Key findings

There was a clinically important difference in the Q2 OSS in favour of TSR over hemiarthroplasty for patients with both normal rotator cuff tendons (5.5 points higher, 95% CI 4.43 to 6.54), and for those with an attenuated rotator cuff (5.5 points higher (95% CI 4.06 to 7.03). Using RTSR in patients with an attenuated rotator cuff may also result in better Q2 OSS outcomes compared to hemiarthroplasty, but the observed benefit was smaller than that for TSR and may not be clinically important.

In the setting of full-thickness tears of the rotator cuff +- CTA, there is a very strong preponderance for the use of RTSRs in nine out of 10 procedures. This is supported by better observed PROMs compared to hemiarthroplasties exceeding the MCID (Q2 OSS 4.6 points higher, 95% CI 2.59 to 6.57). The use of hemiarthroplasties and TSRs was more balanced in the normal rotator cuff group (1/3 vs 2/3). In the attenuated rotator cuff group, one in six patients also underwent RTSR.

Revision surgery was three times more common within two years following hemiarthroplasty than TSR. This estimate was not adjusted and should therefore be interpreted with caution in light of the powerful influence of age on revision risk noted in chapter 5.

8.4.2 Strengths and limitations

This study shares many of the methodological strengths and weaknesses of the prediction modelling studies in chapter 7. The NJR dataset is the largest cohort with pre- and postoperative PROMs for shoulder replacement surgery. The cases are submitted by unselected NHS and independent hospitals and the observed results are therefore widely generalisable to UK practice.

Statistical efficiency was preserved by using multiple imputation to account for missing data,

an approach also known to reduce the risk of bias seen in complete case analysis.[274,376] The risks of misspecification within the imputation model or bias from clustered missing data were discussed in chapter 7.[276]

The gold-standard tool to compare outcomes between different treatment arms is a well-designed RCT. A number of matching methods are commonly used to strengthen causal inferences made from observational studies, including propensity score matching and (coarsened) exact matching. The tuning of these matching processes is a trade-off between the creation of closely matched cohorts with minimal residual confounding and loss of power and generalisability from pruning of unmatched observations. It has been suggested that, in some circumstances, regression analysis may be a weaker method of controlling for confounding.[387] However, simple covariate adjustment with regression analysis allows retention of the data from all patients and can provide more transparent information on the factors influencing outcomes.[388] Treatment effects can be observed across the range of real practice including patients that may typically be excluded from RCTs or by narrow matching criteria. From this study, the uncertainty of treatment effects in younger patients can be quantified and shared with patients. The observations could be strengthened by modelling the outcomes with age as an interaction with treatment type. The sample size required to support valid conclusions from a model with interactions is very much higher.

There are likely to be a number of sources of residual confounding present due to unmeasured factors and treatment selection biases. The study accounts for rotator cuff status, age, preoperative PROMs, prior surgical history, and a coarse measure of comorbidity (ASA grade). Defining rotator cuff status according to macroscopic tendon appearances may overlook important differences in rotator muscle quality (fatty infiltration) and function, which may contribute to subsequent outcomes.[389] The NJR does not collect any data on the morphology and wear pattern of the glenoid. Surgeons commonly use the classification according to Walch et al to describe the glenoid.[390,391] The presence of a B2, B3 or C type glenoid with increased retroversion and/or bone loss may contribute to a decision to offer an RTSR to the patient. It is possible that some of the patients undergoing RTSR in the presence of an attenuated rotator cuff may not have been deemed suitable for a conventional TSR by the surgeon.

The follow-up scores were recorded at a median of 7.4 months (range 5.1 to 12, IQR 6.1 to 9.4). As noted in chapter 7, this may be earlier than when the maximal improvement is achieved, but Simovitch et al have shown that 80% of improvement occurs in the first six months.[377] No differences were seen between the patterns of improvement for different procedure types and so the comparison of outcomes should be fair. A more important limitation than the timing of the early endpoint may be the lack of longer-term follow up to capture deterioration in PROMs over time, rotator cuff degeneration and the need for revision surgery. In a French multicentre study of 704 shoulders undergoing TSR for primary OA, Young et al used Kaplan-Meier analysis to estimate rates of secondary cuff dysfunction at five years (0%), 10 years (16%) and 15 years (56%).[392] This may have particular implications when a hemiarthroplasty or TSR is considered for a patient with an attenuated rotator cuff or partial-thickness tear. However, the functional outcomes of RTSRs may also deteriorate beyond six years and so any decision between hemiarthroplasty, TSR and RTSR based on expected soft-tissue condition and function may not be clear-cut.[119]

8.4.3 Comparison to other studies

The observed benefit of TSR in patients with a normal rotator cuff is similar to pooled estimates from RCTs in chapter 2. Some authors continue to recommend that hemiarthroplasty is a good option in selected patient groups, including high-demand younger manual workers[393]

The NZJR also collects the Q2 OSS at six months post surgery but does not collect baseline scores. From their dataset of 5982 completed questionnaires, unadjusted mean scores at six months following surgery were 39.49 points for TSR, 35.55 points for RTSR, 31.82 for stemmed hemiarthroplasty and 35.65 points for resurfacing hemiarthroplasty.[394] The order of relative performance between implants is the same as that in this current study, but the NZJR annual report encompasses all indications for surgery and, in the absence of stratification or adjustment, no direct comparison can be made. In an independent analysis of patients with CTA from the NZJR, Young et al found that the Q2 OSS was 6.4 points higher following RTSR compared to hemiarthroplasty in 204 patients matched on age, sex, and ASA grade. Similar findings have been mirrored by single-centre series.[298]

Edwards et al assessed the impact on outcomes of partial and full-thickness tears of the supraspinatus tendon in patients treated with TSR and humeral hemiarthroplasty.[389] In 41 patients with partial-thickness tears of the supraspinatus tendon they found a small (3.7 point) but not significant difference in the Constant-Murley score compared to the 472 patients without a tear.

8.4.4 Meaning of the study

Regardless of the condition of the rotator cuff, patient-reported pain and function scores were substantially better after shoulder replacement surgery compared to preoperative scores. Patients with osteoarthritis and intact rotator cuff tendons are likely to report the best early outcomes if treated with a conventional TSR. The trend towards increased use of TSR over hemiarthroplasty shown in chapter 4 (figure 4.5) is supported here by better early postoperative shoulder scores for patients with normal and attenuated rotator cuff tendons. The decision to offer surgery and the type of surgery should also consider the technical feasibility of different treatment options and likely mechanisms of failure, which are not evaluated here.

There is no RCT evidence comparing RTSR to hemiarthroplasty for cuff tear arthropathy. Whilst there may be some circumstances where conventional TSRs or hemiarthroplasties may be appropriate for CTA, they have long-been considered suboptimal for the treatment of most of these patients.[103,395] This study confirms that surgical community has firmly chosen to treat most patients with established CTA using RTSRs. Although the long-term outcomes are not yet clearly understood, it seems unlikely that an RCT in this area would ever be commissioned or recruited to.

Similar benefits of TSRs over hemiarthroplasties were seen in patients with attenuated rotator cuff tendons. RTSRs were associated with smaller but significant benefits over hemiarthroplasties. The average age of patients undergoing RTSRs was slightly higher, suggesting that surgeons are more confident in the results of these implants in older patients. The rate of shoulder replacement surgery in this older group of patients is rapidly expanding and it can be expected that many of the patients considering surgery may have partial-thickness

rotator cuff disease.

8.4.5 Unanswered questions and future research

When asked in the 2015 arthroplasty research priorities survey, BESS members reported a clear preference for TSR in patients older than 55 years with OA and a normal rotator cuff. For older patients with irreparable rotator cuff tears, the preference was for RTSR. For patients with an attenuated rotator cuff or for those younger than 55 years there was greater uncertainty.[300] Of those surveyed, 42% reported willingness to recruit to an RCT comparing TSR to hemiarthroplasty (including resurfacing) in any age group with OA, and 51% reported willingness to recruit to an RCT comparing TSR to RTSR in older patients with OA.

Early functional results from both randomised studies and this current study favour the use of TSR, but still one third of patients underwent hemiarthroplasty. It is perhaps the absence of unbiased longer-term follow-up that leads to persistent uncertainty. Any new RCT to address this question should include long-term PROMs and revision risk to at least five years. Notably, the number of hemiarthroplasties recorded by the NJR has halved in the five years since 2013 and the apparent equipoise expressed in the 2015 BESS survey may no longer exist.

The best type of surgery is not clear for patients with partial-thickness tears or rotator cuff attenuation, particularly in patients aged 65 years and older. The uncertainty between TSR and RTSR expressed by surgeons for this patient group is reinforced by comparable early results between two options. This comparison could be addressed by a well-designed RCT that should plan to report outcomes beyond two years, including pain, function, revision risk, serious adverse events and global perceived success.[143] The nesting of such an RCT within the existing infrastructure of a registry may provide a more cost-effective method of securing higher rates of long-term follow-up and revision reporting.

New implant designs will doubtless lead to a drift in treatment strategies, generating both new research questions and reopening old ones. Pyrocarbon bearings are purported to result in more favourable glenoid wear characteristics and may increase popularity of hemiarthroplasty again for surgeons trying to preserve the native glenoid.[396] Stemless implants and increasingly

modular systems may also impact on patient and surgeon perceptions of the burden of revision surgery and thresholds for such surgery.

As part of its key role in monitoring the safety of devices, future NJR projects need to establish reference standards for PROMs and revision risks within procedure classes, and a framework for testing the equivalence or non-inferiority of new implants against those reference standards.

Chapter 9

Conclusions

The 2015 JLA Priority Setting Partnership outlined the most important uncertainties in the treatment of shoulder problems.[10] The uncertainties relating to shoulder replacement surgery underpinned the key objectives that were set out at the start of this thesis:

1. To produce an updated appraisal of the current state of high-level evidence supporting shoulder replacement surgery practice.
2. To document current patterns and trends of practice, revision rates, adverse events, and patient reported outcomes in the UK using routinely collected national datasets.
3. To develop and validate models predictive of postoperative outcomes following shoulder replacement surgery.

These objectives have been systematically addressed by the studies reported in this thesis. The findings also support a number of the recommendations for practice and research made in the draft guidance on primary joint replacement, including shoulder replacement, released by NICE in October 2019 and due for full publication in March 2020.[397] One of the key recommendations from NICE is that healthcare professionals should support shared decision making including discussion of the potential benefits and risks of the procedure, and the possible need for more surgery in the future. The stated thesis aim of providing better evidence to patients and clinicians on the benefits and risks of shoulder replacement surgery is closely aligned with this principle.

9.1 Summary of findings

9.1.1 Evidence from trials

No evidence was found from RCTs to compare the outcomes of shoulder replacement surgery to any other form of treatment for glenohumeral OA or CTA. In the main comparison, a limited meta-analysis of two trials found low-quality evidence that the level of patient-reported function experienced by patients treated with TSR for OA may be better than those treated with humeral hemiarthroplasty (mean difference 10.57 points, 95% CI 2.11 to 19.02, WOOS Index, 0-100 scale, higher=better). Patients may also have a small but clinically uncertain benefit in their level of pain (mean difference -1.49, 95% CI -2.88 to -0.10, VAS, 0-10 scale, lower=better). Across the 13 different comparisons (20 studies), no other meaningful differences were reliably identified. The trials currently published in the field of shoulder replacement surgery are generally too small, at risk of multiple sources of bias, and too frequently address questions that are unlikely to be of wide importance to patients, surgeons and commissioners of healthcare. The existing trials are barely powered to detect differences in efficacy or effectiveness between different treatment strategies, sample sizes are far too small to reliably detect adverse events, and follow-up is insufficient to report differences in revision risk. Large well-designed trials have the potential to positively change surgical practice, but they are hard to develop and expensive to run and may not be feasible for the investigation of low-frequency secondary endpoints.

9.1.2 Trends in practice

Despite the lack of high-quality evidence to guide the use of shoulder replacement surgery, the standardised incidence of intervention has increased over a 19-year period from $3.89/10^5$ (95% CI 3.24 to 4.52) in 1999 to $13.64/10^5$ (95% CI 12.65 to 14.64) in 2017. The most dramatic increases were observed in the rates of elective surgery performed for patients aged 65 years and older. A major shift occurred since 2010, away from the use of hemiarthroplasty towards the use of TSRs and RTSRs for elective cases, and towards RTSRs for acute trauma. In the elective setting this most likely reflects the increased use of RTSRs in patients with rotator cuff

tendon damage, although HES data does not provide adequate detail on rotator cuff status to check that assumption.

9.1.3 Prediction of patient reported outcomes

The use of shoulder replacement surgery is supported by the PROMs data collected by the NJR. On average, patients undergoing shoulder replacement have poor levels of preoperative function and high levels of pain (median Q1 OSS 16 points). They experience considerable benefit from surgery (median improvement on OSS of 18 points). However, up to 10.5% of patients may achieve no benefit at all from surgery (change score of $\leq$0 points or revision within one year). The prediction modelling studies in chapter 7 aimed to identify this group of patients more clearly in order to support their preoperative decisions. There was marked heterogeneity of outcomes, and the factors recorded within the NJR database had only limited power to explain this variation. The models developed to predict postoperative PROMs and those least likely to benefit from surgery may be useful adjuncts to inform patients regarding the potential benefits of surgery for them, but need to be placed in the context of other sources of evidence and the risks.

Patient-reported outcomes were lowest in patients who were younger, had undergone previous shoulder surgery, had a history of previous trauma, had limiting chronic conditions, and those with lower starting PROMs. Notably, the magnitude of benefit achieved was greater for those with a lower starting PROM. The best PROMs were seen in patients with a normal rotator cuff undergoing TSR for OA and the magnitude of benefit compared to hemiarthroplasty was similar to that reported in RCTs. Patients with a full-thickness tear of the rotator cuff or CTA have worse levels of preoperative pain and function and achieve slightly smaller improvements than those with a normal cuff, but the magnitude of improvement is still large and clinically important (median improvement of 17 points on the OSS). The most important predictors of which patients would have the worst results from shoulder replacement surgery were a history of previous surgery or trauma, and younger age.

9.1.4 Long-term revision risk

Higher long-term risks of revision or reoperation surgery were strongly associated with younger age at the time of primary shoulder replacement surgery. In patients aged less than 60 years, the lifetime risk of needing revision shoulder replacement surgery may be 1 in 4, and this rises to 1 in 3 if other types of further surgery on the same shoulder are considered. Risks of revision surgery were much lower in older patients, but still higher than the rates commonly reported for hip and knee replacement surgery. Patients undergoing surgery for the sequelae of trauma had alarmingly high risks of revision at all time points (10.4% within three years). The likelihood of revision following shoulder replacement for acute trauma was very much lower than elective indications, even in younger patients. This is not explained by the available data. In unadjusted analyses, the risk of revision surgery following hemiarthroplasty was higher than that following TSR but the difference was small (10-year risk 10.9% vs 9.4%). One would expect that revision risk should be falling over time as surgical technology improves, but this was not shown to be the case.

9.1.5 Serious adverse events

Patients undergoing elective shoulder replacement had the same or a lower risk of death within 30 days than age- and sex-matched individuals from the general population. Conversely, patients having surgery for acute trauma had a relative risk of death three times the matched population estimates. Serious systemic adverse events are rarely reported in the shoulder surgery literature and where they have been, the scope is often limited to death and thromboembolic complications, which are rare. Chapter 6 made a broader examination of other types of event including death, pulmonary embolism, myocardial infarction, acute kidney injury, cerebrovascular event, lower respiratory tract infection and urinary tract infection. At 30 days the unadjusted estimates of any of these events was 3.37% in the elective cohort and 9.40% in the acute trauma cohort. Whilst pulmonary emboli, myocardial infarctions and cerebrovascular events were rare, they occurred more frequently than some previous estimates. When death, acute kidney injury, lower respiratory tract infections and urinary tract infections were included,

the risks were much higher than anticipated. The likelihood of experiencing an adverse event was strongly linked to age (particularly those aged 80 years and older) and increasing levels of comorbidity. Trends over time show that both mean age and reported comorbidities of patients undergoing shoulder replacement surgery are increasing over time and, with this trend, the proportion of patients experiencing a serious adverse event is also increasing. With the exception of urinary tract infections, the other adverse events were all strongly associated with and increased likelihood of death within 90 days of surgery.

9.2 Implications for practice

Most patients undergoing shoulder replacement surgery can be reassured that they will benefit from the procedure, and the likelihood of experiencing lasting serious harm from the surgery is low. The commonest group of patients undergoing shoulder replacement surgery are aged 70 to 80, have a diagnosis of primary osteoarthritis or cuff tear arthropathy, have not had previous surgery, and have no limiting chronic conditions. These patients have the highest reported PROMs after surgery, relatively low risks of serious adverse events and a lifetime risk of revision shoulder replacement surgery of less than approximately 1 in 10. Older patients with more comorbidities need to know that their risks of experiencing serious adverse events is much higher. Younger patients, and those with a history of previous trauma or previous surgery should be counselled that the risk that they will not benefit from surgery and/or require revision surgery is considerably higher. The specific and stratified estimates may be used to provide more detailed information regarding individual expected benefits and risks.

Reducing the frequency of common adverse events should be a priority for both local and national quality improvement programmes. These adverse events should be included in the National Clinical Improvement Programme clinician dashboard exercise as key objective markers of patient outcome and quality of care. Based on the National Hip Fracture Database model, there are important opportunities to learn from highly performing units and to optimise the perioperative management of patients undergoing surgery for acute proximal humerus fractures.

9.3 Limitations and future work

9.3.1 Scope for future trials

Whilst the systematic review in chapter 2, observed NJR outcomes in chapter 8, and upcoming NICE guideline all favour the use of TSR over hemiarthroplasty, this is based on PROMs data at a maximum of two years of follow-up. Whether the benefit of TSR is prolonged beyond this is not clear. The NJR is now collecting PROMs data at five years following surgery and future studies should explore these outcomes together with risks of revision surgery. In particular, there should be a detailed analysis of outcomes in younger patients and those with partial-thickness rotator cuff damage/attenuation. In older patients with rotator cuff attenuation there are signs that some surgeons are beginning to offer RTSRs as the primary shoulder replacement procedure. There was no significant difference between the outcomes of RTSR and TSR in this group in the first year, however the later outcomes including deteriorating function and/or revision surgery due to rotator cuff failure have not been quantified. This question should be addressed by a well-designed trial, which is also a research recommendation from the draft NICE guidance. National joint registries have existing infrastructure in place to capture long-term revision risks and subsequent events. Implementation of a future shoulder replacement trial nested within a national registry could be a cost-effective way to improve the reporting of the long-term revision outcomes that have been lacking from previous trials. The effectiveness of RTSRs versus hemiarthroplasties for acute trauma is an important question not directly addressed by this thesis, but this is the subject of an ongoing nationally commissioned trial.[398]

Ideally, shoulder replacement as a whole would be supported by evidence from RCTs against no treatment or non-operative measures. The BESS surgical community has recently proven its ability to delivery high-impact surgical trials including a sham arthroscopy and non-operative arm (CSAW Study).[130] However, it seems unlikely that such a trial would be feasible in shoulder replacement surgery. Compared to the CSAW participants, patients undergoing shoulder replacement surgery recorded in the NJR had a lower starting OSS, higher six-month scores

and larger change scores. It would be challenging to establish equipoise for a study of shoulder replacement versus no treatment amongst patients and surgeons.

9.3.2 Reducing serious adverse events

The most common serious adverse events identified in chapter 6 were lower respiratory tract infections, which occurred within 30 days following 1.44% of procedures performed electively and 3.91% of procedures performed for acute trauma. It is not clear from the available data why patients experience these complications so frequently. In the trauma setting it may be confounded by the presence other injuries, such as rib fractures, and generally impaired mobility following injury. In both settings, the influence of the type of surgery performed, and also the type of anaesthesia performed should be investigated. The draft NICE guidance has made a research recommendation to investigate the clinical- and cost-effectiveness of different anaesthesia strategies, including those using combined or independent general, regional and local anaesthesia methods. Any subsequent studies should assess both the pain-relief profile of the different strategies as well as length of stay and complications. The high frequency of lower respiratory tract infections together with plausible associations with the mode of anaesthesia make it a potentially important secondary study endpoint. Further work needs to be done to reliably identify patients at risk of respiratory complications early, ideally before surgery, and to optimise their care and to reduce the need for additional inpatient care.

9.3.3 Improving data quality

The major study specific limitations of this thesis have been already discussed in detail. The studies in chapters 4 to 8 all relied on routinely collected data. In the case of the NJR this is prospectively maintained and some elements, including procedure type, can be internally validated. The previously cited audit work from the North East Quality Observatory provided reassurance that shoulder replacement procedures were reliably captured by HES data.[261] Revision rates reported by in HES data were higher than those in NJR reports, but comparable to other sources in the literature. In view of this finding and the large proportion of missing

preoperative PROMs data in the NJR, shoulder replacements need to be urgently included in the annual audit process that already occurs for hips and knees. Postoperative PROMs are strongly associated with preoperative PROMs. Outside the setting of a randomised trial, failure to adjust for baseline PROMs may result in invalid conclusions. Strategies are needed to improve the reporting of these scores by hospitals to the NJR.

At the outset of this thesis, data management and cleaning process for the NJR dataset identified important limitations in the existing data structure including inconsistencies between surgeon-defined procedure classes and the implanted prosthetic components. The classification of components within the NJR was redesigned to support the data quality and analyses in this thesis and this work has been shared directly with the NJR database managers and statistical team. The new classification will support improved NJR data quality and will sit behind the system of real-time data entry validation.

9.3.4 Dissemination of findings

This thesis has generated a large amount of information relevant to both healthcare providers and patients considering shoulder replacement surgery. Information needs to be made available to people in a format that they can understand to help them to reach decisions about their care. One of the roles of healthcare professionals is to translate the information from the published literature into appropriate terms for individual patients. Decision aids are tools that may help to facilitate that process and improve patient engagement. The evidence supporting the draft NICE guidance found that decision aids improved shared decision making, but did not identify the optimum type of aid. Further research should engage patients and clinicians to develop an accessible tool to engage and support patients in their decision of whether to undergo shoulder replacement surgery.

Appendix A

Appendices

A.1 Grants, Awards, Publications, Presentations

A.1.1 Related to this thesis

Grants

National Joint Registry/Royal College of Surgeons Research Fellowship (two years) awarded in May 2016.

Awards

1. Lipmann Kessel Award for best podium presentation, British Elbow and Shoulder Society 2019
2. Prize for best poster presentation, International Congress of Arthroplasty Registries 2018

Publications

1. Accepted: Craig RS, Goodier H, Singh JA, Hopewell S, Rees JL. Shoulder replacement surgery for osteoarthritis and rotator cuff tear arthropathy (Review). Cochrane Database

of Systematic Reviews.

2. Craig RS, Lane JCE, Carr AJ, Furniss D, Collins GS, and Rees JL. Serious adverse events and lifetime risk of reoperation after elective shoulder replacement: population based cohort study using hospital episode statistics for England. BMJ 2019;364:l298.

3. Craig RS, Rees JL. Providing better evidence in orthopaedic surgery: RCTs, registries or both? (Opinion piece) BMJ 20 Feb 2019.

4. Craig RS, Goodier H, Singh JA, Hopewell S, Rees JL. Shoulder replacement surgery for osteoarthritis and rotator cuff tear arthropathy (Protocol). Cochrane Database of Systematic Reviews 2017(11) 20 Nov 2017.

Presentations

Podium

1. Craig R et al. Serious adverse events and long-term revision risk following shoulder replacement for acute trauma: a population cohort study of 14 052 patients from hospital episode statistics data. British Elbow and Shoulder Society 2019 (Prize-winning presentation).

2. Craig R et al. The need for pre-operative outcome scores for prediction and case-mix adjustment models in shoulder arthroplasty. British Elbow and Shoulder Society 2019 and International Congress of Arthroplasty Registries 2019.

3. Craig R et al. Application of a new framework for component classification and construct validation within a national shoulder arthroplasty registry. International Congress of Arthroplasty Registries 2019.

4. Craig R et al. The influence of implant type on patient reported outcomes following elective primary shoulder replacement for osteoarthritis and cuff tear arthropathy: an adjusted analysis of procedures from the National Joint Registry. Girdlestone Orthopaedic Society 2019.

5. Craig R et al. Emerging trends, patient characteristics and subsequent outcomes of shoulder arthroplasty for acute trauma: an analysis of 1,524 patients from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man. British Elbow and Shoulder Society 2018.

6. Craig R et al. Patient age and six-month Oxford Shoulder Scores after primary shoulder replacement predict risk of early revision surgery: development of a prediction model using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man. International Congress of Arthroplasty Registries 2018.

7. Craig R et al. Serious adverse events and lifetime risk of reoperation following elective shoulder replacement: a population-based study using Hospital Episode Statistics for England. Duthie Day 2018.

8. Craig R et al. Trends of implant choice and patient characteristics for 21244 primary shoulder replacements from the National Joint Registry. Duthie Day 2017.

Poster

1. Craig R. Implant choice for shoulder osteoarthritis: can matched registry cohorts inform practice? International Congress of Arthroplasty Registries 2019.

2. Craig R et al. Emerging trends, patient characteristics and subsequent outcomes of shoulder arthroplasty for acute trauma: an analysis of 1,524 patients from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man. International Congress of Arthroplasty Registries 2018 (Prize-winning poster).

A.1.2 Additional projects and collaborations

Publications

1. Craig R, Holt T, Rees JL. Acute Rotator Cuff Tears. BMJ 2017 Dec 11;359:j5366.

2. Submitted: Kunetsor S et al Prosthetic joint infection following 649,376 primary total shoulder and elbow replacements: meta-analyses of incidence rates, temporal trends and potential risk factors.

3. Submitted: Alser O et al. Serious adverse events and risk of re-operations after Dupuytren's disease surgery: a longitudinal population-based cohort study using Hospital Episode Statistics for England.

4. Submitted: Lane J et al. Complications and reoperation after carpal tunnel decompression: a nationwide cohort analysis of risk factors in 855 838 surgeries in the NHS in England.

Presentations

Podium

1. Craig R et al. Long-term revision risk following primary total elbow replacement: a nationwide population based cohort study of 8682 cases from Hospital Episode Statistics data. British Elbow and Shoulder Society 2019

2. Lane J et al. Does Revision Surgery signify Disease Persistence or Recurrence? A 19-year Longitudinal Cohort Study of 855,832 Carpal Tunnel Decompression Surgeries in England. British Society for Surgery of the Hand Autumn Scientific Meeting 2018.

3. Lane J et al. What factors influence the need for revision surgery? A 19-year longitudinal cohort study of 855,832 CTDs in England. International Federation of Societies for Surgery of the Hand 2019.

4. Lane J et al. Is Carpal Tunnel Syndrome (CTS) Two Different Diseases? A Joint Analysis of 2 Longitudinal Cohort Studies from USA and England totalling 935,000 Cases. British Society for Surgery of the Hand Autumn Scientific Meeting 2018, American Academy of Orthopaedic Surgery 2019, and International Federation of Societies for Surgery of the Hand 2019.

5. Lane J et al. Low conversion to surgery after intra-articular steroid injection for BTOA: a nineteen-year longitudinal cohort study of over 19,000 injections in England. British Society for Surgery of the Hand Autumn Scientific Meeting 2019.

A.2 Cochrane review additional materials

A.2.1 Hierarchy of outcome measures

Disability/function was measured using shoulder-specific instruments and analysed according to the following hierarchy:

1. Western Ontario Osteoarthritis of the Shoulder Index (WOOS)
2. American Shoulder and Elbow Surgeons Scale (ASES)
3. Oxford Shoulder Score (OSS)
4. Constant-Murley Score
5. Shoulder Pain and Disability Index (SPADI)
6. Disability of the arm, shoulder and hand (DASH) questionnaire

A.2.2 Characteristics of excluded studies

Study ID	Reason for exclusion
Berth 2013	Quasi-randomised study
Ding 2015	Study did not report on outcomes of interest
Edwards 2007	Study did not report on outcomes of interest
Hammond 2013	Non-randomised comparative study
Hendel 2012	Did not report on outcomes of interest
Iannotti 2015	Did not report on outcomes of interest
ISRCTN42881741	This study comparing the conformity of the glenoid and humeral bearing components in conventional TSR was registered with a start date of 2006. No results are available. The inclusion criteria also included some participants with rheumatoid arthritis which does not fit with the criteria for this review.
Kasten 2009	Non-randomised comparative study
Kircher 2009	This study does not appear to be truly randomised; there is no description of any randomised method used to allocate groups. The stated intention was to randomise into two groups of ten patients, however 6 of the 10 patients in the control group had actually crossed over from the intervention group, replaced by 6 new patients in the intervention group. Quote: "To gain the full number of n=10 patients for group 1, the intraoperative navigation was started in n=16 patients. The data of the 6 patients with aborted intraoperative navigation were included in group 2". Although this study was included in the previous version of this review (Singh 2010), the inclusion criteria in the protocol for this updated versions specifically exclude quasi-randomised studies.
Mariotti 2014	Quasi-randomised study
NCT01884077	This study comparing conventional TSR to reverse TSR for the treatment of osteoarthritis (intact cuff) in older adults (aged 70 to 95 years) was terminated due to slow recruitment and poor follow-up.

A.2.3 Characteristics of ongoing studies

Study ID	Study details
NCT01288066	
Study name	A Randomized Multicenter Study Comparing the Effectiveness of Hemi- Versus Total Shoulder Arthroplasty in Patients With a Degenerative Joint Disease
Methods	Randomised controlled trial
Participants	76 patients aged 18 years and older with primary or secondary shoulder arthrosis. Glenoid
Interventions	Standard hemiarthroplasty vs total shoulder replacement
Outcomes	Constant score, Shoulder Pain and Disability Index, duration of surgery, adverse events, revision rate, quality of life (EQ-5D)
Starting date	Sep-11
Contact information	Norbert Suedkamp, Universitätsklinikum Freiburg, Freiburg, Germany
Notes	Estimated study completion date February 2022
NCT01404143	
Study name	Comparison of Two Methods of Subscapularis Management in Shoulder Arthroplasty: Tenotomy Versus Peel: A Multicenter, Randomized Controlled Trial
Methods	Randomised controlled trial
Participants	100 patients aged 18 to 90 years with imaging and intra-operative findings confirming advanced humeral head cartilage loss, and/glenoid cartilage loss. Persistent pain and functional disability for at least 6 months and failure of 6 months of conservative treatment
Interventions	Subscapularis tenotomy vs subscapularis peel for shoulder replacement
Outcomes	Subscapularis strength, WOOS, ASES, Constant score, tendon healing rate on ultrasound
Starting date	Aug-11
Contact information	Peter Lapner, The Ottawa Hospital, Ottawa, Ontario, Canada, K1H 8L6
Notes	Anticipated study completion September 2017. Results not yet publicly available. Further details were requested from investigators.

NCT01587560	
Study name	A Comparison Between a Pyrocarbon and a CoCr Shoulder Resurfacing Implant
Methods	Randomised controlled trial
Participants	80 patients aged 40-75 years with primary or secondary osteoarthritis of the shoulder
Interventions	Pyrocarbon resurfacing shoulder replacement vs cobalt chrome resurfacing shoulder replacement
Outcomes	Fixation to bone measure by Radiostereometric Analysis, EuroQol-5D(EQ-5D), American Shoulder and Elbow Surgeons Score(ASES), Constant score, and Western Ontario Osteoarthritis of the Shoulder index(WOOS)
Starting date	Jan-12
Contact information	Olof Skoldenberg, Director of research, Orthopedic department, Danderyd Hospital, Stockholm, Sweden, 18288
Notes	Estimated study completion date January 2018. Results not yet publicly available.
NCT01697865	
Study name	Reverse Shoulder Arthroplasty With or Without Concomitant Latissimus and Teres Major Transfer for Shoulder Pseudoparalysis With Teres Minor Dysfunction: A Prospective, Randomized Investigation
Methods	Randomised controlled trial
Participants	42 patients with: - Shoulder pseudoparalysis due to chronic rotator cuff dysfunction with or without glenohumeral arthritis. - Chronic rotator cuff tear with severe retraction, atrophy, fatty infiltration. - Active forward elevation of less than 90 degrees - Teres minor dysfunction - Positive lag and hornblower sign - Grade 2 or greater fatty infiltration of the teres minor and infraspinatus seen on MRI - Meet all criteria to have a latissimus and teres major transfer
Interventions	Reverse total shoulder replacement in two arms: 1. Includes a concomitant latissimus and teres major transfer (transfer group), 2. Does not include a concomitant latissimus and teres major transfer (control group).
Outcomes	Activities of Daily Living and External Rotation (ADLER) Score, DASH score, ASES score, SF-12 Score, range of motion, xray measures
Starting date	Oct-12
Contact information	Susan Odum, OrthoCarolina Research Institute, Charlotte, North Carolina, United States, 28207. Susan.Odum@Orthocarolina.com
Notes	Estimated study completion in October 2019
NCT01790113	
Study name	A Prospective, Randomized, Multicenter Study Comparing the Safety and Effectiveness of Arthrex's Eclipse™ Shoulder Prothesis to the Univers™ II Shoulder Prosthesis in Patients With a Degenerative Joint Disease
Methods	Randomised open-label study
Participants	350 adults >21 years with primary or secondary osteoarthritis, avascular necrosis or rheumatoid arthritis
Interventions	Conventional total shoulder replacement using stemless vs stemmed humeral component
Outcomes	Composite score using Constant-Murley score and radiological outcome.
Starting date	Jan-13
Contact information	Melissa Hirschberg, Arthrex, Inc.
Notes	Estimated study completion December 2020. From the clinicaltrials.gov entry, it is ambiguous whether this will be a randomised study due to contradictory statements.
NCT02305966	
Study name	Evaluation of Implant Fixation in Reverse Total Shoulder Arthroplasty
Methods	Randomised controlled trial
Participants	40 patients aged 50 to 85 years with rotator cuff tear arthropathy requiring reverse total shoulder replacement
Interventions	Reverse total shoulder replacement in four arms: 1. Pressfit humerus & non-lateralized glenoid, 2. Pressfit humerus & lateralized glenoid, 3. Cemented humerus & non-lateralized glenoid, 4. Cemented humerus & lateralized glenoid
Outcomes	Component migration measured by Radiostereometric Analysis
Starting date	Aug-17
Contact information	Associate Professor George Athwal, Lawson Health Research Institute, London, Ontario, Canada
Notes	Estimated study completion date November 2019
NCT02768597	
Study name	A Prospective, Randomized Study Comparing The Outcome of Large-Diameter vs Small-Diameter Glenospheres in Primary Reverse Shoulder Arthroplasty Using the ReUnion System
Methods	Randomised controlled trial
Participants	220 participants aged 50-90 requiring a primary reverse total shoulder replacement for a diagnosis of cuff-tear arthropathy (CTA), massive irreparable rotator cuff tear (MRCT) or osteoarthritis (OA) with marked posterior subluxation or bone loss

Interventions	Reverse total shoulder replacement in four arms: 1. Large-Diameter Glenosphere +2 mm offset, 2. Small-Diameter Glenosphere +2 mm offset, 3. Large-Diameter Glenosphere +6 mm offset, 4. Small-Diameter Glenosphere +6 mm offset
Outcomes	Shoulder range of motion, Oxford Shoulder Score, American SHoulder and Elbow Surgeons Scale, Quick-DASH questionnaire
Starting date	Apr-16
Contact information	Mark Morrey, Mayo Clinic, Rochester, Minnesota, United States, 55905
Notes	Estimated study completion date February 2022

NCT02966886	
Study name	Comparison of Techniques in the Management of Glenoid Deficiencies in Shoulder Arthroplasty
Methods	Randomised controlled trial (three parallel arms)
Participants	160 patients aged 18 years and older with shoulder osteoarthritis and B2 type glenoid retroversion.
Interventions	Conventional total shoulder replacement with glenoid technique in three arms: 1. Eccentric reaming, 2. Augented glenoid component 3. Posterior glenoid bone grafting
Outcomes	At 24 months: WOOS Index, Constant-Murley shoulder score, ASES shoulder score, EQ5d-5L, Shoulder Health Utilization Assessment Form.
Starting date	Jan-17
Contact information	Peter Lapner, The Ottawa Hospital, Ontario, Canada. 6137378899 ext 78377, plapner@toh.on.ca
Notes	Estimated study completion date January 2021

NCT03111147	
Study name	The Impact of Humeral Component Version on Outcomes Following Reverse Total Shoulder Arthroplasty: A Prospective Randomized Controlled Trial
Methods	Randomised controlled trial
Participants	70 adults aged 18 years and older undergoing reverse total shoulder replacement for cuff tear arthropathy or osteoarthritis associated with a rotator cuff tear.
Interventions	Reverse total shoulder replacement with the stemmed humeral component implanted in 0 degrees version vs 30 degrees retroversion
Outcomes	Range of motion at 2 years
Starting date	Apr-17
Contact information	Lisa Motowski, Beaumont Helth, Royal Oak, Michigan, United States. 248-551-6679, lisa.stellon@beaumont.org
Notes	Estimated study completion date April 2020

NCT03711175	
Study name	The Influence of Repairing the Sub-scapularis on Outcomes After Reverse Arthroplasty
Methods	Randomised controlled trial
Participants	200 adults aged 21 years and older with severe glenohumeral arthropathy and a grossly deficient rotator cuff
Interventions	Reverse total shoulder replacement (AltiVate system) with repair of subscapularis versus no repair of subscapularis
Outcomes	Physician evaluated strength, range of motion, pain/function using the ASES shoulder score/SST, general health using the VR-12, radiographic assessments, adverse events
Starting date	Sep-18
Contact information	Lisa Holt, Encore Medical, lisa.holt@djoglobal.com
Notes	Estimated study completion date December 2028

NCT03727490	
Study name	The Role of Subscapularis Repair in Reverse Shoulder Arthroplasty
Methods	Randomised controlled trial
Participants	100 adults undergoing reverse total shoulder replacement
Interventions	Reverse total shoulder replacement with bone to bone repair of subscapularis tendon vs no repair of subscapularis tendon
Outcomes	Disability/function measured using shoulder score and Simple Shoulder Test at two years
Starting date	Jan-17
Contact information	Julianne Sefko, 314-747-2496, sefkoj@wudosis.wustl.edu
Notes	Estimated study completion date January 2021

NCT03730597	
Study name	Influenze of Glenosphere Size in Scapular Notch Development. Randomized Study Comparing 42 to 38 ECC
Methods	Randomised controlled trial
Participants	95 patients undergoing reverse shoulder replacement for rotator cuff disorders and acute fractures
Interventions	Reverse total shoulder replacement using 42mm diameter glenosphere vs 38mm diameter glenosphere
Outcomes	Radiological evidence of scapular notching
Starting date	Sep-14
Contact information	Carlos Torrens, Head of Shoulder Unit, Hospital del Mar

Notes	Study completed in September 2016 but only submitted in October 2018. Results not yet available. Inclusion in subsequent versions of this review will depend on availability of subgroup data (excluding trauma cases).

A.2.4 Additional study data

Study ID	Outcome	Measure	Timepoint	Arm 1 Description	n	Outcome	Arm 2 Description	n	Outcome	Notes
Stemmed TSR vs stemmed hemiarthroplasty										
Gartsman 2000	Disability/function	UCLA Shoulder Rating Scale[a]	Intermediate	TSR	27	27.4 ±4.9	HA	24	23.2 ±5.9	p=0.008
Lo 2005	Pain	McGill Pain Questionnaire[a]	Intermediate	TSR	20	0.9 ±1.4	HA	21	2.7 ±6.8	[i]
	Disability/function	ASES Shoulder Score[a]	Intermediate	TSR	20	91.1 ±14.3	HA	21	83.1 ±25.6	p=0.25
		UCLA Shoulder Rating Scale[a]	Intermediate	TSR	20	26.7 ±3.8	HA	21	24.2 ±5.0	p=0.10
		Constant-Murley Score[a]	Intermediate	TSR	20	70.8 ±17.2	HA	21	67.1 ±19.6	p=0.55
Sandow 2013	Pain	Visual analogue scale[b]	Short-term	TSR	16	1 (0-2.8)	HA	13	2 (0-8.8)	p<0.05
			Intermediate	TSR	11	0.2 (0-4)	HA	7	4.6 (0.4-8.5)	p<0.05
	Disability/function	UCLA Shoulder Rating Scale[b]	Short-term	TSR	18	30 (21-35)	HA	13	29 (12-33)	p<0.05
			Intermediate	TSR	11	33 (24-34)	HA	6	18.5 (10-25)	p<0.05
		Constant-Murley Score[b]	Short-term	TSR	15	68 (48-89)	HA	10	59.5 (30-85)	
			Intermediate	TSR	6	77 (67-95)	HA	4	54.5 (43-59)	p<0.05
Humeral head resurfacing vs stemmed hemiarthroplasty										
Rasmussen 2015	Pain	Subdomain of Constant-Murley[c]	Intermediate	Resurfacing	19	11.1 (0-15)	HA	19	8.0 (0-15)	p=0.04
	Disability/function	Constant-Murley Score[c]	Short-term	Resurfacing	19	48.9 (6-80)	HA	19	59.1 (0-88)	p=0.14
	Operating time (Physician-evaluated)	Minutes[c]	Short-term	Resurfacing	20	52 (34-80)	HA	20	80 (56-103)	p<0.001
Stemless TSR vs stemmed TSR										
Uschok 2017	Pain	Subdomain of Constant-Murley[a]	Intermediate	Stemless	15	10.9 ±4.4	Stemmed	18	13.6 ±2.9	p=0.136
			Long-term	Stemless	14	12.7 ±2.4	Stemmed	15	12.4 ±2.1	p=0.590
TSR using cemented polyethylene glenoid component vs uncemented metal-backed glenoid component										
Boileau 2002	Pain	Subdomain of Constant-Murley[d]	Intermediate	Cemented	20	12.5 (4-15)	Uncemented	20	12 (5-15)	
			Long-term	Cemented	17	12 (0-15)	Uncemented	18	13 (3-15)	
	Disability/function	Constant-Murley Score[d]	Intermediate	Cemented	20	67 (6-89)	Uncemented	20	75 (17-89)	
			Long-term	Cemented	17	68 (6-92	Uncemented	18	73 (42-89)	
	Glenoid lucency (Physician-evaluated)	Yes/no	Long-term	Cemented	20	17	Uncemented	20	5	[j]
TSR using pegged glenoid component vs keeled glenoid component										
Nuttall 2007	Pain	Visual analogue scale[e]	Intermediated	Pegged	10	0.6	Keeled	10	0.6	
	Disability/function	ASES Shoulder Score[e]	Intermediate	Pegged	10	78	Keeled	10	84	
		Constant-Murley Score[e]	Intermediate	Pegged	10	65	Keeled	10	62	
Edwards 2010	Disability/function	ASES Shoulder Score[e]	Long-term	Pegged	16	68	Keeled	22	67	p=0.635.[k]
		Constant-Murley Score[e]	Long-term	Pegged	16	59.7	Keeled	22	58.9	p=0.728[k]
		SANE[e]	Long-term	Pegged	16	58.7	Keeled	22	66.6	p=0.247[k]
Rahme 2009	Disability/function	Constant-Murley Score[e]	Intermediate	Pegged	14	70	Keeled	12	70	
		Subjective Shoulder Value[e]	Intermediate	Pegged	14	80	Keeled	12	80	
Gascoyne 2017	Disability/function	WOOS Index[e]	Short-term	Pegged	5	7.15	Keeled	6	34.7	*l*

			Intermediate	Pegged	4	22.3	Keeled	5	18.5	
		ASES Shoulder Score[f]	Short-term	Pegged	5	97.1	Keeled	6	72.5	
			Intermediate	Pegged	4	96.4	Keeled	5	73.5	
		Simple Shoulder Test[f]	Short-term	Pegged	5	11	Keeled	6	7	
			Intermediate	Pegged	4	10.5	Keeled	5	6	
	Radiostereometric analysis	Coronal plane translation[f]	Short-term	Pegged	5	0.267	Keeled	6	1.518	p<0.05
	(Physician-evaluated)		Intermediate	Pegged	4	0.235	Keeled	5	0.99	p<0.05
		Coronal plane rotation[f]	Short-term	Pegged	5	0.601	Keeled	6	0.307	
			Intermediate	Pegged	4	1.074	Keeled	5	-0.624	
Conventional TSR using cemented stemmed humeral component vs uncemented stemmed humeral component										
Litchfield 2011	Disability/function	ASES Shoulder Score[a]	Short-term	Cemented	78	70.2 ±10.3	Uncemented	74	66.2 ±13.9	p=0.09
			Intermediate	Cemented	78	69.2 ±13.3	Uncemented	74	64.74 ±15.7	p=0.2
		MACTAR score[a]	Short-term	Cemented	78	50.6 ±59.1	Uncemented	74	70.1 ±74.1	p=0.19
			Intermediate	Cemented	78	56.1 ±76.6	Uncemented	74	69.2 ±77.7	p=0.49
Conventional TSR with approach via lesser tuberosity osteotomy vs subscapularis tenotomy/peel										
Lapner 2012	Disability/function	ASES Shoulder Score[a]	Short-term	Osteotomy	36	77.1 ±23.7	Peel	37	81.3 ±18.7	
			Intermediate	Osteotomy	36	79.4 ±24.6	Peel	37	83.3 ±19.0	
	Fatty infiltration of cuff (Physician-evaluated)	Goutallier Grade [a]	Short-term	Osteotomy	41	0.90 ±0.89	Peel	41	0.95 ±0.85	
Levine 2019	Pain	Visual analogue scale[e]	Short-term	Osteotomy	29	1.8	Tenotomy	30	1.9	[m]
	Disability/function	ASES Shoulder Score[e]	Short-term	Osteotomy	29	75.6	Tenotomy	30	74.6	
		Simple Shoulder Test[e]	Short-term	Osteotomy	29	9.1	Tenotomy	30	7.6	
	Quality of life	Short Form-36[e]	Short-term	Osteotomy	29	71.1	Tenotomy	30	64.9	
	Operative time	Minutes[e]	Short-term	Osteotomy	29	152.7	Tenotomy	30	129.3	
Humeral head resurfacing using Copeland implant vs Global C.A.P. implant										
Mechlenburg 2014	Disability/function	WOOS Index[g]	Short-term	Copeland	10	298 (81-788)	Global C.A.P.	15	383 (115-822)	
			Intermediate	Copeland	10	128 (53-550)	Global C.A.P.	15	294 (111-477)	
		Constant-Murley Score[h]	Short-term	Copeland	10	71.6 (59.6-87.7)	Global C.A.P.	15	72.7 (58.8-88.2)	
			Intermediate	Copeland	10	76.9 (61.1-81.2)	Global C.A.P.	15	72.6 (64.6-85.7)	
	BMD of humeral head (Physician-evaluated)	Measured in g/cm^3 [h]	Short-term	Copeland	9	0.81 (0.62-0.97)	Global C.A.P.	15	0.83 (0.60-1.04)	
			Intermediate	Copeland	9	0.59 (0.50-0.65)	Global C.A.P.	15	0.57 (0.47-0.73)	
Reverse polarity TSR with neutral glenosphere vs inferior tilted glenosphere										
Edwards 2012	Disability/function	Constant-Murley Score[a]	Intermediate	Neutral	22	71.4 ±14.9	Tilted	20	63.6 ±12.3	p=0.136[n]
		Adjusted Constant-Murley Score[a]	Intermediate	Neutral	22	92.6 ±18.9	Tilted	20	87.7 ±23.6	p=0.129[n]
Reverse polarity TSR using eccentric glenosphere position vs concentric position										
Poon 2014	Disability/function	Oxford Shoulder Score[a]	Intermediate	Eccentric	23	35 ±10.5	Concentric	27	38 ±10.5	p=0.32
Reverse polarity TSR using Bony Increased Offset vs Standard Offset for glenoid component										
Greiner 2015	Pain	Subdomain of Constant-Murley[a]	Intermediate	BIO	16	12.7 ±2.8	STD	15	12.7 ±3.2	
	Disability/function	Adj. Constant-Murley Score[a]	Intermediate	BIO	16	83.3 ±23.4	STD	15	89.4 ±20.8	
		ADLER score[a]	Intermediate	BIO	16	25.7 ±6.9	STD	15	26.1 ±5.0	
		DASH score[a]	Intermediate	BIO	16	40.9 ±23.7	STD	15	34.2 ±20.2	
Reverse polarity TSR using 135 degree humeral neck-shaft angle vs 155 degree humeral neck-shaft angle										

Gobezie 2019	Disability/function	SANE score[a]	Intermediate	135 degree	37	74 ±24.4	155 degree	31	76 ±16.8
		Simple Shoulder Test[a]	Intermediate	135 degree	37	8 ±3.0	155 degree	31	7 ±2.2

[a]Mean+-SD
[b]Median (range)
[c]Mean (range)
[d]"Average" (range)
[e]Mean
[f]Median
[g]Raw score (0-1900 scale, lower=better). Reported as median (IQR)
[h]Median (IQR)
[i]Scale used is unclear. Original questionnaire reference uses 0-78. This is incompatible with the values in this study.
[j]Progression over time only observed in 4 cases in the uncemented group. None in the cemented group.
[k]WOOS Index reported in main analyses
[l]Authors have used WOOS Index in opposite direction to the usual convention
[m]Inconsistencies between text and figures. Numbers reported here are from the text.
[n]ASES Score reported in main analyses
ADLER = Activities of Daily Living Requiring External Rotation
BMD = Bone mineral density

A.2.5 Search strategy

The following strategies were used for MEDLINE and EMBASE. Search terms for other databases were designed to match these will be published in full in the Cochrane Library.

MEDLINE (Ovid)	EMBASE (Ovid)
1 randomized controlled trial.pt.	1 random$.ti.ab.
2 controlled clinical trial.pt.	2 factorial$.ti.ab.
3 randomized.ab.	3 (crossover$ or cross over$ or cross-over$).ti.ab.
4 placebo.ab.	4 placebo$.ti.ab.
5 drug therapy.fs.	5 (doubl$ adj blind$).ti.ab.
6 randomly.ab.	6 (singl$ adj blind$).ti.ab.
7 trial.ab.	7 assign$.ti.ab.
8 groups.ab.	8 allocat$.ti.ab.
9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8	9 volunteer$.ti.ab.
10 exp animals/ not humans.sh.	10 crossover procedure.sh.
11 9 not 10	11 double blind procedure.sh.
12 exp osteoarthritis/	12 randomized controlled trial.sh.
13 osteoarthr$.tw.	13 single blind procedure.sh.
14 (degenerative adj2 arthritis).tw.	14 or/1-13
15 arthrosis.tw.	15 exp osteoarthritis/
16 arthropat$.tw.	16 osteoarthr$.tw.
17 rotator cuff arthro$.tw.	17 (degenerative adj2 arthritis).tw.
18 or/12-17	18 arthrosis.tw.
19 Shoulder/ 19	arthropat$.tw.
20 shoulder joint/	20 rotator cuff arthro$.tw.
21 shoulder pain/	21 or/15-20
22 shoulder$.tw.	22 Shoulder/
23 or/19-22	23 Shoulder Pain/
24 exp surgical Procedures Operative/	24 shoulder$.tw.
25 su.fs.	25 or/22-24
26 (surgery$ or surgeries or surgical or operat$).tw.	26 exp Surgery/
27 (arthroplast$ or hemiarthroplast$ or (joint$ adj2 replace$)).tw.	27 su.fs.
28 (surface$ adj replace$).tw.	28 (surgery$ or surgeries or surgical or operat$).tw.
29 resurfac$.tw.	29 RTSA.tw.
30 RTSA.tw.	30 (arthroplast$ or hemiarthroplast$ or (joint$ adj2 replace$)).tw.
31 glenoid.tw.	31 (surface$ adj replace$).tw.
32 glenosphere.tw.	32 resurfac$.tw.
33 exp "Prostheses and Implants"/	33 glenoid.tw.
34 (glenoid adj2 component).tw.	34 glenosphere.tw.
35 (humor$ adj2 component).tw.	35 (glenoid adj2 component).tw.
36 endopro$.tw.	36 (humor$ adj2 component).tw.
37 reverse.tw.	37 endopro$.tw.
38 or/24-37	38 reverse.tw.
39 and/11,18,23	38 39 or/26-38
40 14 and 21	and 25 and 39

A.3.1 Approvals: HES studies sponsorship letter





Clinical Trials and
Research Governance

RESEARCH SERVICES
Clinical Trials and Research Governance
Joint Research Office
Block 60
Churchill Hospital
Headington
Oxford
OX3 7LE

To whom it may concern

Date: 03/11/2017

Dear Sir/Madam,

Title: The role of patient factors, surgical factors and hospital factors upon patient outcomes and NHS costs in the treatment of upper limb musculoskeletal conditions: spatial and longitudinal analysis of routine data

PID: 12787

Protocol Date/Version: 04.10.2017 Version 1.0

The above study has been designed by Prof Dominic Furniss and colleagues at the University of Oxford and funded by the Wellcome Trust. I confirm that the University will accept the role of Research Sponsor of this study.

Insurance-provided indemnity arrangements are in place for the project:
Newline Underwriting Management Ltd., at Lloyd's of London.

Sponsorship is confirmed subject to the condition that the following are sent to Clinical Trials and Research Governance for review prior to submission to the Research Ethics Committee. Failure to do so may compromise insurance cover for the project.

- Any minor or substantial amendments
- Any extension to the study end date

In addition, annual progress reports, the end of study notification and final report must be copied to Clinical Trials and Research Governance.

Lastly, where they exist, University Core SOPs must be followed by default for all clinical research sponsored by the University. Further information can be found here:
https://www.admin.ox.ac.uk/researchsupport/ctrg/resources/

Any communications relating to Research Sponsorship should be directed to the undersigned, whose contact details are given in this letter.

Yours faithfully

Dr Karen Melham
Clinical Trials and Research Governance

Tel: +44 (0)1865 572221 · Fax: +44 (0)1865 572228 · Web: www.admin.ox.ac.uk/researchsupport/ctrg/
Email: Heather.house@admin.ox.ac.uk or Elaine.Chick@admin.ox.ac.uk

A.3.2 Approvals: HES studies ethics waiver





RESEARCH SERVICES
Clinical Trials and Research Governance
Joint Research Office
Boundary Brook House
Churchill Drive
Headington
Oxford
OX3 7LQ

To whom it may concern 15 June 2018

Title: The role of patient factors, surgical factors and hospital factors upon patient outcomes and NHS costs in the treatment of upper limb musculoskeletal conditions: spatial and longitudinal analysis of routine data

Date/Version: 04.10.2017 Version 1.0

Internal Reference PID: 12787

Further to the letter of 19 April 2017 confirming that the University of Oxford is Research Sponsor of this study and will comply with the requirements of the Department of Health *UK Policy Framework for Health and Social Care Research*, I now confirm that because this research involves analysis of retrospective anonymised data only, it does not require Research Ethics Review.

This is as per section 2.3.10 of the Department of Health *Governance Arrangements for Research Ethics Committees* (GAfREC).

Should you require further information, please do not hesitate to contact me.

Sincerely,



Dr Karen Melham
Senior Clinical Research Support Manager
Clinical Trials and Research Governance
University of Oxford

Tel: +44 (0)1865 616480 · Web: www.admin.ox.ac.uk/researchsupport/ctrg/

A.3.3 Approvals: NJR studies sponsorship letter





RESEARCH SERVICES
Clinical Trials and Research Governance
Joint Research Office
Block 60
Churchill Hospital
Headington
Oxford
OX3 7LE

To whom it may concern

Date: 19 April 2017

Dear Sir/Madam,

Title: Improving Outcomes of Shoulder Replacement Surgery

PID: 12730

Protocol Date/Version: 23 March 2017/ v.1

The above study has been designed by Professor Jonathan Rees and colleagues at the University of Oxford and funded by National Joint Registry and Royal College of Surgeons. I confirm that the University will accept the role of Research Sponsor of this Study and will comply with the requirements of the Department of Health Research Governance Framework for Health and Social Care 2005, in so far as these apply in the United Kingdom.

Insurance-provided indemnity arrangements are in place for the project:
Royal & Sun Alliance plc

Sponsorship is confirmed subject to the condition that the following are sent to Clinical Trials and Research Governance for review prior to submission to the National Joint Register or other bodies. Failure to do so may compromise insurance cover for the project.

- Any amendment
- Any extension to the study end date

Any communications relating to Research Sponsorship should be directed to the undersigned, whose contact details are given in this letter.

Yours Faithfully

E. Chick

Elaine Chick
Deputy Head, Clinical Trials and Research Governance

Tel: +44 (0)1865 572221 · Fax: +44 (0)1865 572228 · Web: www.admin.ox.ac.uk/researchsupport/ctrg/
Email: Heather.house@admin.ox.ac.uk or Elaine.Chick@admin.ox.ac.uk

A.3.4 Approvals: NJR studies ethics waiver





RESEARCH SERVICES
Clinical Trials and Research Governance
Joint Research Office
Boundary Brook House
Churchill Drive
Headington
Oxford
OX3 7LQ

To whom it may concern 15 June 2018

Protocol Title: Improving Outcomes of Shoulder Replacement Surgery

Date/Version: 23 March 2017/ v.1

Internal Reference: 12730

Further to the letter of 19 April 2017 confirming that the University of Oxford is Research Sponsor of this study and will comply with the requirements of the Department of Health *UK Policy Framework for Health and Social Care Research*, I now confirm that because this research involves analysis of retrospective anonymised data only, it does not require Research Ethics Review.

This is as per section 2.3.10 of the Department of Health *Governance Arrangements for Research Ethics Committees* (GAfREC).

Should you require further information, please do not hesitate to contact me.

Sincerely,



Dr Karen Melham
Senior Clinical Research Support Manager
Clinical Trials and Research Governance
University of Oxford

Tel: +44 (0)1865 616480 · Web: www.admin.ox.ac.uk/researchsupport/ctrg/

A.3.5 Acknowledgement and disclaimer statement

I would like to thank the patients and staff of all the hospitals in England, Wales and Northern Ireland who have contributed data to the National Joint Registry and to Hospital Episode Statistics. I am grateful to the Healthcare Quality Improvement Partnership (HQIP), the NJR Research Sub-committee and staff at the NJR Centre for facilitating this work. I have conformed to the NJR and NHS Digital's standard protocols for data access and publication. The views expressed are my own and do not necessarily reflect those of the NHS, NIHR, Department of Health, NJR steering committee, RCS, or Healthcare Quality Improvement Partnership, who do not vouch for how the information is presented. The study sponsors had no role in the design, conduct or reporting of the studies.

A.4 HES database code lists

Table A.4: List of OPCS-4 codes used to identify shoulder replacements in HES database

Primary	Anatomy code	Type	Revision	Anatomy code	Uncertain	Anatomy code
W494		HA	O060		W492	
W504		HA	O063		W498	
W515		HA	O073		W499	
W581	Z691	HA	W430	Z691	W502	
W581	Z813	HA	W433	Z685	W508	
W581	Z814	HA	W433	Z691	W509	
W491		HA	W454	Z814	W512	
W501		HA	W964		W518	
W511		HA	W974		W519	
O061		TSR	W985		O062	
O071		TSR	O070		O068	
O081		TSR	W430	Z685	O069	
W431	Z685	TSR	W433	Z814	O072	
W431	Z691	TSR	W963		O078	
W431	Z814	TSR	O080		O079	
W441	Z814	TSR	O083		O082	
W451	Z814	TSR	O084		O088	
W961		TSR	W430	Z814	O089	
W971		TSR	W440	Z814	W432	Z685
W981		TSR	W443	Z814	W432	Z691
W965		RTSR	W450	Z814	W432	Z814
W975		RTSR	W453	Z814	W438	Z685
W986		RTSR	W490		W438	Z691
W051	Z691	EPR	W493		W439	Z685
W051	Z698	EPR	W500		W439	Z691
W051	Z699	EPR	W503		W439	Z814
W051	Z813	EPR	W510		W442	Z814
W051	Z814	EPR	W513		W448	Z814
W052	Z691	EPR	W514		W449	Z814
W052	Z698	EPR	W580	Z691	W452	Z814
W052	Z699	EPR	W580	Z813	W962	
W052	Z813	EPR	W580	Z814	W969	
W052	Z814	EPR	W582	Z691	W972	
W053	Z691	EPR	W582	Z813	W978	
W053	Z698	EPR	W582	Z814	W979	
W053	Z699	EPR	W960		W982	
W053	Z813	EPR	W966		W458	Z814
W053	Z814	EPR	W970		W459	Z814
			W973		W988	
			W976		W989	
			W980		W438	Z814
			W983		W968	
			W984			
			W987			
			W054	Z691		
			W054	Z698		
			W054	Z699		
			W054	Z813		
			W054	Z814		
			W055	Z691		
			W055	Z698		
			W055	Z699		
			W055	Z813		
			W055	Z814		

Table A.5: List of OPCS-4 codes used to identify reoperations in HES database

Procedure type	**OPCS-4 codes**	**Anatomy codes**
SAD/ACJ excision	O291	
	W572, W844	Z812, Z814, Z891
	T621, T622, T626, T628, T629	Z814, Z891
Cuff repair	T791, T793, T794, T795	
	T641, T642, T643, T744, T648, T649, T67, T68	Z742
Relocation	W652, W658, W689, W662, W668	Z813, Z814, Z891
	W669, W672, W674, W678, W679	Z813, Z814, Z891
Stabilisation	O27	
	W77 (Not W776), W841, W842	
	W72, W73, W74, W75	Z813, Z814, Z891
MUA+-release	W911,W913, W918, W919, W781, W784	Z813, Z814, Z891
	Y421, Y428, Y429	Z813, Z814, Z891
Fracture fixation	W19 (not W191), W20, W21, W22	Z691, Z692, Z693, Z813, Z814, Z891
	W23, W24 (not W241), W25, W26	Z691, Z692, Z693, Z813, Z814, Z891
	W651, W653, W654, W656, W661, W663, W664	Z691, Z692, Z693, Z813, Z814, Z891
	W671, W673, W677, W332	Z691, Z692, Z693, Z813, Z814, Z891
	O172, O173, O175, O178, O179	Z691, Z692, Z693, Z813, Z814, Z891
Arthrodesis	W601, W602, W603, W608, W609	Z813, Z814, Z891
	W611, W612, W613, W618, W619	Z813, Z814, Z891
	W621, W622, W628, W629	Z813, Z814, Z891
	W631, W632, W638, W639	Z813, Z814, Z891
	W641, W642, W648, W649	Z813, Z814, Z891
Excision arthroplasty	W572, W573, W574, W575, W579, Y037	Z813, Z814, Z891
Synovectomy	W691, W692, W693, T711	Z813, Z814, Z891
Debridement/washout	W80, W811, W812, W813, W815, W713	Z813, Z814, Z891
	Y223, Y311, Y318, Y319, Y321	Z813, Z814, Z891
	W18	Z691, Z692, Z693, Z694
Bone resection	W068, W069, W091, W092, W093, W094	Z691, Z692, Z693
	W095, W096, W097, W098, W099	Z691, Z692, Z693
Amputation	X071, X072	
Complex reconstruction	O108, O109	
Other not-specified	Y528, Y767, W816, W817, W818, W819, W83	Z812, Z813, Z814, Z891
	W843, W845, W846, W847, W848, W868, W869	Z812, Z813, Z814, Z891
	W881, W888, W889, W891, W898, W899, O198, O199	Z812, Z813, Z814, Z891
	W714, W718, W719, W694, W695, W698, W699	Z812, Z813, Z814, Z891
	W711, W712, T645, T651, T658, T659, T701, T702	Z812, Z813, Z814, Z891
	W562, W563, W564, W568, W569	Z812, Z813, Z814, Z891

Table A.6: List of ICD-10 codes used to identify serious adverse events in HES database

Event type	ICD-10 codes
Pulmonary embolism	I26
Myocardial infarction	I21, I22
Cerebrovascular event	I60, I61, I62, I63, I64
Acute kidney injury	N17
Lower respiratory tract infection	J12, J13, J14, J15, J16, J18, J22, J86, J440, J851, J690
Urinary tract infection	N10, N300, N308, N309, N390

A.5.1 NJR MDS form, version 6: Primary procedures

National Joint Registry www.njrcentre.org.uk Working for patients, driving forward quality	MDS VERSION 6.0 Shoulder Operation — Form: MDSv6.0 S1 v005
S1 Shoulder Primary	Patient Addressograph
Important: Please tick relevant boxes. All component stickers should be affixed to the accompanying 'Minimum Dataset Form Component Labels Sheet'. Please ensure that all sheets are stapled together.	

All fields are Mandatory unless otherwise indicated

REMEMBER! MAKE A NOTE OF THE NJR REFERENCE NUMBER WHEN YOU ENTER THIS DATA	NJR REF:

PATIENT DETAILS			
NJR Patient Consent Obtained	Yes ☐	No ☐	Not Recorded ☐
Patient Hospital ID			
Handedness	Left ☐	Right ☐	Ambidextrous ☐

PATIENT IDENTIFIERS				
Forename				
Surname				
Gender	Male ☐	Female ☐	Not Known ☐	Not Specified ☐
Date of Birth	DD/MM/YYYY			
Patient Postcode			Overseas Address ☐	
NHS or National Patient Number (if available)				

OPERATION DETAILS					
Hospital					
Operation Date	DD/MM/YYYY				
Anaesthetic Types	General ☐ Regional – Epidural ☐		Regional – Nerve Block ☐		
Patient ASA Grade	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
Operation Funding	NHS ☐	Independent ☐			

SURGEON DETAILS					
Consultant in Charge					
Operating Surgeon					
Operating Surgeon Grade	Consultant ☐	SpR/ST3-8 ☐	F1-ST2 ☐	Specialty Doctor/SAS ☐	Other ☐
First Assistant Grade	Consultant ☐	Other ☐			

SHOULDER PRIMARY PROCEDURE DETAILS						
Side	Left ☐ Right ☐					
Indications for Implantation (select all that apply)	Osteoarthritis	☐	Avascular Necrosis	☐		
	Cuff Tear Arthropathy	☐	Acute Trauma	☐		
	Inflammatory Arthropathy	☐	Trauma Sequelae	☐		
	Metastatic Cancer/Malignancy	☐	Other	☐		
Previous surgery (not arthroplasty) (Select all that apply)	None	☐	Excision ACJ	☐		
	Stabilisation	☐	Subacromial Decompression	☐		
	Cuff Repair	☐	Other	☐		

SURGICAL APPROACH						
Patient Procedure	Primary Total Prosthetic Replacement	☐				
	Primary Hemi-arthroplasty of Joint	☐				
	Primary Resurfacing Arthroplasty of Joint	☐				
	Primary Resurfacing Hemi-arthroplasty of Joint	☐				
	Primary Reverse Polarity Total Prosthetic Replacement	☐				
Fixation Type	Cementless	☐	Cemented	☐	Hybrid	☐
Approach	Delto-Pectoral	☐	Deltoid Detachment	☐		
	Superior (MacKenzie)	☐	Posterior	☐		
	Deltoid Split	☐				
Minimally Invasive Technique Used?	Yes ☐	No ☐				
Computer Guided Surgery Used?	Yes ☐	No ☐				
Biological Resurfacing (Glenoid) (select all that apply)	None	☐	Reaming	☐		
	Microfracture	☐	Interposition	☐		
THROMBOPROPHYLAXIS REGIME (intention to treat)						
Chemical (In Hospital)	Aspirin	☐	Direct Thrombin Inhibitor (eg Dabigatran)	☐		
	LMWH	☐				
	Pentasaccharide (eg Fondaparinux)	☐	Factor Xa Inhibitor (eg Rivaroxaban/Apixaban)	☐		
	Warfarin	☐	Other	☐		
			None	☐		
Mechanical	Foot Pump	☐	Other	☐		
	Intermittent Calf Compression	☐	None	☐		
	TED Stockings	☐				
BONE GRAFT USED						
Humeral Bone Graft	Yes ☐	No ☐				
Glenoid Bone Graft	Yes ☐	No ☐				
SOFT TISSUES						
Long Head Biceps Tenotomy	Yes ☐	No ☐				
Rotator Cuff Condition	Normal ☐	Attenuated ☐	Absent/Torn ☐		Repaired	☐
SURGEON'S NOTES						

INTRA OPERATIVE EVENT						
Untoward Intra Operative Event	None	☐	Fracture Glenoid	☐	Other	☐
	Shaft Penetration	☐	Nerve Injury	☐		
	Fracture Humerus	☐	Vascular Injury	☐		

PRE-OPERATIVE OXFORD SCORES – Tick one box for every question. If no scores available select Pre-operative Oxford Scores Not available					
Pre-operative Oxford Scores Not available					Not available ☐
Pre-operative Oxford Score Date	DD/MM/YYYY				Not available ☐
1.	During the past 4 weeks… How would you describe the **worst** pain you had from your shoulder?				Not available ☐
	None ☐	Mild ☐	Moderate ☐	Severe ☐	Unbearable ☐
2.	During the past 4 weeks… Have you had any trouble dressing yourself because of your shoulder?				Not available ☐
	No trouble at all ☐	A little bit of trouble ☐	Moderate trouble ☐	Extreme difficulty ☐	Impossible to do ☐
3.	During the past 4 weeks… Have you had any trouble getting in and out of a car or using public transport because of your shoulder?				Not available ☐
	No trouble at all ☐	A little bit of trouble ☐	Moderate trouble ☐	Extreme difficulty ☐	Impossible to do ☐
4.	During the past 4 weeks… Have you been able to use a knife and fork at the same time?				Not available ☐
	Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐
5.	During the past 4 weeks… Could you do the household shopping on your own?				Not available ☐
	Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐
6.	During the past 4 weeks… Could you carry a tray containing a plate of food across a room?				Not available ☐
	Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐
7.	During the past 4 weeks… Could you brush/comb your hair with the affected arm?				Not available ☐
	Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐
8.	During the past 4 weeks… How would you describe the pain you usually had from your shoulder?				Not available ☐
	None ☐	Very mild ☐	Mild ☐	Moderate ☐	Severe ☐
9.	During the past 4 weeks… Could you hang your clothes up in a wardrobe, using the affected arm?				Not available ☐
	Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With great difficulty ☐	No, impossible ☐
10.	During the past 4 weeks… Have you been able to wash and dry yourself under both arms?				Not available ☐
	Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐
11.	During the past 4 weeks… How much has pain from your shoulder interfered with your usual work (including housework)?				Not available ☐
	Not at all ☐	A little bit ☐	Moderately ☐	Greatly ☐	Totally ☐
12.	During the past 4 weeks… Have you been troubled by pain from your shoulder in bed at night?				Not available ☐
	No nights ☐	Only 1 or 2 nights ☐	Some nights ☐	Most nights ☐	Every night ☐

Minimum Dataset Form - COMPONENT LABELS

1. Please affix any component labels to this sheet and ensure any extra component label sheets are attached to the main Minimum Dataset Form.
2. Ensure all component details are provided, including cement.
3. The NJR DOES NOT record the following: wire, mesh, cables, screws, surgical tools or endoprostheses.

Humeral stem (if used)	Humeral component
Glenoid component (if used)	Cement (if used)
Accessories	

A.5.2 NJR MDS form, version 7: Primary procedures

National Joint Registry www.njrcentre.org.uk Working for patients, driving forward quality	MDS VERSION 7.0 Shoulder Operation — Form: MDSv7.0 S1 v2.0
S1 Shoulder Primary	Patient Addressograph
Important: Please tick relevant boxes. All component stickers should be affixed to the accompanying 'Minimum Dataset Form Component Labels Sheet'. Please ensure that all sheets are stapled together.	

All fields are Mandatory unless otherwise indicated

REMEMBER! MAKE A NOTE OF THE NJR REFERENCE NUMBER WHEN YOU ENTER THIS DATA	NJR REF:

PATIENT DETAILS				
NJR Patient Consent Obtained	Yes ☐	No ☐	Not Recorded ☐	
Body Mass Index (enter either H&W OR BMI OR tick Not Available box)	Height (IN M) Weight (IN KG)	BMI	Not Available ☐	
Handedness	Left ☐	Right ☐	Ambidextrous ☐	Unknown ☐

PATIENT IDENTIFIERS				
Forename(s)				
Surname				
Gender	Male ☐	Female ☐	Not Known ☐	Not Specified ☐
Date of Birth	DD/MM/YYYY			
Patient Postcode			Overseas Address ☐	
NHS Number OR National Patient Identifier (if available)				
Patient Hospital ID				

OPERATION DETAILS					
Hospital					
Operation Date	DD/MM/YYYY				
Anaesthetic Types	General ☐		Regional – Nerve Block ☐		
Patient ASA Grade	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
Operation Funding	NHS ☐	Independent ☐			

SURGEON DETAILS					
Consultant in Charge					
Operating Surgeon					
Operating Surgeon Grade	Consultant ☐	SpR/ST3-8 ☐	F1-ST2 ☐	Specialty Doctor/SAS ☐	Other ☐
First Assistant Grade	Consultant ☐	Other ☐			

SHOULDER PRIMARY PROCEDURE DETAILS								
Side	Left ☐ Right ☐							
Indications for Implantation (select all that apply)	Osteoarthritis	☐	Avascular Necrosis (AVN)	☐	Cuff Tear Arthropathy	☐	Cuff Tear without Arthropathy	☐
	Acute Fracture	☐	Inflammatory Arthropathy	☐	Trauma Sequelae	☐	Metastatic Cancer/Malignancy	☐
	Dislocation Arthropathy	☐	Other	☐				
Previous surgery (not arthroplasty) (Select all that apply)	None	☐	For Fracture	☐	For Instability	☐	For Impingement	☐
	For Cuff Tear	☐	For Gleno-humeral Arthritis	☐	Previous Arthrodesis	☐	Other	☐

SURGICAL APPROACH								
Patient Procedure	Resurfacing Total Arthroplasty	☐	Resurfacing Hemi-arthroplasty	☐	Stemless Conventional Total Arthroplasty	☐	Stemless Hemi-arthroplasty	☐
	Stemless Total Reverse Arthroplasty	☐	Stemmed Conventional Total Arthroplasty	☐	Stemmed Hemi-arthroplasty	☐	Stemmed Total Reverse Arthroplasty	☐
	Interpositional Arthroplasty (Glenohumeral)	☐						
Fixation Humerus	Uncemented ☐	Cemented ☐	Not applicable ☐					
Fixation Glenoid	Uncemented ☐	Cemented ☐	Not applicable ☐					
Approach	Delto-pectoral ☐	Trans-deltoid ☐	Other ☐					
Patient specific instruments?	Yes ☐	No ☐						
Computer Guided Surgery Used?	Yes ☐	No ☐						
Biological Resurfacing (Glenoid) (select all that apply)	None ☐	Microfracture ☐	Reaming ☐	Interposition ☐				
THROMBOPROPHYLAXIS REGIME (intention to treat)								
Chemical (In Hospital)	Aspirin	☐	LMWH	☐	Pentasaccharide (e.g. Fondaparinux)	☐	Warfarin	☐
	Direct Thrombin Inhibitor (e.g. Dabigatran)	☐	Factor Xa Inhibitor (e.g. Rivaroxaban/Apixaban)	☐	Other	☐	None	☐
Mechanical	Foot Pump	☐	Intermittent Calf Compression	☐	TED Stockings	☐	Other	☐
	None	☐						
BONE GRAFT USED								
Was Humeral Bone graft used?	Yes	☐	No	☐				
Humeral - Form	Structural	☐	Morsellised/chips	☐				
Humeral – Type	Autograft	☐	Allograft	☐	Synthetic	☐	Other	☐
Was Glenoid Bone graft used?	Yes	☐	No	☐				
Glenoid - Form	Structural	☐	Morsellised/chips	☐				
Glenoid - Type	Autograft	☐	Allograft	☐	Synthetic	☐	Other	☐
Rotator Cuff								
Rotator Cuff Condition	Normal ☐	Attenuated ☐	Absent/Torn ☐					
Rotator Cuff Repaired?	Yes ☐	No ☐						
Repair Type	Primary Repair ☐	Augmented Patch Repair ☐						

Other Soft Tissues					
Long Head Biceps (LHB) Present?	Yes	☐	No	☐	
LHB Tenotomy Performed?	Yes	☐	No	☐	
LHB Tenodesis Performed?	Yes	☐	No	☐	
Muscle Transfer?	Yes	☐	No	☐	
Other?	Yes	☐	No	☐	

SURGEON'S NOTES

INTRA-OPERATIVE EVENT						
Untoward Intra-Operative Event	None	☐	Fracture Glenoid	☐	Other	☐
	Fracture Humerus	☐	Vascular Injury	☐		

PRE-OPERATIVE OXFORD SCORES – Tick one box for every question. If no scores available select Pre-operative Oxford Scores Not available

Pre-operative Oxford Score Date	DD/MM/YYYY				Not available ☐
1. During the past 4 weeks… How would you describe the **worst** pain you had from your shoulder?					Not available ☐
None ☐	Mild ☐	Moderate ☐	Severe ☐	Unbearable ☐	
2. During the past 4 weeks… Have you had any trouble dressing yourself because of your shoulder?					Not available ☐
No trouble at all ☐	A little bit of trouble ☐	Moderate trouble ☐	Extreme difficulty ☐	Impossible to do ☐	
3. During the past 4 weeks… Have you had any trouble getting in and out of a car or using public transport because of your shoulder?					Not available ☐
No trouble at all ☐	A little bit of trouble ☐	Moderate trouble ☐	Extreme difficulty ☐	Impossible to do ☐	
4. During the past 4 weeks… Have you been able to use a knife and fork at the same time?					Not available ☐
Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐	
5. During the past 4 weeks… Could you do the household shopping on your own?					Not available ☐
Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐	
6. During the past 4 weeks… Could you carry a tray containing a plate of food across a room?					Not available ☐
Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐	
7. During the past 4 weeks… Could you brush/comb your hair with the affected arm?					Not available ☐
Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐	
8. During the past 4 weeks… How would you describe the pain you usually had from your shoulder?					Not available ☐
None ☐	Very mild ☐	Mild ☐	Moderate ☐	Severe ☐	
9. During the past 4 weeks… Could you hang your clothes up in a wardrobe, using the affected arm?					Not available ☐
Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With great difficulty ☐	No, impossible ☐	
10. During the past 4 weeks… Have you been able to wash and dry yourself under both arms?					Not available ☐
Yes, easily ☐	With little difficulty ☐	With moderate difficulty ☐	With extreme difficulty ☐	No, impossible ☐	
11. During the past 4 weeks… How much has pain from your shoulder interfered with your usual work (including housework)?					Not available ☐
Not at all ☐	A little bit ☐	Moderately ☐	Greatly ☐	Totally ☐	
12. During the past 4 weeks… Have you been troubled by pain from your shoulder in bed at night?					Not available ☐
No nights ☐	Only 1 or 2 nights ☐	Some nights ☐	Most nights ☐	Every night ☐	

Minimum Dataset Form - COMPONENT LABELS

1. Please affix any component labels to this sheet and ensure any extra component label sheets are attached to the main Minimum Dataset Form.
2. Ensure all component details are provided, including cement.
3. The NJR DOES NOT record the following: wire, mesh, cables, screws, surgical tools or endoprostheses.

Humeral stem (if used)	Humeral component
Glenoid component (if used)	Cement (if used)
Accessories	

A.5.3 Component database reclassification

The data structure for the NJR shoulder database was inherited from the hip and knee databases. In section 3.3.9, it was noted that this structure was inadequate to describe the components of many shoulder replacement constructs. Prior to undertaking any analyses of the NJR dataset for this thesis, a new component classification was devised with the following aims:

1. To map procedures recorded on the old five-category MDS forms to the nine new categories
2. To facilitate automated real-time data entry validation
3. To support a component library with an increased number of component characteristics

The classification was designed primarily for this thesis. It has subsequently been shared and discussed within an NJR working group, and used without modification to support the 2018 and 2019 shoulder reports.[11,161] Standardised terminology is used for different components groups, reducing ambiguity from inappropriately used anatomical words, and specifically avoiding the use of manufacturer specific terms. Components are divided into the groups shown in table A.7.

Table A.7: Main component groups for classification

Humeral	Glenoid
Standalone humeral component (includes the bearing)	Standalone glenoid
Humeral bearing	Modular glenoid bearing
Humeral stem	
Proximal humeral stem section (body)	Modular glenoid baseplate
Intermediate humeral stem section	
Distal humeral stem section	
Stemless body	
Humeral accessories	Glenoid accessories
- Adaptor tray	- Peg (Post)
- Adapter Trunion	- Screws
- Spacer	- Spacer
- Suture Collar	- Augment
- Peg (Post)	- Wedge
- Screws	- Other
- Augment	
- Fin	
- Canal plug	
- Centering sleeve	
- Adjusting device	
- Locking ring	
- Other	

The minimum set of components required to identify each valid construct has been defined. Some examples of valid and invalid constructs are shown in table A.8. The new classification was used to map the existing dataset to the nine new categories used in version 7 of the MDS. The matching grid is shown in table A.9.

Table A.8: Examples of valid/invalid constructs

Type	Humeral components	Glenoid components	Valid?	Reason
Stemmed TSR	Conventional bearing Proximal stem section Distal stem section Canal plug	Standalone (poly) bearing	✓	
Stemless TSR	Resurfacing bearing Stemless body	Standalone (poly) bearing	✗	Resurfacing bearing
Stemmed RTSR	Reverse poly bearing Humeral adaptor tray Humeral stem	Reverse bearing Modular baseplate Screws	✓	
Resurfacing TSR	Resurfacing bearing		✗	No glenoid component
Stemmed TSR	Conventional bearing Humeral stem	Standalone (poly) bearing Metal peg	✓	Optional peg ok
Stemmed RTSR	Standalone (reverse)	Standalone (reverse)	✓	Bayley Walker design

Table A.9: Procedure reclassification matrix

Version 7	Resurfacing Hemiarthroplasty	Hemiarthroplasty	Resurfacing Arthroplasty	Total Prosthetic Replacement	Reverse Polarity Total Prosthetic Replacement
Resurfacing Hemiarthroplasty	1997	26	0	0	0
Resurfacing TSR	0	0	341	26	0
Stemless Hemiarthroplasty	112	475	0	0	0
Stemmed Hemiarthroplasty	1	1979	0	0	0
Stemless TSR	0	0	110	1773	5
Stemmed TSR	0	0	2	4438	53
Stemless RTSR	0	0	0	0	73
Stemmed RTSR	0	0	6	72	8909
Pyrocarbon Ball	0	5	0	2	0
Resurfacing (Unclear)	0	0	399	2	0
Unclear	26	32	41	258	4

Green = Matching as expected

Orange = Mismatch between procedure and components - components used to override MDS

Red = Unresolvable mismatch or missing components

Note that 49% of resurfacing arthroplasties did not have a glenoid component recorded.

A.5.4 PROMs completeness and timing of collection

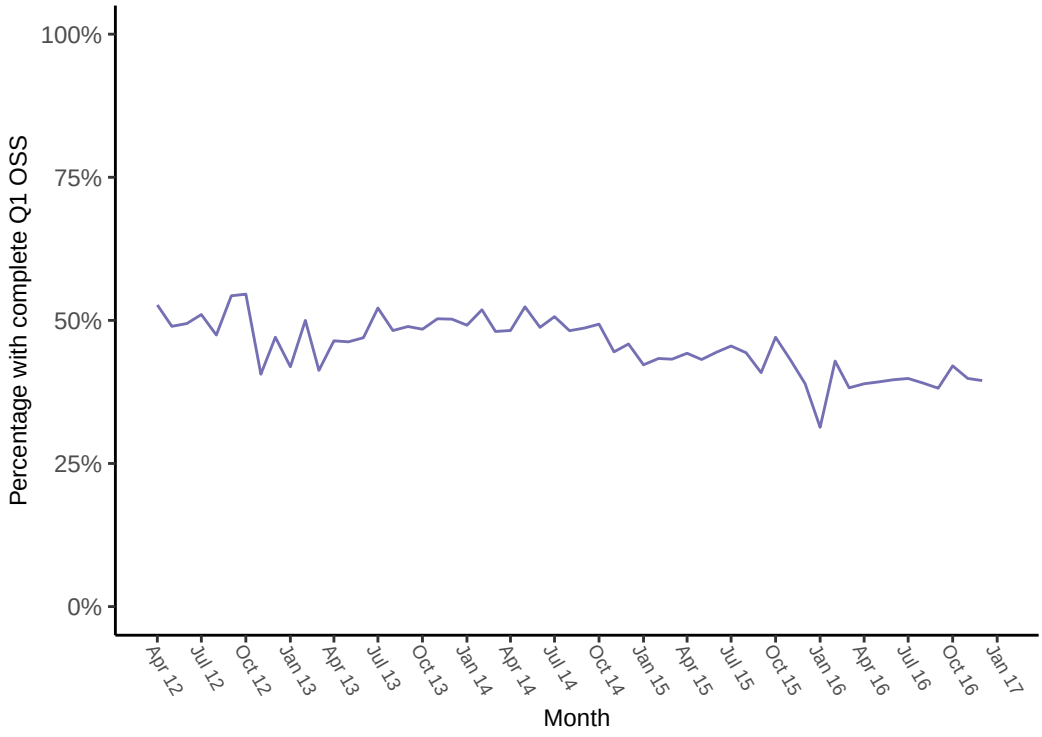


Figure A.1: Proportion of elective procedures with a valid linked Q1 OSS by month of surgery

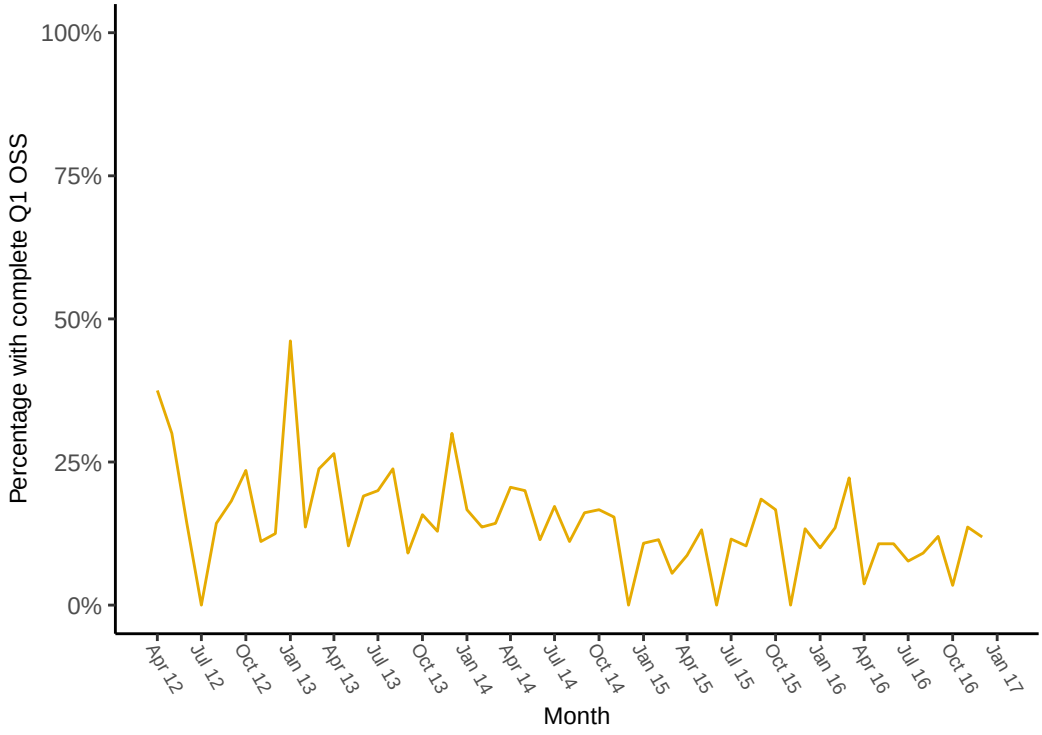


Figure A.2: Proportion of acute trauma procedures with a valid linked Q1 OSS by month of surgery

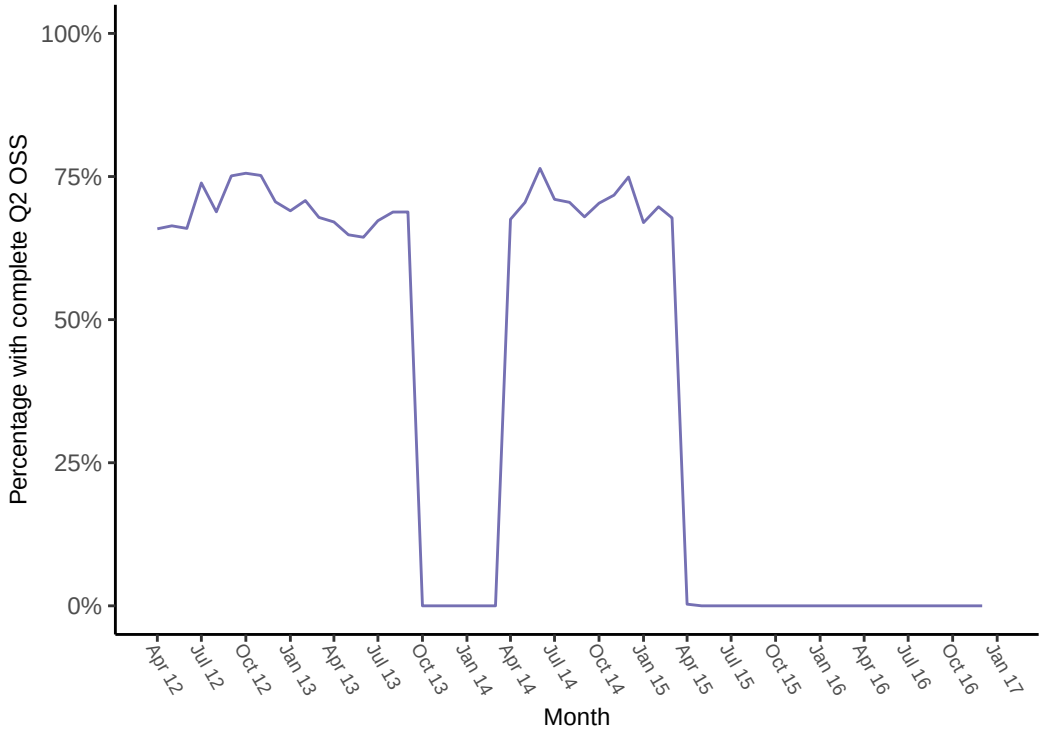


Figure A.3: Proportion of elective procedures with a valid linked Q2 OSS by month of surgery

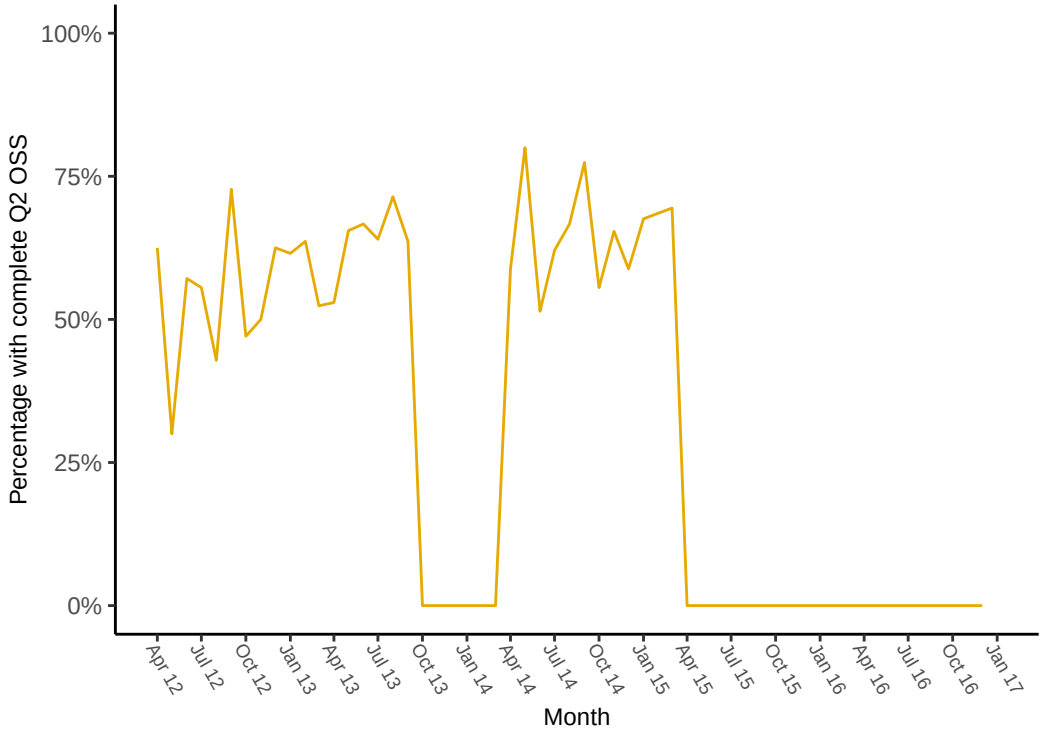


Figure A.4: Proportion of acute trauma procedures with a valid linked Q2 OSS by month of surgery

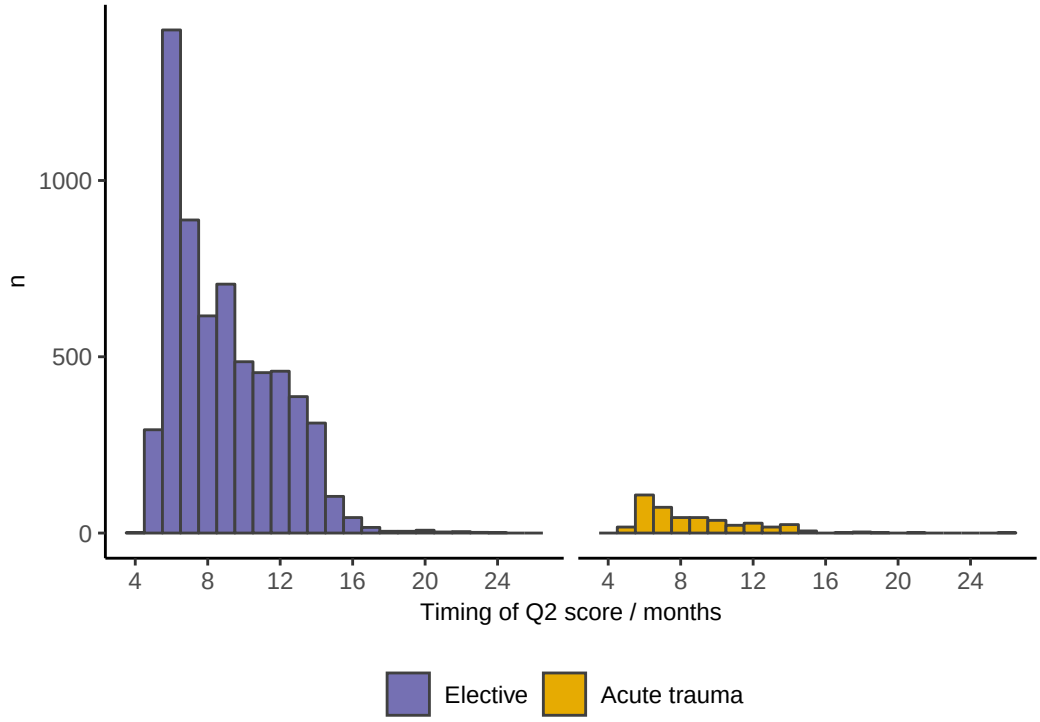


Figure A.5: Timing of Q2 score completion in relation to date of surgery. Includes all scores before the 4 to 12 month limits were applied for analysis

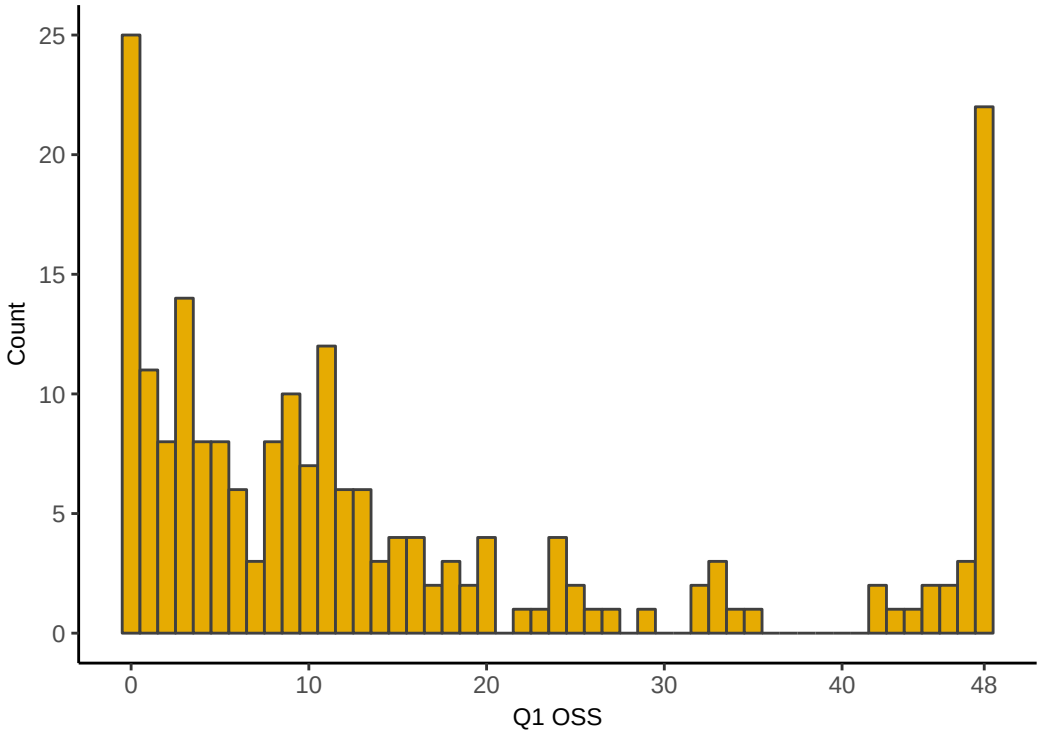


Figure A.6: Histogram of reported Q1 OSS scores for the acute trauma cohort

A.6 Revision: additional materials

A.6.1 Calculation of lifetime risk of revision

Lifetime risk estimates were calculated using a methodology adapted from cancer epidemiology studies.[304,312–314] Patients were grouped in to five-year age-bands and grouped by sex. For each group, a lifetable was constructed for each incremental year. The table included the number of revisions, number of deaths, and number lost to follow-up. The following equations describe how the lifetime risk was calculated using HES data and standardised to the ONS population estimates.

LET:
i=Time in years after primary surgery
r=Number of revisions
q=Number of reoperations
d=Number of deaths
c=Number lost/censored
R=Age specific revision risk
Q=Age specific reoperation risk
D=Age specific mortality risk
n=Number of procedures at risk at start of interval
x=Age group at time of surgery
$\hat{D}$=ONS age standardised mortality risk
$\hat{L}$=Mid-year population estimate
$\hat{N}$=Estimated number at risk at start of interval
$\hat{Y}$=Number of person years lived in interval
$\hat{P}$=Current probability of revision during time interval
LTR=Lifetime risk

CALCULATIONS:
Application of revisions, deaths and censoring: $n_{i+1} = n_i - r_i - d_i - c_i$
Check that revisions, deaths and censoring add up: $n_0 = \sum_0^i (r+d+c)$
Revision risk for time interval i: $R_i = \frac{r_i}{n_i - c_i/2}$
Apply revision risk to standard population: $\hat{N}_x = \hat{N}_{x-1}(1-\hat{D}_{x-1})(1-R_{x-1})$
Number of person years at risk: $\hat{Y}_i = \frac{\hat{N}_i + \hat{N}_{i+1}}{2}$
Current probability of revision: $\hat{P}_i = \hat{Y}_i R_i$
Lifetime risk: $LTR = \sum_0^i \hat{P}$
Variance of $\hat{P}_i$): $Var(\hat{P}_i) = \frac{\hat{Y}_i^2 R_i}{\hat{N}^2}$
Confidence interval of LTR: $LTR \pm 1.96\sqrt{\sum_0^i Var(\hat{P}_i)}$

A.6.2 Table of lifetime risk estimates

Table A.10: Lifetime risk of revision and reoperation by age and sex. 95% CIs shown.

	Revision	Reoperation
Male		
50-54	19.34 (18.47 to 20.21)	29.75 (28.67 to 30.82)
55-59	23.61 (22.65 to 24.58)	31.86 (30.74 to 32.99)
60-64	15.86 (15.06 to 16.65)	24.22 (23.24 to 25.21)
65-69	12.94 (12.22 to 13.66)	20.36 (19.45 to 21.28)
70-74	10.55 (9.90 to 11.21)	16.94 (16.10 to 17.79)
75-79	8.03 (7.46 to 8.60)	13.14 (12.38 to 13.90)
80-84	7.06 (6.52 to 7.60)	9.08 (8.44 to 9.71)
85+	3.80 (3.43 to 4.17)	4.36 (3.96 to 4.77)
Female		
50-54	19.79 (18.92 to 20.67)	28.01 (26.96 to 29.05)
55-59	19.02 (18.16 to 19.88)	27.87 (26.82 to 28.91)
60-64	15.20 (14.43 to 15.98)	22.61 (21.66 to 23.55)
65-69	12.42 (11.72 to 13.13)	19.41 (18.53 to 20.30)
70-74	9.28 (8.67 to 9.89)	14.08 (13.32 to 14.84)
75-79	6.42 (5.91 to 6.93)	9.88 (9.24 to 10.53)
80-84	4.25 (3.83 to 4.66)	7.63 (7.02 to 8.24)
85+	2.68 (2.35 to 3.02)	3.70 (3.28 to 4.12)

A.6.3 Sensitivity analyses accounting for death as a competing risk

Table A.11: Risk of revision surgery up to 19 years following primary shoulder replacement surgery. Cumulative incidence of revision and death reported alongside Kaplan-Meier estimates.

Time (Years)	Number at risk	Number revised	Censored (Death)	Censored (Study end)	KM	CI	CI
					Revision/%	Revision/%	Death/%
Elective							
0	61013						
1	52793	1306	1217	5697	2.25	2.23	2.02
2	45576	967	1261	4989	4.15	4.07	4.35
3	38951	575	1335	4715	5.44	5.29	7.07
4	32847	420	1435	4253	6.54	6.29	10.32
5	27593	274	1279	3697	7.38	7.02	13.6
6	22869	211	1229	3284	8.15	7.67	17.17
7	18723	140	1190	2816	8.76	8.16	21.12
8	15004	109	1053	2567	9.36	8.60	25.10
9	11738	92	892	2272	9.98	9.05	29.11
10	8974	51	753	1960	10.43	9.34	33.16
11	6821	42	576	1535	10.89	9.64	36.88
12	5118	28	475	1206	11.32	9.88	40.63
13	3816	22	379	895	11.75	10.11	44.39
14	2785	9	284	738	11.99	10.23	47.90
15	1954	16	206	609	12.59	10.51	51.24
16	1335	5	150	464	12.85	10.62	54.28
17	795	2	101	437	13.00	10.67	57.19
18	349	1	35	410	13.18	10.74	58.92
19	0	0	8	341	13.18	10.74	59.68
Trauma							
0	14624						
1	12641	357	774	852	2.55	2.49	5.24
2	11029	162	564	886	3.87	3.72	9.40
3	9575	70	537	847	4.51	4.29	13.63
4	8150	39	565	822	4.93	4.64	18.58
5	6897	19	523	710	5.17	4.82	23.64
6	5792	7	464	634	5.27	4.90	28.59
7	4853	10	375	554	5.45	5.02	33.00
8	4039	11	308	497	5.70	5.17	37.06
9	3312	10	272	443	5.96	5.33	41.10
10	2685	3	243	381	6.05	5.38	45.05
11	2147	3	205	330	6.16	5.44	48.92
12	1717	5	144	283	6.42	5.55	52.07
13	1331	2	114	268	6.53	5.61	54.96
14	1020	1	89	221	6.61	5.64	57.72
15	745	2	74	199	6.83	5.72	60.49
16	519	0	49	177	6.83	5.72	62.94
17	314	0	37	168	6.83	5.72	65.59
18	143	0	21	150	6.83	5.72	68.13
19	0	0	3	140	6.83	5.72	69.19

A.7 Serious adverse events: additional data

A.7.1 Age and Charlson score distributions

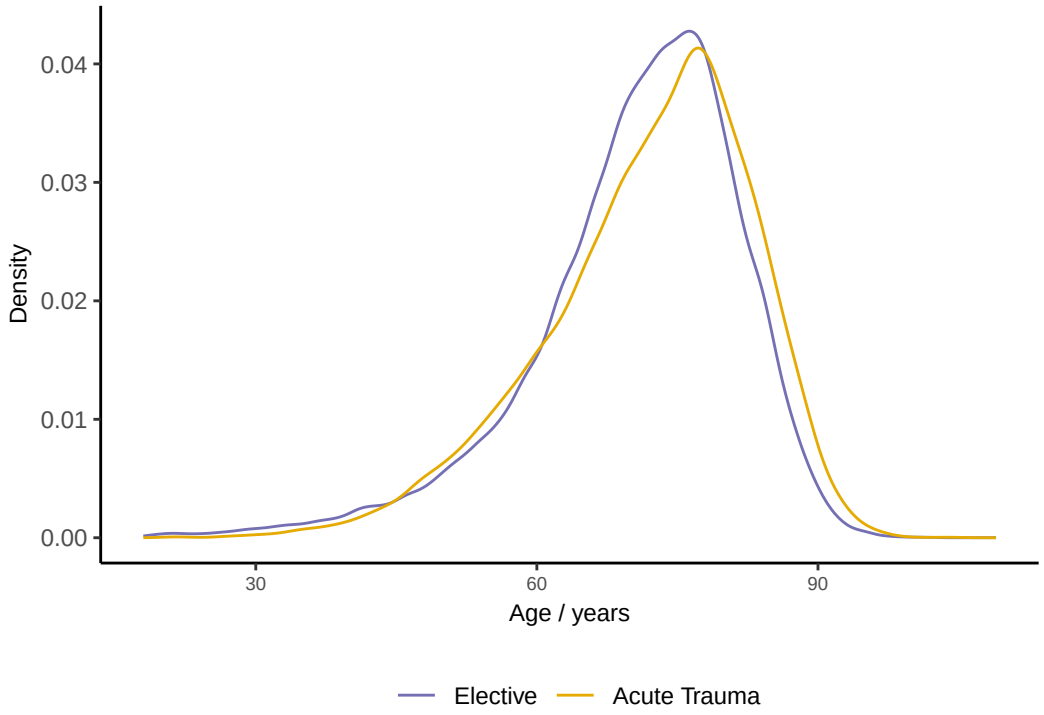


Figure A.7: Density plot for age of patients in the HES revision and serious adverse event cohorts

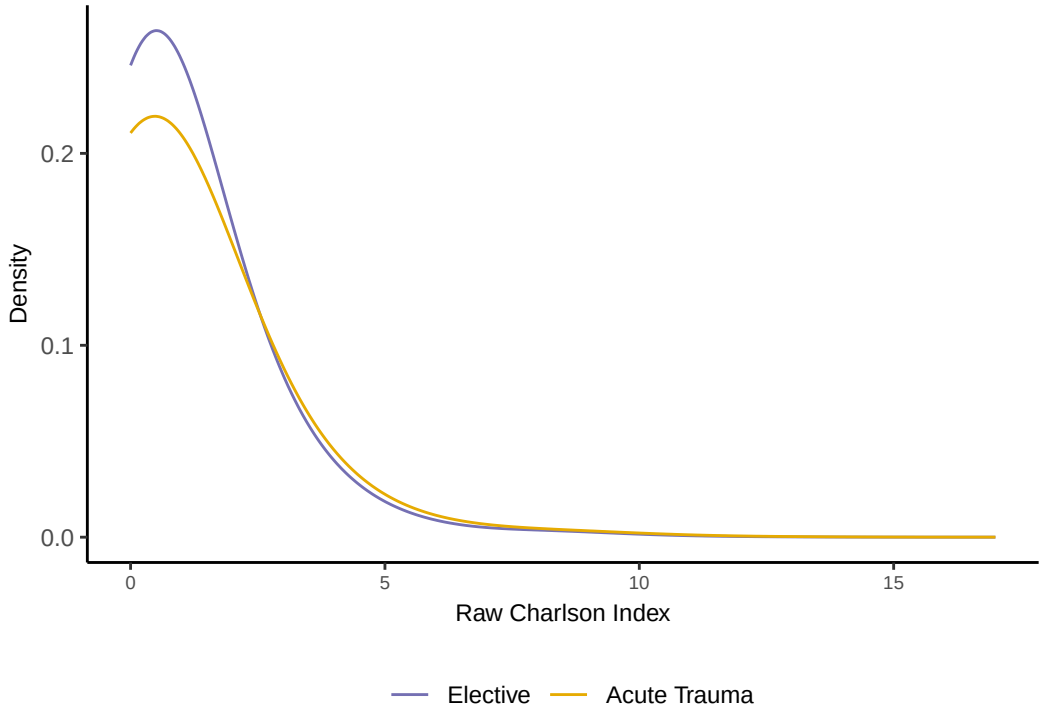


Figure A.8: Density plot for Charlson Index of patients in the HES revision and serious adverse event cohorts

A.7.2 Risks of serious adverse events by age and sex

Table A.12: Risks of each serious adverse event for elective cohort: age- and sex-stratified (%)

		Male		Female	
Age	Event	30 days	90 days	30 days	90 days
18-49	Any	0.72 (0.40 to 1.30)	1.18 (0.74 to 1.87)	1.22 (0.76 to 1.97)	1.94 (1.33 to 2.83)
18-49	Any (+ death)	0.78 (0.45 to 1.38)	1.31 (0.84 to 2.03)	1.22 (0.76 to 1.97)	1.94 (1.33 to 2.83)
18-49	Death	0.07 (0.01 to 0.46)	0.20 (0.06 to 0.61)	0	0
18-49	PE	0.13 (0.03 to 0.52)	0.13 (0.03 to 0.52)	0.07 (0.01 to 0.51)	0.07 (0.01 to 0.51)
18-49	MI	0	0.13 (0.03 to 0.52)	0	0
18-49	CVE	0	0	0.07 (0.01 to 0.51)	0.07 (0.01 to 0.51)
18-49	AKI	0	0	0.07 (0.01 to 0.51)	0.07 (0.01 to 0.51)
18-49	LRTI	0.59 (0.31 to 1.13)	0.85 (0.49 to 1.46)	0.43 (0.19 to 0.96)	0.93 (0.54 to 1.61)
18-49	UTI	0	0.07 (0.01 to 0.46)	0.58 (0.29 to 1.15)	0.86 (0.49 to 1.52)
50-64	Any	1.50 (1.19 to 1.90)	2.00 (1.63 to 2.45)	1.50 (1.23 to 1.83)	2.03 (1.72 to 2.41)
50-64	Any (+ death)	1.53 (1.21 to 1.93)	2.08 (1.71 to 2.54)	1.56 (1.28 to 1.89)	2.19 (1.86 to 2.58)
50-64	Death	0.06 (0.02 to 0.20)	0.17 (0.09 to 0.34)	0.08 (0.03 to 0.18)	0.20 (0.12 to 0.34)
50-64	PE	0.13 (0.06 to 0.29)	0.19 (0.10 to 0.37)	0.15 (0.08 to 0.28)	0.21 (0.13 to 0.36)
50-64	MI	0.09 (0.03 to 0.23)	0.19 (0.10 to 0.37)	0.03 (0.01 to 0.12)	0.05 (0.01 to 0.14)
50-64	CVE	0.09 (0.03 to 0.23)	0.13 (0.06 to 0.29)	0	0
50-64	AKI	0.45 (0.29 to 0.69)	0.54 (0.36 to 0.80)	0.18 (0.10 to 0.32)	0.29 (0.19 to 0.46)
50-64	LRTI	0.71 (0.50 to 1.00)	0.88 (0.65 to 1.20)	0.87 (0.67 to 1.13)	1.16 (0.93 to 1.45)
50-64	UTI	0.21 (0.12 to 0.40)	0.28 (0.16 to 0.48)	0.40 (0.27 to 0.58)	0.57 (0.41 to 0.78)
65-74	Any	2.23 (1.89 to 2.64)	2.90 (2.51 to 3.36)	2.24 (2.01 to 2.50)	2.96 (2.70 to 3.25)
65-74	Any (+ death)	2.31 (1.96 to 2.73)	3.13 (2.72 to 3.61)	2.30 (2.07 to 2.56)	3.12 (2.85 to 3.42)
65-74	Death	0.11 (0.05 to 0.24)	0.31 (0.20 to 0.49)	0.12 (0.08 to 0.19)	0.32 (0.24 to 0.42)
65-74	PE	0.08 (0.03 to 0.20)	0.15 (0.08 to 0.28)	0.16 (0.11 to 0.24)	0.22 (0.16 to 0.31)
65-74	MI	0.15 (0.08 to 0.28)	0.26 (0.16 to 0.43)	0.16 (0.11 to 0.24)	0.17 (0.12 to 0.26)
65-74	CVE	0.15 (0.08 to 0.28)	0.23 (0.14 to 0.39)	0.05 (0.03 to 0.11)	0.15 (0.10 to 0.23)
65-74	AKI	0.65 (0.48 to 0.89)	0.70 (0.52 to 0.95)	0.51 (0.41 to 0.64)	0.62 (0.51 to 0.76)
65-74	LRTI	0.99 (0.77 to 1.28)	1.26 (1.00 to 1.57)	0.97 (0.83 to 1.15)	1.28 (1.11 to 1.48)
65-74	UTI	0.52 (0.37 to 0.74)	0.82 (0.62 to 1.08)	0.67 (0.55 to 0.82)	0.96 (0.81 to 1.13)
75-84	Any	5.60 (4.92 to 6.37)	7.36 (6.58 to 8.23)	3.98 (3.68 to 4.30)	5.07 (4.73 to 5.43)
75-84	Any (+ death)	5.86 (5.17 to 6.65)	7.82 (7.01 to 8.72)	4.09 (3.79 to 4.41)	5.39 (5.04 to 5.75)
75-84	Death	0.68 (0.47 to 0.98)	1.33 (1.02 to 1.73)	0.28 (0.21 to 0.38)	0.63 (0.52 to 0.77)
75-84	PE	0.27 (0.15 to 0.48)	0.39 (0.24 to 0.63)	0.22 (0.16 to 0.31)	0.34 (0.26 to 0.44)
75-84	MI	0.29 (0.16 to 0.51)	0.58 (0.39 to 0.86)	0.25 (0.19 to 0.34)	0.32 (0.24 to 0.42)
75-84	CVE	0.29 (0.16 to 0.51)	0.55 (0.37 to 0.84)	0.18 (0.13 to 0.26)	0.29 (0.22 to 0.38)
75-84	AKI	1.79 (1.42 to 2.24)	1.93 (1.55 to 2.40)	0.86 (0.73 to 1.01)	1.07 (0.92 to 1.24)
75-84	LRTI	2.53 (2.09 to 3.07)	3.31 (2.80 to 3.91)	1.64 (1.45 to 1.84)	2.12 (1.91 to 2.36)
75-84	UTI	1.13 (0.85 to 1.51)	1.69 (1.34 to 2.13)	1.33 (1.16 to 1.52)	1.84 (1.64 to 2.06)
85+	Any	14.80 (12.05 to 18.17)	18.21 (15.13 to 21.92)	8.34 (7.42 to 9.38)	9.76 (8.76 to 10.88)
85+	Any (+ death)	17.07 (14.10 to 20.67)	20.98 (17.65 to 24.93)	9.05 (8.09 to 10.13)	11.12 (10.05 to 12.31)
85+	Death	4.72 (3.28 to 6.79)	6.67 (4.91 to 9.05)	1.89 (1.48 to 2.42)	3.05 (2.51 to 3.70)
85+	PE	0.49 (0.16 to 1.51)	0.49 (0.16 to 1.51)	0.24 (0.12 to 0.47)	0.41 (0.25 to 0.70)
85+	MI	1.14 (0.54 to 2.39)	1.46 (0.76 to 2.81)	0.65 (0.43 to 0.99)	0.65 (0.43 to 0.99)
85+	CVE	0.49 (0.16 to 1.51)	0.65 (0.24 to 1.73)	0.44 (0.27 to 0.74)	0.59 (0.38 to 0.92)
85+	AKI	3.41 (2.23 to 5.24)	4.55 (3.14 to 6.59)	1.69 (1.30 to 2.19)	2.01 (1.59 to 2.55)
85+	LRTI	10.24 (8.00 to 13.11)	12.68 (10.16 to 15.83)	3.40 (2.83 to 4.08)	4.08 (3.46 to 4.82)
85+	UTI	3.74 (2.49 to 5.63)	5.04 (3.54 to 7.17)	3.52 (2.94 to 4.21)	4.32 (3.67 to 5.08)

Table A.13: Risks of each serious adverse event for acute trauma cohort: age- and sex-stratified (%)

		Male		Female	
Age (yrs)	Event	30 days	90 days	30 days	90 days
18-49	Any (+ death)	2.53 (1.36 to 4.69)	3.54 (2.09 to 5.97)	1.78 (0.57 to 5.50)	4.14 (1.97 to 8.69)
18-49	Any	2.53 (1.36 to 4.69)	3.03 (1.72 to 5.34)	1.78 (0.57 to 5.50)	4.14 (1.97 to 8.69)
18-49	Death	0	0.51 (0.13 to 2.02)	0	0.59 (0.08 to 4.20)
18-49	PE	0	0	0	0
18-49	MI	0.25 (0.04 to 1.79)	0.25 (0.04 to 1.79)	0	0
18-49	CVE	0.25 (0.04 to 1.79)	0.51 (0.13 to 2.02)	0	1.18 (0.30 to 4.73)
18-49	AKI	1.26 (0.53 to 3.03)	1.26 (0.53 to 3.03)	0	0.59 (0.08 to 4.20)
18-49	LRTI	0.76 (0.24 to 2.35)	1.01 (0.38 to 2.69)	0.59 (0.08 to 4.20)	1.18 (0.30 to 4.73)
18-49	UTI	0.25 (0.04 to 1.79)	0.25 (0.04 to 1.79)	1.18 (0.30 to 4.73)	1.18 (0.30 to 4.73)
50-64	Any (+ death)	6.13 (4.79 to 7.84)	7.10 (5.65 to 8.93)	3.54 (2.75 to 4.55)	4.30 (3.42 to 5.41)
50-64	Any	5.74 (4.45 to 7.41)	6.42 (5.04 to 8.17)	3.30 (2.54 to 4.29)	4.01 (3.16 to 5.08)
50-64	Death	0.49 (0.20 to 1.17)	0.97 (0.52 to 1.81)	0.35 (0.16 to 0.79)	0.65 (0.36 to 1.17)
50-64	PE	0.49 (0.20 to 1.17)	0.49 (0.20 to 1.17)	0.29 (0.12 to 0.71)	0.29 (0.12 to 0.71)
50-64	MI	0.29 (0.09 to 0.90)	0.29 (0.09 to 0.90)	0.24 (0.09 to 0.63)	0.24 (0.09 to 0.63)
50-64	CVE	0.58 (0.26 to 1.30)	0.68 (0.32 to 1.43)	0.18 (0.06 to 0.55)	0.29 (0.12 to 0.71)
50-64	AKI	0.88 (0.46 to 1.68)	0.97 (0.52 to 1.81)	0.59 (0.32 to 1.10)	0.65 (0.36 to 1.17)
50-64	LRTI	3.40 (2.44 to 4.74)	3.89 (2.85 to 5.30)	1.24 (0.81 to 1.90)	1.59 (1.09 to 2.32)
50-64	UTI	1.26 (0.73 to 2.18)	1.36 (0.81 to 2.30)	1.00 (0.62 to 1.61)	1.24 (0.81 to 1.90)
65-74	Any (+ death)	9.90 (8.11 to 12.08)	12.14 (10.15 to 14.53)	6.91 (6.08 to 7.86)	7.95 (7.05 to 8.97)
65-74	Any	9.29 (7.56 to 11.40)	11.22 (9.31 to 13.53)	6.76 (5.94 to 7.70)	7.60 (6.72 to 8.59)
65-74	Death	0.82 (0.41 to 1.63)	1.63 (1.00 to 2.67)	0.45 (0.27 to 0.74)	0.95 (0.67 to 1.35)
65-74	PE	0.41 (0.15 to 1.09)	0.61 (0.28 to 1.36)	0.66 (0.43 to 1.00)	0.66 (0.43 to 1.00)
65-74	MI	0.61 (0.28 to 1.36)	0.71 (0.34 to 1.50)	0.36 (0.20 to 0.63)	0.42 (0.25 to 0.70)
65-74	CVE	0.82 (0.41 to 1.63)	1.33 (0.77 to 2.28)	0.39 (0.22 to 0.67)	0.48 (0.29 to 0.78)
65-74	AKI	2.55 (1.72 to 3.78)	2.76 (1.89 to 4.02)	0.86 (0.60 to 1.24)	1.19 (0.87 to 1.62)
65-74	LRTI	5.00 (3.78 to 6.62)	5.71 (4.40 to 7.43)	2.83 (2.31 to 3.46)	3.40 (2.83 to 4.08)
65-74	UTI	2.04 (1.32 to 3.16)	2.65 (1.81 to 3.90)	2.80 (2.29 to 3.43)	3.22 (2.66 to 3.88)
75-84	Any (+ death)	12.43 (10.33 to 14.96)	15.43 (13.06 to 18.22)	10.70 (9.77 to 11.72)	12.73 (11.71 to 13.84)
75-84	Any	11.21 (9.22 to 13.62)	13.65 (11.44 to 16.29)	10.17 (9.27 to 11.17)	11.69 (10.72 to 12.76)
75-84	Death	2.11 (1.35 to 3.31)	4.00 (2.88 to 5.54)	1.17 (0.89 to 1.54)	2.12 (1.73 to 2.60)
75-84	PE	0.33 (0.11 to 1.03)	0.44 (0.17 to 1.18)	0.46 (0.30 to 0.71)	0.55 (0.37 to 0.82)
75-84	MI	1.11 (0.60 to 2.06)	1.55 (0.92 to 2.62)	0.48 (0.32 to 0.74)	0.62 (0.43 to 0.91)
75-84	CVE	0.44 (0.17 to 1.18)	0.78 (0.37 to 1.63)	0.44 (0.28 to 0.69)	0.67 (0.46 to 0.96)
75-84	AKI	3.44 (2.42 to 4.89)	4.00 (2.88 to 5.54)	1.77 (1.42 to 2.22)	1.98 (1.60 to 2.45)
75-84	LRTI	5.33 (4.01 to 7.07)	6.55 (5.07 to 8.45)	4.24 (3.67 to 4.89)	5.06 (4.44 to 5.78)
75-84	UTI	3.44 (2.42 to 4.89)	4.77 (3.54 to 6.44)	4.74 (4.14 to 5.44)	5.55 (4.89 to 6.29)
85+	Any (+ death)	21.34 (16.22 to 28.08)	28.45 (22.43 to 36.09)	20.45 (18.09 to 23.10)	25.06 (22.44 to 27.99)
85+	Any	17.57 (12.99 to 23.78)	20.92 (15.86 to 27.60)	18.77 (16.53 to 21.33)	21.88 (19.44 to 24.62)
85+	Death	6.28 (3.78 to 10.41)	12.55 (8.78 to 17.95)	3.82 (2.88 to 5.07)	7.72 (6.32 to 9.42)
85+	PE	0	0	0.40 (0.17 to 0.96)	0.56 (0.27 to 1.17)
85+	MI	2.09 (0.87 to 5.03)	2.51 (1.13 to 5.59)	1.27 (0.78 to 2.08)	1.35 (0.84 to 2.18)
85+	CVE	1.26 (0.40 to 3.89)	1.26 (0.40 to 3.89)	0.80 (0.43 to 1.48)	1.03 (0.60 to 1.78)
85+	AKI	5.86 (3.47 to 9.89)	7.11 (4.42 to 11.44)	3.66 (2.74 to 4.89)	4.14 (3.15 to 5.43)
85+	LRTI	9.62 (6.39 to 14.48)	10.46 (7.07 to 15.48)	8.19 (6.76 to 9.94)	9.71 (8.13 to 11.59)
85+	UTI	7.11 (4.42 to 11.44)	9.62 (6.39 to 14.48)	8.51 (7.04 to 10.29)	10.50 (8.85 to 12.45)

A.7.3 Model equation for risks of any serious advent within 30 days

$$\text{Prob}\{\text{Any complication within 30 days} = 1\} = \frac{1}{1+\exp(-X\beta)}, \text{ where}$$

$$\begin{aligned} X\hat{\beta} = \\ & -5.66 \\ & +0.0181 * \text{Age} + 5.52\times10^{-5} * (\text{Age}-56)^3_+ \\ & -0.000149 * (\text{Age}-73)^3_+ + 9.38\times10^{-5} * (\text{Age}-83)^3_+ \\ & -0.22 * \text{Female} \\ & +0.67 * (\text{Charlson} = 1) + 1.29 * (\text{Charlson} = 2+) \\ & +1.11 * \text{Acute trauma} \end{aligned}$$

Table A.14: Rates of complications at 30 days from NSQIP studies

Study	Years	Subgroups	N	Mean age	Elec/Trauma	Death	PE	DVT	VTE	MI	Stroke	Sepsis	AKI	UTI	LRTI
Shields 2014[341]	2005-2010	HA	523	65.6	490/33				0.40%						
		TSR/RTSR	1195	68.6	1153/42				0.40%						
Waterman 2015[342]	2006-2011		2004	68.8	?	0.25%	0.25%	0.30%		0.20%	0.20%	0.30%	0%	1.50%	0.50%
Westermann 2016[344]	2013		2779	70.4	?	0.18%			0.22%						
Belmont 2017[345]	2011-2013		3547		?										
Newman 2018[171]	2008-2015	<80 years	8802		?		0.30%	0.30%		0.10%	0.10%	0.20%	0.02%	0.60%	0.40%
		≥80 years	1551		?		0.80%	0.30%		0.60%	0.30%	0.30%	0.20%	1.70%	0.90%

Jiang 2016[343] reported stratified by BMI and not included here. All estimates are pooled TSR/RTSR except where stated. HA=Hemiarthroplasty

A.8 Prediction model equations

Model 1

$$
\begin{aligned}
Q2OSS = \\
&15.64 \\
&+0.19 * \text{Age} - 3.49 \times 10^{-5} * (\text{Age} - 53.17)^3 \\
&-0.00111 * (\text{Age} - 69.86)^3 + 0.00206 * (\text{Age} - 76.67)^3 \\
&-0.000916 * (\text{Age} - 85.81)^3 \\
&+0.492 * \text{Q1OSS} - 0.0001342289 * (\text{Q1OSS} - 5)^3 \\
&-0.000163 * (\text{Q1OSS} - 13)^3 + 0.000584 * (\text{Q1OSS} - 20)^3 \\
&-0.000287 * (\text{Q1OSS} - 31)^3 \\
&+0.34 * \text{Male} \\
&-3.62 * \text{Post} - \text{trauma} - 0.95 * \text{Inflammatory} - 1.60 * \text{AVN} - 2.39 * \text{Other diagnosis} \\
&+0.022 * \text{Attenuated} - 1.77 * \text{Repaired} - 1.96 * \text{Absent/Torn} \\
&-2.48 * \text{CTA} \\
&-2.59 * \text{Previous surgery} \\
&-0.45 * \text{ASA 2} - 2.66 * \text{ASA 3} - 5.44 * \text{ASA 4}+ \\
&-0.58 * \text{Non} - \text{dominant side}
\end{aligned}
$$

Model 2 (Cox with imputation)

$$\text{Prob}\{T \geq t\} = S_0(t)^{e^{X\beta}}, \text{ where}$$

$$\begin{aligned}
X\hat{\beta} = \\
&0.69 \\
&-0.0187 * \text{Age} - 8.05! \times 10^{-6} * (\text{Age} - 53.47)^3 \\
&-2.11 \times 10^{-5} * (\text{Age} - 69.94)^3 \\
&+6.64 \times 10^{-5} * (\text{Age} - 76.82)^3 \\
&-3.73! \times 10^{-5} * (\text{Age} - 85.74)^3 \\
&+0.0450 * \text{Q1OSS} - 0.000225 * (\text{Q1OSS} - 5)^3 \\
&+0.000610 * (\text{Q1OSS} - 13)^3 - 0.000466 * (\text{Q1OSS} - 20)^3 \\
&+8.14 \times 10^{-5} * (\text{Q1OSS} - 31)^3 \\
&+0.34 * \text{Male} \\
&+0.34 * \text{Post} - \text{trauma} + 0.006 * \text{Inflammatory} - 0.66 * \text{AVN} + 0.46 * \text{Other diagnosis} \\
&+0.09 * \text{Attenuated} - 0.05 * \text{Repaired} + 0.28 * \text{Absent/Torn} \\
&+0.12 * \text{CTA} \\
&+0.43 * \text{Previous surgery} \\
&+0.005 * \text{ASA 2} + 0.14 * \text{ASA 3} + 0.22 * \text{ASA 4+} \\
&+0.18 * \text{Non} - \text{dominant side}
\end{aligned}$$

t	$S_0(t)$
0	1.000
2	0.980
4	0.968

Model 3

$$\text{Prob}\{\text{Poor outcome} = 1\} = \frac{1}{1+\exp(-X\beta)}, \text{ where}$$

$$\begin{aligned} X\hat{\beta} = \\ &-0.69 \\ &-0.0268 * \text{Age} - 2.00 \times 10^{-6} * (\text{Age} - 53.17)^3 \\ &+0.000163 * (\text{Age} - 69.86)^3 \\ &-0.000278 * (\text{Age} - 76.67)^3 \\ &+0.000117 * (\text{Age} - 85.81)^3 \\ &+0.29 * \text{Male} \\ &+0.54 * \text{Post} - \text{trauma} + 0.07 * \text{Inflammatory} + 0.10 * \text{AVN} + 0.09 * \text{Other diagnosis} \\ &-0.11 * \text{Attenuated} + 0.26 * \text{Repaired} + 0.20 * \text{Absent/Torn} \\ &+0.28 * \text{CTA} \\ &+0.49 * \text{Previoussurgery} \\ &-0.06 * \text{ASA2} + 0.14 * \text{ASA3} + 0.62 * \text{ASA4+} \\ &+0.13 * \text{Non} - \text{dominanthand} \end{aligned}$$

A.9 PROMS: additional materials

A.9.1 Distribution of Q1 OSS by rotator cuff status and implant type

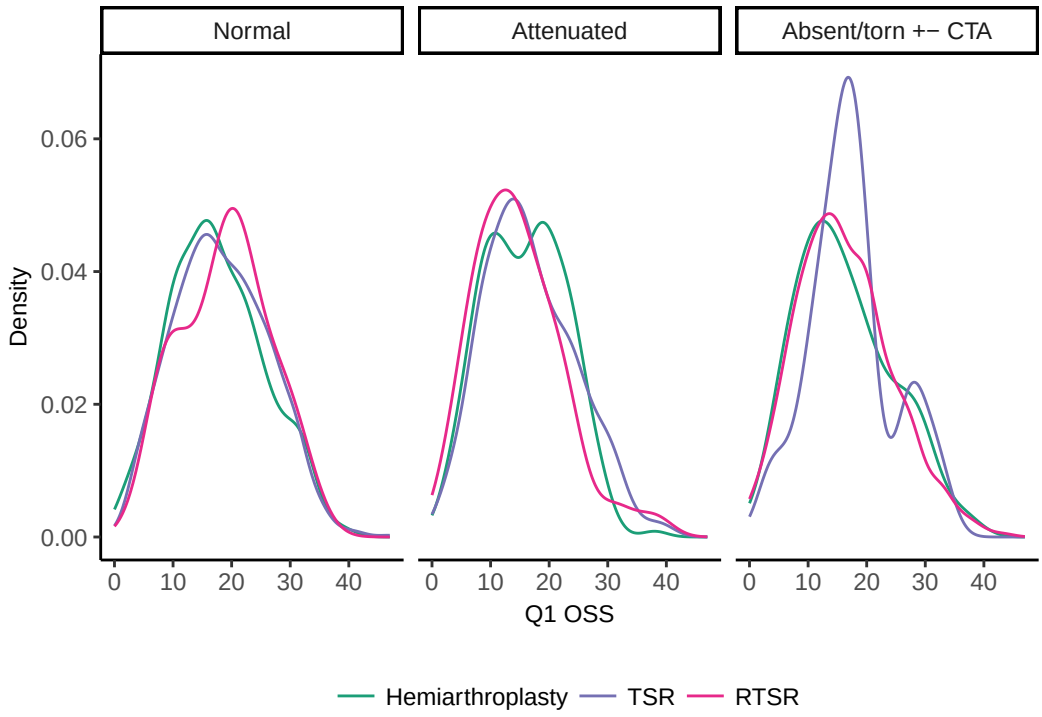


Figure A.9: Q1 OSS distributions by state of rotator cuff and type of surgery performed

Bibliography

1. Izquierdo R, Voloshin I, Edwards S, Freehill MQ, Stanwood W, Wiater JM, Watters WC, Goldberg MJ, Keith M, Turkelson CM, Wies JL, Anderson S, Boyer K, Raymond L, Sluka P and American Academy of Orthopedic Surgeons. Treatment of glenohumeral osteoarthritis. J Am Acad Orthop Surg 2010;18:375–82.
2. Bull D, Tai Kie A, Hanusch B, Kulkarni R, Rees J and Rangan A. Is there sufficient evidence to support intervention to manage shoulder arthritis? Shoulder Elbow 2016;8:77–89.
3. Thomas M, Bidwai A, Rangan A, Rees JL, Brownson P, Tennent D, Connor C and Kulkarni R. Glenohumeral osteoarthritis. Shoulder Elbow 2016;8:203–14.
4. Boorman RS, Kopjar B, Fehringer E, Churchill R, Smith K and Matsen FA. The effect of total shoulder arthroplasty on self-assessed health status is comparable to that of total hip arthroplasty and coronary artery bypassgrafting. J Shoulder Elbow Surg 2003;12:158–63.
5. Fevang BS, Lygre SHL, Bertelsen G, Skredderstuen A, Havelin LI and Furnes O. Good function after shoulder arthroplasty. Acta Orthop 2012;83:467–73.
6. Cho CH, Song KS, Hwang I, Coats-Thomas MS and Warner JJ. Changes in Psychological Status and Health-Related Quality of Life Following Total Shoulder Arthroplasty. J Bone Joint Surg Am 2017;99:1030–5.
7. GMC. Good Medical Practice. URL: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice Last accessed on 26/07/2018.
8. GMC. Consent: patients and doctors making decisions together. URL: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/consent Last accessed on 26/07/2018.
9. Obama B. State of the Union Address. 2015.
10. Rangan A, Upadhaya S, Regan S, Toye F and Rees JL. Research priorities for shoulder surgery: results of the 2015 James Lind Alliance patient and clinician priority setting partnership. BMJ Open 2016;6:e010412.
11. NJR. National Joint Registry for England, Wales and Northern Ireland. 15th Annual Report. 2018.
12. Kim SH, Wise BL, Zhang Y and Szabo RM. Increasing Incidence of Shoulder Arthroplasty in the United States. J Bone Joint Surg Am 2011;93:2249–54.
13. Lübbeke A, Rees JL, Barea C, Combescure C, Carr AJ and Silman AJ. International variation in shoulder arthroplasty. Acta Orthop 2017;88:592–9.
14. Dillon MT, Chan PH, Inacio MCS, Singh A, Yian EH and Navarro RA. Yearly Trends in Elective Shoulder Arthroplasty, 2005 through 2013. Arthritis Care Res (Hoboken) 2017;69:1574–81.

15. Singh JA, Sperling J, Buchbinder R and McMaken K. Surgery for shoulder osteoarthritis. Cochrane database Syst Rev 2010;6:CD008089.

16. Chaudhury S, Gwilym SE, Moser J and Carr AJ. Surgical options for patients with shoulder pain. Nat Rev Rheumatol 2010;6:217–26.

17. Namdari S, Horneff JG and Baldwin K. Comparison of Hemiarthroplasty and Reverse Arthroplasty for Treatment of Proximal Humeral Fractures. J Bone Joint Surg Am 2013;95:1701–8.

18. Neer CS, Craig EV and Fukuda H. Cuff-tear arthropathy. J Bone Joint Surg Am 1983;65:1232–44.

19. Burkhart SS. Fluoroscopic comparison of kinematic patterns in massive rotator cuff tears. A suspension bridge model. Clin Orthop Relat Res 1992:144–52.

20. Burkhart SS. Reconciling the paradox of rotator cuff repair versus debridement: a unified biomechanical rationale for the treatment of rotator cuff tears. Arthroscopy 1994;10:4–19.

21. Nam JL, Takase-Minegishi K, Ramiro S, Chatzidionysiou K, Smolen JS, Heijde D van der, Bijlsma JW, Burmester GR, Dougados M, Scholte-Voshaar M, Vollenhoven R van and Landewé R. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2016 update of the EULAR recommendations for the management of rheumatoid arthritis. Ann Rheum Dis 2017;76:1113–36.

22. Cordtz R, Hawley S, Prieto-Alhambra D, Højgaard P, Zobbe K, Kristensen LE, Overgaard S, Odgaard A, Soussi BG and Dreyer L. Reduction in Upper Limb Joint Surgery among Rheumatoid Arthritis Patients: An Interrupted Time Series Analysis using Danish Health Care Registers. Arthritis Care Res (Hoboken) 2019.

23. Leroux TS, Basques BA, Saltzman BM, Nicholson GP, Romeo AA and Verma NN. Shoulder Arthroplasty in Patients with Rheumatoid Arthritis: A Population-Based Study Examining Utilization, Adverse Events, Length of Stay, and Cost. Am J Orthop 2018;47.

24. Sarris I, Weiser R and Sotereanos DG. Pathogenesis and treatment of osteonecrosis of the shoulder. Orthop Clin North Am 2004;35:397–404, xi.

25. Gruson KI and Kwon YW. Atraumatic osteonecrosis of the humeral head. Bull NYU Hosp Jt Dis 2009;67:6–14.

26. Hattrup SJ and Cofield RH. Osteonecrosis of the humeral head: relationship of disease stage, extent, and cause to natural history. J Shoulder Elbow Surg 1999;8:559–64.

27. Schoch BS, Barlow JD, Schleck C, Cofield RH and Sperling JW. Shoulder arthroplasty for post-traumatic osteonecrosis of the humeral head. J Shoulder Elbow Surg 2016;25:406–12.

28. Boileau P, Trojani C, Walch G, Krishnan SG, Romeo A and Sinnerton R. Shoulder arthroplasty for the treatment of the sequelae of fractures of the proximal humerus. J Shoulder Elbow Surg 2001;10:299–308.

29. Audigé L, Graf L, Flury M, Schneider MM and Müller AM. Functional improvement is sustained following anatomical and reverse shoulder arthroplasty for fracture sequelae: a registry-based analysis. Arch Orthop Trauma Surg 2019.

30. Nowak LL, Vicente MR, McKee MD, Hall JA, Nauth A and Schemitsch EH. Orthopaedic surgeons' opinions surrounding the management of proximal humerus fractures: an international survey. Int Orthop 2017;41:1749–55.

31. Handoll HHG and Brorson S. Interventions for treating proximal humeral fractures in adults. Cochrane database Syst Rev 2015;11:CD000434.

32. Sebastia-Forcada E, Cebrian-Gomez R, Lizaur-Utrilla A and Gil-Guillen V. Reverse shoulder arthroplasty versus hemiarthroplasty for acute proximal humeral fractures. A blinded, randomized, controlled, prospective study. J Shoulder Elbow Surg 2014;23:1419–26.

33. Torchia MT, Austin DC, Cozzolino N, Jacobowitz L and Bell JE. Acute versus delayed reverse total shoulder arthroplasty for the treatment of proximal humeral fractures in the elderly population: a systematic review and meta-analysis. J Shoulder Elbow Surg 2019;28:765–73.

34. Handoll H, Brealey S, Rangan A, Torgerson D, Dennis L, Armstrong A, Chuang LH, Cross B, Dumville J, Gardner S, Goodchild L, Hamilton S, Hewitt C, Madhok R, Maffulli N, Micklewright L, Wadsworth V, Wallace A, Williams J and Worthy G. Protocol for the ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial: a pragmatic multi-centre randomised controlled trial of surgical versus non-surgical treatment for proximal fracture of the humerus in adults. BMC Musculoskelet Disord 2009;10:140.

35. Kaa AKS, Jørgensen PH, Søjbjerg JO and Johannsen HV. Reverse shoulder replacement after resection of the proximal humerus for bone tumours. Bone Joint J 2013;95-B:1551–5.

36. Nota S, Teunis T, Kortlever J, Ferrone M, Ready J, Gebhardt M, Raskin K, Hornicek F, Schwab J and Lozano Calderon S. Functional Outcomes and Complications After Oncologic Reconstruction of the Proximal Humerus. J Am Acad Orthop Surg 2018;26:403–9.

37. Nowak LL, Hall J, McKee MD and Schemitsch EH. A higher reoperation rate following arthroplasty for failed fixation versus primary arthroplasty for the treatment of proximal humeral fractures. Bone Joint J 2019;101-B:1272–9.

38. Cowling PD, Holland P, Kottam L, Baker P and Rangan A. Risk factors associated with intraoperative complications in primary shoulder arthroplasty. Acta Orthop 2017;88:587–91.

39. Clark JC, Simon P, Clark RE, Christmas KN, Allert JW, Streit JJ, Mighell MA, Hess A, Stone J and Frankle MA. The influence of patient- and surgeon-specific factors on operative duration and early postoperative outcomes in shoulder arthroplasty. J Shoulder Elbow Surg 2017;26:1011–6.

40. Ventafridda V, Saita L, Ripamonti C and De Conno F. WHO guidelines for the use of analgesics in cancer pain. Int J Tissue React 1985;7:93–6.

41. Ballantyne JC, Kalso E and Stannard C. WHO analgesic ladder: a good concept gone astray. BMJ 2016;352:i20.

42. Leopoldino AO, Machado GC, Ferreira PH, Pinheiro MB, Day R, McLachlan AJ, Hunter DJ and Ferreira ML. Paracetamol versus placebo for knee and hip osteoarthritis. Cochrane database Syst Rev 2019;2:CD013273.

43. Zhang W, Jones A and Doherty M. Does paracetamol (acetaminophen) reduce the pain of osteoarthritis?: a meta-analysis of randomised controlled trials. Ann Rheum Dis 2004;63:901–7.

44. Lanza FL, Codispoti JR and Nelson EB. An endoscopic comparison of gastroduodenal injury with over-the-counter doses of ketoprofen and acetaminophen. Am J Gastroenterol 1998;93:1051–4.

45. Nash DM, Markle-Reid M, Brimble KS, McArthur E, Roshanov PS, Fink JC, Weir MA and Garg AX. Nonsteroidal anti-inflammatory drug use and risk of acute kidney injury and hyperkalemia in older adults: a population-based study. Nephrol Dial Transplant 2019;34:1145–54.

46. McNicol ED, Ferguson MC and Hudcova J. Patient controlled opioid analgesia versus non-patient controlled opioid analgesia for postoperative pain. Cochrane database Syst Rev 2015.

47. Wiffen PJ, Wee B and Moore RA. Oral morphine for cancer pain. Cochrane database Syst Rev 2013.

48. Toupin April K, Bisaillon J, Welch V, Maxwell LJ, Jüni P, Rutjes AW, Husni ME, Vincent J, El Hindi T, Wells GA and Tugwell P. Tramadol for osteoarthritis. Cochrane database Syst Rev 2019;5:CD005522.

49. Els C, Jackson TD, Hagtvedt R, Kunyk D, Sonnenberg B, Lappi VG and Straube S. High-dose opioids for chronic non-cancer pain: an overview of Cochrane Reviews. Cochrane database Syst Rev 2017;10:CD012299.

50. Els C, Jackson TD, Kunyk D, Lappi VG, Sonnenberg B, Hagtvedt R, Sharma S, Kolahdooz F and Straube S. Adverse events associated with medium- and long-term use of opioids for chronic non-cancer pain: an overview of Cochrane Reviews. Cochrane database Syst Rev 2017;10:CD012509.

51. Dowell D, Haegerich TM and Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain—United States, 2016. JAMA 2016;315:1624.

52. Straube S, Derry S, Moore RA, Wiffen PJ and McQuay HJ. Single dose oral gabapentin for established acute postoperative pain in adults. Cochrane database Syst Rev 2010:CD008183.

53. Wiffen PJ, Derry S, Bell RF, Rice AS, Tölle TR, Phillips T and Moore RA. Gabapentin for chronic neuropathic pain in adults. Cochrane database Syst Rev 2017;6:CD007938.

54. Torjesen I. Pregabalin and gabapentin: what impact will reclassification have on doctors and patients? BMJ 2019;364:l1107.

55. Green S, Buchbinder R and Hetrick SE. Physiotherapy interventions for shoulder pain. Cochrane database Syst Rev 2003:CD004258.

56. Lin I, Wiles L, Waller R, Goucke R, Nagree Y, Gibberd M, Straker L, Maher CG and O'Sullivan PPB. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: systematic review. Br J Sports Med 2019:bjsports–2018–099878.

57. Guo JJ, Wu K, Guan H, Zhang L, Ji C, Yang H and Tang T. Three-Year Follow-up of Conservative Treatments of Shoulder Osteoarthritis in Older Patients. Orthopedics 2016;39:e634–e641.

58. Buchbinder R, Green S and Youd JM. Corticosteroid injections for shoulder pain. Cochrane database Syst Rev 2003:CD004016.

59. Jüni P, Hari R, Rutjes AW, Fischer R, Silletta MG, Reichenbach S and Costa BR da. Intra-articular corticosteroid for knee osteoarthritis. Cochrane database Syst Rev 2015:CD005328.

60. Gross C, Dhawan A, Harwood D, Gochanour E and Romeo A. Glenohumeral joint injections: a review. Sports Health 2013;5:153–9.

61. Brockmeier SF and Shaffer BS. Viscosupplementation therapy for osteoarthritis. Sports Med Arthrosc Rev 2006;14:155–62.

62. Zhang B, Thayaparan A, Horner N, Bedi A, Alolabi B and Khan M. Outcomes of hyaluronic acid injections for glenohumeral osteoarthritis: a systematic review and meta-analysis. J Shoulder Elbow Surg 2019;28:596–606.

63. Schoch B, Schleck C, Cofield RH and Sperling JW. Shoulder arthroplasty in patients younger than 50 years: minimum 20-year follow-up. J Shoulder Elbow Surg 2015;24:705–10.

64. Sperling JW, Antuna SA, Sanchez-Sotelo J, Schleck C and Cofield RH. Shoulder arthroplasty for arthritis after instability surgery. J Bone Joint Surg Am 2002;84:1775–81.

65. Cole BJ, Yanke A and Provencher MT. Nonarthroplasty alternatives for the treatment of glenohumeral arthritis. J Shoulder Elbow Surg 2007;16:S231–S240.

66. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Ashton CM and Wray NP. A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee. N Engl J Med 2002;347:81–8.

67. Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ, Feagan BG, Donner A, Griffin SH, D'Ascanio LM, Pope JE and Fowler PJ. A Randomized Trial of Arthroscopic Surgery for Osteoarthritis of the Knee. N Engl J Med 2008;359:1097–107.

68. Katz JN, Brophy RH, Chaisson CE, Chaves L de, Cole BJ, Dahm DL, Donnell-Fink LA, Guermazi A, Haas AK, Jones MH, Levy BA, Mandl LA, Martin SD, Marx RG, Miniaci A, Matava MJ, Palmisano J, Reinke EK, Richardson BE, Rome BN, Safran-Norton CE, Skoniecki DJ, Solomon DH, Smith MV, Spindler KP, Stuart MJ, Wright J, Wright RW and Losina E. Surgery versus Physical Therapy for a Meniscal Tear and Osteoarthritis. N Engl J Med 2013;368:1675–84.

69. Sihvonen R, Paavola M, Malmivaara A, Itälä A, Joukainen A, Nurmi H, Kalske J, Järvinen TL and Finnish Degenerative Meniscal Lesion Study (FIDELITY) Group. Arthroscopic Partial Meniscectomy versus Sham Surgery for a Degenerative Meniscal Tear. N Engl J Med 2013;369:2515–24.

70. Sihvonen R, Paavola M, Malmivaara A, Itälä A, Joukainen A, Nurmi H, Kalske J, Ikonen A, Järvelä T, Järvinen TAH, Kanto K, Karhunen J, Knifsund J, Kröger H, Kääriäinen T, Lehtinen J, Nyrhinen J, Paloneva J, Päiväniemi O, Raivio M, Sahlman J, Sarvilinna R, Tukiainen S, Välimäki VV, Äärimaa V, Toivonen P, Järvinen TLN and FIDELITY (Finnish Degenerative Meniscal Lesion Study) Investigators. Arthroscopic partial meniscectomy versus placebo surgery for a degenerative meniscus tear: a 2-year follow-up of the randomised controlled trial. Ann Rheum Dis 2018;77:188–95.

71. Millett PJ and Gaskill TR. Arthroscopic Management of Glenohumeral Arthrosis: Humeral Osteoplasty, Capsular Release, and Arthroscopic Axillary Nerve Release as a Joint-Preserving Approach. Arthroscopy 2011;27:1296–303.

72. Mitchell JJ, Horan MP, Greenspoon JA, Menge TJ, Tahal DS and Millett PJ. Survivorship and Patient-Reported Outcomes After Comprehensive Arthroscopic Management of Glenohumeral Osteoarthritis. Am J Sports Med 2016;44:3206–13.

73. Weinstein DM, Bucchieri JS, Pollock RG, Flatow EL and Bigliani LU. Arthroscopic debridement of the shoulder for osteoarthritis. Arthroscopy 2000;16:471–6.

74. Richards DP and Burkhart SS. Arthroscopic Debridement and Capsular Release for Glenohumeral Osteoarthritis. Arthroscopy 2007;23:1019–22.

75. Van Thiel GS, Sheehan S, Frank RM, Slabaugh M, Cole BJ, Nicholson GP, Romeo AA and Verma NN. Retrospective analysis of arthroscopic management of glenohumeral degenerative disease. Arthroscopy 2010;26:1451–5.

76. Barlow J and Abboud J. Surgical options for the young patient with glenohumeral arthritis. Int J Shoulder Surg 2016;10:28.

77. Puskas GJ, Meyer DC, Lebschi JA and Gerber C. Unacceptable failure of hemiarthroplasty combined with biological glenoid resurfacing in the treatment of glenohumeral arthritis in the young. J Shoulder Elbow Surg 2015;24:1900–7.

78. Hammond LJ, Lin EC, Harwood DP, Juhan TW, Gochanour E, Klosterman EL, Cole BJ, Nicholson GP, Verma NN and Romeo AA. Clinical outcomes of hemiarthroplasty and biological resurfacing in patients aged younger than 50 years. J Shoulder Elbow Surg 2013;22:1345–51.

79. Matsen FA, Whitson A, Jackins SE, Neradilek MB, Warme WJ and Hsu JE. Ream and run and total shoulder: patient and shoulder characteristics in five hundred forty-four concurrent cases. Int Orthop 2019.

80. Sellers TR, Abdelfattah A and Frankle MA. Massive Rotator Cuff Tear: When to Consider Reverse Shoulder Arthroplasty. Curr Rev Musculoskelet Med 2018;11:131–40.

81. Eichinger JK. Editorial Commentary: Treating Irreparable Rotator Cuff Tears With a Patch or Balloon: Is It All a Bunch of Hot Air? Arthroscopy 2019;35:390–1.

82. Moon AS, Patel HA, Ithurburn MP, Brabston EW, Ponce BA and Momaya AM. Subacromial Spacer Implantation for the Treatment of Massive Irreparable Rotator Cuff Tears: A Systematic Review. Arthroscopy 2019;35:607–14.

83. Mihata T, Lee TQ, Watanabe C, Fukunishi K, Ohue M, Tsujimura T and Kinoshita M. Clinical Results of Arthroscopic Superior Capsule Reconstruction for Irreparable Rotator Cuff Tears. Arthroscopy 2013;29:459–70.

84. Burkhart SS and Hartzler RU. Superior Capsular Reconstruction Reverses Profound Pseudoparalysis in Patients With Irreparable Rotator Cuff Tears and Minimal or No Glenohumeral Arthritis. Arthroscopy 2019;35:22–8.

85. Kane TP, Rogers P, Hazelgrove J, Wimsey S and Harper GD. Pulsed radiofrequency applied to the suprascapular nerve in painful cuff tear arthropathy. J Shoulder Elbow Surg 2008;17:436–40.

86. Richards RR, Beaton D and Hudson AR. Shoulder arthrodesis with plate fixation: Functional outcome analysis. J Shoulder Elbow Surg 1993;2:225–39.

87. Vastamäki M. Shoulder arthrodesis for paralysis and arthrosis. Acta Orthop Scand 1987;58:549–53.

88. Wagner ER, McLaughlin R, Sarfani S, Cofield RH, Sperling JW, Sanchez-Sotelo J and Elhassan BT. Long-Term Outcomes of Glenohumeral Arthrodesis. J Bone Joint Surg Am 2018;100:598–604.

89. Rispoli DM, Sperling JW, Athwal GS, Schleck CD and Cofield RH. Pain relief and functional results after resection arthroplasty of the shoulder. J Bone Joint Surg Br 2007;89-B:1184–7.

90. Muh SJ, Streit JJ, Wanner JP, Lenarz CJ, Shishani Y, Rowland DY, Riley C, Nowinski RJ, Edwards TB and Gobezie R. Early follow-up of reverse total shoulder arthroplasty in patients sixty years of age or younger. J Bone Joint Surg Am 2013;95:1877–83.

91. Eynon-Lewis NJ, Ferry D and Pearse MF. Themistocles Gluck: an unrecognised genius. BMJ 1992;305:1534–6.
92. Péan JE. On Prosthetic Methods Intended to Repair Bone Fragments. Clin Orthop Relat Res 1973;94.
93. Baron R and Senn L. [Acrylic prosthesis for the shoulder]. Presse Med 1951;59:1480.
94. Richard A. [Malformation of unknown origin of the right humeral head; loss of function; resection and acrylic prosthesis]. Mémoires l'Académie Chir 1950;76:821–3.
95. Richard A, Judet R and Rene L. [Reconstruction of the humerus with acrylic prosthesis with special reference to fractures-luxations]. J Chir (Paris) 1952;68:537–47.
96. MacAusland W. Nylon prosthesis in lesions of the shoulder, elbow and finger. Am J Surg 1953;85:164–73.
97. Jackson Burrows H. Major prosthetic replacement of bone: lessons learnt in seventeen years. J Bone Joint Surg Br 1968;50-B:225–6.
98. Spencer R and Skirving AP. Silastic interposition arthroplasty of the shoulder. J Bone Joint Surg Br 1986;68:375–7.
99. Krueger F. A vitallium replica arthroplasty on the shoulder; a case report of aseptic necrosis of the proximal end of the humerus. Surgery 1951;30:1005–11.
100. Neer CS, Brown Jr. TH and McLaughlin HL. Fracture of the neck of the humerus with dislocation of the head fragment. Am J Surg 1953;85:252–8.
101. Neer C. Articular replacement for the humeral head. J Bone Joint Surg Am 1955;37-A:215–28.
102. Neer CS. Replacement arthroplasty for glenohumeral osteoarthritis. J Bone Joint Surg Am 1974;56:1–13.
103. Neer CS, Watson KC and Stanton FJ. Recent experience in total shoulder replacement. J Bone Joint Surg Am 1982;64:319–37.
104. Franklin JL, Barrett WP, Jackins SE and Matsen FA. Glenoid loosening in total shoulder arthroplasty. Association with rotator cuff deficiency. J Arthroplasty 1988;3:39–46.
105. Schoch BS, Wright TW, Zuckerman JD, Bolch C, Flurin PH, Roche C and King JJ. Glenoid component lucencies are associated with poorer patient-reported outcomes following anatomic shoulder arthroplasty. J Shoulder Elbow Surg 2019;28:1956–63.
106. McLendon P, Schoch B, Cofield R, Sanchez-Sotelo J and Sperling J. Does Prolonged Use of Walkers in Shoulder Arthroplasty Patients Lead to Accelerated Failure Rates? Cureus 2019;11:e5890.
107. Welsher A, Gohal C, Madden K, Miller B, Bedi A, Alolabi B and Khan M. A comparison of pegged vs. keeled glenoid components regarding functional and radiographic outcomes in anatomic total shoulder arthroplasty: a systematic review and meta-analysis. JSES open access 2019;3:136–144.e1.
108. Boileau P, Moineau G, Morin-Salvo N, Avidor C, Godenèche A, Lévigne C, Baba M and Walch G. Metal-backed glenoid implant with polyethylene insert is not a viable long-term therapeutic option. J Shoulder Elbow Surg 2015;24:1534–43.
109. Post M, Haskell SS and Jablon M. Total shoulder replacement with a constrained prosthesis. J Bone Joint Surg Am 1980;62:327–35.
110. Post M and Jablon M. Constrained total shoulder arthroplasty. Long-term follow-up observations. Clin Orthop Relat Res 1983:109–16.

111. Lettin AW, Copeland SA and Scales JT. The Stanmore total shoulder replacement. J Bone Joint Surg Br 1982;64:47–51.

112. Kessel L and Bayley I. Prosthetic replacement of shoulder joint: preliminary communication. J R Soc Med 1979;72:748–52.

113. Wretenberg PF and Wallensten R. The Kessel total shoulder arthroplasty. A 13- to 16-year retrospective followup. Clin Orthop Relat Res 1999:100–3.

114. Middleton C, Uri O, Phillips S, Barmpagiannis K, Higgs D, Falworth M, Bayley I and Lambert S. A reverse shoulder arthroplasty with increased offset for the treatment of cuff-deficient shoulders with glenohumeral arthritis. Bone Joint J 2014;96-B:936–42.

115. Maclean S, Malik SS, Evans S, Gregory J and Jeys L. Reverse shoulder endoprosthesis for pathologic lesions of the proximal humerus: a minimum 3-year follow-up. J Shoulder Elbow Surg 2017;26:1990–4.

116. Grammont PM and Baulot E. Delta shoulder prosthesis for rotator cuff rupture. Orthopedics 1993;16:65–8.

117. Wall B, Nove-Josserand L, O'Connor DP, Edwards TB and Walch G. Reverse total shoulder arthroplasty: a review of results according to etiology. J Bone Joint Surg Am 2007;89:1476–85.

118. Favard L, Levigne C, Nerot C, Gerber C, De Wilde L and Mole D. Reverse Prostheses in Arthropathies With Cuff Tear: Are Survivorship and Function Maintained Over Time? Clin Orthop Relat Res 2011;469:2469–75.

119. Guery J, Favard L, Sirveaux F, Oudet D, Mole D and Walch G. Reverse total shoulder arthroplasty. Survivorship analysis of eighty replacements followed for five to ten years. J Bone Joint Surg Am 2006;88:1742–7.

120. Jónsson E, Egund N, Kelly I, Rydholm U and Lidgren L. Cup arthroplasty of the rheumatoid shoulder. Acta Orthop Scand 1986;57:542–6.

121. Levy O and Copeland SA. Cementless surface replacement arthroplasty of the shoulder. 5- to 10-year results with the Copeland mark-2 prosthesis. J Bone Joint Surg Br 2001;83:213–21.

122. Huguet D, DeClercq G, Rio B, Teissier J, Zipoli B and Group T. Results of a new stemless shoulder prosthesis: radiologic proof of maintained fixation and stability after a minimum of three years' follow-up. J Shoulder Elbow Surg 2010;19:847–52.

123. Garret J, Godeneche A, Boileau P, Molé D, Etzner M, Favard L, Levigne C, Sirveaux F, Gauci MO, Dezaly C and Walch G. Pyrocarbon interposition shoulder arthroplasty: preliminary results from a prospective multicenter study at 2 years of follow-up. J Shoulder Elbow Surg 2017;26:1143–51.

124. Craig RS and Rees JL. Providing better evidence in orthopaedic surgery: RCTs, registries or both? (Opinion piece). BMJ 2019.

125. Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gibbons CLM, Ostlere S, Athanasou N, Gill HS and Murray DW. Pseudotumours associated with metal-on-metal hip resurfacings. J Bone Joint Surg Br 2008;90-B:847–51.

126. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, Nicholl J, Balliol Collaboration, Aronson JK, Barkun JS, Blazeby JM, Boutron IC, Campbell WB, Clavien PA, Cook JA, Ergina PL, Feldman LS, Flum DR, Maddern GJ, Nicholl J, Reeves BC, Seiler CM, Strasberg SM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, Leval M de, Deeks J, Ergina PL, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Meakins J, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U and Vandenbroucke J. No surgical innovation without evaluation: the IDEAL recommendations. Lancet 2009;374:1105–12.

127. Wright JG, Swiontkowski MF and Heckman JD. Introducing levels of evidence to the journal. J Bone Joint Surg Am 2003;85-A:1–3.

128. Rangan A, Handoll H, Brealey S, Jefferson L, Keding A, Martin BC, Goodchild L, Chuang LH, Hewitt C, Torgerson D and PROFHER Trial Collaborators. Surgical vs Nonsurgical Treatment of Adults With Displaced Fractures of the Proximal Humerus. JAMA 2015;313:1037.

129. Carr A, Cooper C, Campbell MK, Rees J, Moser J, Beard DJ, Fitzpatrick R, Gray A, Dawson J, Murphy J, Bruhn H, Cooper D and Ramsay C. Effectiveness of open and arthroscopic rotator cuff repair (UKUFF). Bone Joint J 2017;99-B:107–15.

130. Beard DJ, Rees JL, Cook JA, Rombach I, Cooper C, Merritt N, Shirkey BA, Donovan JL, Gwilym S, Savulescu J, Moser J, Gray A, Jepson M, Tracey I, Judge A, Wartolowska K, Carr AJ, Ahrens P, Baldwick C, Brinsden M, Brownlow H, Burton D, Butt MS, Carr A, Charalambous CP, Conboy V, Dennell L, Donaldson O, Drew S, Dwyer A, Gidden D, Hallam P, Kalogrianitis S, Kelly C, Kulkarni R, Matthews T, McBirnie J, Patel V, Peach C, Roberts C, Robinson D, Rosell P, Rossouw D, Senior C, Singh B, Sjolin S, Taylor G, Venkateswaran B and Woods D. Arthroscopic subacromial decompression for subacromial shoulder pain (CSAW): a multicentre, pragmatic, parallel group, placebo-controlled, three-group, randomised surgical trial. Lancet 2017.

131. Sacks H, Chalmers TC and Smith H. Randomized versus historical controls for clinical trials. Am J Med 1982;72:233–40.

132. Concato J, Shah N and Horwitz RI. Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs. N Engl J Med 2000;342:1887–92.

133. Horwitz RI. Complexity and contradiction in clinical trial research. Am J Med 1987;82:498–510.

134. Pibouleau L, Boutron I, Reeves BC, Nizard R and Ravaud P. Applicability and generalisability of published results of randomised controlled trials and non-randomised studies evaluating four orthopaedic procedures: methodological systematic review. BMJ 2009;339:b4538.

135. Inacio MC, Paxton EW and Dillon MT. Understanding Orthopaedic Registry Studies. J Bone Joint Surg Am 2016;98:e3.

136. Graves SE. The value of arthroplasty registry data. Acta Orthop 2010;81:8–9.

137. Paxton EW, Prentice HA, Inacio MC, Dillon MT, Page RS, Rasmussen JV, Salomonsson B, Steiger R de and Navarro RA. Are we throwing the baby out with the bath water? J Shoulder Elbow Surg 2017.

138. Kuhn JE and Mallon WJ. Quantitative Observational Database Research in the Journal of Shoulder and Elbow Surgery. J Shoulder Elbow Surg 2016;25:1733–4.

139. Zhan C and Miller MR. Administrative data based patient safety research: a critical review. Qual Saf Health Care 2003;12:58ii–63.

140. Khan T. Survival analysis of time-to-event data in orthopaedic surgery. Bone & joint 360 2017;6:37–9.

141. Rampal R and Cowling PD. Outcomes of shoulder surgery. Orthop Trauma 2018;32:200–6.

142. Buchbinder R, Page MJ, Huang H, Verhagen AP, Beaton D, Kopkow C, Lenza M, Jain NB, Richards B, Richards P, Voshaar M, Windt D van der, Gagnier JJ and Shoulder Core Outcome Set Special Interest Group. A Preliminary Core Domain Set for Clinical Trials of Shoulder Disorders: A Report from the OMERACT 2016 Shoulder Core Outcome Set Special Interest Group. J Rheumatol 2017;44:1880–3.

143. Ramiro S, Page MJ, Whittle SL, Huang H, Verhagen AP, Beaton D, Richards P, Scholte-Voshaar M, Shea B, Windt D van der, Kopkow C, Lenza M, Jain N, Richards B, Hill CL, Gill TK, Koes B, Foster N, Conaghan PG, Smith TO, Malliaras P, Roe Y, Gagnier JJ and Buchbinder R. The Omeract Core Domain Set for Clinical Trials of Shoulder Disorders. J Rheumatol 2019:jrheum.181070.

144. Page MJ, Huang H, Verhagen AP, Gagnier JJ and Buchbinder R. Outcome Reporting in Randomized Trials for Shoulder Disorders: Literature Review to Inform the Development of a Core Outcome Set. Arthritis Care Res (Hoboken) 2018;70:252–9.

145. Constant CR and Murley AH. A clinical method of functional assessment of the shoulder. Clin Orthop Relat Res 1987:160–4.

146. Amstutz HC, Sew Hoy AL and Clarke IC. UCLA anatomic total shoulder arthroplasty. Clin Orthop Relat Res 1981:7–20.

147. Varghese M, Lamb J, Rambani R and Venkateswaran B. The use of shoulder scoring systems and outcome measures in the UK. Ann R Coll Surg Engl 2014;96:590–2.

148. Dawson J, Fitzpatrick R and Carr A. Questionnaire on the perceptions of patients about shoulder surgery. J Bone Joint Surg Br 1996;78:593–600.

149. Dawson J, Rogers K, Fitzpatrick R and Carr A. The Oxford shoulder score revisited. Arch Orthop Trauma Surg 2009;129:119–23.

150. Dawson J, Harris KK, Doll H, Fitzpatrick R and Carr A. A comparison of the Oxford shoulder score and shoulder pain and disability index: factor structure in the context of a large randomized controlled trial. Patient Relat Outcome Meas 2016;Volume 7:195–203.

151. Lo I, Griffin S and Kirkley A. The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder (WOOS) index. Osteoarthritis Cartilage 2001;9:771–8.

152. Booker S, Alfahad N, Scott M, Gooding B and Wallace WA. Use of scoring systems for assessing and reporting the outcome results from shoulder surgery and arthroplasty. World J Orthop 2015;6:244.

153. Bafus BT, Hughes RE, Miller BS and Carpenter JE. Evaluation of utility in shoulder pathology: Correlating the American Shoulder and Elbow Surgeons and Constant scores to the EuroQoL. World J Orthop 2012;3:20–4.

154. Lippitt S, Harryman II D and Matsen III F. A practical tool for evaluating function: the Simple Shoulder Test. In: The shoulder: a balance of mobility and stability. Ed. by Matsen III F, Fu F and Hawkins R. Rosemont IL: The American Academy of Orthopaedic Surgeons, 1993:545–59.

155. Hudak PL, Amadio PC, Bombardier C, Beaton D, Cole D, Davis A, Hawker G, Katz JN, Makela M, Marx RG, Punnett L and Wright J. Development of an upper extremity outcome measure: The DASH (disabilities of the arm, shoulder, and hand). The Upper Extremity Collaborative Group (UECG). Am J Ind Med 1996;29:602–8.

156. Hapuarachchi KS and Poon PC. A correlation study of the American Shoulder and Elbow Society Score and the Oxford Shoulder Score with the use of regression analysis to predict one score from the other in patients undergoing reverse shoulder joint arthroplasty for cuff tear arthropathy. Shoulder Elbow 2014;6:81–9.

157. Baker P, Nanda R, Goodchild L, Finn P and Rangan A. A comparison of the Constant and Oxford shoulder scores in patients with conservatively treated proximal humeral fractures. J Shoulder Elbow Surg 2008;17:37–41.

158. Amundsen A, Rasmussen JV, Olsen BS and Brorson S. Mortality after shoulder arthroplasty: 30-day, 90-day, and 1-year mortality after shoulder replacement—5853 primary operations reported to the Danish Shoulder Arthroplasty Registry. J Shoulder Elbow Surg 2016;25:756–62.

159. Liebs TR, Splietker F and Hassenpflug J. Is a Revision a Revision? An Analysis of National Arthroplasty Registries' Definitions of Revision. Clin Orthop Relat Res 2015;473:3421–30.

160. NJR/BESS Joint Statement on the definition of revision joint replacement. 2018. URL: http://www.odep.org.uk/Portals/0/ODEPStatements/ODEP%7B%5C_%7DShoulders-Revision % 7B % 5C _ %7DDefinition % 7B % 5C _ %7D2018 . pdf Last accessed on 20/08/2019.

161. NJR. National Joint Registry for England, Wales and Northern Ireland. 16th Annual Report. 2019.

162. Singh JA, Sperling JW and Cofield RH. Revision surgery following total shoulder arthroplasty: analysis of 2588 shoulders over three decades (1976 to 2008). J Bone Joint Surg Br 2011;93:1513–7.

163. Singh JA, Sperling JW and Cofield RH. Risk factors for revision surgery after humeral head replacement: 1,431 shoulders over 3 decades. J Shoulder Elbow Surg 2012;21:1039–44.

164. US FDA: Definition of serious adverse event. URL: https : / / www . fda . gov / safety / reporting - serious - problems - fda / what - serious - adverse - event Last accessed on 21/08/2019.

165. NIHR: Definition of serious adverse event. URL: https : / / www . invo . org . uk / posttypejargon / serious - adverse - event - sae - or - serious - adverse - reaction - sar/ Last accessed on 21/08/2019.

166. Jiang JJ, Toor AS, Shi LL and Koh JL. Analysis of perioperative complications in patients after total shoulder arthroplasty and reverse total shoulder arthroplasty. J Shoulder Elbow Surg 2014;23:1852–9.

167. Young BL, Menendez ME, Baker DK and Ponce BA. Factors associated with in-hospital pulmonary embolism after shoulder arthroplasty. J Shoulder Elbow Surg 2015;24:e271–8.

168. Ponce BA, Oladeji LO, Rogers ME and Menendez ME. Comparative analysis of anatomic and reverse total shoulder arthroplasty: in-hospital outcomes and costs. J Shoulder Elbow Surg 2015;24:460–7.

169. Villacis D, Sivasundaram L, Pannell WC, Heckmann N, Omid R and Hatch GFR3. Complication rate and implant survival for reverse shoulder arthroplasty versus total shoulder arthroplasty: results during the initial 2 years. J Shoulder Elbow Surg 2016;25:927–35.

170. Cancienne JM, Brockmeier SF, Gulotta LV, Dines DM and Werner BC. Ambulatory Total Shoulder Arthroplasty: A Comprehensive Analysis of Current Trends, Complications, Readmissions, and Costs. J Bone Joint Surg Am 2017;99:629–37.

171. Newman JM, Stroud SG, Yang A, Sodhi N, Dixit A, Doran JP, Hayden AJ, Casagrande DJ and Mont MA. Total shoulder arthroplasty in octogenarians: Is there a higher risk of adverse outcomes? J Orthop 2018;15:671–5.

172. Jameson SS, James P, Howcroft DWJ, Serrano-Pedraza I, Rangan A, Reed MR and Candal-Couto J. Venous thromboembolic events are rare after shoulder surgery: Analysis of a national database. J Shoulder Elbow Surg 2011;20:764–70.

173. Anakwenze OA, O'Donnell EA, Jobin CM, Levine WN and Ahmad CS. Medical Complications and Outcomes After Total Shoulder Arthroplasty: A Nationwide Analysis. Am J Orthop 2018;47.

174. Audigé L, Flury M, Müller AM and Durchholz H. Complications associated with arthroscopic rotator cuff tear repair: definition of a core event set by Delphi consensus process. J Shoulder Elbow Surg 2016;25:1907–17.

175. Cowling P, Holland P, Kottam L, Baker P and Rangan A. Risk factors associated with intraoprative complications during primary shoulder arthroplasty surgery: a study using the national joint registry dataset. Shoulder Elbow 2016;8:S15.

176. Simmen B, Bachmann L, Drerup S, Schwyzer HK, Burkhart A and Goldhahn J. Development of a predictive model for estimating the probability of treatment success one year after total shoulder replacement – cohort study. Osteoarthritis Cartilage 2008;16:631–4.

177. Matsen FA, Russ SM, Vu PT, Hsu JE, Lucas RM and Comstock BA. What Factors are Predictive of Patient-reported Outcomes? A Prospective Study of 337 Shoulder Arthroplasties. Clin Orthop Relat Res 2016;474:2496–510.

178. Wong SE, Zhang AL, Berliner JL, Ma CB and Feeley BT. Preoperative patient-reported scores can predict postoperative outcomes after shoulder arthroplasty. J Shoulder Elbow Surg 2016;25:913–9.

179. Gowd AK, Agarwalla A, Amin NH, Romeo AA, Nicholson GP, Verma NN and Liu JN. Construct validation of machine learning in the prediction of short-term postoperative complications following total shoulder arthroplasty. J Shoulder Elbow Surg 2019.

180. Chen RE, Papuga MO, Nicandri GT, Miller RJ and Voloshin I. Preoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores predict postoperative outcome in total shoulder arthroplasty patients. J Shoulder Elbow Surg 2019;28:547–54.

181. Schwartz DG, Cottrell BJ, Teusink MJ, Clark RE, Downes KL, Tannenbaum RS and Frankle MA. Factors that predict postoperative motion in patients treated with reverse shoulder arthroplasty. J Shoulder Elbow Surg 2014;23:1289–95.

182. Werner BC, Wong AC, Mahony GT, Craig EV, Dines DM, Warren RF and Gulotta LV. Causes of poor postoperative improvement after reverse total shoulder arthroplasty. J Shoulder Elbow Surg 2016;25:e217–e222.

183. Swarup I, Henn CM, Nguyen JT, Dines DM, Craig EV, Warren RF, Gulotta LV and Henn III RF. Effect of pre-operative expectations on the outcomes following total shoulder arthroplasty. Bone Joint J 2017;99-B:1190–6.

184. Mahony GT, Werner BC, Chang B, Grawe BM, Taylor SA, Craig EV, Warren RF, Dines DM and Gulotta LV. Risk factors for failing to achieve improvement after anatomic total shoulder arthroplasty for glenohumeral osteoarthritis. J Shoulder Elbow Surg 2018;27:968–75.

185. Rauck RC, Swarup I, Chang B, Dines DM, Warren RF, Gulotta LV and Henn RF. Effect of preoperative patient expectations on outcomes after reverse total shoulder arthroplasty. J Shoulder Elbow Surg 2018;27:e323–e329.

186. Fevang BTS, Lie SA, Havelin LI, Skredderstuen A and Furnes O. Risk factors for revision after shoulder arthroplasty: 1,825 shoulder arthroplasties from the Norwegian Arthroplasty Register. Acta Orthop 2009;80:83–91.

187. Farng E, Zingmond D, Krenek L and Soohoo NF. Factors predicting complication rates after primary shoulder arthroplasty. J Shoulder Elbow Surg 2011;20:557–63.

188. Dillon MT, Inacio MCS, Burke MF, Navarro RA and Yian EH. Shoulder arthroplasty in patients 59 years of age and younger. J Shoulder Elbow Surg 2013;22:1338–44.

189. Wagner ER, Houdek MT, Schleck CD, Harmsen WS, Sanchez-Sotelo J, Cofield R, Elhassan BT and Sperling JW. The role age plays in the outcomes and complications of shoulder arthroplasty. J Shoulder Elbow Surg 2017;26:1573–80.

190. Lehtimaki K, Rasmussen JV, Mokka J, Salomonsson B, Hole R, Jensen SL and Aarimaa V. Risk and risk factors for revision after primary reverse shoulder arthroplasty for cuff tear arthropathy and osteoarthritis: a Nordic Arthroplasty Register Association study. J Shoulder Elbow Surg 2018.

191. Wagner ER, Houdek MT, Schleck C, Harmsen WS, Sanchez-Sotelo J, Cofield R, Sperling JW and Elhassan BT. Increasing Body Mass Index Is Associated with Worse Outcomes After Shoulder Arthroplasty. J Bone Joint Surg Am 2017;99:929–37.

192. Gauci MO, Bonnevialle N, Moineau G, Baba M, Walch G and Boileau P. Anatomical total shoulder arthroplasty in young patients with osteoarthritis. Bone Joint J 2018;100-B:485–92.

193. Chalmers PN, Gupta AK, Rahman Z, Bruce B, Romeo AA and Nicholson GP. Predictors of early complications of total shoulder arthroplasty. J Arthroplasty 2014;29:856–60.

194. McCormick F, Nwachukwu B, Kiriakopoulos ES, Schairer W, Provencher M and Levy J. In-hospital mortality risk for total shoulder arthroplasty: A comprehensive review of the medicare database from 2005 to 2011. Int J Shoulder Surg 2015;9:110.

195. Johnson CC, Sodha S, Garzon-Muvdi J, Petersen SA and McFarland EG. Does preoperative American Society of Anesthesiologists score relate to complications after total shoulder arthroplasty?. Clin Orthop Relat Res 2014;472:1589–96.

196. Anthony CA, Westermann RW, Gao Y, Pugely AJ, Wolf BR and Hettrich CM. What Are Risk Factors for 30-day Morbidity and Transfusion in Total Shoulder Arthroplasty? A Review of 1922 Cases. Clin Orthop Relat Res 2015;473:2099–105.

197. Singh A, Yian EH, Dillon MT, Takayanagi M, Burke MF and Navarro RA. The effect of surgeon and hospital volume on shoulder arthroplasty perioperative quality metrics. J Shoulder Elbow Surg 2014;23:1187–94.

198. Singh JA and Ramachandran R. Does hospital volume predict outcomes and complications after total shoulder arthroplasty in the US? Arthritis Care Res (Hoboken) 2015;67:885–90.

199. Mollon B, Mahure SA, Ding DY, Zuckerman JD and Kwon YW. The influence of a history of clinical depression on peri-operative outcomes in elective total shoulder arthroplasty: a ten-year national analysis. Bone Joint J 2016;98-B:818–24.

200. Anakwenze OA, Yehyawi T, Dillon MT, Paxton E, Navarro R and Singh A. Effect of Age on Outcomes of Shoulder Arthroplasty. Perm J 2017;21.

201. Li X, Veltre DR, Cusano A, Yi P, Sing D, Gagnier JJ, Eichinger JK, Jawa A and Bedi A. Insurance status affects postoperative morbidity and complication rate after shoulder arthroplasty. J Shoulder Elbow Surg 2017;26:1423–31.

202. Kim CY, Sivasundaram L, LaBelle MW, Trivedi NN, Liu RW and Gillespie RJ. Predicting adverse events, length of stay, and discharge disposition following shoulder arthroplasty: a comparison of the Elixhauser Comorbidity Measure and Charlson Comorbidity Index. J Shoulder Elbow Surg 2018.

203. Modi CS, Veillette CJH, Gandhi R, Perruccio AV and Rampersaud YR. Factors that influence the choice to undergo surgery for shoulder and elbow conditions. Clin Orthop Relat Res 2014;472:883–91.

204. National Institute for Health and Care Excellence. NICE Clinical Guideline 138: Patient experience in adult NHS services. 2012. URL: https://www.nice.org.uk/guidance/cg138.

205. Torrens C, Miquel J and Santana F. Do we really allow patient decision-making in rotator cuff surgery? A prospective randomized study. J Orthop Surg Res 2019;14:116.

206. Zipkin DA, Umscheid CA, Keating NL, Allen E, Aung K, Beyth R, Kaatz S, Mann DM, Sussman JB, Korenstein D, Schardt C, Nagi A, Sloane R and Feldstein DA. Evidence-Based Risk Communication. Ann Intern Med 2014;161:270.

207. Fortin JM, Hirota LK, Bond BE, O'Connor AM and Col NF. Identifying patient preferences for communicating risk estimates: a descriptive pilot study. BMC Med Inform Decis Mak 2001;1:2.

208. Godlee F. Balancing benefits and harms. BMJ 2013;346:f3666.

209. American Academy of Orthopaedic Surgeons. The Diagnosis and Treatment of Glenohumeral Arthritis: Guideline and Evidence Report. 2009. URL: https://www.aaos.org/uploadedFiles/PreProduction/Quality/Guidelines%7B%5C_%7Dand%7B%5C_%7DReviews/guidelines/glenohumal-joint-osteoarthritis-clinical-practice-guideline-4-24-19.pdf.

210. Smith R. The Cochrane collaboration at 20. BMJ 2013;347:f7383.

211. Higgins J, Lasserson T, Chandler J, Tovey D and Churchill R. Methodological Expectations of Cochrane Intervention Reviews. July 2019. 2019.

212. Craig RS, Goodier H, Singh JA, Hopewell S and Rees JL. Shoulder replacement surgery for osteoarthritis and rotator cuff tear arthropathy (Protocol). Cochrane database Syst Rev 2017:CD012879.

213. Veritas Health Innovation. Covidence systematic review software. Melbourne, Australia. URL: https://www.covidence.org/.

214. Higgins JP, Altman DG and Sterne JA. Assessing risk of bias in included studies. In: Cochrane handbook for systematic reviews of interventions. Ed. by Higgins JP and Green S. The Cochrane Collaboration, 2011. Chap. 8.

215. Rohatgi A. WebPlotDigitizer. San Francisco, CA, 2019. URL: https://automeris.io/WebPlotDigitizer.

216. Schünemann H, Oxman AD, Vist GE, Higgins JP, Deeks JJ and Glasziou P. Interpreting results and drawing conclusions. In: Cochrane handbook for systematic reviews of interventions. Ed. by Higgins JP and Green S. 5.1.0. The Cochrane Collaboration, 2011. Chap. 12.

217. Hao Q, Devji T, Zeraatkar D, Wang Y, Qasim A, Siemieniuk RAC, Vandvik PO, Lähdeoja T, Carrasco-Labra A, Agoritsas T and Guyatt G. Minimal important differences for improvement in shoulder condition patient-reported outcomes: a systematic review to inform a BMJ Rapid Recommendation. BMJ Open 2019;9:e028777.

218. Polk A, Rasmussen JV, Brorson S and Olsen BS. Reliability of patient-reported functional outcome in a joint replacement registry. Acta Orthop 2013;84:12–7.

219. Werner BC, Chang B, Nguyen JT, Dines DM and Gulotta LV. What Change in American Shoulder and Elbow Surgeons Score Represents a Clinically Important Change After Shoulder Arthroplasty? Clin Orthop Relat Res 2016;474:2672–81.

220. Higgins JP and Deeks JJ. Selecting studies and collecting data. In: Cochrane handbook for systematic reviews of interventions. Ed. by Higgins JP and Green S. 5.1.0. The Cochrane Collaboration, 2011. Chap. 7.

221. Deeks JJ, Higgins JP and Altman DG. Analysing data and undertaking meta-analyses. In: Cochrane handbook for systematic reviews of interventions. Ed. by Higgins JP and Green S. 5.1.0. Th Cochrane Collaboration, 2011. Chap. 9.

222. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ and GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–6.

223. Schünemann HJ, Oxman AD, Higgins JP, Vist GE, Glasziou P and Guyatt GH. Presenting results and 'Summary of findings' tables. In: Cochrane handbook for systematic reviews of interventions. Ed. by Higgins JP and Green S. 5.1.0. The Cochrane Collaboration, 2011. Chap. 11.

224. Kircher J, Wiedemann M, Magosch P, Lichtenberg S and Habermeyer P. Improved accuracy of glenoid positioning in total shoulder arthroplasty with intraoperative navigation: A prospective-randomized clinical study. J Shoulder Elbow Surg 2009;18:515–20.

225. Gartsman GM, Roddey TS and Hammerman SM. Shoulder arthroplasty with or without resurfacing of the glenoid in patients who have osteoarthritis. J Bone Joint Surg Am 2000;82:26–34.

226. Lo IK, Litchfield RB, Griffin S, Faber K, Patterson SD and Kirkley A. Quality-of-life outcome following hemiarthroplasty or total shoulder arthroplasty in patients with osteoarthritis. A prospective, randomized trial. J Bone Joint Surg Am 2005;87:2178.

227. Sandow MJ, David H and Bentall SJ. Hemiarthroplasty vs total shoulder replacement for rotator cuff intact osteoarthritis: how do they fare after a decade? J Shoulder Elbow Surg 2013;22:877–85.

228. Rasmussen JV, Olsen BS, Sorensen AK, Hróbjartsson A and Brorson S. Resurfacing hemiarthroplasty compared to stemmed hemiarthroplasty for glenohumeral osteoarthritis: a randomised clinical trial. Int Orthop 2015;39:263–9.

229. Uschok S, Magosch P, Moe M, Lichtenberg S and Habermeyer P. Is the stemless humeral head replacement clinically and radiographically a secure equivalent to standard stem humeral head replacement in the long-term follow-up? A prospective randomized trial. J Shoulder Elbow Surg 2017;26:225–32.

230. Boileau P, Avidor C, Krishnan SG, Walch G, Kempf JF and Molé D. Cemented polyethylene versus uncemented metal-backed glenoid components in total shoulder arthroplasty: A prospective, double-blind, randomized study. J Shoulder Elbow Surg 2002;11:351–9.

231. Edwards TB, Labriola JE, Stanley RJ, O'Connor DP, Elkousy HA and Gartsman GM. Radiographic comparison of pegged and keeled glenoid components using modern cementing techniques: A prospective randomized study. J Shoulder Elbow Surg 2010;19:251–7.

232. Gartsman GM, Elkousy HA, Warnock KM, Edwards TB and O'Connor DP. Radiographic comparison of pegged and keeled glenoid components. J Shoulder Elbow Surg 2005;14:252–7.

233. Gascoyne TC, McRae SM, Parashin SL, Leiter JR, Petrak MJ, Bohm ER and MacDonald PB. Radiostereometric analysis of keeled versus pegged glenoid components in total shoulder arthroplasty: a randomized feasibility study. Can J Surg 2017;60:273–9.

234. Nuttall D, Haines JF and Trail II. A study of the micromovement of pegged and keeled glenoid components compared using radiostereometric analysis. J Shoulder Elbow Surg 2007;16:S65–S70.

235. Rahme H, Mattsson P, Wikblad L, Nowak J and Larsson S. Stability of Cemented in-Line Pegged Glenoid Compared with Keeled Glenoid Components in Total Shoulder Arthroplasty. J Bone Joint Surg Am 2009;91:1965–72.

236. Litchfield RB, McKee MD, Balyk R, Mandel S, Holtby R, Hollinshead R, Drosdowech D, Wambolt SE, Griffin SH and McCormack R. Cemented versus uncemented fixation of humeral components in total shoulder arthroplasty for osteoarthritis of the shoulder: a prospective, randomized, double-blind clinical trial-A JOINTs Canada Project. J Shoulder Elbow Surg 2011;20:529–36.

237. Kwon YW and Zuckerman JD. Subscapularis-Sparing Total Shoulder Arthroplasty: A Prospective, Double-Blinded, Randomized Clinical Trial. Orthopedics 2019;42:e61–e67.

238. Lapner PLC, Sabri E, Rakhra K, Bell K and Athwal GS. Comparison of lesser tuberosity osteotomy to subscapularis peel in shoulder arthroplasty: a randomized controlled trial. J Bone Joint Surg Am 2012;94:2239–46.

239. Levine WN, Munoz J, Hsu S, Byram IR, Bigliani LU, Ahmad CS, Kongmalai P and Shillingford JN. Subscapularis tenotomy versus lesser tuberosity osteotomy during total shoulder arthroplasty for primary osteoarthritis: a prospective, randomized controlled trial. J Shoulder Elbow Surg 2019;28:407–14.

240. Mechlenburg I, Klebe TM, Døssing KV, Amstrup A, Søballe K and Stilling M. Evaluation of periprosthetic bone mineral density and postoperative migration of humeral head resurfacing implants: two-year results of a randomized controlled clinical trial. J Shoulder Elbow Surg 2014;23:1427–36.

241. Edwards TB, Trappey GJ, Riley C, O'Connor DP, Elkousy HA and Gartsman GM. Inferior tilt of the glenoid component does not decrease scapular notching in reverse shoulder arthroplasty: results of a prospective randomized study. J Shoulder Elbow Surg 2012;21:641–6.

242. Greiner S, Schmidt C, Herrmann S, Pauly S and Perka C. Clinical performance of lateralized versus non-lateralized reverse shoulder arthroplasty: a prospective randomized study. J Shoulder Elbow Surg 2015;24:1397–404.

243. Poon PC, Chou J, Young SW and Astley T. A Comparison of Concentric and Eccentric Glenospheres in Reverse Shoulder Arthroplasty. J Bone Joint Surg Am 2014;96:e138.

244. Gobezie R, Shishani Y, Lederman E and Denard PJ. Can a functional difference be detected in reverse arthroplasty with 135 degree versus 155 degree prosthesis for the treatment of rotator cuff arthropathy: a prospective randomized study. J Shoulder Elbow Surg 2019.

245. Kilian CM, Press CM, Smith KM, O'Connor DP, Morris BJ, Elkousy HA, Gartsman GM and Edwards TB. Radiographic and clinical comparison of pegged and keeled glenoid components using modern cementing techniques: midterm results of a prospective randomized study. J Shoulder Elbow Surg 2017;26:2078–85.

246. Lapner PL, Sabri E, Rakhra K, Bell K and Athwal GS. Healing rates and subscapularis fatty infiltration after lesser tuberosity osteotomy versus subscapularis peel for exposure during shoulder arthroplasty. J Shoulder Elbow Surg 2013;22:396–402.

247. Lazarus MD, Jensen KL, Southworth C and Matsen FA III. The radiographic evaluation of keeled and pegged glenoid component insertion. J Bone Joint Surg Am 2002;84A:1174–82.

248. Valenti P, Boutens D and Nerot C. Delta 3 reversed prosthesis for osteoarthritis with massive rotator cuff tear: long term results (>5 years)[French]. In: 2000 prothèses d'Epaule Recul 2 à 10 ans. Ed. by Walch G Molé D BP. Sauramps Médical, 2001.

249. Bryant D, Litchfield R, Sandow M, Gartsman GM, Guyatt G and Kirkley A. A comparison of pain, strength, range of motion, and functional outcomes after hemiarthroplasty and total shoulder arthroplasty in patients with osteoarthritis of the shoulder: a systematic review and meta-analysis. J Bone Joint Surg Am 2005;87:1947–56.

250. Duan X, Zhang W, Dong X, Liu M, Gao Y, Huang F, Li J and Xiang Z. Total shoulder arthroplasty versus hemiarthroplasty in patients with shoulder osteoarthritis: A meta-analysis of randomized controlled trials. Semin Arthritis Rheum 2013;43:297–302.

251. Radnay CS, Setter KJ, Chambers L, Levine WN, Bigliani LU and Ahmad CS. Total shoulder replacement compared with humeral head replacement for the treatment of primary glenohumeral osteoarthritis: a systematic review. J Shoulder Elbow Surg 2007;16:396–402.

252. Bekerom MPJ van den, Geervliet PC, Somford MP, Borne MPJ van den and Boer R. Total shoulder arthroplasty versus hemiarthroplasty for glenohumeral arthritis: A systematic review of the literature at long-term follow-up. Int J Shoulder Surg 2013;7:110–5.

253. Choate WS, Kwapisz A, Momaya AM, Hawkins RJ and Tokish JM. Outcomes for subscapularis management techniques in shoulder arthroplasty: a systematic review. J Shoulder Elbow Surg 2018;27:363–70.

254. Papadonikolakis A and Matsen FAI. Metal-Backed Glenoid Components Have a Higher Rate of Failure and Fail by Different Modes in Comparison with All-Polyethylene Components: A Systematic Review. J Bone Joint Surg Am 2014;96:1041–7.

255. Vavken P, Sadoghi P, Keudell A von, Rosso C, Valderrabano V and Muller AM. Rates of radiolucency and loosening after total shoulder arthroplasty with pegged or keeled glenoid components. J Bone Joint Surg Am 2013;95:215–21.

256. Erickson BJ, Frank RM, Harris JD, Mall N and Romeo AA. The influence of humeral head inclination in reverse total shoulder arthroplasty: a systematic review. J Shoulder Elbow Surg 2015;24:988–93.

257. NHS Digital. Hospital Episode Statistics. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics Last accessed on 26/07/2018.

258. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D and Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). Int J Epidemiol 2017;46:1093–1093i.

259. NJR. National Joint Registry for England and Wales. 6th Annual Report. 2009. URL: http://www.njrreports.org.uk.

260. NJR. National Joint Registry for England and Wales. 9th Annual Report. 2012. URL: www.njrreports.org.uk.

261. Lingard L and Goffe K. National Shoulder Replacement Dashboard: Findings from the Audit of the Quality of Shoulder Replacement Coding (personal communication). 2015.

262. NHS Digital. HES Data Dictionaries. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/hospital-episode-statistics-data-dictionary Last accessed on 29/08/2019.

263. NHS Digital. National data opt-out. URL: https://digital.nhs.uk/about-nhs-digital/our-work/keeping-patient-data-safe/how-we-look-after-your-health-and-care-information/your-information-choices/opting-out-of-sharing-your-confidential-patient-information Last accessed on 30/08/2019.

264. NHS Digital. Care Information Choices - March 2017. 2017.

265. Busby J, Purdy S and Hollingworth W. Calculating hospital length of stay using the Hospital Episode Statistics; a comparison of methodologies. BMC Health Serv Res 2017;17:347.

266. Charlson ME, Pompei P, Ales KL and MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.

267. Stagg V. CHARLSON: Stata module to calculate Charlson index of comorbidity. 2006.

268. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE and Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130–9.

269. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, Januel JM and Sundararajan V. Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries. Am J Epidemiol 2011;173:676–82.

270. Piel FB, Parkes BL, Daby H, Hansell AL and Elliott P. The challenge of opt-outs from NHS data: a small-area perspective. J Public Health (Oxf) 2018;40:e594–e600.

271. Porteous M. Highlights: Data quality audit. 2018. URL: http://www.njrreports.org.uk/NJR-Data-Quality-Audit-update Last accessed on 28/08/2019.

272. NHS Digital. National PROMs Programme. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms Last accessed on 28/08/2019.

273. Little R and Rubin D. Statistical Analysis with Missing Data. 2nd ed. Hoboken, NJ: Wiley, 2002.

274. Janssen KJ, Donders ART, Harrell FE, Vergouwe Y, Chen Q, Grobbee DE and Moons KG. Missing covariate data in medical research: To impute is better than to ignore. J Clin Epidemiol 2010;63:721–7.

275. Harrell Jr FE. Hmisc: Harrell Miscellaneous. R package version 4.1-1. 2018. URL: https://cran.r-project.org/package=Hmisc.

276. Morris TP, White IR and Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Med Res Methodol 2014;14:75.

277. Sullivan TR, Salter AB, Ryan P and Lee KJ. Bias and Precision of the "Multiple Imputation, Then Deletion" Method for Dealing With Missing Outcome Data. Am J Epidemiol 2015;182:528–34.

278. Sullivan TR, Lee KJ, Ryan P and Salter AB. Multiple imputation for handling missing outcome data when estimating the relative risk. BMC Med Res Methodol 2017;17:134.

279. Kontopantelis E, White IR, Sperrin M and Buchan I. Outcome-sensitive multiple imputation: a simulation study. BMC Med Res Methodol 2017;17:2.

280. White IR, Royston P and Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;30:377–99.

281. Rubin D. Multiple Imputation for Nonresponse in Surveys. New York, NY: Wiley, 1987.

282. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. 2017. URL: https://cran.r-project.org/package=tidyverse.

283. Thernau T. Survival: A Package for Survival Analysis in S. R package version 2.38. 2015. URL: https://cran.r-project.org/package=survival.

284. Jackson C. flexsurv: A Platform for Parametric Survival Modeling in R. J Stat Softw 2016;70.

285. Harrell Jr FE. rms: Regression Modeling Strategies. R package version 5.1-2. 2018. URL: https://cran.r-project.org/package=rms.

286. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag, 2016.

287. Kassambara A, Kosinski M and Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.5.999. 2019. URL: http://www.sthda.com/english/rpkgs/survminer/.

288. Neuwirth E. RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. 2014. URL: https://cran.r-project.org/package=RColorBrewer.

289. Office for National Statistics. Population Estimates for England. URL: www.nomisweb.org.uk.

290. Eurostat. Revision of the European Standard Population: Report of Eurostat's task force. 2013. URL: https://ec.europa.eu/eurostat/documents/3859598/5926869/KS-RA-13-028-EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f.

291. Esteve J, Benhamou E and Raymond L. Descriptive Epidemiology (IARC Scientific Publications No.128). Lyon: International Agency for Research on Cancer, 1994:67–8.

292. Office for National Statistics. Overview of the UK population: July 2017. 2017. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/july2017%7B%5C#%7Dthe-uks-population-is-getting-older-but-its-not-the-same-in-all-areas-of-the-uk.

293. Appleby J, Raleigh V, Frosini F, Bevan G, Gao H and Lyscom T. Variations in health care. London, 2011. URL: https://www.kingsfund.org.uk/sites/files/kf/Variations-in-health-care-good-bad-inexplicable-report-The-Kings-Fund-April-2011.pdf.

294. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ 2002;325:961–4.

295. Birkmeyer JD, Reames BN, McCulloch P, Carr AJ, Campbell WB and Wennberg JE. Understanding of regional variation in the use of surgery. Lancet 2013;382:1121–9.

296. Singh JA and Ramachandran R. Age-related differences in the use of total shoulder arthroplasty over time: use and outcomes. Bone Joint J 2015;97-B:1385–9.

297. Gadea F, Alami G, Pape G, Boileau P and Favard L. Shoulder hemiarthroplasty: outcomes and long-term survival analysis according to etiology. Orthop Traumatol Surg Res 2012;98:659–65.

298. Leung B, Horodyski M, Struk AM and Wright TW. Functional outcome of hemiarthroplasty compared with reverse total shoulder arthroplasty in the treatment of rotator cuff tear arthropathy. J Shoulder Elbow Surg 2012;21:319–23.

299. Young SW, Zhu M, Walker CG and Poon PC. Comparison of Functional Outcomes of Reverse Shoulder Arthroplasty with Those of Hemiarthroplasty in the Treatment of Cuff-Tear Arthropathy. J Bone Joint Surg Am 2013;95:910–5.

300. Holland P, Ingoe H, Rees JL and Rangan A. BESS arthroplasty research priorities survey. Shoulder Elbow 2016;8:S15.

301. Kontakis G, Koutras C, Tosounidis T and Giannoudis P. Early management of proximal humeral fractures with hemiarthroplasty: a systematic review. J Bone Joint Surg Br 2008;90:1407–13.

302. Craig RS, Lane JCE, Carr AJ, Furniss D, Collins GS and Rees JL. Serious adverse events and lifetime risk of reoperation after elective shoulder replacement: population based cohort study using hospital episode statistics for England. BMJ 2019;364:l298.

303. BESS Abstracts. Shoulder Elbow 2019;11:S5–S30.

304. Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, Judge A, Cooper C, Carr AJ, Arden NK, Beard DJ and Price AJ. Articles The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. Lancet 2017.

305. Levy O, Tsvieli O, Merchant J, Young L, Trimarchi A, Dattani R, Abraham R, Copeland SA, Narvani A and Atoun E. Surface replacement arthroplasty for glenohumeral arthropathy in patients aged younger than fifty years: results after a minimum ten-year follow-up. J Shoulder Elbow Surg 2015;24:1049–60.

306. Samuelsen BT, Wagner ER, Houdek MT, Elhassan BT, Sanchez-Sotelo J, Cofield R and Sperling JW. Primary reverse shoulder arthroplasty in patients aged 65 years or younger. J Shoulder Elbow Surg 2017;26:e13–e17.

307. Otto RJ, Clark RE and Frankle MA. Reverse shoulder arthroplasty in patients younger than 55 years: 2- to 12-year follow-up. J Shoulder Elbow Surg 2017;26:792–7.

308. Roberson TA, Bentley JC, Griscom JT, Kissenberth MJ, Tolan SJ, Hawkins RJ and Tokish JM. Outcomes of total shoulder arthroplasty in patients younger than 65 years: a systematic review. J Shoulder Elbow Surg 2017;26:1298–306.

309. Leathers MP, Ialenti MN, Feeley BT, Zhang AL and Ma CB. Do younger patients have better results after reverse total shoulder arthroplasty? J Shoulder Elbow Surg 2018;27:S24–S28.

310. Liu JN, Garcia GH, Wong AC, Sinatro A, Wu HH, Dines DM, Warren RF and Gulotta LV. Return to Work After Anatomic Total Shoulder Arthroplasty for Patients 55 Years and Younger at Average 5-Year Follow-up. Orthopedics 2018;41:e310–e315.

311. NJR. National Joint Registry for England, Wales and Northern Ireland. 14th Annual Report. 2017. URL: http://www.njrreports.org.uk.

312. Schouten LJ, Straatman H, Kiemeney LA and Verbeek AL. Cancer incidence: life table risk versus cumulative risk. J Epidemiol Community Health 1994;48:596–600.

313. Sasieni PD and Adams J. Standardized lifetime risk. Am J Epidemiol 1999;149:869–75.

314. Cancer Research UK. Lifetime risk calculations. URL: https://www.cancerresearchuk.org/sites/default/files/cs%7B%5C_%7Dlifetimerisk.xlsx Last accessed on 04/04/2018.

315. Kaplan EL and Meier P. Nonparametric Estimation from Incomplete Observations. J Am Stat Assoc 1958;53:457.

316. Kalbfleisch J and Prentice R. The statistical analysis of failure time data. New York, NY: John Wiley & Sons, 1980.

317. Lambert PC, Dickman PW, Nelson CP and Royston P. Estimating the crude probability of death due to cancer and other causes using relative survival models. Stat Med 2010;29:885–95.

318. Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT and Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics 1978;34:541–54.

319. Australian Orthopaedic Association National Joint Replacement Registry. Annual Report. 2018. URL: https://aoanjrr.sahmri.com.

320. Cheung EV, Sarkissian EJ, Sox-Harris A, Comer GC, Saleh JR, Diaz R and Costouros JG. Instability after reverse total shoulder arthroplasty. J Shoulder Elbow Surg 2018;27:1946–52.

321. Gallinet D, Cazeneuve JF, Boyer E, Menu G, Obert L, Ohl X, Bonnevialle N, Valenti P, Boileau P and Société Française de Chirurgie Orthopédique et Traumatologique (SoFCOT). Reverse shoulder arthroplasty for recent proximal humerus fractures: Outcomes in 422 cases. Orthop Traumatol Surg Res 2019;105:805–11.

322. Ranstam J, Kärrholm J, Pulkkinen P, Mäkelä K, Espehaug B, Pedersen AB, Mehnert F, Furnes O and group N study. Statistical analysis of arthroplasty data: I. Introduction and background. Acta Orthop 2011;82:253–7.

323. Ranstam J, Kärrholm J, Pulkkinen P, Mäkelä K, Espehaug B, Pedersen AB, Mehnert F, Furnes O and group N study. Statistical analysis of arthroplasty data: II. Guidelines. Acta Orthop 2011;82:258–67.

324. Sayers A, Evans JT, Whitehouse MR and Blom AW. Are competing risks models appropriate to describe implant failure? Acta Orthop 2018;89:256–8.

325. Norwegian National Advisory Unit on Arthroplasty and Hip Fractures Annual Report. 2018.

326. Rahme H, Salomonsson B and Ahlborg H. ÅRSRAPPORT 2018: Svenska Skulder och Armbågsregistret. Stockholm, 2018.

327. Partridge T, Jameson S, Baker P, Deehan D, Mason J and Reed MR. Ten-Year Trends in Medical Complications Following 540,623 Primary Total Hip Replacements from a National Database. J Bone Joint Surg Am 2018;100:360–7.

328. Office for National Statistics. Mortality statistics - underlying cause, sex and age. URL: https://www.nomisweb.co.uk Last accessed on 04/09/2018.

329. Punjani N, Lanting B, McClure JA, Winick-Ng J and Welk B. The Impact of Common Urologic Complications on the Risk of a Periprosthetic Joint Infection. J Bone Joint Surg Am 2018;100:1517–23.

330. Parthasarathy M, Reid V, Pyne L and Groot-Wassink T. Are we recording postoperative complications correctly? Comparison of NHS Hospital Episode Statistics with the American College of Surgeons National Surgical Quality Improvement Program. BMJ Qual Saf 2015;24:594–602.

331. Lévesque LE, Hanley JA, Kezouh A and Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ 2010;340:b5087.

332. Wright FL, Green J, Canoy D, Cairns BJ, Balkwill A, Beral V and Million Women Study Collaborators. Vascular disease in women: comparison of diagnoses in hospital episode statistics and general practice records in England. BMC Med Res Methodol 2012;12:161.

333. Daniel P, Bewick T, Welham S and Lim WS. British Thoracic Society. Adult Community Acquired Pneumonia (CAP) Audit Report. National Audit Period: 1 December 2014 – 31 January 2015. 2016. URL: https://www.brit-thoracic.org.uk/document-library/quality-improvement/audit-reports/adult-community-acquired-pneumonia-201415/.

334. Tomlinson LA, Riding AM, Payne RA, Abel GA, Tomson CR, Wilkinson IB, Roland MO and Chaudhry AN. The accuracy of diagnostic coding for acute kidney injury in England - a single centre study. BMC Nephrol 2013;14:58.

335. Bouras G, Burns EM, Howell AM, Bottle A, Athanasiou T and Darzi A. Risk of Post-Discharge Venous Thromboembolism and Associated Mortality in General Surgery: A Population-Based Cohort Study Using Linked Hospital and Primary Care Data in England. PLoS One 2015;10. Ed. by Ahmad F:e0145759.

336. Kuan V, Denaxas S, Gonzalez-Izquierdo A, Direk K, Bhatti O, Husain S, Sutaria S, Hingorani M, Nitsch D, Parisinos CA, Lumbers RT, Mathur R, Sofat R, Casas JP, Wong ICK, Hemingway H and Hingorani AD. A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. Lancet Digit Heal 2019;1:e63–e77.

337. Petersen I, Douglas I and Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. BMJ 2016;354:i4515.

338. Oladeji LO, Raley JA, Menendez ME and Ponce BA. Risk factors for in-hospital myocardial infarction after shoulder arthroplasty. Am J Orthop 2015;44:E142–7.

339. Murphy AB, Menendez ME, Watson SL and Ponce BA. Metabolic syndrome and shoulder arthroplasty: epidemiology and peri-operative outcomes. Int Orthop 2016;40:1927–33.

340. Healthcare Cost and Utilization Project. National Inpatient Sample Overview. URL: https://www.hcup-us.ahrq.gov/nisoverview.jsp Last accessed on 03/10/2019.

341. Shields E, Iannuzzi JC, Thorsness R, Noyes K and Voloshin I. Perioperative complications after hemiarthroplasty and total shoulder arthroplasty are equivalent. J Shoulder Elbow Surg 2014;23:1449–53.

342. Waterman BR, Dunn JC, Bader J, Urrea L, Schoenfeld AJ and Belmont PJJ. Thirty-day morbidity and mortality after elective total shoulder arthroplasty: Patient-based and surgical risk factors. J Shoulder Elbow Surg 2015;24:24–30.

343. Jiang JJ, Somogyi JR, Patel PB, Koh JL, Dirschl DR and Shi LL. Obesity is Not Associated with Increased Short-term Complications After Primary Total Shoulder Arthroplasty. Clin Orthop Relat Res 2016;474:787–95.

344. Westermann RW, Anthony CA, Duchman KR, Pugely AJ, Gao Y and Hettrich CM. Incidence, Causes and Predictors of 30-Day Readmission After Shoulder Arthroplasty. Iowa Orthop J 2016;36:70–4.

345. Belmont PJJ, Kusnezov NA, Dunn JC, Bader JO, Kilcoyne K and Waterman BR. Predictors of Hospital Readmission After Total Shoulder Arthroplasty. Orthopedics 2017;40:e1–e10.

346. National Institute for Health and Care Excellence. NICE Clinical Guideline 89: Venous thromboembolism in over 16s. 2018. URL: https://www.nice.org.uk/guidance/ng89.

347. BESS. BESS VTE Guidelines for Shoulder and Elbow Surgery. 2013. URL: https://tinyurl.com/yyv8wfff.

348. Soong J, Poots AJ, Scott S, Donald K and Bell D. Developing and validating a risk prediction model for acute care based on frailty syndromes. BMJ Open 2015;5:e008457.

349. Soong JTY, Kaubryte J, Liew D, Peden CJ, Bottle A, Bell D, Cooper C and Hopper A. Dr Foster global frailty score: an international retrospective observational study developing and validating a risk prediction model for hospitalised older persons from administrative data sets. BMJ Open 2019;9:e026759.

350. Urmey WF, Talts KH and Sharrock NE. One Hundred Percent Incidence of Hemidiaphragmatic Paresis Associated With Interscalene Brachial Plexus Anesthesia as Diagnosed by Ultrasonography. Anesth Analg 1991;72:498???503.

351. Vorobeichik L, Brull R, Bowry R, Laffey JG and Abdallah FW. Should continuous rather than single-injection interscalene block be routinely offered for major shoulder surgery? A meta-analysis of the analgesic and side-effects profiles. Br J Anaesth 2018;120:679–92.

352. Metcalfe D, Zogg CK, Judge A, Perry DC, Gabbe B, Willett K and Costa ML. Pay for performance and hip fracture outcomes. Bone Joint J 2019;101-B:1015–23.

353. NJR. NJR Surgeon and Hospital Profile. URL: http://www.njrsurgeonhospitalprofile.org.uk Last accessed on 04/09/2019.

354. Shmueli G. To Explain or to Predict? Stat Sci 2010;25:289–310.

355. Allison PD. Change Scores as Dependent Variables in Regression Analysis. Sociol Methodol 1990;20:93.

356. Vickers AJ and Altman DG. Statistics Notes: Analysing controlled trials with baseline and follow up measurements. BMJ 2001;323:1123–4.

357. Glymour MM, Weuve J, Berkman LF, Kawachi I and Robins JM. When Is Baseline Adjustment Useful in Analyses of Change? An Example with Education and Cognitive Change. Am J Epidemiol 2005;162:267–78.

358. Wilkinson JM, Kadirkamanathan V, Aram P, Blom AW, Sayers A, Trela-Larsen L, Hills A and Zotov E. NJR Patient Decision Support Tool for Joint Replacement (JointCalc). URL: https://jointcalc.shef.ac.uk/.

359. NJR. National Joint Registry for England, Wales and Northern Ireland. 13th Annual Report. 2016. URL: http://www.njrreports.org.uk.

360. Collins GS, Reitsma JB, Altman DG and Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med 2015;162:55.

361. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF and Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015;162:W1.

362. Perperoglou A, Sauerbrei W, Abrahamowicz M and Schmid M. A review of spline function procedures in R. BMC Med Res Methodol 2019;19:46.

363. Royston P and Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med 2002;21:2175–97.

364. Van Smeden M, De Groot JAH, Moons KGM, Collins GS, Altman DG, Eijkemans MJC and Reitsma JB. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med Res Methodol 2016;16:163.

365. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Moons KG and Collins GS. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. Stat Med 2019;38:1262–75.

366. Riley RD, Snell KI, Ensor J, Burke DL, Harrell Jr FE, Moons KG and Collins GS. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med 2019;38:1276–96.

367. Nagelkerke N. A Note on a General Definition of the Coefficient of Determination. Biometrika 1991;78:691.

368. Harrell FE. Regression modeling strategies : with applications to linear models, logistic and ordinal regression, and survival analysis. 2nd Ed. 2015:582.

369. Brier GW. Verification of forecasts expressed in terms of probability. Mon Weather Rev 1950;78:1–3.

370. Graf E, Schmoor C, Sauerbrei W and Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. Stat Med 1999;18:2529–45.

371. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ and Kattan MW. Assessing the Performance of Prediction Models. Epidemiology 2010;21:128–38.

372. Janssens A. A more intuitive interpretation of the area under the ROC curve. PeerJ Preprints 2017;5:e3468v1.

373. Copas JB. Regression, Prediction and Shrinkage. J R Stat Soc Series B Stat Methodol 1983;45:311–35.

374. Van Houwelingen JC and Le Cessie S. Predictive value of statistical models. Stat Med 1990;9:1303–25.

375. Harrell FE, Lee KL and Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.

376. Steyerberg EW, Uno H, Ioannidis JP, Calster B van, Ukaegbu C, Dhingra T, Syngal S and Kastrinos F. Poor performance of clinical prediction models: the harm of commonly applied methods. J Clin Epidemiol 2018;98:133–43.

377. Simovitch RW, Friedman RJ, Cheung EV, Flurin PH, Wright T, Zuckerman JD and Roche C. Rate of Improvement in Clinical Outcomes with Anatomic and Reverse Total Shoulder Arthroplasty. J Bone Joint Surg Am 2017;99:1801–11.

378. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T and Clinical Significance Consensus Meeting Group. Estimating clinically significant differences in quality of life outcomes. Qual life Res an Int J Qual life Asp Treat care Rehabil 2005;14:285–95.

379. Simovitch R, Flurin PH, Wright T, Zuckerman JD and Roche CP. Quantifying success after total shoulder arthroplasty: the substantial clinical benefit. J Shoulder Elbow Surg 2018;27:903–11.

380. Coles J. Risk Adjustment models for Patient Reported Outcome Measures. 2010. URL: https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2013/07/proms-ris-adj-meth-sur-orth.pdf.

381. NJR. National Joint Registry fo England, Wales and Northern Ireland. 10th Annual Report. 2013.

382. Jayakumar P, Overbeek CL, Lamb S, Williams M, Funes C, Gwilym S, Ring D and Vranceanu AM. What Factors Are Associated With Disability After Upper Extremity Injuries? A Systematic Review. Clin Orthop Relat Res 2018;476:2190–215.

383. Jayakumar P, Teunis T, Williams M, Lamb SE, Ring D and Gwilym S. Factors associated with the magnitude of limitations during recovery from a fracture of the proximal humerus. Bone Joint J 2019;101-B:715–23.

384. Werner BC, Wong AC, Chang B, Craig EV, Dines DM, Warren RF and Gulotta LV. Depression and Patient-Reported Outcomes Following Total Shoulder Arthroplasty. J Bone Joint Surg Am 2017;99:688–95.

385. Smith CD, Guyver P and Bunker TD. Indications for reverse shoulder replacement. J Bone Joint Surg Br 2012;94-B:577–83.

386. Kutner MH, Nachtsheim CJ, Neter J and Li W. Single-Factor Studies. In: Applied Linear Statistical Models. 5th. New York, NY, 2005. Chap. 16:716–22.

387. Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K and Robins JM. Results of Multivariable Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-based Weighting under Conditions of Nonuniform Effect. Am J Epidemiol 2006;163:262–70.

388. Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, Nichols M, Stone GW and Pocock SJ. Comparison of Propensity Score Methods and Covariate Adjustment. J Am Coll Cardiol 2017;69:345–57.

389. Edwards TB, Boulahia A, Kempf JF, Boileau P, Nemoz C and Walch G. The influence of rotator cuff disease on the results of shoulder arthroplasty for primary osteoarthritis: results of a multicenter study. J Bone Joint Surg Am 2002;84-A:2240–8.

390. Walch G, Badet R, Boulahia A and Khoury A. Morphological Study of the Glenoid in Primary Glenohumeral Osteoarthritis. J Arthroplasty 1999;14:756–60.

391. Bercik MJ, Kruse K, Yalizis M, Gauci MO, Chaoui J and Walch G. A modification to the Walch classification of the glenoid in primary glenohumeral osteoarthritis using three-dimensional imaging. J Shoulder Elbow Surg 2016;25:1601–6.

392. Young AA, Walch G, Pape G, Gohlke F and Favard L. Secondary rotator cuff dysfunction following total shoulder arthroplasty for primary glenohumeral osteoarthritis: results of a multicenter study with more than five years of follow-up. J Bone Joint Surg Am 2012;94:685–93.

393. Gowd AK, Garcia GH, Liu JN, Malaret MR, Cabarcas BC and Romeo AA. Comparative analysis of work-related outcomes in hemiarthroplasty with concentric glenoid reaming and total shoulder arthroplasty. J Shoulder Elbow Surg 2019;28:244–51.

394. The New Zealand Joint Registry Nineteen Year Report - January 1999 to December 2017. 2018. URL: https://nzoa.org.nz/system/files/DH8152%7B%5C_%7DNZJR%7B%5C_%7D2018%7B%5C_%7DReport%7B%5C_%7Dv6%7B%5C_%7D4Decv18.pdf.

395. Matsen FA, Somerson JS, Hsu JE, Lippitt SB, Russ SM and Neradilek MB. Clinical effectiveness and safety of the extended humeral head arthroplasty for selected patients with rotator cuff tear arthropathy. J Shoulder Elbow Surg 2019;28:483–95.

396. Garret J, Harly E, Le Huec JC, Brunner U, Rotini R and Godenèche A. Pyrolytic carbon humeral head in hemi-shoulder arthroplasty: preliminary results at 2-year follow-up. JSES open access 2019;3:37–42.

397. National Institute for Health and Care Excellence. NICE Clinical Guideline. Joint replacement (primary): hip, knee and shoulder (Draft consultation). 2019. URL: https://www.nice.org.uk/guidance/indevelopment/gid-ng10084.

398. Rangan A, Gwilym S, DiMascio L, Danjoux G, Jones V, Kassam J, Torgeson D, Hewitt C, Keding A, Corbacho B, McDaid C, Tharmanathan P, Kottam L, Arundel C and Anderson D. Effectiveness and cost-effectiveness of Reverse Shoulder Arthroplasty versus Hemiarthroplasty versus Non-surgical care for acute 3 and 4 part fractures of the proximal humerus in older adults - The PROFHER-2 Randomised Clinical Trial. NIHR HTA Protocol. 2017. URL: https://www.journalslibrary.nihr.ac.uk/programmes/hta/167303/%7B%5C#%7D/.