AARHUS
UNIVERSITY
HEALTH

Low-grade infections in patients with shoulder replacements -diagnose and outcome

PhD thesis

Thomas Falstie-Jensen

Faculty of Health
University of Aarhus
2019

Low-grade infections in patients with shoulder replacements -diagnose and outcome

PhD thesis

Thomas Falstie-Jensen

Faculty of Health
University of Aarhus
and
Department of Orthopaedic Surgery,
Section for shoulder and elbow surgery
Aarhus University Hospital

Supervisors

Kjeld Søballe, MD MDSc, Professor, Aarhus University, Department of Orthopaedics, Aarhus University Hospital, Denmark (main supervisor)

Jeppe Lange, MD, PhD, Associate Professor, Department of Orthopaedics, Horsens Regional Hospital, Denmark

Janne Ovesen, MD, PhD, Department of Orthopaedics, Aarhus University Hospital, Denmark

Henrik Daugaard, MD, PhD, Department of Orthopaedics, Herlev-Gentofte Hospital, Denmark

Evaluation committee

Kirsten Bouchelouche, MD, MDSc, Associate professor (chairman), Department of Nuclear Medicine and PET-Center, Aarhus University Hospital, Denmark

Marjan Wouthuyzen-Bakker, MD, PhD, Department of Medical Microbiology and Infection Prevention, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands

Jeppe Vejlgaard Rasmussen, MD, PhD, Associate professor, Department of Orthopaedics, Herlev-Gentofte hospital, Denmark

Correspondence
Thomas Falstie-Jensen, MD
Ortopædkirurgisk afd.
Palle Juul Jensens Boulevard 99
8200 Aarhus N, Denmark
E-mail: thofal@rm.dk

Acknowledgement

Inspired by the patients we treated for periprosthetic infections at the Orthopaedic department, Aarhus University Hospital I wondered if we could treat some of these patients differently. In 2013, we began to contemplate a study investigating treatment of infected shoulder replacements. Ideas evolved and we ended up with studies on both diagnoses and treatment which lead to my enrolment as a PhD student at Aarhus University in 2015.

This thesis was made possible by a scholarship from Aarhus University, support from Fonden til lægevidenskabens fremme, Lundbeckfonden and most of all the flexibility of my colleagues at the section of shoulder and elbow surgery.

I would like to express my deepest appreciation to my supervisors. Main supervisor Kjeld Søballe, for providing me with the opportunity to conduct my research in collaboration with Orthopaedic Research unit and for always encouraging me when perplexity set in. My co-supervisor Henrik Daugaard, for inspiration and meticulous reviews of my manuscripts. Janne Ovesen, for constantly being open for professional and scientific discussions and always allowing me days off from clinical work when science required it. Jeppe Lange, for many fruitful discussions, guidance and endless support.

I thank the Department of clinical physiology and PET-centre at Aarhus University Hospital and Herlev Hospital for supporting these studies by making scintigraphy and PET-scans widely accessible. I am especially grateful for the support and patience Lars Christian Gormsen demonstrated when giving me insight in new and complex matters previous unknown to me.

I warmly thank my fellow-surgeons at Gentofte-Herlev hospital and Aarhus University Hospital for great collaboration. I hope we can continue our work on future studies.

Special appreciation goes to Charlotte Diana Nielsen and Annette Isling Jørgensen at Gentofte-Herlev Hospital for retrieval of questionnaires from the patients and for invaluable administration of all my study-binders in your office.

A very special thank goes to the staff in my unit for many enjoyable days and their support during the whole course of this PhD. Helle Koch without your Excel-sheets keeping track of patients would have been so much more difficult. Christina Rasmussen, Helle Hahn, Lisbeth Lind-Byrialsen and Hanne Pedersen for providing immense help in scheduling and caring for "my" patients.

Lastly, I would like to express my deepest and most sincere gratitude to my wife Anne Mette and our children Jeppe, Signe and Asger for their encouragement and continuous support during the last 5 years.

Thomas Falstie-Jensen, 2019

List of Studies

The thesis is based on the following three studies:

Labeled white blood cell/bone marrow single-photon emission computed tomography with computed tomography fails in diagnosing chronic periprosthetic shoulder joint infection.

Falstie-Jensen T, Daugaard H, Søballe K, Ovesen J, Arveschou A K, Lange J on behalf of the ROSA study-group.

Journal of Shoulder and Elbow surgery 2019, Jun;28(6):1040-1048

^{18}F FDG-PET/CT has poor diagnostic accuracy in diagnosing shoulder PJI.

Falstie-Jensen T, Lange J, Daugaard H, Sørensen AKB, Zerahn B, Ovesen J, Søballe K, Gormsen L C on behalf of the ROSA study-group.

European Journal of Nuclear Medicine and Molecular imaging 2019, Sep;46(10):2013-2022

Unexpected positive cultures after revision shoulder arthroplasty -does it impact outcome?

Falstie-Jensen T, Lange J, Daugaard H, Sørensen AKB, Ovesen J, Søballe K on behalf of the ROSA study-group.

Manuscript in preparation

The papers of this thesis will be referred to in the text by their Roman numerals (I-III). Those currently in print have been reprinted by permission of the copyright holder.

Abbreviations

BM	Bone marrow
C acnes	Cutibacterium acnes
CI	Confidence interval
CMS	Constant Murley Score
CoNS	Coagulase negative staphylococcus
CT	Computed tomography
DSR	Danish Shoulder Arthroplasty Register
FDG	2-fluoro-2-deoxyglucose
HA	Hemiarthroplasty
ICM	International Consensus on Musculoskeletal Infection
NJR	The National Joint Register of England, Wales, Northern Ireland and Isle of Man
NPV	Negative predictive value
PET	Positron emission tomography
PJI	Periprosthetic joint infection
PPV	Positive predictive value
PRO	Patient reported outcome (refers to score)
PROM	Patient reported outcome measure (refers to instrument)
RSA	Reverse total shoulder replacement
SPECT	Single photon emission computed tomography
TSR	Anatomic total shoulder replacement
UPC	Unexpected positive cultures
WBC	White blood cell

Contents

1 English summary

Infection after any joint replacement can cause detrimental effects to function. Diagnose is often straightforward if the infection is acute, but a clear-cut diagnosis of chronic periprosthetic joint infection (PJI) can be a challenge. Further, up to 30% of aseptic revised shoulder replacements exhibit positive cultures. The interpretation and impact of such Unexpected Positive Cultures (UPC) are still debated.

To extend the knowledge of diagnosis of PJI and the impact of UPC on outcome, this thesis evaluated the diagnostic accuracy of ^{111}Indium-labelled White blood cells/^{99}Technesium labelled bone marrow SPECT CT (Study I) and ^{18}F-FDG PET/CT (Study II) to diagnose chronic PJI. Furthermore, patient reported outcome after revisions with emergence of UPC was assessed (Study III). All three studies were based on a prospective cohort of patients referred with a failed shoulder replacement from April 1st, 2014 to September 30th, 2017. All patients were followed at least two years after revision.

During a one-year period 28 patients were included in Study I of which 11 were infected. The scans were true positive in two cases and false positive in zero cases, false negative in 9 cases and true negative in 17 cases. Sensitivity was 0.18 (95% CI: 0.00-0.41), specificity 1.00 (95% CI: 1.00-1.00), PPV 1.00 (95% CI: 1.00-1.00) and NPV 0.67 (95% CI: 0.49-0.84). Despite the study was prematurely ceased due to reduced scanning capacity, results showed a clear trend of low diagnostic performance.

In Study II 86 patients with FDG-PET prior to revision surgery were included. Three scans were true positive, 6 false positive, 51 true negative and 19 false negative; corresponding to a sensitivity of 0.14 (95% CI: 0.03–0.35), specificity 0.91 (95% CI: 0.81–0.97), PPV 0.38 (95% CI: 0.15–0.70) and NPV 0.71 (95% CI: 0.67–0.75). To increase the homogeneity of the FDG-PET diagnoses, a criterion of three distinct patterns of tracer uptake each defining infection were used. Despite these well-defined criteria, the interobserver agreement between the three reviewers was only moderate.

Study III compared outcome described by Oxford Shoulder Scores (OSS) of 126 patients who underwent a standard revision. UPC was found in 28 patients and all cases were due to low-virulent bacteria. At baseline an identical OSS score of 20 (95% CI: 18-22) in the culture-negative group and UPC group 20 (95% CI: 17-23) was found. The 2-year follow-up did also reveal nearly identical scores; 33 (95% CI: 31-36) in the culture-negative and 34 (95% CI: 29-38) in the UPC group. Regardless of culture result a clinically relevant increase in OSS score and decrease in pain were seen. Similarly, most patients experienced a gain in forward elevation. In contrast, the external rotation was unchanged in all groups.

In conclusion, neither WBC/BM SPECT CT nor FDG-PET performs adequately to detect PJI to justify a routine use in the preoperative workup. If no preoperative suspicion of infection exists most patients gain from the revision. If UPC emerges with bacteria of low virulence outcome of the procedure does not seem to be affected.

2 Danish summary

En inficeret led-protese (PJI) er en potentiel katastrofe, da det kan resultere i varig nedsat funktion og kroniske smerter. Akutte infektioner volder sjældent diagnostisk besvær, hvorimod kroniske infektioner kan være vanskelige at påvise. Flere forfattere har endda påvist, at der i ca. 30% af alle operationer, hvor en skulderprotese udskiftes, kan påvises bakterie i vævsprøver – på trods af der ikke var den mindste mistanke om PJI. Hvordan sådanne uventede bakteriefund (på engelsk: Unexpected Postive Cultures, UPC) skal tolkes, eller om de har betydning for operationsresultatet er ikke endeligt klarlagt.

Formålet med denne afhandling var at undersøge brugbarheden af kombineret leukocyt/knoglemarvs-skintigrafi (Studie I) samt ^{18}F-FDG PET/CT (Studie II) i udredningen af kronisk PJI samt undersøge hvorledes UPC påvirker patienternes selvvurderede resultat efter udskiftning af en skulderprotese (Studie III). Alle deltagende patienter var henvist til behandling for en smertende protese i perioden 1. april 2014 til 30. september 2017, og alle oplysninger blev indsamlet fortløbende.

I Studie I deltog 28 patienter, hvoraf 11 var inficerede. Scanningen var sandt positiv i tre tilfælde, falsk positiv i nul tilfælde og falsk negativ i 9 tilfælde. Dette resulterede i en sensitivitet på 0.18 (95% CI: 0.00-0.41), specificitet på 1.00 (95% CI: 1.00-1.00), PPV 1.00 (95% CI: 1.00-1.00) og NPV 0.67 (95%CI: 0.49-0.84). På grund af udfordringer med scannerkapacitet måtte undersøgelsen stoppes før tid. Resultaterne pegede dog klart i retning af begrænset brugbarhed af scanningen.

I Studie II deltog 86 patienter. Scanningen var sandt positiv i 3 tilfælde, falsk positiv i 6, sandt negativ i 51 og falsk negativ i 19 tilfælde. Dette giver en sensitivitet på 0.14 (95% CI: 0.03–0.35), specificitet 0.91 (95% CI: 0.81–0.97), PPV 0.38 (95% CI: 0.15–0.70) og NPV 0.71 (95% CI: 0.67–0.75). Da vurdering af PET-undersøgelser baseres på optagelsesmønster af sporstoffet, blev der udviklet et kriterie bestående af tre mønstre som hver især definerede infektion. På trods af dette kriterie, var der kun moderat overensstemmelse i PET-diagnosen mellem de tre læger, der vurderede billederne.

I Studie III indgik 126 patienter, der fik udskiftet en protese, og i 28 tilfælde fandt man UPC. Forud for operationen var OSS-værdierne ens; 20 (95% CI: 18-22) hos patienter uden UPC og 20 (95% CI: 17-23) hos de patienter, der senere udviklede UPC. To år efter operationen var scorerne fortsat næsten ens, 33 (95% CI: 31-36) i gruppen uden UPC og 34 (95% CI: 29-38) i gruppen med UPC. En analyse af smerteniveau og skulderens bevægelighed viste, at de fleste patienter angav færre smerter og bedre fremadføring efter operationen. Det skal dog bemærkes, at ingen af patientgrupperne var smertefrie; ligesom udadrotationen var uændret i alle grupper efter operationen.

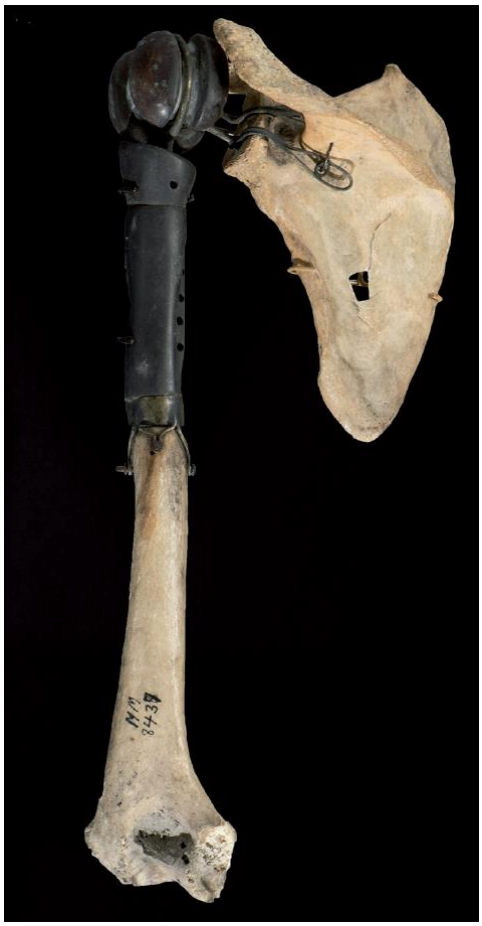
På baggrund af resultaterne i de tre studier kan det derfor konkluderes: hverken kombineret leukocyt/knoglemarvsskintigrafi eller ^{18}F-FDG PET/CT synes i vores kohorte at kunne diagnosticere kronisk PJI i skulderen med acceptabel sikkerhed. Resultaterne viser også, at protese-udskiftnings -operationer med UPC formentlig ikke giver et dårligere resultat sammenlignet med en tilsvarende operation uden bakteriefund.

3 Introduction

The desire to replace a degenerated and painful joint with an artificial but well-functioning replica is the basis of all joint replacement surgery. Many joints can be replaced ranging from small joints like finger and jaw to larger joints like elbow, shoulder, knee and hip. The latter is by far the most common joint replacement. However, shoulder replacement is becoming increasingly frequent.

The first shoulder replacement was allegedly performed by the French surgeon Jules-Émile Péan in 1893 [1]. The patient, Mr. Perdoux, was a 37-year old baker who suffered from bony tuberculosis. The prosthesis was made in two parts; the shaft was made of platinum and the head was made of hardened rubber. The initial results were encouraging but unfortunately Dr. Péan had to remove the prosthesis two years later owing to infection.

Figure 1: Replica of Dr. Péan's shoulder replacement [1]



Since then much has changed regarding implant designs, materials and preventions of complication. Even so, many of the basic principles and risks of shoulder replacements are still valid and infection continues to be one of the most disabling complications a patient can suffer from.

To Dr. Peán periprosthetic joint infection (PJI) of Mr. Perdoux's shoulder was obvious; a draining sinus was present. Today, in absence of such findings low-grade PJI may pose a significant challenge to diagnose. Even with an array of advanced diagnostic tools available, a clear diagnosis of chronic PJI cannot always be made. Despite thorough preoperative investigations a considerable number of revisions thought to be aseptic turn out with unexpected positive cultures (UPC) after revision.

This thesis aims at establishing the diagnostic accuracy of two nuclear medicine modalities in diagnosing low-grade PJI. Furthermore, the thesis will investigate results of revisions with UPC by patient reported outcome measures.

4 Background

4.1 Periprosthetic joint infections

Worldwide the incidence of shoulder replacements has increased and in the U.S. the number of patients treated with a shoulder replacement has more than doubled from 2002 to 2012 [2]. In Denmark approximately 1200 patients are treated with a shoulder replacement every year. The most frequent indications for shoulder replacement are arthrosis (42%), fracture (26%), and severe rotator cuff insufficiency with or without concomitant arthrosis (23%) [3]. Since the number of shoulder replacements increases, the number of complications like periprosthetic joint infections (PJI) is expected to increase concurrently. The incidence of shoulder PJI is reported to range from approximately 1% in primary shoulder replacements and up to 15% in revision cases [2, 4-6].

Unfortunately, not all patients have a successful outcome after a shoulder replacement and failure may be due to PJI, pain, instability or limited ROM. This group of patients is referred to as having a "failed shoulder replacement" and some will subsequently be revised to obtain a better result. Especially PJI can be detrimental to joint function and may result in residual limitations and pain [7-10]. Although PJI is a rare complication, according to the Danish Shoulder arthroplasty Registry 25% of the 110 annual revisions are attributed to PJI [3].

4.1.1 Pathways for patients referred with a failed shoulder replacement

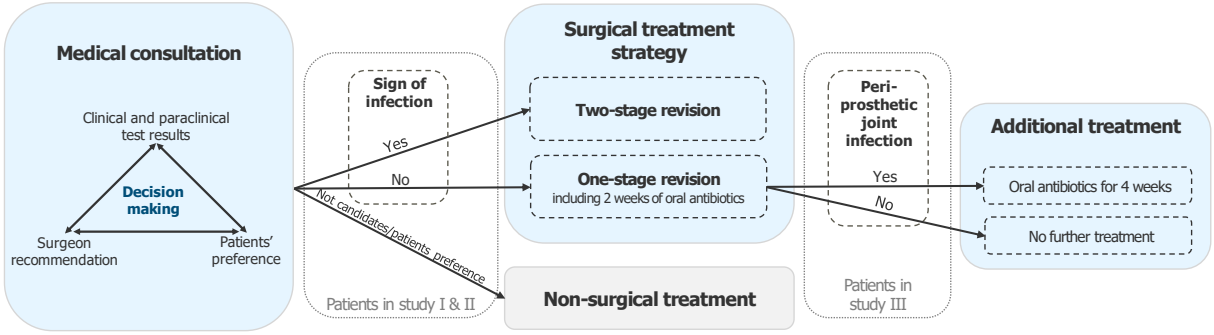
In Denmark, all patients with a failed shoulder replacement are referred to one of the three public medical centres approved by the Danish National Board of Health to perform revision of shoulder replacements [11]. At the first medical consultation investigations to identify patients with a PJI are initiated. The surgeon obtains a thorough patient history and physical examination. Furthermore, paraclinical tests are performed including blood tests, x-ray and in some cases joint aspiration.

If a patient is presenting with high suspicion of PJI, he/she is scheduled for a two-stage revision. High suspicion of PJI is usually prompted by findings such as a draining sinus, fever, raised infectious markers or marked osteolysis on x-ray. A two-stage revision implies removal of implant and placement of an antibiotic-releasing cement-spacer during the first operation; after an interim period, the final implant is inserted at a second operation. In studies of shoulders, hips and knees, this technique has proven reliable to eradicate PJI with a success rate around 80-90% [8, 12, 13]. A two-staged approach has, however, the disadvantage of multiple operations, a prolonged period of morbidity for the patient and thus a high draw on patient, hospital and societal resources.

If no suspicion of PJI is present, the decision for further surgical treatment is a triage of 'test results', 'surgeon recommendation', and 'patient choice'. Thus, the choice of treatment strategy depends on individual patient factors like time of onset of symptoms, preoperative identification of infection causing organism, patient co-morbidities and technical factors like bone stock, need for cement-removal and options of implants. Some patients do not receive any further surgical treatment based on i) their own

choice not to undergo surgery and/or ii) the surgeon's decision of not judging the patient fit for surgery and rehabilitation or that indication for revision is not found. The remaining patients undergo a standard revision in which the implant and all cement are removed followed by implantation of the new prothesis during the same operation. The patients are treated with oral antibiotics for 2 weeks until the microbiological test results reveal whether the patient had PJI. Patients with microbiological established PJI are further treated with either IV or oral antibiotics for at least four weeks.

Thus, this treatment strategy can be labelled as a one-stage revision for PJI. Figure 2 depicts the treatment pathway of a referred patient.



The use of single-staged revisions has proved effective to eliminate PJI in selected patients with hip replacements [14, 15]. The literature on one-stage revisions of shoulder replacements is sparse, yet the few studies showed promising results [16-20]. Furthermore, the most recent literature reviews suggest that one-stage revisions result in comparable infection eradication to two-stage revisions, but only when proper patient selection is achieved [7, 21].

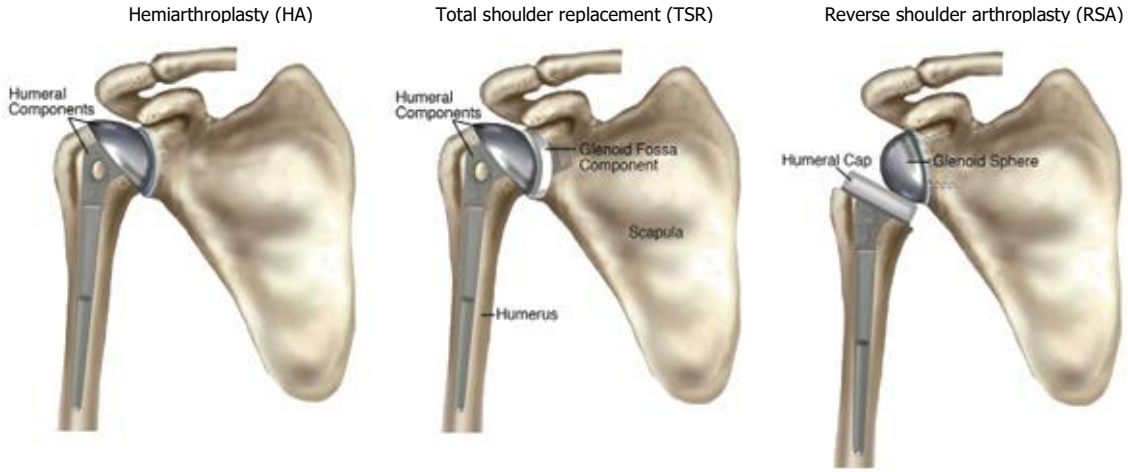
4.2 Shoulder replacement(s) designs

The shoulder joint consists of a ball (humoral head) and socket (glenoid) with a high degree of range of motion (ROM). Because of the anatomic structure with the glenoid being almost flat the shoulder has limited static stability. The major stabilising component in the shoulder is the rotator-cuff which consists of four tendons arranged in a harness-like fashion around the humoral head. Roughly, each muscle is responsible for either internal rotation, external rotation or abduction. When the muscles and tendons are intact balanced interaction keeps the humoral head centred in the glenoid (dynamic stability) at all times. However, when this mechanism fails it can lead to pain, instability or both.

Three arthroplasty designs are used in revision surgery as depicted in Figure 3. The first design is a hemiarthroplasty (HA) in which only the humoral side is replaced. The second design is a total anatomic shoulder replacement (TSR) in which the humoral side is replaced in combination with a glenoid replacement. The third design is a reverse shoulder arthroplasty (RSA) in which the anatomy of the shoulder is flipped 180 degrees resulting in a socket on the humeral side and a ball on the glenoid side. Both the HA and TSR mimic anatomy and are unconstrained with the consequence that no static stability is built into the design. Thus, intact dynamic stabilisers are required to achieve optimal function. In

contrast, the RSA is a semi-constrained design where the two components have a more pronounced ball-and-bowl shape which increases static stability and good function and a stable joint can be achieved in RSA joint replacements despite lack of dynamic stabilizers. Thus, the state of the rotator cuff is important in the choice of implant design.

Figure 3: Three implant designs used in shoulder replacement surgery. Image courtesy of Marie Dauenheimer [22].



Small rotator-cuff tears have shown to be insignificant whereas larger or multiple ruptured tendons yield inferior results with an unconstrained prosthesis [23]. In revision surgery, the problem of insufficient rotator cuff is frequently encountered and is often caused by scarring from prior surgery, wear or sequela from fracture. Consequently, RSA is the most frequently used design in revision cases owing to the ability to deliver acceptable function and pain relief despite failure of dynamic stabilisers [24]. TSR may be the choice of revision implant in cases with glenoid attrition due to a HA leaving the bone stock good and dynamic stabilisers intact. TSR has shown to be superior to both RSA and HA in terms of outcome in pain relief and ROM [25]. HA is declining in use owing to the versatility of the RSA [24]. Currently, HA is chosen in selected patients with conditions preventing implantation of an RSA.

4.3 Diagnosing periprosthetic joint infection

Since choice of treatment strategy is highly dependent of the absence/presence of PJI, it is of utmost importance to identify these patients with PJI prior to surgery. However, diagnosing PJI is influenced by several factors like microbiology including types of bacteria and possible formation of biofilm, if definition of PJI by either cultures or composite criteria used and the diagnostic tools available. These factors will be outlined in the following paragraphs.

4.3.1 Microbiology

Isolation of bacteria is paramount in diagnosing PJI since surgery is always followed by antibiotic treatment. When the bacterium is known the most effective drug can be administered. The most frequent isolated bacteria in shoulder PJI are C acnes, S epidermidis, CoNS and S aureus [7, 26-28]. However, these bacteria can be expected under different conditions. The high-virulent bacteria, like S

aureus, are more likely to cause PJI with acute onset of symptoms either shortly after surgery or due to haematogenous seeding. In contrast, low-virulent bacteria, like C acnes and CoNS, are more likely to cause chronic PJI with gradually increasing symptoms over time. Trampuz et al. described this correlation between virulence, time and route of infection by classifying high-virulent PJI as likely early (<3 months after surgery) or late PJI (>24 months) and low-grade PJI as likely delayed (2-24 months after surgery) [29]. This knowledge can influence treatment since high-virulent PJI may be treated successfully by debridement and implant retention and low-grade PJI by two-stage revisions due to higher risk of mature biofilm (see below) owing to longstanding presence of bacteria in the joint [30].

The most frequent isolated low-virulent bacterium is C acnes [26-28, 31]. Since C acnes is slow-growing anaerobe but aerotolerant bacteria, special attention must be paid to avoid missing such infections [32]. Most bacteria will exhibit growth during the first 4 days of culturing, however, detection of C acnes has proven to require both aerobic and anaerobic cultures in combination with prolonged time [33]. Currently, it is recommended to culture aerobe and anaerobe for at least 14 days to avoid false negative results [33-35].

The use of prolonged cultures and growing interest in low-grade PJI in general have consolidated the category of UPC. This entity covers cultures which show growth with bacteria after apparently aseptic revisions. Investigation of routine cultures after aseptic shoulder revisions has revealed growth in up to 56% of the cases and C acnes was present in more than 2/3 [32]. Surprisingly, C acnes has also appeared in cultures after primary joint replacements, primary arthroscopic surgery and plate removal after successful osteosyntheses of the clavicle [36-38]. Several studies have confirmed that C acnes can cause a fulminant PJI which will need treatment, however, studies of UPC after both hip and shoulder revisions (of which many were caused by C acnes) have not documented increased short to mid-term increase in risk of re-revision [39-42]. The interpretation of UPC and impact on outcome is not clear but illustrates the difficulty routine biopsies obtained during every revision may pose to the subsequent treatment [43]. The conclusions some authors have drawn of no impact on outcome after identification of bacteria in a presumed sterile environment spurn a question if tissue cultures alone are enough to define an infection [44].

Biofilm is an important aspect to consider in the diagnosis and treatment of implant infections, since it has been shown to reduce the bactericidal effect of the innate immune system and significantly impair the effect of a majority of available antibiotics [45]. Furthermore, bacteria residing in biofilm have proven difficult to detect by normal cultures and many authors recommend the use of sonication to dislodge the biofilm from implant to attain higher success of positive yield [34, 46, 47]. Biofilm is a sessile accumulation of bacteria embedded in an extracellular matrix attached to a surface. When bacteria engage in biofilm formation it changes phenotype from a free-floating planktonic type to a more dormant and phenotypical distinct biofilm type [48]. The biofilm matrix consists of polysaccharides, proteins and DNA of which at least some components are produced by the bacteria itself. After bacteria are exposed to a susceptible surface, e.g. shoulder replacement, the biofilm can rapidly start to form but will likely

take from days to weeks to mature [49]. Once biofilm has formed, the colony of bacteria resides in an environment well-protected from the immune system and many antibiotics. Many bacteria, both high- and low-virulent, are able to form biofilm among others S aureus, C acnes, S Epi and CoNS [45].

4.3.2 Preoperative diagnostic tests

The degree of suspicion of PJI is based on a combination of presented symptoms and objective findings. High suspicion of PJI are evident when patients present with classical symptoms such as local swelling, erythema, a draining sinus or increased secretion from a wound [50]. Equally, loosening of the humeral component should raise suspicion of PJI because aseptic loosening is rare [32, 51]. But absence of such findings does not exclude PJI, hence preoperative tests including serum inflammatory markers, x-ray and/or analysis of synovial fluid are used to identify patients with high suspicion of PJI [32, 52, 53].

The most widespread used serum inflammatory markers to screen for PJI are C-reactive protein (CRP), erythrocyte sedimentation rate (ESR) and white blood cell count (WBC) since elevation of these markers can signal infection. In cases with high-grade bacteria, they would most likely be elevated [54], however, all markers have also shown to be elevated due to a multitude of other causes, thus interpretation is not clear-cut [55]. To add to the complexity, the false negative rate is reported high in cases with low-grade bacteria [56].

In synovial fluid elevated cell count, increased polymorphonuclear percentage or presence of leukocyte esterase can be signs of infection [57-59]. However, in shoulders it is often not possible to aspirate any fluid from the joint (so-called dry-taps) [60] and combined with potential risk of introducing an infection by the aspiration-procedure itself considerations must be taken before attempting aspiration prior to surgery. Furthermore, leukocyte esterase has documented low diagnostic performance in shoulders [60, 61]. Several other biomarkers in the synovial fluid are of interest with especially alpha-defensin gaining attention [62]. However, in a selected cohort of shoulder PJI, the test has only showed a fair sensitivity of 63% and specificity of 95% [63].

4.3.3 Perioperative diagnostic tests

During surgery tissue-specimens are obtained for culturing. In Denmark, the most frequently used method was forwarded by Kamme and Linberg in 1981 [64]. Five individual specimens are obtained at the same time and from the same area in close proximity to cement and implant. All specimens are obtained with clean utensils to reduce the risk of contamination. This number of acquired specimens was recently supported by Kheir et al, who demonstrated that at least five biopsy-specimens were needed to maximise culture yield and diagnose PJI. However, they also noted that in cases with C acnes more biopsies may be needed to detect this microbe in at least two cultures. For details regarding culture protocol please see section 6.3.

4.4 Defining PJI

There are several ways of defining PJI. In this thesis we use two, either by cultures alone or by the definition proposed by the Musculoskeletal Infection Society (MSIS).

The definition of PJI based solely on cultures is based on the assumption that the joint is sterile, hence suspicion of infection is raised when the same bacteria repletely is isolated from the joint. The definition by Kamme and Lindberg proposed growth in four or five of the cultures, with one or two identical strains of bacteria to represent PJI. Others recommended only three positives cultures in a set of five to represent significant growth [65, 66] which is the threshold typically used in Denmark.

The main limitation of relying solely on microbiology to define PJI is the risk of missing infections because sensitivity of cultures in shoulders is reported as low as 55% [46]. The term "culture-negative" PJI covers cases in which a clinical suspicion of PJI exists and patients are treated as such, but no microbes can be identified. These cases are reported to account for up to 20% of patients treated for PJI [67]. Novel techniques such as ultrasonic wash (sonication) to dislodge biofilm from the explanted prostheses and gene sequencing have been proposed to increase yield and decrease the rate of culture negative PJI [47, 68, 69]. However, the efficacy of sonication in diagnosing shoulder PJI has been questioned and it remains to find its exact place in the diagnostic workup [46, 68].

The MSIS definition addressed the need for a more universally adopted definition and in 2014 the composite criteria to define PJI which is outlined in Figure 4 were presented [70] [71]. It is notably that the definition contains a mix of pre- and perioperative information, thus it is not possible solely to use this definition in the diagnostic workup to decide further course of treatment. Furthermore, the definition classifies patients as either infected or not-infected not taking into account patients with absence of clinically signs of infection despite isolation bacteria from several cultures ('grey zone' patients).

Figure 4: MSIS definition adopted ad ICM 2014

Figure 4: Shoulder PJI definition as adopted at ICM 2014
Major criteria:
Presence of a sinus tract from the skin surface to the prosthesis
A pathogen is isolated by culture from at least two separate tissue or fluid samples obtained from the affected prosthetic joint
Minor criteria:
Elevated serum C-reactive protein (CRP) AND erythrocyte sedimentation rate (ESR) *
Elevated synovial fluid white blood cell (WBC) count OR ++change on leukocyte esterase * test strip
Elevated synovial fluid polymorphonuclear neutrophil percentage (PMN%) *
Positive histological analysis of periprosthetic tissue *
A single positive culture
INTERPRETATION:
PJI is present if one of the major criteria OR at least three of the five minor criteria exist.

**Threshold for chronic PJI (>90 days from surgery): ESR >30 mm/hr, CRP >10 ml/L, Synovial WBC count >3000 cell/µL, leucocyte esterase elevated if + OR ++, histological analysis positive if >5 PMN in at least 5 high power fields*

To address this challenge, other authors have introduced shoulder PJI definitions by adding categories with intermediate risk into their definition [72]. To avoid fragmentation of a diagnosing shoulder PJI a new MSIS-shoulder definition was introduced in 2018 (a detailed description can be found in the Appendix) [73]. This definition also accounted for the challenge imposed by the high rates of positive cultures with C acnes in absence of other signs of infection [32, 34, 74]. The MSIS-shoulder definition uses a scoring system to allocate patients into four categories: "unlikely infection", "possible PJI", "probable PJI" and "definite PJI" according to the summarised score. However, the four categories are not followed by treatment guidance, thus studies are needed to clarify the usefulness of the four categories for treatment purpose.

To summarise both definitions, either solely based on cultures or by the MSIS composite criteria, require analysis of perioperative acquired material to obtain a definitive diagnosis of PJI. Thus, the definitions are not able to guide surgeons in the decision-making process and choose the most optimal treatment strategy. Today all patients including those with little suspicion of PJI undergo standard revision including oral antibiotic treatment for two weeks after surgery due to the need for prolonged culturing. This limbo between operation and culture-result may result in overtreatment of patients with no PJI and undertreatment of patients with PJI, as they might have been better off with a two-stage revision. This challenge highlights the need for new methods for non-invasive and preoperative diagnostic modalities to be incorporated into our assessments preoperatively.

4.5 Nuclear medicine

When objective symptoms of PJI are vague and symptoms are found in combination with malpositioning of implant components, failure of rotator cuff or glenoid attrition it becomes virtually impossible to attribute the subtle symptoms like stiffness and pain to infection or to mechanical reasons. In such cases new high performing diagnostic modalities are welcomed to aid the preoperative decision-making.

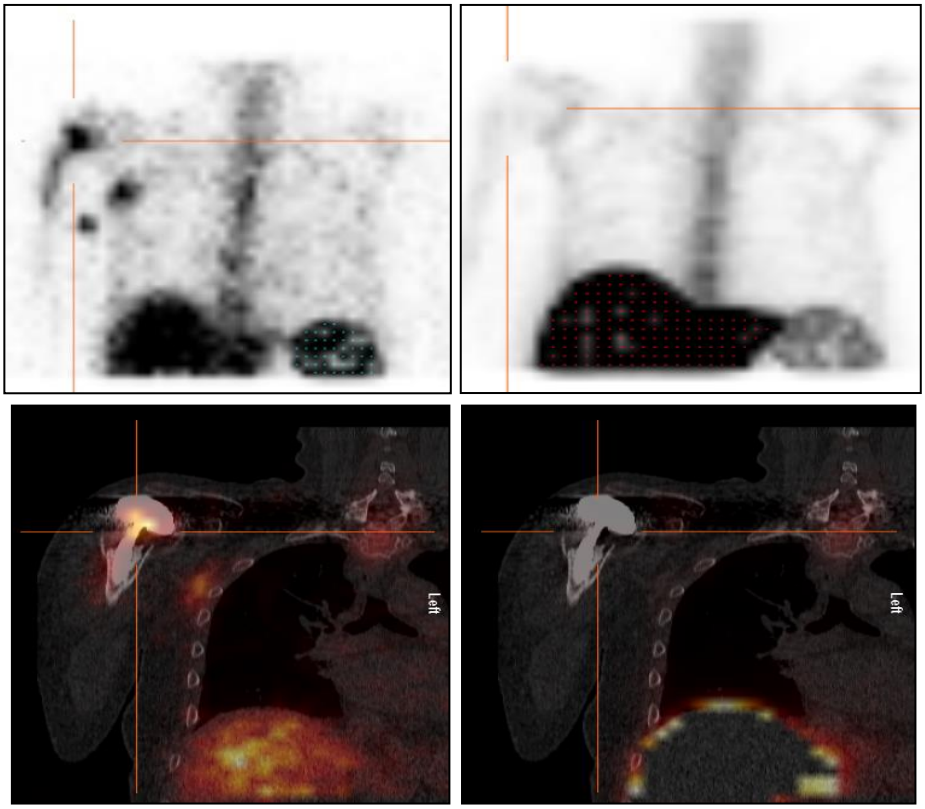
Nuclear imaging such as scintigraphy or PET relies on labelling or uptake of a tracer in specific tissue and shows physiologic activity rather than snapshot of the current bone and prosthesis components depicted on a normal radiograph. WBC-scintigraphy has in several studies shown sensitivity and specificity over 0.8 [75, 76] in detecting PJI of the hip. Similarly, FDG-PET has shown sensitivity and specificity over 0.8 [77-79] in detecting PJI of the hip; values which surpass many of the more common tools of PJI workup described in section 4.3. Despite the reported high diagnostic accuracy in diagnosing PJI of both WBC-scintigraphy and FDG-PET routine use seems relatively limited; probably in part due to sparse experience among both surgeons and nuclear medicine departments. One review from 2015 even questioned the use of such imaging in the diagnostic workup of PJI [80]. Nevertheless, a recently published paper endorsed by many infection oriented societies proposed a diagnostic strategy including nuclear imaging in cases of suspicion of PJI after standard workup [81].

4.5.1 SPECT CT

A scintigraphy is a nuclear medicine investigation in which a molecule labelled by a radioactive isotope (the radiotracer) is injected into the circulation and subsequently decays in the tissues in which it has accumulated [82]. Radiotracers have been developed to accumulate in all sorts of tissues and may be used to depict both loss of function (as in perfusion imaging) or increased activity (as in "hot" thyroid nodules). As the tracer decays, gamma-radiation is emitted as photons with isotope specific energies (Single Photon Emission), e.g. 174 keV and 247 keV by ^{111}Indium and 140 keV for ^{99m}Technetium. The decaying radiotracers can be detected by a gamma camera in either two dimensions termed planar scintigraphy or in three dimensions termed a Single Photon Emission Computerized Tomography (SPECT). If isotopes decaying by different energies are used simultaneously, the gamma camera can discriminate between the radiotracers by setting an energy level threshold of e.g. 140 keV and 247 keV. Three-dimensional SPECT enables a more precise localisation of the activity compared to a planar scintigraphy with a resolution of approximately 8 mm [83, 84]. Adding to the clinical use of scintigraphy, newer SPECT scanners are in most cases equipped with a conventional CT scanner. When obtaining a conventional CT scan simultaneously with the SPECT scan images can afterwards be combined to pinpoint the activity to specific organs visible on the conventional CT scan. This procedure is called a SPECT CT scan.

Evaluation of an infection/inflammation on planar or SPECT scan is often done by pattern recognition. Activity in areas not expected to show activity usually represents a pathologic finding. A more elaborate setup may be needed in instances like diagnosis of infection; either by scanning at several timepoints and/or labelling of both leucocytes and bone marrow. In a white blood cell scan (WBC scan) leucocytes are labelled ex-vivo, re-injected into the patient and subsequently they migrate to the infected site [85]. To increase the diagnostic accuracy of the WBC scan, sequential scans may be performed as early (1 hour after re-injection), delayed (3 hours after re-injection) and late (24 hours after re-injection) imaging. If activity peaks after 3 hours and declines to zero at 24 hours, it indicates inflammation. If the activity increases from early or delayed to late images it is indicative of infection. Frequently used isotopes such as ^{111}Indium has a half-life of 2,8 days and ^{99m}Technetium a half-life of 6 hours. Consequently, if ^{99m}Technetium is used for late image acquisition, a protocol correcting for decreasing activity (decay corrected) is often used. Some pitfalls in the WBC should be addressed though. Labelled leucocytes accumulate in areas of infection but can also accumulate in areas with active bone marrow. In such cases, non-infected areas can be misinterpreted as infected. To overcome this challenge, a dual-isotope approach can be utilised to enhance diagnostic performance [86]. One isotope e.g. ^{111}Indium is used to label leucocytes and another isotope (^{99m}Tc) is used to label colloids that accumulate in the bone marrow. Activity from the radiotracers is detected results in two images, one depicting leucocyte accumulation and one depicting areas with active bone marrow. Finally, the bone marrow image is "subtracted" from the WBC image resulting in an image showing only pathological leucocyte accumulation.

Figure 5: Example of a positive WBM/BM SPECT CT scan. Images showing mismatch between leucocyte and bone marrow activity. Upper left coronal SPECT showing accumulation of labelled leucocytes near the arthroplasty and in lymph nodes near the axilla. Upper right coronal SPECT showing bone marrow activity. Lower left SPECT CT of the same patient showing leucocyte activity. Lower right SPECT CT of the same patient showing bone marrow activity.

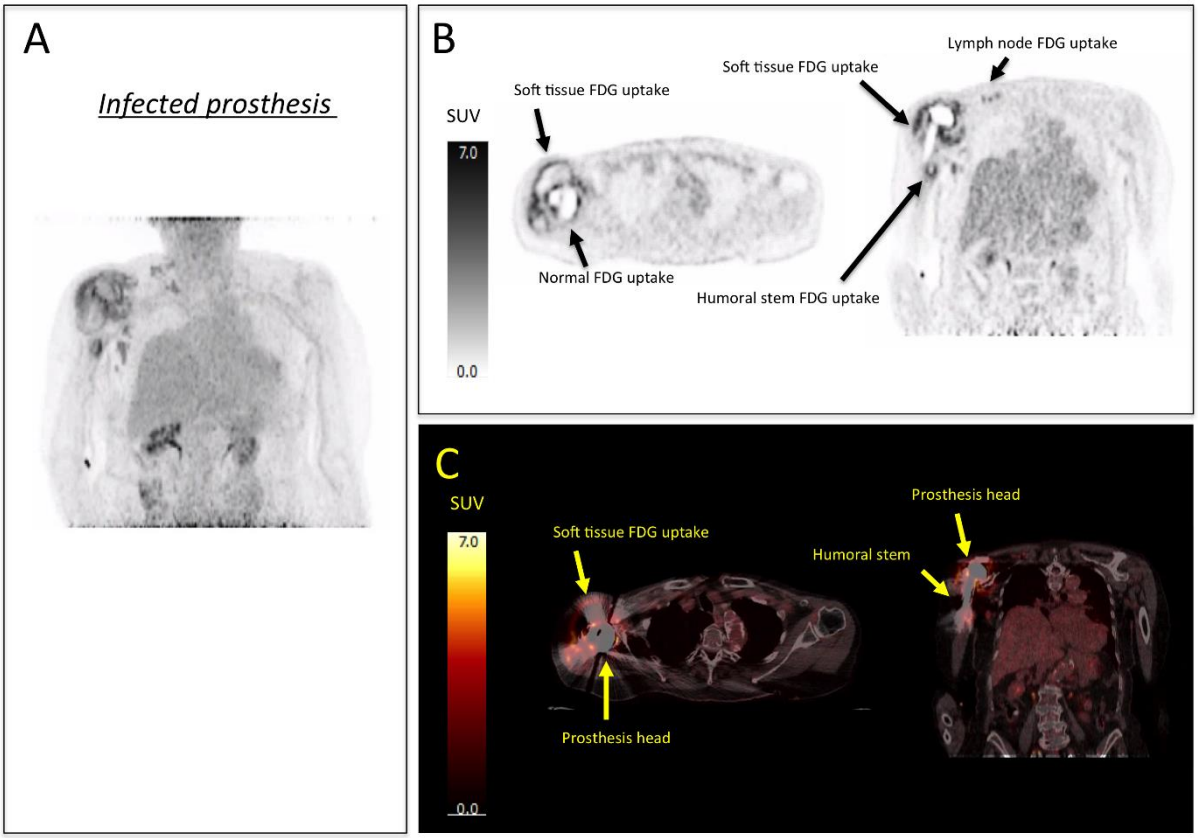


4.5.2 ^{18}F-FDG-PET/CT

Positron Emission Tomography (PET) is a nuclear imaging technique that differs somewhat from SPECT CT. In PET the radioactive isotopes used emit positrons, a positively charged particle which is antiparticle to the electron [82]. The positron travels a few millimetres in tissue before it collides and interacts with an electron. The interaction produces a pair of photons moving in opposite directions and with the same energy (511 keV, annihilation photons). For image acquisition the patient is placed on a gurney encompassed by a circular detector capable of detecting the annihilation photons. However, activity is counted only when the two photons traveling in opposite direction reach the detector simultaneously (coincidence detection). If all detections were to be counted, location of source of emission is not possible, consequently, all counts with no coincidence detection are discarded. In this way a dataset of millions of coincidence counts is collected and these data are then processed to a three-dimensional image of the tracer activity. Figure 6 illustrates FDG-PET and fused FDG-PET/CT images.

The most commonly used PET isotope is ^{18}Fluor (^{18}F) with a half-life of 110 minutes. This relatively long half-life allows for production of large doses of radiotracer and consequently also for implementation in the clinic. When the hydroxyl group in a glucose-molecule is substituted with ^{18}F, the result is the glucose-analogue PET radiotracer: 2-fluoro-2-deoxyglucose (FDG). FDG is taken up by the same glucose transporters (GLUTs) as endogenous glucose but does not undergo full glycolysis. It is therefore "trapped" in all energy consuming tissues such as the brain, heart and liver but also notably in malignant and inflammatory cells. Since FDG competes for uptake by the same GLUTs as endogenous glucose, it is recommended that serum-glucose levels be kept below 11 mmol/L [87, 88]. Today, FDG accounts for nearly 60% of all PET imaging in Denmark and is by far the most important and useful PET radiotracer. However, other more specific PET radiotracers than FDG have been developed and imaging of bone activity, peptide synthesis, hypoxia or angiogenesis is possible [82].

Figure 6: Example of a positive FDG-PET/CT. Images A and B show increased uptake near a left shoulder replacement. Image C showing the fused PET and CT scans with evident tracer activity near the implant.



FDG-PET has a resolution of approximately 5 mm [84] and evaluation of images is primarily qualitative and based on pattern of activity. A semiquantitative measure of magnitude of activity can be represented by tissue-to-background ratios or Standardized Uptake Value (SUV). The latter is a measure of local tracer activity normalised for injected dose and body weight [89]. It is often reported as the maximal uptake measured, the SUV_{max}, for a given region of interest (ROI, applied on 2D images) or volume of interest (VOI, applied on 3D images). Such values can be utilised to monitor treatment response and

result in less observer-variation compared to a visual diagnosis only. However, the semiquantitative measures cannot necessarily be used as a threshold for disease, as evidenced by the poor correlation between SUV_{max} and PJI [90, 91].

4.6 Functional outcome after revision

Functional outcome can be assessed by either objective measures e.g. ROM or strength or by a subjective perspective e.g. pain or ability to cope with activities of daily living. The latter are important not only to the patients but also to the clinicians, thus patient reported outcome (PRO) is increasingly used. In this thesis the term patient reported outcome measures (PROM) refers to the instrument used to collect data and PRO to the result or score [92]. More than 30 shoulder PROMs have been described [93, 94] of which several have been translated, validated and found to be reliable and responsive in Danish among those Oxford Shoulder Score (OSS), Constant Murley Score (CMS) and Western Ontario Osteoarthritis Score (WOOS) [95-99]. Many shoulder specific PROMs are general e.g. American Shoulder and Elbow Society Standardized Shoulder Assessment Form (ASES) and University of California at Los Angeles Shoulder Score (UCLA), but several PROMS are also condition specific e.g. OSS for outcome after shoulder surgery [100] or disease specific e.g. WOOS for arthritis [101]. However, in the literature there seems to be a lack of consensus in choice of PROM; to illustrate 11 publications reporting on outcome after shoulder revisions used 12 different scores. Direct comparison of scores of different PROMs cannot be performed and although some have been correlated interpretation should be performed with caution [102]. No PROM has been designed specifically to shoulder replacement revision surgery [94]. However, ASES, OSS and WOOS are frequently used to cover both primary and revision joint replacements, the two latter being used as PROM in The New Zealand Joint Registry, The National Joint Register of England, Wales, Northern Ireland and The National Joint Register of England, Wales, Northern Ireland and Isle of Man of Man (NJR), and The Danish Shoulder arthroplasty Register (DSR).

When PRO scores are evaluated one must pay attention to that a change may be statistically significant but not clinically relevant to the patient. To overcome this the minimal clinically important difference (MCID) describes how large a difference (calculated in points) that is needed before patients actually experience change in outcome. However, MCID is not always available. Furthermore, it is important to note in which context the MCID is established, especially regarding the more general shoulder PROMs. To give an example, the same number of points may not translate to experienced change in a group of patients treated with rotator cuff surgery as it would in a group of patients treated with joint replacement despite that the same PROM is used.

Most studies presenting PRO report on outcome of aseptic revisions [9, 10, 103-108]. In contrast, studies selectively assessing outcome in patients with UPC have mainly focused on frequency of UPC [52, 109], risk-factors [32, 33, 40, 110] or used hard outcomes like re-infection or re-revision as primary endpoints [42, 111, 112], only few have used PRO scores as outcome [41, 42]. Furthermore, some of the UPC studies have been limited by small sample size [42], long time span of inclusion [112] or the lack of complete preoperative outcome measures [42, 112, 113].

To summarise, the consensus of which PROM best describes the outcome of revision surgery is non-existing. However, ASES, WOOS and OSS have all been used in this setting hereby providing some literature to compare results. The knowledge of functional outcome in patients with UPC is limited and based on few studies.

5 Aim of the thesis

The findings from previous studies identified a need for studies that investigate whether nuclear medicine can provide a useful tool to diagnose chronic shoulder PJI. Furthermore, there is a lack of studies that investigate change in functional outcome after revised shoulder joint replacement in patients with emergence of UPC.

The thesis is based on three studies with the following designs and aims:

The diagnostic accuracy of WBC/BM SPECT CT (Study I)

The aim of this diagnostic accuracy study was to establish the diagnostic performance of ^{111} Indium-labelled White blood cell/^{99}Technesium labelled bone marrow SPECT CT in diagnosing PJI of failed shoulder replacements. We hypothesised that WBC/BM SPECT CT has similar sensitivity and specificity in detecting shoulder PJI as seen in lower-limb PJI.

The diagnostic accuracy of FDG-PET (Study II)

The aim of this diagnostic accuracy study was to establish the diagnostic performance of ^{18}F-FDG-PET/CT diagnosing PJI of failed shoulder replacements. We hypothesised that FDG-PET has similar sensitivity and specificity in detecting shoulder PJI as seen in lower-limb PJI.

Functional outcome after revision surgery (Study III)

The aim of this prospective cohort study was to assess change in short-term functional outcome in patients with and without UPC after revision surgery. We hypothesised that patients with emergence of UPC resulted in inferior short-term outcome after a standard revision compared with patients without emergence of UPC.

6 Materials & methods

6.1 Ethical considerations

In all three studies, we obtained approval to store and handle patient data from the Danish Data Protection Agency (ref. no. 1-16-02-567-13). According to the local ethical committee, no formal approval was required for Study I and III, because the studies were conducted as a quality control study at orthopaedic departments (ref. no. 217/2013). Ethical approval was obtained for Study II by the ethical committee in the Central Region Denmark (ref. no. 1-10-72-229-15) and the study was conducted in accordance with the Declaration of Helsinki [114].

For Study I and II in general, patients did not directly benefit from participating since results from the scans were used only as additional information and not as a definitive diagnosis in the process of diagnosing PJI. This was owing to the aim of the studies to clarify the diagnostic accuracy of FDG-PET in a setting of patients with failed shoulder replacements.

When patients are examined with WBC/BM SPECT CT (Study I), they are exposed to approximately 10 milliSievert (mS) radiation. In comparison, the background radiation in Denmark ranges from 3 mS - 20 mS/year depending on location. In a population with a mean age of 65 years, the radiation exposed to participant of the study equals an induced risk of cancer of approximately 0.03% corresponding to an increase from 50% to 50,03% (results estimated via xrayrisk.com). After discussions among orthopaedic surgeons and nuclear medicine physicians, the minor increase in cancer risk induced by adding the WBC/BM SPECT CT to the standard preoperative investigation was found acceptable as the increased risk was considered to be outweighed by the potential impact on treatment.

When patients are examined with FDG-PET (Study II), they are exposed to approximately 12 mSv radiation. This exposure was rated a class III study according to the guidelines for use of ionizing radiation in medical research (dose exceeding 10 mSv) [115]. Similar to SPECT CT, the risk of inducing cancer is increased from approximately 50% to 50,03% in an average study participant. In FDG-PET, an intravenous access was needed to inject the radiotracer and this procedure was accompanied with minimal pain and risk of infection. The potential risks and adverse events related to the FDG-PET scan are very small and counterbalanced by the benefit future patients can expect from the study. Before enrolment, patients received oral and written information and only patients with written consents were enrolled in the study.

6.2 Patients

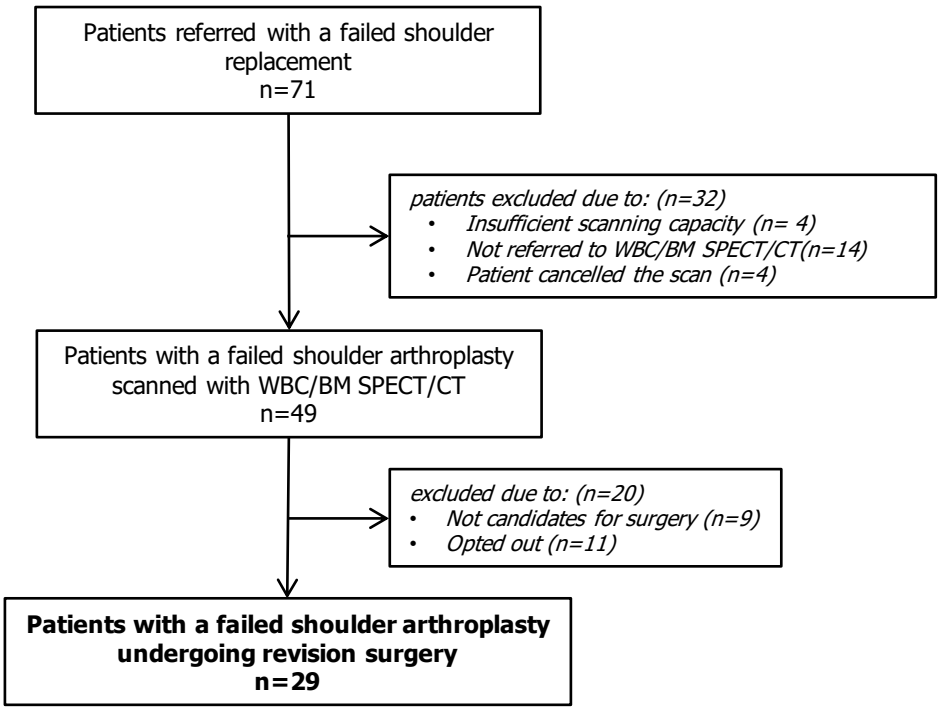
At the time of initiation of the study, October 1st, 2014 at Herlev Hospital (Centre One) and April 1st, 2014 at Aarhus University Hospital (Centre Two) the two participating centres were the only centres approved by the Danish National Board of Health to perform revisions of shoulder replacements. During a 3-year study period at each centre all patients referred with complications or an unsatisfactory result after a shoulder replacement were registered. Referrals came from either general practitioner, another

orthopaedic department or were "own patients". This group formed the basis cohort from which patients to the three studies were recruited.

The diagnostic accuracy of WBC/BM SPECT CT (Study I)

During the period of April 1st, 2014 to May 15th, 2015 patients were enrolled at Centre Two. As part of the preoperative workup a WBM/BM SPECT CT was planned for all referred patients. However, patients were excluded from participation in the study if they were suffering from acute postoperative infection (<4 weeks after last surgery in the affected shoulder), had an acute periprosthetic fracture, were unable to corporate or communicate in Danish, had received any prior treatment due to chronic infection in the ipsilateral shoulder or were not scanned. Patients who completed the scan, but decided to opt out or were found unfit for surgery were also excluded. This resulted in a study population of 29 patients with a WBC/BM SPECT CT and a true infection-diagnosis based on cultures of biopsy-specimens after revision of a joint replacement. Flow-chart illustrating patient inclusion found in Figure 7.

Figure 7: Flow-chart of patient inclusion in Study I



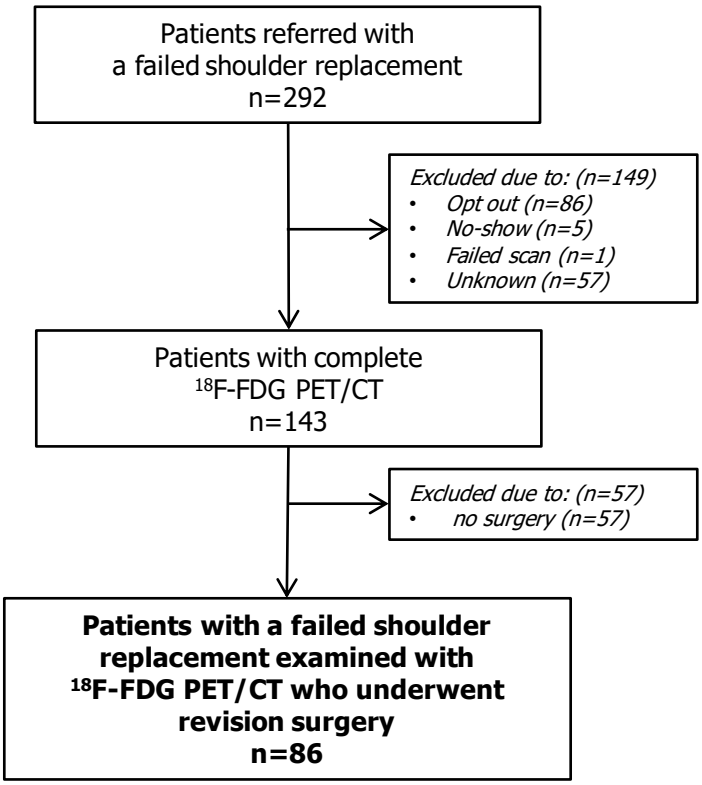
The diagnostic accuracy of FDG-PET (Study II)

For Study II, we enrolled consecutive patients with written consent from both Centres One and Two from October 1st, 2015 to September 30th, 2017. We excluded patients, if they i) suffered from acute postoperative infection (<4 weeks after last surgery in the affected shoulder), ii) had an acute periprosthetic fracture, iii) had prior revisions due to chronic infection in the ipsilateral shoulder, or iiii) were unable to corporate or communicate in Danish. Furthermore, we excluded patients not undergoing

surgery (despite of complete scan) as tissue cultures would not be available for validation of the FDG-PET diagnosis.

In total, the study population consisted of 86 patients with a FDG-PET and an infection-diagnose diagnosis based on cultures of biopsy-specimens after revision surgery. Flow-chart of patient inclusion is illustrated in Figure 8.

Figure 8: Flow-chart of patient inclusion in Study II



Functional outcome after revision surgery (Study III)

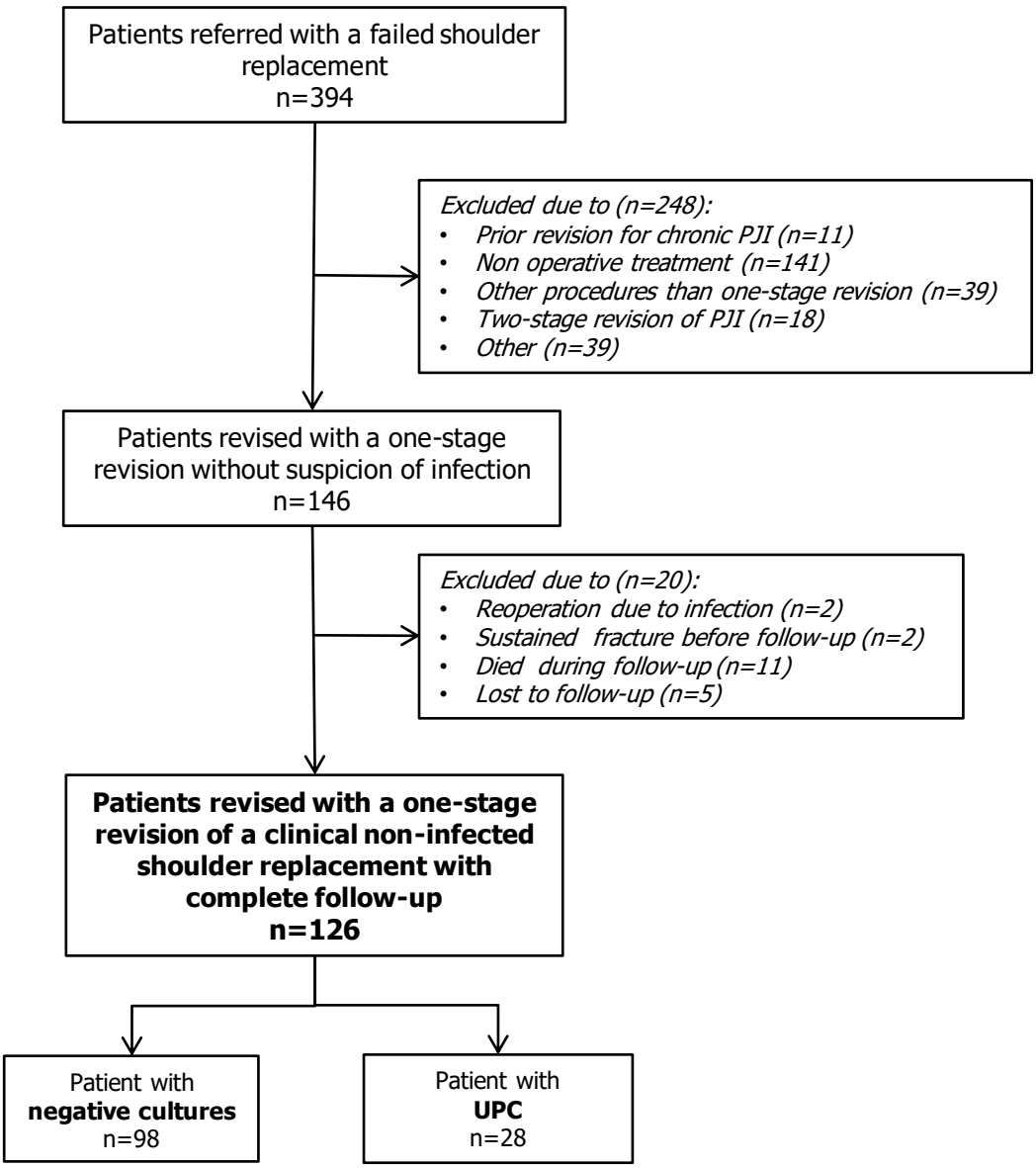
For Study III, we included consecutive patients from the target population from both Centres in the period from April 1st, 2014 to September 30th, 2017.

After completion of baseline assessment and preoperative workup, we enrolled patients without i) acute postoperative infection (<4 weeks after last surgery in the affected shoulder), ii) prior treatment due to chronic PJI of the affected shoulder, iii) a diagnose of infection using the MSIS PJI definition, or iiii) other procedures planned than one-staged component exchange (e.g. two-staged revision, open/arthroscopic release, isolated bone-grafting of glenoid bone-loss), or iiiii) the ability to corporate or communicate in Danish. At time of retrieval and analysis of case-report-forms, we furthermore excluded patients with missing data. In addition, patients who died during the study period or experienced reinfection or factures were excluded.

At the time of writing this thesis 9 patients awaits 2-year follow-up with the last scheduled in December 2019. Since all these patients have baseline registrations, I chose to include them in the analysis of Study III. Consequently, the current number of included patients is 117 but the expected is 126.

In total the study population consisted of 126 patients undergoing a standard revision of a failed shoulder replacement. Figure 9 shows flowchart of patient inclusion.

Figure 9: Flow-chart of patient inclusion in Study III



6.3 Definition of variables

Imaging protocol and analysis of WBC/BM SPECT CT (Study I)

Image acquisition was done according to the following protocol. To label the WBC, 50 ml of whole blood was extracted and ex-vivo labelled with 30 Mega Becquerel (MBq) ^{111}Indium-Tropolone, and subsequently reinjected. Since time is needed for the WBC to migrate to the infected areas, the image acquisition was scheduled approximately 24 hours later. At day of imaging (day two) 400 MBq ^{99m}Tc – sulfur colloid was injected intravenously to label the bone marrow. Patients were scanned within one hour after injection using a gamma-camera with simultaneous dual isotope imaging. A Siemens Symbia (T16)-scanner (Erlangen, Germany) equipped with medium-energy collimators was used with 15% window centred on 140 keV, a 15% window centred on 174 keV and a 15% window centred on 247 keV in a 128 X 128 matrix. Planar imaging of thorax and shoulders in anterior and posterior projection was performed with a 128*128 matrix with 20 minutes pr. view. The same collimator and energy windows were used to perform both planar and tomographic imaging. To locate pathological foci in 3D, a SPECT with low-dose CT (50 mAs) (SPECT CT) was performed in a 64*64 matrix with 32 views of 40 seconds pr. view. The estimated dose of radiation per patient was 10 mSv.

Tracer distribution was analysed visually by recognition of tracer-uptake-pattern. The WBC and BM images were visually compared to identify discordant uptake. Increased focal uptake on the leucocyte scan and decreased or normal uptake on the bone scan is suggestive of infection. If no infection is present uptake will show identical distribution on the two scans. Hence, a scan positive for infection was defined as any mismatch of uptake showing increased leucocyte activity in close proximity to the arthroplasty regardless of the intensity or location. The final WBC/BM SPECT CT diagnosis was dichotomous (positive/negative).

Imaging protocol and analysis of FDG-PET

Image acquisition was done according to the following protocol. All scans were performed on a Siemens Biograph 64 FDG-PET scanner (Erlangen, Germany). Patients fasted at least 6 hours before the procedure, to ensure a bedside fasting blood glucose of no more than 11 mmol/L. Patients with a higher level were rescheduled. An initial low-dose CT without contrast enhancement (50 mAs) was performed in order to correct for photon attenuation and to co-localize FDG uptake and anatomical structures. The imaging was performed 60 minutes after intravenous administration of the radiotracer. At Centre One the FDG-PET was performed from elbow to basis cranii (4 MBq 18F-FDG kg-1; 2 min per bed position in three-dimensional mode) and at Centre Two from the mid-thigh to the skull (5 MBq 18F-FDG kg-1; 3 min per bed position in three-dimensional mode). Reconstruction of attenuation-corrected images was done using visually comparable, ordered subset expectation maximization algorithms with point-spread function (PSF) (Siemens Biograph: Four iterations, 21 subsets, 3-mm Gaussian post-processing filter, matrix size 336 x 336). The estimated dose of radiation per patient was 9.5 mSv at Centre One and 12.8 mSv at Centre Two. Tracer distribution was analysed visually based on recognition of tracer-uptake-

pattern. Since visual interpretation of the images is a cornerstone of the FDG-PET diagnosis, a standardised assessment was developed since evaluation of FDG-PET of failed shoulder replacements has not previously been described. First, two experienced ortho-nuclear consultants reviewed all scans based on best experience and gave a dichotomous diagnosis (infection/no infection). Second, the Shoulder specific assessment was developed by comparing results from cultures, scans, knowledge of rate of loosening of the different prosthesis components and previous developed FDG-PET criteria for PJI of hip replacements [77, 116-118]. This Shoulder specific assessment consists of three criteria of which at least one had to be present to define infection. Examples of each criterion are found in Figure 10. Subsequently, all scans were graded according to the Shoulder specific assessment by three blinded reviewers. The final FDG-PET diagnosis was dichotomous (positive/negative) and established by majority decision e.g. if two reviewers rated a scan negative and one positive, the final FDG-PET diagnosis was negative.

Figure 10: Shoulder specific assessment of FDG-PET. Criterium 1: Increased FDG uptake in soft tissue adjoining joint cavity. Criterium 2: Increased FDG uptake along the humerus stem. Criterium 3: Increased FDG uptake in regional lymph nodes

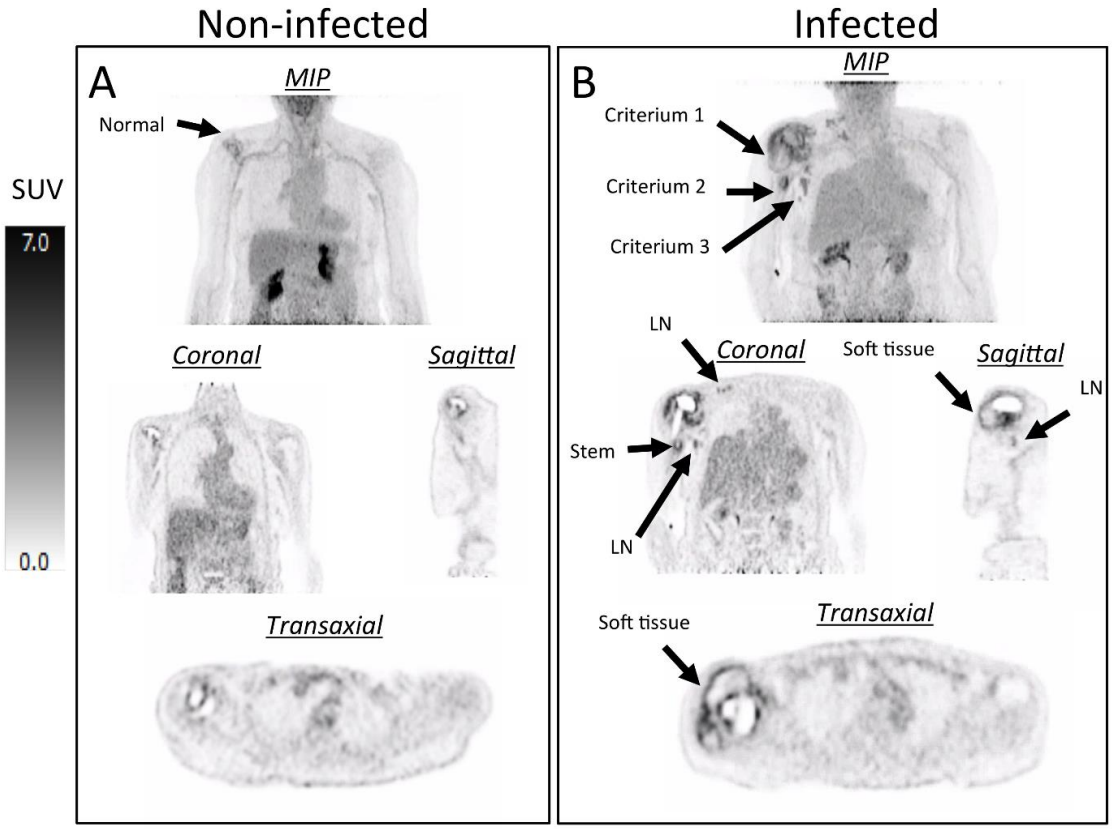


Image A: Non-infected patient with right sided shoulder replacement and no areas with pathologic FDG uptake. Image B: Infected patient with increased FDG uptake in soft tissue, along humerus stem and in regional lymph nodes.

Definition of infection (all three studies)

In all three studies, we used the same method for identification of bacterial growth. During revision, at least five biopsy-specimens were obtained with at least one specimen obtained from the bone-prosthesis surface. The remaining specimens were obtained from the joint synovia and areas with clinical suspicion of infection. All five specimens were taken immediately after surgical exposure to minimize the risk of perioperative contamination. Each specimen was obtained using clean utensils and transferred directly to sterile containers.

At the microbiological department, each specimen was divided in five. One sub-sample was cultured on 5% Blood Agar, on Anaerobic Agar plates (SSI, Denmark), in BLL Chromagar Orientation Medium (Becton Dickinson, Germany), one inoculated in Semisolid Agar+Pepsis Blood+Thioglycollate (SSI, Denmark) and one Serum Broth (SSI, Denmark). Agar plates were inspected for growth up to four days and were discarded at day four if no growth was observed. Semisolid Agar and Serum Broth were visually inspected for signs of growth on day four and day fourteen. Matrix-assisted laser desorption/ionization coupled with time-of-flight mass spectrometer (Bruker, Germany) (MALDI-TOF) was used for identification of strains.

In Study I and II, patients were classified as having PJI if growth was observed in minimum three of the five specimens OR if patients presented with draining sinus at initial examination (irrespectively of number of specimens with growth).

In Study III, only patients with growth in minimum three of five specimens were classified as PJI. Due to the preoperative lack of suspicion of PJI, we designated these patients as having UPC.

6.3.1 Exposure and outcomes

Study III was designed as a prospective cohort study with UPC as exposure and OSS as primary outcome and ROM and OSS-pain score as secondary outcomes.

Oxford Shoulder Score (OSS)

The OSS is a patient reported outcome measure (PROM) developed to measure the outcome after shoulder surgery due to degenerative conditions such as rotator cuff disease, arthritis and subacromial impingement syndrome in a single value. The OSS was developed to assess group and not individual change, before and after surgery and has shown to be highly responsive in this context (excluding stability surgery for which a separate questionnaire has been developed) [119]. The questionnaire consists of 12-items describing function (8-items) and pain (4 items) over the last four weeks. Extraction of measures on a subscale level can provide selective and meaningful information on pain or disability. As such, both the total OSS value and subscale values of pain (item 1,8,11,12) and function (item 2-7,9,10) can be utilised to evaluate the results of surgical intervention of the shoulder [120]. Using the current recommendation for scoring, each item is awarded from zero to four points. The total value represents the OSS with a maximum of 48 points. Higher score equals better functional result [121].

A Minimal Clinically Important Difference (MCID) is important to bear in mind, when changes on score are translated to a clinical setting. However, a MCID has not been established for shoulder replacement surgery. Studies have demonstrated a MCID of approximately 5 in a cohort of patients surgically treated for rotator cuff and subacromial decompression surgery. In a cohort of patients with proximal humeral fracture treated both operatively and non-operatively the MCID was 5 or 11 depending on the statistical method used [122, 123]. Further, authors of the OSS propose an estimation of MCID to be calculated as half of the standard deviation of the difference between the pre- and postoperative score ($½*SD_{oss\_diff}$) [121]. Comparison to other PROMS should always be performed with caution, but the OSS has been found to correlate with ASES, WOOS and CMS [95, 96, 124].

When first published, the scoring system assigned a score from 12 to 60, a higher value represented a worse function. However, in 2009 a change in scoring was recommended from the developers of the OSS. The change prompted a new system with a score from 0 to 48 points, a higher score represents better result. This change was issued to make scoring more intuitive and avoid confusion for surgeons using similar scores e.g. Oxford Hip Score. As a consequence, it is important to note if the old or the new scoring system was used when comparing published results. To convert the old score to the new score, simply subtract the old score from 60.

The questionnaire used in this study was the official translation (with minor changes due to misspellings in the official version) from Oxford University Innovations which has been validated and translated into Danish [96]. A copy is available in the appendix. Patients were asked to fill out a paper-version of the OSS at the end of the interview at baseline and 2-year follow-up. They were instructed to answer according to the affected shoulder for which they were receiving treatment. Assistance to answering was provided if a patient requested it. Attention was given not to rephrase questions or suggest where to tick off an item. If patients did not return for follow-up, they were contacted by phone and at their discretion the questionnaire was forwarded either online or by mail.

Measurement of forward elevation and external rotation (Study III)

The arc of motion was evaluated using a manual goniometer. The patients were standing and asked to elevate the shoulder in the scapula plane. The arc of motion was measured in 10 degrees intervals with a range of 0-180° for forward elevation and -90° to 90° for external rotation.

Forward elevation was measured by the maximal active pain-tolerable height with free scapula and the elbow in full extension. External rotation was measured with the elbow flexed 90°, placed to the side of the thorax and pointing in sagittal plane. This position represents 0°. Using the forearm as dial, rotation towards abdomen equals negative arc-of-motion and rotation away from abdomen represents positive arc-of-motion. Rotation was measured as the maximal active pain-tolerable rotation without thoracoscapular compensation or rotation in the spine.

6.4 Statistics

In all three studies, normality was checked visually by plotting data and using Shapiro-Wilk test. Parametric data were described with means and 95% confidence intervals (CI). Two independent means were compared with two-sided Student t-test. Paired means were compared with paired t-test. Non-parametric data were described with medians and 25-75% interquartile range (IQR) or range. Comparison of two independent medians was performed with Wilcoxon rank test, paired medians with Wilcoxon Signed Rank test. Comparison of means between 3 or more groups was performed using one- or two-way analysis of variance (ANOVA). Assumption for equal variance was checked with Levene's test of variances. Means were subsequent compared with the abovementioned method if significance was found using in either main- or interaction-effect. Distribution of binomial or categorical data was tested using Chi^2 statistics.

In Study I and II, sensitivity, specificity, accuracy, PPV and NPV were reported as fractions with 95% CI. We performed a sensitivity analysis using the MSIS major criteria as reference to investigate the impact of applying a less strict definition of PJI on the test results.

Sample size calculation for Study I: Prior studies of WBC/BM scintigraphy used to diagnose PJI in hip and knee replacements have consistently reported specificity in the range of 0.92-1.00 [125]. The prevalence of positive cultures was assumed to be 25% based on analysis of earlier years revision cases at the two centres. The maximal margin of error was set to 0.1. Setting level of significance to 0.05 and using the above assumptions minimum 46 patients were needed to estimate the specificity. A similar calculation based on sensitivity showed minimum 139 patients needed.

Sample size calculation for Study II: Prior studies evaluating diagnostic accuracy of FDG-PET to diagnose PJI of hip and knee joint replacements have reported sensitivity and specificity in the range from 0.33-1.00. When calculating the sample size for our study we estimated a specificity and sensitivity of 0.8. From analysis of earlier years revision cases at the two centres, an estimated prevalence of positive cultures was assumed to be 25%. The maximal margin of error was set to 0.1. Setting level of significance to 0.05 and using the above assumptions minimum 81 patients were needed for the study. A similar calculation based on sensitivity showed 246 patients needed in the study.

In Study II we furthermore assessed interrater agreement between the three independent reviewers using Fleiss' kappa for multiple reviewers [126]. A kappa value of zero indicates an agreement as expected by chance, values 0.00-0.20 were graded as poor, 0.21-0.40 as fair, 0.41-60 as moderate, 0.61-0.80 as good and values above 0.80 as excellent [127].

In Study III, missing data corresponding to one or two unanswered items were handled by imputing the mean of the remaining 10 or 11 values, as stated in the OSS manual [128]. The entire questionnaire was omitted if more than two items were unanswered. If patients had indicated two answers for one item, the worst score was adopted.

To evaluate if UPC had an impact on patient disabilities a comparison of OSS and ROM between patients with and without UPC at both baseline and 2-year follow-up was conducted. This comparison was planned using ANOVA, however, OSS at follow-up (in contrast to OSS at baseline) was not normal-distributed; hereby not allowing analysis with two-way ANOVA. Since an appropriate non-parametric alternative is not available the comparisons of OSS at follow-up were conducted pairwise with Wilcoxon rank sum test; having the increased risk of type 1-error due to multiple testing in mind. Further, to assess the outcome of the revision, analysis of the mean-change in OSS, OSS-pain score, elevation and external rotation from baseline to 2-year follow-up were conducted. The subscale OSS-pain score was calculated as the sum of scores from items 1,8,11 & 12.

7 Results

7.1 Patient characteristics

The diagnostic accuracy of SPECT CT (Study I)

In total, 29 patients were included as shown in table 1. There were no statistical differences between patients included in the study and the target population including referred patients with a failed shoulder replacement according to the characteristics shown in Table 1, except for the infection rate. Table 1 shows baseline characteristics for the referred population and the study group.

Table 1: Demographics of patients in Study I

		Referred patients	Study group
No. of patients		71	29
Mean age, (range)		65 (29-88)	63 (29-84)
Gender, % (n)	Female	46% (33)	34% (10)
	Male	54% (38)	46% (19)
Prosthesis age, median (range)		3y (0-29y)	2,5y (0-13y)
Revision rate, % (n)		62% (44)	59% (29/49)
Infection rate, % (n)		27% (14/51[x]) *	38% (11/29)

*statistical difference compared to study group,

[x] Denominator refers to all revised patients from the referred cohort.

The diagnostic accuracy of DGT-CT (Study II)

In total, 89 patients were examined with FDG-PET and included in the study. There were no statistical differences between patients included in the study and the target population including referred patients with a failed shoulder replacement. Table 2 shows baseline characteristics for the referred population and the study group.

Table 2: Demographics of referred and included patients in Study II

		Referred patients	Study group
No. of patients		292	86
Mean age (range)		68 (34-89)	67 (42-85)
Gender n (%)	Female	62% (182)	73% (63)
	Male	38% (110)	27% (23)
Prosthesis age, mean		4.2 y	4.5 y
Revision rate, % (n)		47% (137/292)	100% (86/86)
Infection rate, % (n)		27% (38/137)[x]	26% (22/86)

[x]Denominator refers to all revised patients from the referred cohort.

Functional outcome after revision surgery (Study III)

In total, 126 patients underwent a standard revision without any suspicion of infection and were included in the study. The baseline demographics are presented in table 3 according to type of UPS. Comparison between the culture negative and UPC groups revealed that significantly more males than females had UPC.

Table 3: Demographics of included patients in Study III			
		Patients without UPC	Patients with UPC
No. of patients n, (%)		98 (77%)	28 (23%)
Mean Age (years, range)		67 (29-85)	68 (44-81)
Gender, n (%)	Female	71% (70)	25% (7) *
	Male	29% (28)	75% (21) *
Prosthesis age (years, range)		5 (0-19)	5 (0-15)
Primary arthroplasty			
Hemiarthroplasty, with stem		52% (51)	54% (15)
Hemiarthroplasty, resurfacing		37% (36)	39% (11)
Total Shoulder replacement		8% (8)	4% (1)
Reverse prosthesis		3% (3)	4% (1)
Indication for primary arthroplasty			
Fracture/fracture seq.		32% (31)	32% (9)
Arthrosis		50% (49)	64 % (18)
Rheumatoid arthritis		6% (6)	4% (1)
Other/unknown		12% (12)	0% (0)
Serum markers			
C-reactive protein (mg/l, reference <8)		3.5 (95% CI: 2.5-4.4)	2.7 (95% CI: 0.9-4.4)
White blood cell count (10^9/l, reference <10.0)		7.3 (95% CI: 7.0-7.6)	7.2 (95% CI: 6.3-8.1)
Erythrocyte sedimentation ratio (mm/hour, reference <30)		13 (95% CI: 10-16)	9 (95% CI: 6-11)

*=statistical difference compared to culture negative group (p>0.05)

7.2 Main results

7.2.1 The diagnostic accuracy of SPECT CT (Study I)

Of the 29 included patients, the WBM/BM SPECT CT diagnosed two patients as positive and 27 patients as negative. Infection defined by cultures diagnosed 11 patients as infected and 18 patients as non-infected. Consequently, two scans were true positive, 18 true negative, 9 false negative and zero scans were false negative. Distribution of scanning results and infection-diagnose are presented in table 4.

Table 4: Summarised results of WBC/BM SPECT CT and infections status

	WBC/BM SPECT CT		
Infection diagnose by cultures	Negative	Positive	Total
No infection	18	0	18
Infection	9	2	11
In total	27	2	29

The diagnostic accuracy of WBC/BM SPECT CT showed a sensitivity of 0.18 (95% CI: 0.00-0.41), a specificity of 1.00 (95% CI: 1.00-1.00), a PPV of 1.00 (95% CI: 1.00-1.00) and a NPV of 0.67 (95% CI: 0.49-0.84). Applying MSIS definition as reference did not impact the overall findings, as only minor changes were observed in the sensitivity, NPV and accuracy. Table 5 presents diagnostic performances with different infection-criteria as reference standard.

Table 5: Diagnostic performance of WBC/BM SPECT CT based on cultures and MSIS definition

	By cultures	By MSIS
Sensitivity	0.18 (95% CI: 0.00-0.41)	0.15 (95% CI: 0.0-0.35)
Specificity	1.00 (95% CI: 1.00-1.00)	1.00 (95% CI: 1.00-1.00)
PPV	1.00 (95% CI: 1.00-1.00)	1.00 (95% CI: 0.00-0.00)
NPV	0.67 (95% CI: 0.49-0.84)	0.59 (95% CI: 0.41-0.78)
Accuracy	0.69 (95% CI: 0.52-0.86)	0.62 (95% CI: 0.44-0.80)

7.2.2 The diagnostic accuracy of FDG-PET (Study II)

Of the 89 patients included, the FDG-PET diagnosed 9 patients as positive and 77 patients as negative. Infection defined by cultures diagnosed 22 patients as infected and 64 as non-infected. Of the 9 patients with positive scan, three patients had infection and thus true positive, whereas six patients were non-infected and thus false positive. For the rest of the patients, 19 were classified as false negative and 58 as true negative. The distribution of scanning results are presented in table 6.

Table 6: FDG-PET Diagnose using Shoulder specific assessment

	FDG-PET DIAGNOSE		
Infection diagnose by cultures	Negative	Positive	Total
No infection	58	6	64
Infection	19	3	22
In total	77	9	86

Using the Shoulder-specific-assessment, the diagnostic accuracy of FDG-PET showed a sensitivity of 0.14 (95% CI: 0.03-0.35), a specificity of 0.91 (95% CI: 0.81-0.97), a PPV of 0.38 (95% CI: 0.67-0.76) and a NPV of 0.71 (95% CI: 0.61-0.81). Applying MSIS criteria (+one growth) as reference did impact the overall findings, as only minor changes were observed in the sensitivity, NPV and accuracy. Table 7 presented diagnostic performances with different infection-criteria as reference standard.

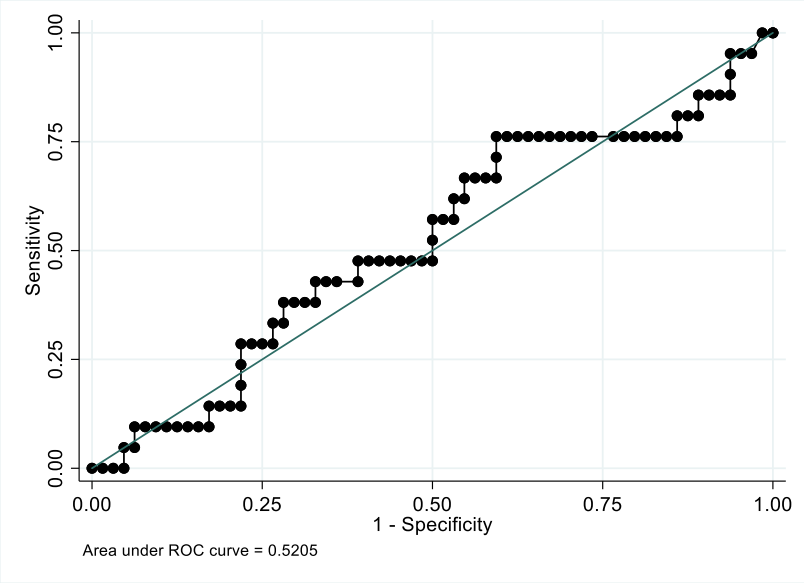
Table 7: Diagnostic performance of FDG-PET using Shoulder specific assessment

	By cultures	MSIS criteria
Sensitivity	0.14 (95% CI: 0.03-0.35)	0.21 (95% CI: 0.06-0.35)
Specificity	0.91 (95% CI: 0.81-0.97)	0.95 (95% CI: 0.89-1.00)
PPV	0.38 (95% CI: 0.15-0.70)	0.67 (95% CI: 0.36-0.97)
NPV	0.71 (95% CI: 0.67-0.75)	0.70 (95% CI: 0.60-0.80)
Accuracy	0.71 (95% CI: 0.61-0.81)	0.70 (95% CI: 0.60-0.79)

Based on the Shoulder specific assessment, the interrater agreement on the FDG-PET diagnose was moderate (0.56 Fleiss' kappa). When investigating the agreement on the sub-criteria of the Shoulder specific assessment, the highest rate of agreement was 0.61 (good) for pathologic uptake along the stem followed by 0.57 (moderate) for uptake in regional lymph nodes and 0.34 (fair) for tracer uptake in soft tissue adjoining the shoulder joint.

A receiver operating characteristic (ROC) curve is presented in Figure 11. The area under the curve is 0.52, indicating SUV_{max} is uninformative as a predicter of PJI.

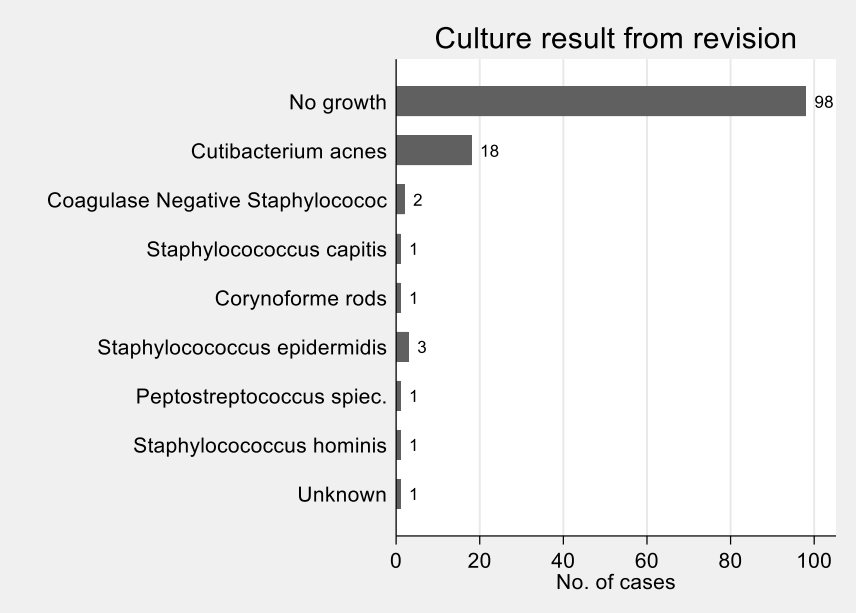
Figure 11: ROC curve of SUV_{max}.



7.2.3 Functional outcome after revision surgery (Study III)

In total, 126 patients were included in the study and thus thought to be aseptic. Nevertheless, 23% (28/126) patients were diagnosed with infection by cultures and subsequently designated as patients with UPC. The most frequently isolated bacterium found was C acnes in 64% of the patients (18/28). A frequency graph of the isolated bacteria from the revision is presented in Figure 12.

Figure 12: Bar-chart of isolated bacteria in culture positive patients



Both patients with and without UPC had a median OSS at baseline of 22 points. At 2 years follow-up the patients without UPC had a median OSS of 37 (IQR: 26-44) and those with UPC slightly lower 35 (IQR: 24-45). There was a significant increase in the change in OSS in both groups of patients corresponding to a mean increase in OSS of 14 points (95% CI: 11-16) for patients without UPC and 13 points (95% CI: 8-17) for patients with UPC. Stratification by final implant showed an increase in OSS at 2-year follow-up score for all implants. The MCID for OSS was estimated by ½*SD to 6. Table 8 shows the OSS for all patients and stratified by implant.

Table 8: Oxford Shoulder Score before and after revision

		Oxford Shoulder Score			
	No. of Patients	Baseline		2-year follow-up***	Mean difference
All patients		Mean (95% CI)	Median (IQR)	Median (IQR)	(95% CI)
Culture negative	98	20 (18-22)	22 (13-26)	37 (26-44)	14 (11;16)
UPC	28	20 (17-23)	22 (16-24)	35 (24-45)	13 (8;17)
Patients by final joint implant					
Reverse					
Culture negative	67	18 (16-20)	17 (12-24)	31 (43-35)	15 (11;18)
UPC	18	19 (15-23)	21 (15-25)	26 (21-40)	9 (3;15)
TSR					
Culture negative	21	27 (25-29) x	28 (24-31)	31 (43-35)	11 (8;15)
UPC	4	20 (14-25) x	20 (17-23)	45 (37-48)	23 (11;34)
Hemi					
Culture negative	10	18 (12-23)	16 (16-25)	29 (23-34) x	12 (3;21)
UPC	6	24 (16-32)	24 (21-26)	45 (38-46) x	17 (3;30)

x =significant difference between culture negative and UPC group, **=significant difference between baseline and 2-year follow up, ***Due to non-parametric distribution of data, scores are only presented as median and IQR (25th and 75th percentile)

At baseline and 2-year follow-up single parameters differed statistically between patients with and without UPC (baseline OSS in the TSR group, 2-year OSS, baseline and 2-year external rotation in the HA group). The two latter favouring patients with UPC. However, no overall pattern in outcome measures, either before or after surgery, could be found to support a hypothesis of negative impact of UPC. Detailed results are presented in table 8 and 9.

Patients with and without UPC had similar forward elevation with a mean of 78°at baseline. Patients without UPC had a slightly higher forward elevation at 2-year follow-up compared to patients with UPC corresponding to 123° (95% CI: 114-132) and 115° (95% CI: 94-137), respectively. Thus, the range of forward elevation increased significantly in both groups with no statistical difference between groups

with and without UPC were found. Stratifying according to final implant, patients revised to a RSA experienced a significant gain in forward elevation regardless of culture status. Equally, TSR patients without UPC had a significant increase in forward elevation. The remaining patients did also demonstrate a trend towards increased forward elevation, but without being statistically significant. In contrast, neither patients with or without UPC showed signs of improvement in external rotation. Range of external rotation is presented in Table 9 for all patients and stratified by implant.

At baseline patients without UPC had a mean pain-score of 4.2 (95% CI: 3.7-4.7) and at 2-year follow-up the score was increased to 9.7 (95% CI: 8.8-10.6) corresponding to a mean difference of 5.6 95% (CI:4.6-6.5). Similarly, the UPC group had a mean score of 4.4 (95% CI: 3.7-5.1) at baseline and 9.8 (95% CI:8.0-11.6) at 2-year follow-up, corresponding to a mean difference of 5.4 (95% CI:0.3.4-7.4). The MCID of OSS-pain score was estimated by ½*SD to 2. A comparison of the mean difference in OSS pain-score showed a significant reduction in level of pain across all implants and regardless of culture status. No significant differences in level of pain were found among the implant designs, neither at baseline nor at 2-year follow-up. Pain scores are shown in Figure 13 by a Forrest plot with stratification by implant design and culture status.

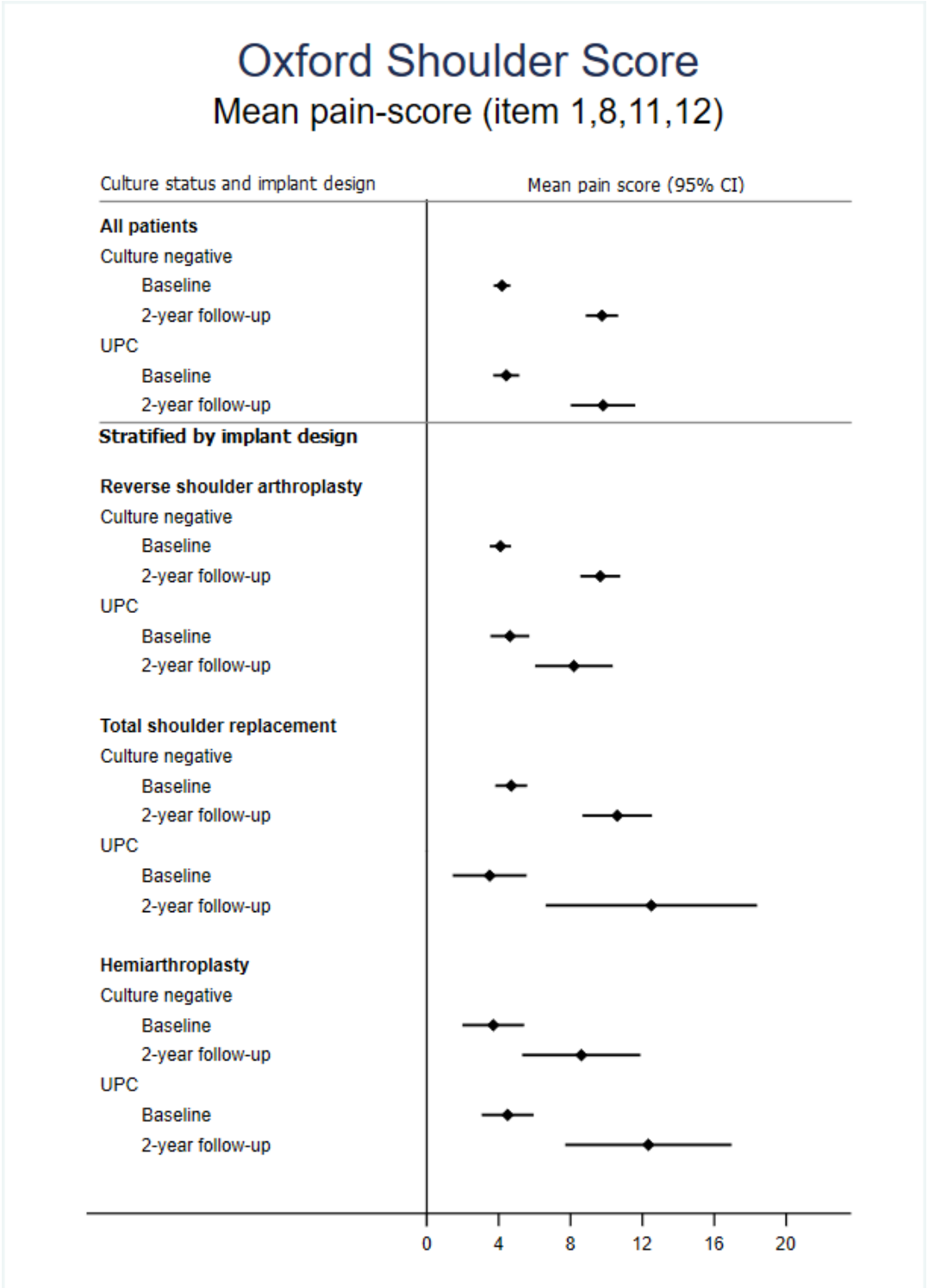
Table 9: Elevation and external rotation before and after revision

		Mean forward elevation				Mean external rotation			
		No. of Patients	Baseline	2-year follow-up	Mean difference	No. of Patients	Baseline	2-year follow-up	Mean difference
All patients			Mean (95% CI)	Mean (95% CI)	Mean (95% CI)		Mean (95% CI)	Mean (95% CI)	(95% CI)
	Negative cultures	85	78 (69-87)	123 (114-132)	45 (37-54) **	65	35 (29-40)	31 (27-45)	4 (-2-9)
	UPC	24	78 (60-95)	115 (94-137)	38 (20-56) **	15	40 (31-49)	39 (33-45)	1 (-10-12)
Patients by final joint implant									
Reverse	Negative cultures	55	66 (57-74)	120 (110-131)	55 (46-65) **	41	29 (23-36)	31 (23-39)	-1 (-7-5)
	UPC	14	56 (36-76)	101 (73-128)	45 (20-69) **	7	38 (28-47)	41 (21-60)	3 (-20-26)
TSR	Negative cultures	20	121 (102-140)	149 (133-164)	28 (8-48) **	19	37 (31-42)	44 (35-54)	10 (-3-23)
	UPC	4	125 (83-167)	158 (117-199)	33 (-3-68)	4	36 (14-58)	33 (6-59)	-5 (-29-20)
Hemi	Negative cultures	10	61 (32-89)	85 (60-110)	25 (0-49)	5	25 (13-37)[x]	26 (9-43)[x]	1 (-24-26)
	UPC	6	97 (64-130)	122 (63-181)	25 (-32-82)	4	43 (26-59)[x]	45 (45-45)[x]	2.5 (-14-19)

[x]=significant difference between culture negative and UPC group

**=significant difference between baseline and 2-year follow up

Figure 13: Forrest plot of mean pain score stratified by implant design and culture status. Higher score=less pain



8 Discussion

8.1 Summary of findings

Our studies are the first to investigate the diagnostic performance of nuclear medicine modalities to diagnose chronic PJI in patients with failed shoulder replacements. The performance of both modalities was low as our studies showed a sensitivity of 0.18, specificity of 1.00, PPV of 1.00 and NPV 0.67 for WBC/BM SPECT CT, and a sensitivity of 0.14, specificity of 0.91, PPV of 0.38 and NPV of 0.71 for FDG-PET. Thus, results are difficult to act upon in a clinical setting. Furthermore, our final study is also among the first to assess change in functional outcome after revision surgery in a cohort of patients with or without UPC. In our cohort, we found that 23% of the patients had UPC with C acnes accounting for 64% of the positive cultures. The study showed equal and clinically relevant gain in functional outcome and reduction of pain after surgery between patients with and without UPC. The presence of UPC did not impact functional outcome assessed by OSS since identical OSS at baseline and 2-year follow-up was found. Similar findings were shown for forward elevation and external rotation.

8.2 Reference test

To determine the diagnostic accuracy of our included modalities, we based our definition of PJI on the major criteria of the MSIS definition as no gold standard is available [68, 129, 130]. In the effort to diagnose PJI prior to surgery, we planned to performed ultrasound guided aspirations on all patients. However, we experienced a high rate of dry-taps; an observation in line with others [57]. Thus, we stopped performing the procedure routinely owing to limited diagnostic benefit at expense of the risk of causing an iatrogenic infection. Furthermore, perioperative histopathology of PJI cases was not routinely performed at the participating institutions. Owing to limited support from literature of shoulder PJI [131, 132] combined with the limited expertise at the centres, we did not incorporate histopathology in our evaluation. Consequently, we disregarded MSIS minor criteria, except from blood-markers. In addition, we raised the number of identified organisms needed to define infection from two (used in most composite PJI criteria) to three to reduce the risk of misinterpreting contamination as infection; this increase relied on an expectation of high incidence of C acnes in our cohort.

In 2018, the MSIS-shoulder definition was adopted in order to include 'grey zone' patients in the classification (please see section 4.3.4). By introducing categories of 'probable' and 'possible' PJI, we can highlight these patients in order to make further judgements in relation to further treatment. Applying this PJI definition on our cohort would cause 3 patients to change from not infected to infected in Study I (from 11 to 14 patients) and cause a change of 6 patients in Study II (from 22 to 28 infected patients). Despite an increased number of infected patients, our results would have remained almost unchanged. In Study I, the sensitivity of WBC/BM SPECT CT would increase from 0.18 to 0.25 and specificity decreased from 1.00 to 0.96. In Study II, the sensitivity of FDG-PET would decrease from 0.14 to zero and specificity from 0.91 to 1.00. The improved and more detailed definition of shoulder PJI is welcomed, however, the lack of appertaining treatment guidelines for grey zone patients must be addressed to ensure adequate treatment to minimise over- and undertreatment.

8.3 Origin of bacteria

Our studies exclusively deal with patients undergoing revision shoulder surgery. During the 3-year study period, we found a positive culture rate of 30% of all the revised patients. The high incidence is in line with other studies investigating shoulder surgery with revision of joint replacements or removal of hardware; mainly with isolation of C acnes [34, 36, 52, 109]. Most low-grade PJI are caused by bacteria part of the skin microbiome. Studies have shown that C acnes is especially abundant around the shoulder on males and can be isolated from the skin and dermis in 40% of patients undergoing shoulder surgery, despite surgical preparation with alcoholic chlorhexidine [133-136]. A recent study by Falconer et al. found that instruments and gloves posed the most likely route of transfer of C acnes into the surgical wound [137]. Thus, contamination during primary surgery and subsequent infection of the implant seem likely. This possible route of transfer implies that organisms could also be introduced during the revision procedure. The paradox that the procedure of obtaining biopsies itself might introduce microbes is yet to be solved. We tried to compensate using our stricter definition of infection.

8.4 Diagnostic accuracy of WBC/BM SPECT CT (Study I)

The diagnostic accuracy of WBC/BM imaging has a reported sensitivity between 0.5 and 1.0 and a specificity from 0.6 to 1.0 in patients undergoing revision of hip or knee replacements [76, 78, 79, 138-142]. The low sensitivity of 0.5 found by Pill el al was attributed to limited neutrophil recruitment in chronic PJI compared to acute infections [78] whereas the low sensitivity of 0.6 shown by Brammen et al was explained by inexperience in interpreting images [143]. Subsequently the authors demonstrated that the sensitivity could be raised to 1.00 when images were reviewed by a "leading nuclear medicine specialist within the field". We used SPECT CT to achieve the best spatial resolution [83, 142] and imposed that all scans were reviewed by a senior nuclear specialist, nevertheless, our sensitivity is inferior to any study. We attribute the main cause of discrepancy in results to differences in microbial distribution between shoulder and low-limp PJI [31, 144] and find the low sensitivity is unlikely to be explained by inexperience. The cellular response of chronic low-grade PJI is dominated by monocytes and studies have shown that C acnes can survive phagocytosis, which results in a limited activation of the immune system [145]. Furthermore, less leucocytes migrate to the infected area compared to infections caused by more virulent organisms e.g. Staph aureus [132, 146, 147]. Consequently, the labelled leucocytes injected might not aggregate around the joint replacement in quantities large enough to exhibit substantial increase in tracer activity.

Mainly two isotopes ^{111}In and ^{99m}Tc have been used in the previous studies to label the WBC. Each isotope has its advantages and disadvantages. ^{111}In has a longer half-life, resulting in more radiotracers available to detect at time of late imaging (20-24 hours after injection) but leading to a higher exposure of radiation in the patients. ^{99m}Tc on the other hand is cheaper to produce, more readily available and has more favourable physical characteristics, properties which may have lead the European Association of Nuclear Medicine (EANM) to recommend ^{99m}Tc as the first choice isotope for investigation of orthopaedic infections [125]. However, it should be noted, that ^{99m}Tc has not been proven superior to ^{111}In scans and the EANM-guideline also states that ^{111}In may be preferable in PJI with low-grade

infections [125]. Studies investigating different acquisition protocols found the use of delayed and late ^{99m}Tc WBC imaging superior to diagnose PJI. In equivocal cases addition of BM imaging proved most sensitive and specific [75, 76]. In 2018 an EANM-guideline was published recommending this protocol [125]. In contrast, we performed a dual isotope technique with late ^{111}In labelled WBC scan and ^{99m}Tc colloid labelled bone marrow imaging, hence WBC images and bone marrow scan were acquired simultaneously on all patients. This approach was chosen based on the expected overweight of chronic low-grade PJI and potential need for bone marrow imaging in many patients. Studies with only late WBC imaging have shown acceptable diagnostic accuracy [138, 139, 141] but it inherently induces a risk of misinterpreting inflammation as infection, since only presence of tracer activity can be assessed, and not if activity is increased/unchanged (indicative of PJI) or decreased (indicative of inflammation) from a 4-hour time-point. But, by definition, negative scans can only be the consequence of either unchanged or decreasing activity. Since none of the negative scans in our study showed any activity at the 24-hour timepoint, images were unlikely to be false positive. In addition, the two patients with positive scans were found clinically infected at revision and thus true positive. Had many false positive scans been encountered, a delayed scan would have proved beneficial to distinguish inflammation from infection. However, the opposite seemed to be the case in our study with 9 false negative results. Due to the lack of any activity on the negative scans an underestimation of the sensitivity of WBC/BM SPECT CT was unlikely. Furthermore, we believe the lack of observed tracer activity in our infected patients is a result of limited inflammatory response evoked by low-grade PJI rather than choice of scanning protocol.

8.5 Diagnostic accuracy of ^{18}F-FDG-PET/CT (Study II)

Our study is the first to investigate the diagnostic accuracy of FDG-PET to diagnose PJI of the shoulder. Recent meta-analyses on utility of FDG-PET to diagnose lower-limb PJI reported a specificity of 0.93 and sensitivity of 0.93 of hip PJI, and somewhat lower specificity of 0.84 and sensitivity 0.70 of knee PJI [77, 148]. Despite other single studies have reported lower sensitivities [90, 118, 141] only one study showed a sensitivity as low as ours. Mayer-Wagner et al. investigated a cohort of hip- and knee-replacements and found FDG-PET feasible for determining septic loosening of a hip replacement (sensitivity 0.75), but not a knee replacement (sensitivity was 0.14) [149]. The low sensitivity was attributed to small intramedullary areas of implant-bone contact due to very short stems in knee-implants compared to longer stemmed hip implants. This argument could potentially apply to our results, since 30% of our patients had a resurfacing HA, thus an implant without stem. However, we did not observe increased tracer-uptake at the implant-bone interface in these patients nor did we find a single loose implant at revision; hence, it is an unlikely explanation to our results. Other authors have concluded that it is difficult to differentiate between aseptic and septic loosening using FDG-PET because uptake along bone-stem interface is increased in both conditions [90, 116, 150]. Furthermore, it is shown that a diffuse uptake near the calcar or tip of a hip-replacement can be expected in well-functioning hip replacements but not in the soft-tissue surrounding the prosthesis [151]. We found a similar observation with high uptake at the neck of the prosthesis when we evaluated our scans. Other authors have recommended PJI to be characterised by diffuse uptake in the periprosthetic tissue with

or without uptake near the prosthesis [116, 141, 152]. However, the main finding in shoulder revisions is low-grade PJI without loosening of the implant. We experienced a loose humerus stem in 10% (9/86) of the patients; and 5 exhibited increased FDG-uptake along bone-stem interface. Infection was confirmed by cultures in 4 patients, however, further two patients exhibited strong clinical suspicion of infection during revision despite negative cultures. The difference in uptake patterns of loose implants in shoulders compared to hips may be due to different load-bearing characteristics and consequently reduced bone activity in shoulders owing to less mechanical stress.

Many shoulder replacements consist of both a glenoid- and a humerus component. Although component loosening may be suggestive of infection, the glenoid component is known to be prone for loosening due to mechanical stress and is as such not highly associated with infections [117]. In our cohort we found 10 patients with loosened glenoid components, of which 2 were infected. None of the 10 patients showed increased activity on the FDG-PET scan, hence, incorporating glenoid loosening in our criterion would not have changed the result.

Another way of diagnosing PJI by FDG-PET is to pursue a more quantitative definition of infection. Thus several authors have tried to correlate the measured intensity of uptake, e.g. such as SUV_{max} or a target-to-tissue uptake rate but they could not demonstrate a good correlation to PJI [91, 118]. We performed an analysis of SUV_{max} and found similar poor correlation between SUV_{max} and PJI.

Most recently, Verberne et al. stressed the importance of the applied criteria to diagnose PJI by comparing different patterns of uptake and found increased uptake at the stem-bone interface to be the most reliable criterion [77]. However, due to the different uptake patterns suggested for PJI and since no research on FDG-PET of shoulder replacements has been published, we defined our own Shoulder specific assessment for the diagnose of shoulder PJI to attain a uniform evaluation of the scans. However, only moderate interobserver agreement of diagnosis was obtained and with agreement according to the first sub-criterion was the lowest. This is not surprising since implant-near soft tissue activity varies tremendously, resulting in wider reader-interpretation [151]. The increased activity in absence of infection can be attributed to inflammation induced by foreign particles due to wear of the polyethylene component or synovitis from mechanical stress [91]. What was more surprising was the moderate agreement on assessment of lymph node activity - a relatively well-defined entity - and we have no obvious explanation for this. Overall, we found no increased tracer activity on the majority of scans from the infected patients. The low or lack of tracer uptake seen in infected patients is attributed to the limited inflammatory response low-grade PJI often elicit [132].

8.6 Functional outcome after revision surgery (Study III)

Few papers report outcome by PRO after revisions with UPC and only one study compares outcome between groups with and without UPC [41]. Hsu et al. used Simple Shoulder Test (SST) and found the score to increase from 3.2 to 7.8 in the UPC group and from 2.6 to 6.1 in the culture negative groups, no difference between groups were demonstrated. They concluded that patients experienced a clinically relevant gain in function and that UPC did not affect the functional outcome, a conclusion in line with

ours. The study by Grosso et al. investigated 17 patients with UPC and reported change in Penn Shoulder Score [42]. They found a change from 22 at baseline to 50 at 3-year follow-up, however, the postoperative score was based on only 70% of the patients, thus potential selection bias must be taken into consideration. Although correlation between OSS and SST or Penn score has not been investigated, a doubling in score in both studies from baseline to follow-up is fairly similar to our observations.

Three recent studies have described functional outcome after aseptic shoulder revisions by using PRO. First, a British study investigated patients converted from a failed resurfacing HA to either TSR or RSA and found an improvement in OSS from 13 to 39 at 3-years follow-up [103]. Second, a Dutch study also measured outcome by OSS after revision of failed HA, despite no change in OSS reported the postoperative OSS was 42 and 28 for TSR and RSA, respectively [10]. Third, Seth et al. evaluated 28 patients with glenoid arthrosis who were revised from HA to TSR. At 5-years follow-up the mean ASES score was 78 and mean pain (Visual analogue score) was 2.3 [104].

Further two recent studies investigated outcome of one-stage revisions by PRO. First, Stone et al. investigated revision of different implant designs and found an increase in ASES from 35 to 64 points [16]. Second, Sevelda et al. measured function by CMS and found an increased score of 38 (from 27 to 65) after one-stage PJI revisions (most patients with C acnes detected in one aspiration of joint fluid only) [20].

In comparison the DSR reported a median WOOS one-year postoperative stratified by indication for revision [3]. Patients revised due to glenoid erosion performed best with a WOOS of 65, followed by aseptic revision of other indications with a WOOS of 49 and revisions due to PJI had a WOOS of 35. Thus, the findings in all six studies correlate well with the notable increase in OSS shown in both groups in our study, despite differences between study populations. The somewhat lower outcome scores due to PJI reported by DSR may be explained by the registration process. Patients with UPC are likely to be categorised as aseptic revisions due the lack of suspicion of PJI at time of revision which often corresponds to the time of reporting to the register.

A recent multicentre study assessed OSS in a population with no history of shoulder pathology [153]. They found that patients before the age of 80 years had a mean OSS of approximately 46 regardless of sex and ethnicity; with a trend towards lower scores among patients over the age of 80. In our study most patients were under the age of 80 and here 14 patients reported an OSS of at least 46 at two-year follow-up; an OSS which clinically seems high for patients with a history of multiple shoulder operations.

Moreover, the increase in OSS shown in our study is surprisingly comparable to data from patients treated with elective primary shoulder replacement presented in the 13th annual NJR report [154]. The median preoperative OSS was 16 points with an increase of 18 points 6 months after surgery. Our baseline OSS is slightly higher and the change slightly smaller but taken into consideration that our patients had multiple prior surgeries our increase seems relatively high. This might represent some sort

of adaptation or acceptance of limitations in patients which must be clarified further to have a full understanding of the mechanisms behind.

The majority of our UPC patients (90%) were treated with antibiotics for six weeks. As such, the procedure performed could be categorised as one-stage revisions for PJI. A recent systematic review on results of revisions of shoulder PJI reported a postoperative mean ASES of 60 after one-stage revisions [7]. This score falls in line with ours. The two most predominant prescribed antibiotics were oral Penicillin and Amoxicillin due to susceptibility of C Acnes to these drugs. Whether the type of drug, duration of treatment or route of administration influences outcome cannot be answered by this study. Stratification to compare outcome between treatment groups would lead to very small groups with imprecise estimates rendering conclusions nugatory.

8.7 Methodological considerations

Study I and II were designed to investigate the diagnostic accuracy of two nuclear medicine modalities to identify PJI whereas Study III used an observational design to reflect the association between UPC and functional outcome. Before drawing any conclusions, methodological considerations must be taken into account to assess potential systematic or random errors implications on the studies' findings. Hence, internal validity of the studies is discussed in the following paragraphs.

Bias of diagnostic accuracy studies (Study I and II)

Results of diagnostic accuracy studies can be affected by specific types of bias which individually can impact accuracy estimates.

In our study most patients were diagnosed with chronic low-grade PJI, thus spectrum bias must be taken into account because it relates to the assumption that it is easier to detect PJI caused by more virulent organisms compared to low-grade PJI. Consequently, a better performance of modalities is expected in studies including a population with a higher incidence of aggressive PJI compared with ours. In most comparable literature a pre-scan selection due to signs of PJI led to the nuclear scan performed. Thus, these patients are expected to have more pronounced disease compared to our cohort [139, 155]. In essence, spectrum bias does not affect the internal validity of our studies but must be considered when comparing our results to other studies.

In Study I, two patients presented with intraarticular purulence at time of surgery. These two patients were classified as having PJI by the PET CT, but no growth was evident in any of the five cultures leading to a classification of negative PJI. This discrepancy has been discussed by others as some argues that purulence is not equal to infection within the joint[156]. But given the poor finding in our study, the magnitude of this potential misclassification of PJI has minimal impact on the presented result.

Not only is the reference standard susceptible for misclassification so is the reading of the index tests. Since pattern recognition is subjective and forms the basis of readings, it is important that the same patterns are utilized by multiple reviewers. In Study I, only one reviewer read the WBC/BM scans. But since the mismatch of WBC and BM activity classifying PJI is well-defined and relatively easy to judge,

we consider the risk of misclassification of the WBC/BM SPECT CT results to be minimal. However, in Study II, the more subtle changes on the FDG-PET scans may lead to misclassification. To minimise the risk of such bias, we trained the reviewers in our Shoulder specific assessment prior to reading. Despite training, the interrater agreement among the three reviewers was only fair to moderate (Fleiss' kappa 0.3-0.6) indicating a possible source of misclassification. Because we classified the result of FDG-PET on basis of the majority of findings the low interrater agreement is considered to have little impact on our result.

To avoid diagnostic review bias, the reviewers were blinded for clinical information in all cases, hence the threshold for positive findings on the index tests was unaffected by the patients' and surgeons' preoperative suspicion of infection.

Partial verification bias can be introduced if patients with positive scans were revised more frequently than patients with negative scans, leading to potential overestimation of sensitivity and underestimation of specificity. However, we do not believe that our results are affected by verification bias because patients undergoing revision surgery were revised irrespectively of the result of the WBC/BM SPECT CT or FDG-PET.

The group of referred patients that were treated non-operatively proved more difficult to define than expected. Not all patients agree to surgery, even if it was recommended. Many reasons for patients to opt out exist, ranging from "I-know-what-I-have-and-not-what-I-get" to inability to cope with a demanding rehabilitation after surgery. Similarly, surgeons may apply a "gut-feeling" on top of the objective assessment, leading to a somewhat more subjective reasoning for offering or denying revision-surgery than according to the in-/exclusion criteria. Unfortunately, such reasoning was not documented in our studies and it is indeed likely than two similar patients received different treatments. One can only speculate of the infection status of these patients, but due to the "stealth" nature of most low-grade PJI we do not expect the prevalence of PJI to be different in this group compared to the two study groups. The impact to our results would be a sample size reduction and more imprecise estimates.

Bias in follow-up studies (Study III)

The risk of selection problems was addressed by including relevant Danish patients referred with a failed shoulder replacement to participating centres. This was possible because all patients have unfettered access to healthcare, thus all patients with this condition are exclusively treated at public centres. In addition, we applied a-priori well-defined in- and exclusion criteria for patient inclusion. Thus, these factors reduce the probability of a systematic exclusion of patients. In the 3-year study period, we experienced a drop-out due to missing OSS baseline data in 24 patients. During the study period, we recorded that patients were more than willing to fill out the OSS, thus we do not contribute the missing data to unwillingness. On the other hand, we did not ask the patients to fill in the baseline OSS retrospectively due to the risk of introducing severe recall bias. A plausible explanation to missing baseline OSS is the fact that administration of questionnaires was handled by the surgeons whereas 2-year follow-up were handled by the project manager. In light of surgeons handled relatively few

administrations yearly, we contribute the missing data to oversight rather than systematic error. Looking deeper into this issue, we found an overweight of patients with missing data at the centre not hosting the project manager which highlights the need of thorough project management at centres without a dedicated person in charge.

No questionnaires have been designed to address functional outcome after shoulder revision arthroplasty or for that matter shoulder arthroplasty surgery. Despite the lack of validation in this setting, the OSS has been used for several years by the NJR after both primary and revision shoulder joint replacement. The registry has reported a pronounced ceiling effect at six months [157], which questions the responsiveness of OSS in this population. However, this phenomenon was not evident in our study at either baseline or 2-year follow-up, thus OSS was able to illuminate functional outcome in our population.

Incorrect rating of OSS by the patients is possible and may affect the accuracy of the data used. However, this misclassification is believed to be of non-differential nature because the OSS is based on the average perception of symptoms during the preceding four weeks prior to the time of completion. Thus, we expect that patients with or without UPC are equally able to rate symptoms as the real picture of perceived functional limitations at time of completion. In addition, data are collected prospectively at baseline prior to classification of UPC status and after 2 years when patients are believed to be fully recovered, thus patients were unaware/unaffected of culture status at time of completion.

Confounding is another limitation of the study with unmeasured confounding influencing the result. A potential confounder is previous ipsilateral shoulder operations including number and types of surgery. Previous operations may be associated with increased risk of contamination emerging as UPC and may lead to reduced function of the shoulder. Unfortunately, we have no valid data on previous shoulder operations due to imprecise wording in the case report form, thus the impact of this factor is unknown. In order to address effect modification, we performed stratified analyses by final implant design to accommodate the potential association to functional outcome. However, no trend was evident for neither OSS nor ROM leaving out effect modification by final prothesis types.

Evaluation of generalisability (Study I, II & III)

Patients participating in the three studies represent a highly selected cohort. Patients were referred from either general practitioner or other orthopaedic departments with surgeons almost exclusively dealing with shoulder surgery. However, in Denmark health care is tax-paid leading all patients to receive the same treatment, regardless of socio-economic status; at least in theory. Indeed, this is not the case in all countries which might affect the generalisability to the populations of individual surgeons or orthopaedic centres. Nevertheless, treating failed shoulders is worldwide considered a highly specialised function and it seems fair to assume that most surgeons performing revision shoulder surgery are affiliated to departments performing advanced orthopaedic surgery. Thus, our cohort will probably reflect the average patient in such a setting.

9 Conclusions

Study I

We found a poor diagnostic performance of ^{111} Indium-labelled white blood cells/^{99}Technesium labelled bone marrow SPECT CT in diagnosing chronic low-grade PJI of painful shoulder replacements. A negative result of the scan could not in rule out infection with acceptable certainty our study group, rendering results difficult to act upon in a clinical setting.

We cannot recommend this modality as a tool in the routine preoperative workup for diagnosing chronic PJI of the shoulder.

Study II

We found a poor diagnostic performance of ^{18}F-FDG PET/CT in diagnosing chronic low-grade PJI in failed shoulder replacements. The tracer-uptake in infected joint replacements was highly variable and despite a predefined criterion of pathological uptake patterns, only moderate agreement of diagnosis between multiple reviewers was obtained.

^{18}F-FDG-PET/CT cannot be recommended as a part of routine preoperative diagnostics to diagnose chronic low-grade PJI of painful shoulder replacements.

Study III

This study found 24% of presumably aseptic revision to produce UPC, and C acnes was the most frequent bacteria isolated. Despite the emergence of UPC, we found no pattern in PRO or change in ROM to suggest that UPC caused by low-virulent bacteria impacts the short-term result. Patients can expect an increase in overall perceived function, increased forward elevation to approximately shoulder height and unchanged external rotation. The level of pain will likely be significantly reduced; however, the majority of patients will still experience residual shoulder pain.

UPC does not seem to impact the sort-term outcome of a standard revision of a failed shoulder replacement.

10 Perspectives

Any existing PJI should preferably be diagnosed prior to revision surgery, however, even the definition of PJI poses a challenge. Infections after joint replacement surgery can be a source of frustration and fear to the patient. How is my prognosis now? Can the infection be eradicated? Do I need further surgery? Many questions will inadvertently surface in a patient's mind. Unfortunately, the treating physicians cannot always give a firm evidenced based answer.

This thesis investigated the diagnostic accuracy of two different nuclear medicine modalities, of which WBC/BM imaging was considered a gold standard in the setting of diagnosing PJI. Surprisingly, both WBC/BM and FDG-PET proved ineffective in diagnosing PJI. The data obtained during the study did also reveal that PJI diagnosed by UPC alone did not reduce the outcome of the revision compared to patients without UPC.

The thesis does not provide unambiguous answers, on the contrary, it contributes to unfold aspects of our treatment of shoulder PJI and spurn several questions which remain unanswered and should be of focus in future research.

First, the results could question if the large proportion of UPC reflects an infection in our traditional understanding. Extensive research has gone into a defining PJI, but so far, an exhaustive definition has not been presented. In addition, different microbiology seen in different joints has led to the development of a shoulder-specific MSIS definition which handles our incomplete knowledge, in particular of low-grade infections, by incorporating a large "grey zone" in the definition e.g. UPC patients. A group of patients which nobody really knows how to treat! Our results can justify future randomized studies of UPC comparing no antibiotic to antibiotic treatment. In this way we can truly evaluate the impact of UPC.

Second, a major challenge in this field of research is how to assess the presence of bacteria inside the joint, without the risk of contamination during the biopsy-process itself. Currently, this paradox has not been solved. However, despite the lack of diagnostic power in our study, PET could prove a viable solution. If a tracer specific to e.g. C acnes or biofilm could be developed, a possible "non-invasive" and very specific test could be the result.

Third, several countries have well-organised registries with high completeness. However, a way to identify the UPC revisions is needed since classification of a revision as infected/uninfected is usually done at time of surgery - when knowledge of culture result can be 14 days away. In Denmark, besides the Shoulder Arthroplasty Register, a register of all perioperative obtained biopsies exists (Danish Microbiology Database). A survival analysis of revisions with UPC could be performed by merging information from the two registries, hence a more precise estimate of implant survival could be produced compared to smaller cohort studies.

Last, the analysis of PRO data has underscored that most of the patients we treat with revision surgery get better. A valuable knowledge to the patients who entrust a surgeon with their well-being and

shoulder function. But what about all the patients we do not revise? Little is known about this group of patients. Is the condition stable, do they experience further deterioration of function or maybe even improvement? Since this thesis is based on a cohort of patients referred due to a failed shoulder replacement, we do have a considerable amount of both PRO and clinical data from patients treated conservatively. A future study could survey this group years after the initial referral and compare initial to current PRO scores providing knowledge never published.

11 References

1. Mehmood, A. and W. Everett, *Shoulder replacements: a review.* Br J Hosp Med (Lond), 2016. **77**(2): p. 84-9.
2. Padegimas, E.M., et al., *Periprosthetic shoulder infection in the United States: incidence and economic burden.* Journal of shoulder and elbow surgery, 2015. **24**(5): p. 741-746.
3. RKKP, D.o., *Dansk Skulderalloplastik Register (DSR) Årsrapport 2019.* 2019.
4. Singh, J.A., et al., *Periprosthetic infections after total shoulder arthroplasty: a 33-year perspective.* Journal of shoulder and elbow surgery, 2012. **21**(11): p. 1534-1541.
5. Favard, L., *Revision of total shoulder arthroplasty.* Orthopaedics & traumatology, surgery & research : OTSR, 2013. **99**(1 Suppl): p. S12-21.
6. Florschutz, A.V., P.D. Lane, and L.A. Crosby, *Infection after primary anatomic versus primary reverse total shoulder arthroplasty.* J Shoulder Elbow Surg, 2015. **24**(8): p. 1296-301.
7. Mercurio, M., et al., *Outcomes of revision surgery after periprosthetic shoulder infection: a systematic review.* J Shoulder Elbow Surg, 2019.
8. Nelson, G.N., D.E. Davis, and S. Namdari, *Outcomes in the treatment of periprosthetic joint infection after shoulder arthroplasty: a systematic review.* Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ...[et al.], 2016.
9. Dines, J.S., et al., *Outcomes analysis of revision total shoulder replacement.* J Bone Joint Surg Am, 2006. **88**(7): p. 1494-500.
10. Hartel, B.P., et al., *Difference in clinical outcome between total shoulder arthroplasty and reverse shoulder arthroplasty used in hemiarthroplasty revision surgery.* Int J Shoulder Surg, 2015. **9**(3): p. 69-73.
11. SST, *Specialevejledning for ortopædisk kirurgi*. 2015: 08.01.2015.
12. Assenmacher, A.T., et al., *Two-stage reimplantation for the treatment of deep infection after shoulder arthroplasty.* J Shoulder Elbow Surg, 2017. **26**(11): p. 1978-1983.
13. Leonard, H.A., et al., *Single- or two-stage revision for infected total hip arthroplasty? A systematic review of the literature.* Clinical orthopaedics and related research, 2014. **472**(3): p. 1036-1042.
14. Svensson, K., et al., *Similar Risk of Re-Revision in Patients after One- or Two-Stage Surgical Revision of Infected Total Hip Arthroplasty: An Analysis of Revisions in the Swedish Hip Arthroplasty Register 1979(-)2015.* J Clin Med, 2019. **8**(4).
15. Zahar, A., et al., *Ten-Year Results Following One-Stage Septic Hip Exchange in the Management of Periprosthetic Joint Infection.* J Arthroplasty, 2019.
16. Stone, G.P., et al., *Surgical management of periprosthetic shoulder infections.* J Shoulder Elbow Surg, 2017. **26**(7): p. 1222-1229.
17. Beekman, P.D., et al., *One-stage revision for patients with a chronically infected reverse total shoulder replacement.* The Journal of bone and joint surgery.British volume, 2010. **92**(6): p. 817-822.
18. Ince, A., et al., *One-stage exchange shoulder arthroplasty for peri-prosthetic infection.* The Journal of bone and joint surgery.British volume, 2005. **87**(6): p. 814-818.
19. Klatte, T.O., et al., *Single-stage revision for peri-prosthetic shoulder infection: outcomes and results.* The bone & joint journal, 2013. **95-B**(3): p. 391-395.
20. Sevelda, F. and B. Fink, *One-stage exchange of septic shoulder arthroplasty following a standardized treatment algorithm.* J Shoulder Elbow Surg, 2018.
21. Kunutsor, S.K., et al., *One- and two-stage surgical revision of infected shoulder prostheses following arthroplasty surgery: A systematic review and meta-analysis.* Sci Rep, 2019. **9**(1): p. 232.
22. Dauenheimer, M.
23. Boileau, P., et al., *Arthroplasty of the shoulder.* J Bone Joint Surg Br, 2006. **88**(5): p. 562-75.
24. Wagner, E.R., et al., *The impact of the reverse prosthesis on revision shoulder arthroplasty: analysis of a high-volume shoulder practice.* J Shoulder Elbow Surg, 2019. **28**(2): p. e49-e56.
25. Radnay, C.S., et al., *Total shoulder replacement compared with humeral head replacement for the treatment of primary glenohumeral osteoarthritis: a systematic review.* J Shoulder Elbow Surg, 2007. **16**(4): p. 396-402.

26. Zimmerli, W., A. Trampuz, and P.E. Ochsner, *Prosthetic-joint infections.* The New England journal of medicine, 2004. **351**(16): p. 1645-1654.
27. Bonnevialle, N., et al., *Periprosthetic shoulder infection: an overview.* EFORT Open Rev, 2017. **2**(4): p. 104-109.
28. Rangan, A., et al., *Investigation and Management of Periprosthetic Joint Infection in the Shoulder and Elbow: Evidence and consensus based guidelines of the British Elbow and Shoulder Society.* Shoulder Elbow, 2018. **10**(1 Suppl): p. S5-S19.
29. Trampuz, A. and W. Zimmerli, *Prosthetic joint infections: update in diagnosis and treatment.* Swiss Med Wkly, 2005. **135**(17-18): p. 243-51.
30. Osmon, D.R., et al., *Diagnosis and management of prosthetic joint infection: clinical practice guidelines by the Infectious Diseases Society of America.* Clin Infect Dis, 2013. **56**(1): p. e1-e25.
31. Mook, W.R., et al., *The Incidence of Propionibacterium acnes in Open Shoulder Surgery: A Controlled Diagnostic Study.* The Journal of bone and joint surgery.American volume, 2015. **97**(12): p. 957-963.
32. Pottinger, P., et al., *Prognostic factors for bacterial cultures positive for Propionibacterium acnes and other organisms in a large series of revision shoulder arthroplasties performed for stiffness, pain, or loosening.* J Bone Joint Surg Am, 2012. **94**(22): p. 2075-83.
33. Butler-Wu, S.M., et al., *Optimization of periprosthetic culture for diagnosis of Propionibacterium acnes prosthetic joint infection.* J Clin Microbiol, 2011. **49**(7): p. 2490-5.
34. Renz, N., et al., *Orthopedic implant-associated infections caused by Cutibacterium spp. - A remaining diagnostic challenge.* PLoS One, 2018. **13**(8): p. e0202639.
35. Ascione, T., et al., *General Assembly, Diagnosis, Pathogen Isolation - Culture Matters: Proceedings of International Consensus on Orthopedic Infections.* J Arthroplasty, 2019. **34**(2S): p. S197-S206.
36. Both, A., et al., *Growth of Cutibacterium acnes is common on osteosynthesis material of the shoulder in patients without signs of infection.* Acta Orthop, 2018. **89**(5): p. 580-584.
37. Levy, O., et al., *Propionibacterium acnes: an underestimated etiology in the pathogenesis of osteoarthritis?* Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ...[et al.], 2013. **22**(4): p. 505-511.
38. Patzer, T., et al., *Prevalence of Propionibacterium acnes in the glenohumeral compared with the subacromial space in primary shoulder arthroscopies.* J Shoulder Elbow Surg, 2018. **27**(5): p. 771-776.
39. Boot, W., et al., *Missed low-grade infection in suspected aseptic loosening has no consequences for the survival of total hip arthroplasty.* Acta orthopaedica, 2015. **86**(6): p. 678-683.
40. Padegimas, E.M., et al., *Future surgery after revision shoulder arthroplasty: the impact of unexpected positive cultures.* Journal of shoulder and elbow surgery, 2017. **26**(6): p. 975-981.
41. Hsu, J.E., et al., *Single-Stage Revision Is Effective for Failed Shoulder Arthroplasty with Positive Cultures for Propionibacterium.* The Journal of bone and joint surgery.American volume, 2016. **98**(24): p. 2047-2051.
42. Grosso, M.J., et al., *Reinfection rates after 1-stage revision shoulder arthroplasty for patients with unexpected positive intraoperative cultures.* Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ...[et al.], 2012. **21**(6): p. 754-758.
43. Brolin, T.J., et al., *Routine cultures for seemingly aseptic revision shoulder arthroplasty: are they necessary?* Journal of shoulder and elbow surgery, 2017. **26**(11): p. 2060-2066.
44. Moroder, P., A. Trampuz, and M. Scheibel, *Propionibacterium: We Found It, Now We Have to Deal with It: Commentary on an article by Jason E. Hsu, MD, et al.: "Single-Stage Revision Is Effective for Failed Shoulder Arthroplasty with Positive Cultures for Propionibacterium".* The Journal of bone and joint surgery.American volume, 2016. **98**(24): p. e112.
45. Achermann, Y., et al., *Propionibacterium acnes: from commensal to opportunistic biofilm-associated implant pathogen.* Clin Microbiol Rev, 2014. **27**(3): p. 419-40.
46. Piper, K.E., et al., *Microbiologic diagnosis of prosthetic shoulder infection by use of implant sonication.* J Clin Microbiol, 2009. **47**(6): p. 1878-84.
47. Trampuz, A., et al., *Sonication of removed hip and knee prostheses for diagnosis of infection.* N Engl J Med, 2007. **357**(7): p. 654-63.
48. Donlan, R.M., *New approaches for the characterization of prosthetic joint biofilms.* Clin Orthop Relat Res, 2005(437): p. 12-9.
49. Koseki, H., et al., *Early staphylococcal biofilm formation on solid orthopaedic implant materials: in vitro study.* PLoS One, 2014. **9**(10): p. e107588.

50. Updegrove, G.F., A.D. Armstrong, and H.M. Kim, *Preoperative and intraoperative infection workup in apparently aseptic revision shoulder arthroplasty.* J Shoulder Elbow Surg, 2015. **24**(3): p. 491-500.
51. Keener, J.D., P.N. Chalmers, and K. Yamaguchi, *The Humeral Implant in Shoulder Arthroplasty.* J Am Acad Orthop Surg, 2017. **25**(6): p. 427-438.
52. Kelly, J.D., 2nd and E.R. Hobgood, *Positive culture rate in revision shoulder arthroplasty.* Clin Orthop Relat Res, 2009. **467**(9): p. 2343-8.
53. Marcheggiani Muccioli, G.M., et al., *Diagnosis and Treatment of Infected Shoulder Arthroplasty: Current Concepts Review.* Joints, 2018. **6**(3): p. 173-176.
54. Fink, B. and F. Sevelda, *Periprosthetic Joint Infection of Shoulder Arthroplasties: Diagnostic and Treatment Options.* Biomed Res Int, 2017. **2017**: p. 4582756.
55. Higuera, C.A., *Serum Erythrocyte Sedimentation Rate and C-Reactive Protein Level Have Weaknesses but Are the Best Screening So Far: Commentary on an article by Michael M. Kheir, MD, et al.: "Routine Diagnostic Tests for Periprosthetic Joint Infection Demonstrate a High False-Negative Rate and Are Influenced by the Infecting Organism".* J Bone Joint Surg Am, 2018. **100**(23): p. e150.
56. Kheir, M.M., et al., *Routine Diagnostic Tests for Periprosthetic Joint Infection Demonstrate a High False-Negative Rate and Are Influenced by the Infecting Organism.* J Bone Joint Surg Am, 2018. **100**(23): p. 2057-2065.
57. Strahm, C., et al., *Accuracy of Synovial Leukocyte and Polymorphonuclear Cell Count in Patients with Shoulder Prosthetic Joint Infection.* J Bone Jt Infect, 2018. **3**(5): p. 245-248.
58. Garrigues, G.E., et al., *Proceedings from the 2018 International Consensus Meeting on Orthopedic Infections: evaluation of periprosthetic shoulder infection.* J Shoulder Elbow Surg, 2019. **28**(6S): p. S32-S66.
59. Wyatt, M.C., et al., *The Alpha-Defensin Immunoassay and Leukocyte Esterase Colorimetric Strip Test for the Diagnosis of Periprosthetic Infection: A Systematic Review and Meta-Analysis.* J Bone Joint Surg Am, 2016. **98**(12): p. 992-1000.
60. Unter Ecker, N., et al., *What Is the Diagnostic Accuracy of Alpha-Defensin and Leukocyte Esterase Test in Periprosthetic Shoulder Infection?* Clin Orthop Relat Res, 2019. **477**(7): p. 1712-1718.
61. Pearl, M.L., *CORR Insights(R): What Is the Diagnostic Accuracy of Alpha-Defensin and Leukocyte Esterase Test in Periprosthetic Shoulder Infection?* Clin Orthop Relat Res, 2019. **477**(7): p. 1719-1721.
62. Deirmengian, C., et al., *Diagnosing periprosthetic joint infection: has the era of the biomarker arrived?* Clin Orthop Relat Res, 2014. **472**(11): p. 3254-62.
63. Frangiamore, S.J., et al., *alpha-Defensin as a predictor of periprosthetic shoulder infection.* Journal of shoulder and elbow surgery, 2015. **24**(7): p. 1021-1027.
64. Kamme, C. and L. Lindberg, *Aerobic and anaerobic bacteria in deep infections after total hip arthroplasty: differential diagnosis between infectious and non-infectious loosening.* Clinical orthopaedics and related research, 1981. **(154)**(154): p. 201-207.
65. Atkins, B.L., et al., *Prospective evaluation of criteria for microbiological diagnosis of prosthetic-joint infection at revision arthroplasty. The OSIRIS Collaborative Study Group.* J Clin Microbiol, 1998. **36**(10): p. 2932-9.
66. Mikkelsen, D.B., et al., *Culture of multiple peroperative biopsies and diagnosis of infected knee arthroplasties.* APMIS, 2006. **114**(6): p. 449-52.
67. Tan, T.L., et al., *Culture-Negative Periprosthetic Joint Infection: An Update on What to Expect.* JB JS Open Access, 2018. **3**(3): p. e0060.
68. Grosso, M.J., et al., *Performance of implant sonication culture for the diagnosis of periprosthetic shoulder infection.* Journal of shoulder and elbow surgery, 2017.
69. Tarabichi, M., et al., *Diagnosis of Periprosthetic Joint Infection: The Potential of Next-Generation Sequencing.* The Journal of bone and joint surgery.American volume, 2018. **100**(2): p. 147-154.
70. Parvizi, J., et al., *New definition for periprosthetic joint infection: from the Workgroup of the Musculoskeletal Infection Society.* Clinical orthopaedics and related research, 2011. **469**(11): p. 2992-2994.
71. Parvizi, J., T. Gehrke, and I. International Consensus Group on Periprosthetic Joint, *Definition of periprosthetic joint infection.* J Arthroplasty, 2014. **29**(7): p. 1331.
72. Frangiamore, S.J., et al., *Synovial fluid interleukin-6 as a predictor of periprosthetic shoulder infection.* The Journal of bone and joint surgery.American volume, 2015. **97**(1): p. 63-70.

73. Garrigues, G.E., et al., *Proceedings from the 2018 International Consensus Meeting on Orthopedic Infections: the definition of periprosthetic shoulder infection.* J Shoulder Elbow Surg, 2019. **28**(6S): p. S8-S12.
74. Wong, J.C., et al., *Culture positivity in primary total shoulder arthroplasty.* J Shoulder Elbow Surg, 2018. **27**(8): p. 1422-1428.
75. Glaudemans, A.W., et al., *A large retrospective single-centre study to define the best image acquisition protocols and interpretation criteria for white blood cell scintigraphy with (9)(9)mTc-HMPAO-labelled leucocytes in musculoskeletal infections.* Eur J Nucl Med Mol Imaging, 2013. **40**(11): p. 1760-9.
76. Erba, P.A., et al., *Image acquisition and interpretation criteria for 99mTc-HMPAO-labelled white blood cell scintigraphy: results of a multicentre study.* Eur J Nucl Med Mol Imaging, 2014. **41**(4): p. 615-23.
77. Verberne, S.J., et al., *Fluorodeoxyglucose positron emission tomography imaging for diagnosing periprosthetic hip infection: the importance of diagnostic criteria.* Int Orthop, 2018. **42**(9): p. 2025-2034.
78. Pill, S.G., et al., *Comparison of fluorodeoxyglucose positron emission tomography and (111)indium-white blood cell imaging in the diagnosis of periprosthetic infection of the hip.* J Arthroplasty, 2006. **21**(6 Suppl 2): p. 91-7.
79. Basu, S., et al., *FDG PET for diagnosing infection in hip and knee prostheses: prospective study in 221 prostheses and subgroup comparison with combined (111)In-labeled leukocyte/(99m)Tc-sulfur colloid bone marrow imaging in 88 prostheses.* Clinical nuclear medicine, 2014. **39**(7): p. 609-615.
80. Diaz-Ledezma, C., et al., *Diagnosis of Periprosthetic Joint Infection: The Role of Nuclear Medicine May Be Overestimated.* The Journal of arthroplasty, 2015. **30**(6): p. 1044-1049.
81. Signore, A., et al., *Consensus document for the diagnosis of prosthetic joint infections: a joint paper by the EANM, EBJIS, and ESR (with ESCMID endorsement).* Eur J Nucl Med Mol Imaging, 2019.
82. Ziessman, H.A., MD; O'Malley, Janis P., MD; Thrall, James H., MD, *Nuclear Medicine: The Requisites*. 2014.
83. Filippi, L. and O. Schillaci, *Usefulness of hybrid SPECT/CT in 99mTc-HMPAO-labeled leukocyte scintigraphy for bone and joint infections.* Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2006. **47**(12): p. 1908-1913.
84. Glaudemans, A.W., et al., *Leukocyte and bacteria imaging in prosthetic joint infection.* European cells & materials, 2013. **25**: p. 61-77.
85. Cohen, J., *Infectious Diseases*. fourth ed. 2017: Elsevier Limited.
86. Love, C. and C.J. Palestro, *Radionuclide imaging of infection.* J Nucl Med Technol, 2004. **32**(2): p. 47-57; quiz 58-9.
87. Jamar, F., et al., *EANM/SNMMI guideline for 18F-FDG use in inflammation and infection.* J Nucl Med, 2013. **54**(4): p. 647-58.
88. Boellaard, R., et al., *FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0.* Eur J Nucl Med Mol Imaging, 2015. **42**(2): p. 328-54.
89. Boellaard, R., *Standards for PET image acquisition and quantitative data analysis.* J Nucl Med, 2009. **50 Suppl 1**: p. 11S-20S.
90. Delank, K.S., et al., *The implications of 18F-FDG PET for the diagnosis of endoprosthetic loosening and infection in hip and knee arthroplasty: results from a prospective, blinded study.* BMC Musculoskelet Disord, 2006. **7**: p. 20.
91. Reinartz, P., *FDG-PET in patients with painful hip and knee arthroplasty: technical breakthrough or just more of the same.* Q J Nucl Med Mol Imaging, 2009. **53**(1): p. 41-50.
92. Weldring, T. and S.M. Smith, *Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs).* Health Serv Insights, 2013. **6**: p. 61-8.
93. Wright, R.W. and K.M. Baumgarten, *Shoulder outcomes measures.* J Am Acad Orthop Surg, 2010. **18**(7): p. 436-44.
94. Schmidt, S., et al., *Evaluation of shoulder-specific patient-reported outcome measures: a systematic and standardized comparison of available evidence.* Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ...[et al.], 2014. **23**(3): p. 434-444.
95. Rasmussen, J.V., et al., *Translation and validation of the Western Ontario Osteoarthritis of the Shoulder (WOOS) index - the Danish version.* Patient Relat Outcome Meas, 2013. **4**: p. 49-54.
96. Frich, L.H., P.M. Noergaard, and S. Brorson, *Validation of the Danish version of Oxford Shoulder Score.* Dan Med Bull, 2011. **58**(11): p. A4335.

97. Christensen, K.B., et al., *Rasch validation of the Danish version of the shoulder pain and disability index (SPADI) in patients with rotator cuff-related disorders.* Qual Life Res, 2019. **28**(3): p. 795-800.
98. Eshoj, H., et al., *Translation, adaptation and measurement properties of an electronic version of the Danish Western Ontario Shoulder Instability Index (WOSI).* BMJ Open, 2017. **7**(7): p. e014053.
99. Moeller, A.D., et al., *The Danish version of the modified Constant-Murley shoulder score: reliability, agreement, and construct validity.* J Orthop Sports Phys Ther, 2014. **44**(5): p. 336-40.
100. Dawson, J., R. Fitzpatrick, and A. Carr, *Questionnaire on the perceptions of patients about shoulder surgery.* The Journal of bone and joint surgery.British volume, 1996. **78**(4): p. 593-600.
101. Lo, I.K., S. Griffin, and A. Kirkley, *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**(8): p. 771-8.
102. Oh, J.H., et al., *Comparative evaluation of the measurement properties of various shoulder outcome instruments.* Am J Sports Med, 2009. **37**(6): p. 1161-8.
103. Thangarajah, T., et al., *Revision shoulder arthroplasty for failed humeral head resurfacing hemiarthroplasty.* Shoulder Elbow, 2019. **11**(2 Suppl): p. 35-41.
104. Sheth, M.M., et al., *Revision of failed hemiarthroplasty for painful glenoid arthrosis to anatomic total shoulder arthroplasty.* J Shoulder Elbow Surg, 2018. **27**(10): p. 1884-1890.
105. Rasmussen, J.V., et al., *Outcome of Revision Shoulder Arthroplasty After Resurfacing Hemiarthroplasty in Patients with Glenohumeral Osteoarthritis.* J Bone Joint Surg Am, 2016. **98**(19): p. 1631-1637.
106. Ortmaier, R., et al., *Reverse shoulder arthroplasty in revision of failed shoulder arthroplasty-outcome and follow-up.* Int Orthop, 2013. **37**(1): p. 67-75.
107. Sassoon, A.A., et al., *Revision total shoulder arthroplasty for painful glenoid arthrosis after humeral head replacement: the nontraumatic shoulder.* J Shoulder Elbow Surg, 2012. **21**(11): p. 1484-91.
108. Sajadi, K.R., Y.W. Kwon, and J.D. Zuckerman, *Revision shoulder arthroplasty: an analysis of indications and outcomes.* J Shoulder Elbow Surg, 2010. **19**(2): p. 308-13.
109. McGoldrick, E., et al., *Substantial cultures of Propionibacterium can be found in apparently aseptic shoulders revised three years or more after the index arthroplasty.* J Shoulder Elbow Surg, 2015. **24**(1): p. 31-5.
110. Lucas, R.M., et al., *Loose glenoid components in revision shoulder arthroplasty: is there an association with positive cultures?* J Shoulder Elbow Surg, 2016. **25**(8): p. 1371-5.
111. Cheung, E.V., J.W. Sperling, and R.H. Cofield, *Revision shoulder arthroplasty for glenoid component loosening.* J Shoulder Elbow Surg, 2008. **17**(3): p. 371-5.
112. Foruria, A.M., et al., *Clinical meaning of unexpected positive cultures (UPC) in revision shoulder arthroplasty.* Journal of shoulder and elbow surgery, 2013. **22**(5): p. 620-627.
113. Piggott, D.A., et al., *Characteristics and Treatment Outcomes of Propionibacterium acnes Prosthetic Shoulder Infections in Adults.* Open Forum Infect Dis, 2016. **3**(1): p. ofv191.
114. Association, W.M., *Declaration of Helsinki*. 2018.
115. ethics, N.C.o.H.r. *Appendix 2: Guidelines on the use of ionising radiation in health research projects*. [Guideline] 2011; Available from: http://en.nvk.dk/how-to-notify/useful-resources/appendix-2-guidelines-on-the-use-of-ionising-radiation-in-health-research-projects.
116. Reinartz, P., et al., *Radionuclide imaging of the painful hip arthroplasty: positron-emission tomography versus triple-phase bone scanning.* J Bone Joint Surg Br, 2005. **87**(4): p. 465-70.
117. Hsu, J.E., et al., *What can be learned from an analysis of 215 glenoid component failures?* J Shoulder Elbow Surg, 2018. **27**(3): p. 478-486.
118. Stumpe, K.D., et al., *FDG PET for differentiation of infection and aseptic loosening in total hip replacements: comparison with conventional radiography and three-phase bone scintigraphy.* Radiology, 2004. **231**(2): p. 333-341.
119. Dawson, J., R. Fitzpatrick, and A. Carr, *Questionnaire on the perceptions of patients about shoulder surgery.* J Bone Joint Surg Br, 1996. **78**(4): p. 593-600.
120. Dawson, J., et al., *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. **7**: p. 195-203.

121. Dawson, J., et al., *The Oxford shoulder score revisited.* Arch Orthop Trauma Surg, 2009. **129**(1): p. 119-23.
122. van de Water, A.T., et al., *Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture.* Disabil Rehabil, 2014. **36**(13): p. 1072-9.
123. Christiansen, D.H., et al., *Responsiveness and Minimal Clinically Important Change: A Comparison Between 2 Shoulder Outcome Measures.* J Orthop Sports Phys Ther, 2015. **45**(8): p. 620-5.
124. Hapuarachchi, K.S. and P.C. Poon, *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**(2): p. 81-9.
125. Signore, A., et al., *Clinical indications, image acquisition and data interpretation for white blood cells and anti-granulocyte monoclonal antibody scintigraphy: an EANM procedural guideline.* Eur J Nucl Med Mol Imaging, 2018. **45**(10): p. 1816-1831.
126. Fleiss, J.L., *Measuring nominal scale agreement among many raters.* Psychological Bulletin, 1971. **76(5)**: p. 378-382.
127. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
128. Limited, O.U.I., *Scoring System for the Oxford Shoulder Score*. 2011.
129. Shohat, N., et al., *Hip and Knee Section, What is the Definition of a Periprosthetic Joint Infection (PJI) of the Knee and the Hip? Can the Same Criteria be Used for Both Joints?: Proceedings of International Consensus on Orthopedic Infections.* J Arthroplasty, 2019. **34**(2S): p. S325-S327.
130. Parvizi, J. and C.J. Della Valle, *AAOS Clinical Practice Guideline: diagnosis and treatment of periprosthetic joint infections of the hip and knee.* J Am Acad Orthop Surg, 2010. **18**(12): p. 771-2.
131. Topolski, M.S., et al., *Revision shoulder arthroplasty with positive intraoperative cultures: the value of preoperative studies and intraoperative histology.* Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ...[et al.], 2006. **15**(4): p. 402-406.
132. Grosso, M.J., et al., *Sensitivity of frozen section histology for identifying Propionibacterium acnes infections in revision shoulder arthroplasty.* The Journal of bone and joint surgery.American volume, 2014. **96**(6): p. 442-447.
133. Phadnis, J., et al., *Frequent isolation of Propionibacterium acnes from the shoulder dermis despite skin preparation and prophylactic antibiotics.* J Shoulder Elbow Surg, 2016. **25**(2): p. 304-10.
134. Hsu, J.E., et al., *Preoperative skin cultures are predictive of Propionibacterium load in deep cultures obtained at revision shoulder arthroplasty.* J Shoulder Elbow Surg, 2018.
135. Patel, A., et al., *Propionibacterium acnes colonization of the human shoulder.* Journal of shoulder and elbow surgery, 2009. **18**(6): p. 897-902.
136. Scheer, V.M., et al., *Topical benzoyl peroxide application on the shoulder reduces Propionibacterium acnes: a randomized study.* J Shoulder Elbow Surg, 2018. **27**(6): p. 957-961.
137. Falconer, T.M., et al., *Contamination of the Surgical Field with Propionibacterium acnes in Primary Shoulder Arthroplasty.* J Bone Joint Surg Am, 2016. **98**(20): p. 1722-1728.
138. Simonsen, L., et al., *White blood cell scintigraphy for differentiation of infection and aseptic loosening: a retrospective study of 76 painful hip prostheses.* Acta Orthop, 2007. **78**(5): p. 640-7.
139. Trevail, C., et al., *An evaluation of the role of nuclear medicine imaging in the diagnosis of periprosthetic infections of the hip.* Clin Radiol, 2016. **71**(3): p. 211-9.
140. Teiler, J., et al., *Is 99mTc-HMPAO-leukocyte imaging an accurate method in evaluating therapy result in prosthetic joint infection and diagnosing suspected chronic prosthetic joint infection?* Q J Nucl Med Mol Imaging, 2019.
141. Love, C., et al., *Diagnosing infection in the failed joint replacement: a comparison of coincidence detection 18F-FDG and 111In-labeled leukocyte/99mTc-sulfur colloid marrow imaging.* Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2004. **45**(11): p. 1864-1871.
142. Kim, H.O., et al., *Usefulness of adding SPECT/CT to 99mTc-hexamethylpropylene amine oxime (HMPAO)-labeled leukocyte imaging for diagnosing prosthetic joint infections.* J Comput Assist Tomogr, 2014. **38**(2): p. 313-9.

143. Brammen, L., et al., *Accuracy of radioactively labeled autologous leukocytes in patients with infected prosthetic joints.* Molecular imaging and radionuclide therapy, 2015. **24**(1): p. 39.
144. Flurin, L., K.E. Greenwood-Quaintance, and R. Patel, *Microbiology of polymicrobial prosthetic joint infection.* Diagn Microbiol Infect Dis, 2019. **94**(3): p. 255-259.
145. Webster, G.F., et al., *Susceptibility of Propionibacterium acnes to killing and degradation by human neutrophils and monocytes in vitro.* Infect Immun, 1985. **49**(1): p. 116-21.
146. Nilsdotter-Augustinsson, A., et al., *Inflammatory response in 85 patients with loosened hip prostheses: a prospective study comparing inflammatory markers in patients with aseptic and septic prosthetic loosening.* Acta Orthop, 2007. **78**(5): p. 629-39.
147. Pajarinen, J., et al., *Profile of toll-like receptor-positive cells in septic and aseptic loosening of total hip arthroplasty implants.* J Biomed Mater Res A, 2010. **94**(1): p. 84-92.
148. Verberne, S.J., et al., *What is the Accuracy of Nuclear Imaging in the Assessment of Periprosthetic Knee Infection? A Meta-analysis.* Clin Orthop Relat Res, 2017. **475**(5): p. 1395-1410.
149. Mayer-Wagner, S., et al., *Use of 18F-FDG-PET in the diagnosis of endoprosthetic loosening of knee and hip implants.* Arch Orthop Trauma Surg, 2010. **130**(10): p. 1231-8.
150. Aksoy, S.Y., et al., *FDG and FDG-labelled leucocyte PET/CT in the imaging of prosthetic joint infection.* European journal of nuclear medicine and molecular imaging, 2014. **41**(3): p. 556-564.
151. Gelderman, S.J., et al., *(18)F-FDG-PET uptake in non-infected total hip prostheses.* Acta Orthop, 2018. **89**(6): p. 634-639.
152. Zhuang, H., et al., *The promising role of 18F-FDG PET in detecting infected lower limb prosthesis implants.* J Nucl Med, 2001. **42**(1): p. 44-8.
153. McLean, J.M., et al., *An international, multicenter cohort study comparing 6 shoulder clinical scores in an asymptomatic population.* J Shoulder Elbow Surg, 2018. **27**(2): p. 306-314.
154. National Joint Registry for England, W., Northern Ireland and Isle of Man., *13th Annual Report.* 2016.
155. Blanc, P., et al., *The use of labelled leucocyte scintigraphy to evaluate chronic periprosthetic joint infections: a retrospective multicentre study on 168 patients.* Eur J Clin Microbiol Infect Dis, 2019.
156. Alijanipour, P., et al., *Intraoperative Purulence Is Not Reliable for Diagnosing Periprosthetic Joint Infection.* J Arthroplasty, 2015. **30**(8): p. 1403-6.
157. National Joint Registry for England, W., Northern Ireland and Isle of Man., *15th Annual Report. 2018.* 2018.

12 Appendix

Oxford Shoulder Score Problemer med din opererede skulder	□ Venstre □ Højre

Sæt et × ved hvert spørgsmål!

1. I løbet af de sidste 4 uger …

Hvordan vil du beskrive den værste smerte du har haft i din skulder?

Ingen	Mild	Moderat	Svær	Uudholdelig
□	□	□	□	□

2. I løbet af de sidste 4 uger …

Har du haft svært ved at tage tøj på, på grund af din skulder?

Intet besvær	Lidt besvær	Besvær	Meget besvær	Umuligt
□	□	□	□	□

3. I løbet af de sidste 4 uger …

Har du haft svært ved at komme ind og ud af en bil eller ved at bruge offentlig transport på grund af din skulder?

Intet besvær	Lidt besvær	Besvær	Meget besvær	Umuligt
□	□	□	□	□

4. I løbet af de sidste 4 uger …

Har du været i stand til at bruge kniv og gaffel - på samme tid?

Ja, let	Lidt besvær	Besvær	Meget Besvær	Umuligt
□	□	□	□	□

5. I løbet af de sidste 4 uger …

Kunne du selv klare de daglige indkøb?

Ja, let	Lidt besvær	Besvær	Meget Besvær	Umuligt
□	□	□	□	□

6. I løbet af de sidste 4 uger …

Kunne du bære en bakke med en tallerken med mad gennem et lokale?

Ja, let	Lidt besvær	Besvær	Meget Besvær	Umuligt
□	□	□	□	□

7. I løbet af de sidste 4 uger …

Kunne du børste/rede dit hår med den dårlige arm?

Ja, let	Lidt besvær	Besvær	Meget Besvær	Umuligt
□	□	□	□	□

8. I løbet af de sidste 4 uger …

Hvordan vil du beskrive den smerte, du normalt har haft i din skulder?

Ingen	Mild	Moderat	Svær	Uudholdelig
□	□	□	□	□

9. I løbet af de sidste 4 uger …

Kunne du hænge dit tøj op i en garderobe, med din dårlige arm?

Ja, let	Lidt besvær	Besvær	Meget Besvær	Umuligt
□	□	□	□	□

10. I løbet af de sidste 4 uger …

Har du været i stand til at vaske og tørre dig selv under begge arme?

Ja, let	Lidt besvær	Besvær	Meget Besvær	Umuligt
□	□	□	□	□

11. I løbet af de sidste 4 uger …

Hvor meget har smerten fra din skulder forstyrret dit normale arbejde (inkl. husligt arbejde)?

Slet ikke	En lille smule	Moderat	Meget	Totalt
□	□	□	□	□

12. I løbet af de sidste 4 uger …

Har du været besværet af smerter i din skulder i din seng om natten?

Ingen nætter	1 til 2 nætter	Nogle nætter	Fleste nætter	Hver nat
□	□	□	□	□

MSIS-shoulder definition

Definition of shoulder PJI adopted at the 2nd ICM	
Primary signs:	
Presence of a sinus tract from the skin surface to the prosthesis	
Gross intra-articular pus	
Two positive tissue cultures with phenotypically identical high-virulent organisms	
Secondary signs:	Weight
Unexpected Wound drainage	4
Single positive tissue culture (virulent organism)	3
Single positive tissue culture (low-virulence organism)	1
Second positive tissue culture (identical low-virulence organism)	3
Humoral loosening	3
Positive frozen section (5 PMN in at least 5 high power fields)	3
Positive preoperative aspirate culture (low or high-virulence organism)	3
Elevated synovial neutrophil percentage (>80%) *	2
Elevated synovial WBC (>3000 cell/µL)	2
Elevated ESR (>3mm/hr)	2
Elevated CRP (>10 mg/L)	2
Elevated synovial alpha-defensin	2
Cloudy fluid	2
INTERPRETATION:	
Meeting one of the primary signs is diagnostic of definite shoulder PJI. If none of the primary signs are met, an evaluation of score is appropriate to obtain diagnose.	
≥6 points with identified organism	Probable PJI
≥6 *without* identified organism	Possible PJI
<6 points:	
- Single positive culture with virulent organism	Possible PJI
- Two positive cultures with low-virulent organism	Possible PJI
- Negative cultures or only single positive culture for low-virulent organism	PJI unlikely

13 Papers

Labeled white blood cell/bone marrow single-photon emission computed tomography with computed tomography fails in diagnosing chronic periprosthetic shoulder joint infection	PAPER I
^{18}F FDG-PET/CT has poor diagnostic accuracy in diagnosing shoulder PJI	PAPER II
Unexpected positive cultures after revision shoulder arthroplasty -does it impact outcome	PAPER III

PAPER I

Labeled white blood cell/bone marrow single-photon emission computed tomography with computed tomography fails in diagnosing chronic periprosthetic shoulder joint infection

J Shoulder Elbow Surg (2019) 28, 1040–1048

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

Labeled white blood cell/bone marrow single-photon emission computed tomography with computed tomography fails in diagnosing chronic periprosthetic shoulder joint infection

Thomas Falstie-Jensen, MD[a,*], Henrik Daugaard, MD, PhD[b], Kjeld Søballe, MD, DMSc[a], Janne Ovesen, MD, PhD[a], Anne Kirstine Arveschoug, MD[c], Jeppe Lange, MD, PhD[d] the ROSA study group[†]

[a]*Department of Orthopedics, Aarhus University Hospital, Aarhus, Denmark*
[b]*Department of Orthopedics, Herlev Hospital, Herlev, Denmark*
[c]*Department of Nuclear Medicine, Aarhus University Hospital, Aarhus, Denmark*
[d]*Department of Orthopedics, Horsens Regional Hospital, Horsens, Denmark*

Background: Shoulder periprosthetic joint infections (PJI) caused by low-virulent bacteria pose a diagnostic challenge. Combined labeled leukocyte (WBC) and technetium 99m sulfur colloid bone marrow imaging (WBC/BM) is considered the radionuclide imaging gold standard for diagnosing lower limb PJI. However, it is laborious and expensive to perform, and documentation on shoulder arthroplasties is lacking. This study investigated WBC/BM single-photon emission computed tomography-computed tomography diagnostic performance in shoulder PJI.

Method: All patients with a failed arthroplasty referred to a highly specialized shoulder department were scheduled for a diagnostic program including a WBC/BM. If an arthroplasty was revised, biopsy specimens were obtained and cultured for 14 days. The diagnostic performance of WBC/BM imaging was determined using biopsy specimens as a reference.

Results: Of the 49 patients who underwent a WBC/BM scan, 29 (59%) were revised. Infection was present in 11 patients, in whom 2 WBC/BM scans were true positive. The WBC/BM scan in 9 patients was false negative. The remaining 18 patients all had a true negative WBC/BM scan. WBC/BM showed a sensitivity 0.18 (95% confidence interval [CI], 0.00-0.41) and specificity 1.00 (95% CI, 1.00-1.00) in detecting shoulder PJI. The positive predictive value was 1.00 (95% CI, 1.00-1.00), and negative predictive value

[†]ROSA study group investigators: B.S. Olsen, MD, PhD (Department of Orthopedics, Herlev Hospital); A.K. Sørensen, MD, PhD (Department of Orthopedics, Herlev Hospital); L.C. Gormsen, MD, PhD (Department of Nuclear Medicine, Aarhus University Hospital); B. Zerahn, MD, PhD (Department of Clinical Physiology, Herlev Hospital); H.V.S. Johanssen, MD (Department of Orthopedics, Aarhus University Hospital); B. Elmengaard, MD, PhD (Department of Orthopedics, Aarhus University Hospital); T.M. Thillemann, MD, PhD (Department of Orthopedics, Aarhus University Hospital); L. Bolvig, MD (Department of Radiology, Aarhus University Hospital).

The Local Ethical Committee of Central Region Denmark approved this study (Ethical Board approval: ref. no. 217/2013).

*Reprint requests: Thomas Falstie-Jensen, MD, Department of Orthopedics, Shoulder and Elbow Section, Aarhus University Hospital, Palle Juul Jensens Blvd 99, 8200 Aarhus N, Denmark.

E-mail address: thomfals@rm.dk (T. Falstie-Jensen).

1058-2746/$ - see front matter
https://doi.org/10.1016/j.jse.2018.10.024

was 0.67 (95% CI, 0.49-0.84). No patients infected with *Cutibacterium* (formerly *Propionibacterium*) *acnes* resulted in a positive WBC/BM, nor had they preoperative or perioperative signs of infection.
Conclusion: A positive WBC/BM was found only in patients with obvious PJI. Hence, the scan added nothing to the preoperative diagnosis. The WBC/BM single-photon emission computed tomography–computed tomography scan cannot be recommended as a screening procedure when evaluating failed shoulder arthroplasties for possible infection.
Level of evidence: Level III; Diagnostic Study

Keywords: Revision; shoulder arthroplasty; periprosthetic infection; nuclear medicine; cultures; revision shoulder arthroplasty; hybrid imaging

When patients present with a failed shoulder arthroplasty, surgeons are often put in a difficult position. Are symptoms caused by infection or is it aseptic failure? Fulminant infections caused by high-virulent bacteria presenting with swelling, erythema, fever, or a draining sinus will often reveal the diagnosis. Unfortunately, symptoms of chronic low-grade infection of a shoulder arthroplasty caused by low-virulent bacteria, such as *Cutibacterium* (formerly *Propionibacterium*) *acnes*, can be equally subtle. Patients often present with pain and stiffness as the only complaints. Such symptoms mimic aseptic failure and rarely are any obvious symptoms of infection observed.[2,13] In such situations, the diagnosis of infection can be extremely difficult to obtain preoperatively.

Accurately distinguishing between the different causes of failure preoperatively is important because the choice of treatment and prognosis after revision differ considerably.[10]

No exact diagnostic approach has yet been established despite extensive research in biomarkers, radionuclide imaging, and the use of consensus infectious criteria such as developed by the Musculoskeletal Infection Society (MSIS) or the American Academy of Orthopaedic Surgeons (AAOS).[5,8,19]

Combined labeled leukocyte (WBC) and technetium 99m sulfur colloid bone marrow (WBC/BM) imaging is currently regarded as the gold standard of radionuclide imaging for detecting periprosthetic joint infections (PJI).[6] Most studies investigating WBC/BM imaging have been performed using conventional scintigraphy (planar imaging), which detects tracer activity in 2-dimensions. Technical innovation has made single-photon emission computed tomography (SPECT) possible. A SPECT scan is a scintigraphic technique in which tracer activity is detected in 3-dimensions, making the spatial resolution superior to the planar detection. Furthermore, when SPECT is coupled to a low-dose computed tomography (CT) scan (SPECT CT), the tracer activity can be located both spatially and related to specific anatomic structures.

This enhanced spatial and anatomic resolution makes SPECT CT the current technique of choice in diagnosing infections of the skeletal system. Based on lower extremity PJI, several studies have reported sensitivity and specificity above 90%.[12,14] Despite the promising results, WBC/BM has drawbacks that limit its usefulness. It involves handling of blood products, labeling of blood cells requires 2 work days to complete, exposes the patient to a high level of radiation, and is expensive.

If the high diagnostic precision reported in lower limb PJI can be achieved in shoulder PJI, it could, despite the limitations of the WBC/BM, be a valuable help to shoulder surgeons. No formal study using WBC/BM on failed shoulder arthroplasties has yet been published.

The objective of this study was to establish the diagnostic performance of WBC/BM SPECT CT when used to diagnose chronic infection in a consecutive cohort referred with a failed shoulder arthroplasty. We hypothesized that WBC/BM SPECT CT would have similar specificity and sensitivity in detecting shoulder PJI as seen in lower limb PJI.

Materials and methods

This prospective cohort study followed the Strengthening the Reporting of Observational Studies in Epidemiology guideline in designing and reporting of observational studies.[21]

In Denmark, all citizens (5.7 million) have free access to health care because of tax funding; thus, only public hospitals treat patients with a failed shoulder arthroplasty. Two departments are appointed by the Danish National Board of Health as the only units to perform revision shoulder arthroplasty surgery in Denmark. Approximately 140 patients are annually referred to the 2 centers with this condition. We included all patients regularly referred with a failed shoulder arthroplasty for evaluation of a possible revision arthroplasty. A failed shoulder arthroplasty was defined as patient-reported unsatisfactory result of any cause after shoulder replacement causing the patient to consult a physician and resulting in a referral to the highly specialized The plan was to include 50 revised patients in the study starting April 1, 2014.

A standardized preoperative workup consisting of 3 visits was set up to identify potential infected arthroplasties in the referred patients. At the first visit, a number of diagnostic tests were scheduled, including blood tests (erythrocyte sedimentation rate, WBC count, and C-reactive protein), ultrasonography with aspiration, x-ray imaging, and a thorough clinical assessment conducted by the surgeon.

At the second visit, a WBC/BM SPECT CT was performed. If the patient was on antibiotic treatment, it was paused, and the scan scheduled after at least 14 days to reduce the risk of false negative scans due to suppression of bacterial activity. At the last visit, the patient was informed of the results of the diagnostic tests, upon which the surgeon recommended revision surgery or nonsurgical treatment. Revision surgery was recommended if our modified MSIS criteria indicated infection (Fig. 1), or if signs of component loosening, obvious mechanical failure (eg, cranial migration), glenoid

attrition, or tubercular resorption were observed. The surgeon made an individual recommendation in cases with severe pain or significant functional limitation without the aforementioned findings or if the patient suffered from severe comorbidities.

The decision of surgery was ultimately left to the patient, even when the surgeon found surgery was indicated. Only when the surgeon suspected deep infection were the patients strongly advised to undergo surgery. Some patients choose not to undergo surgery after receiving information of pros, cons, and expected results after revision surgery. If the latter was the case, and the patient decided to live with the current limitations, no further actions were taken and the patient was omitted from the study.

Patients planned for surgical revision were paused for any antibiotic treatment for at least 14 days before surgery, and perioperative antibiotics were withheld until all biopsy specimens were obtained. During the revision procedure, biopsy specimens were obtained according to the definition of Kamme-Linberg.[11] Five individual samples were acquired with clean instruments from areas showing sign of infection and from the prosthesis surface. If no infection or suspicious signs were observed, samples were obtained from different intra-articular spots. The samples were cultured 4 days for aerobic bacteria and 14 days for anaerobic bacteria to detect bacteria such as *C acnes*.

After 1 year, the results were reviewed. Studies have reported that 1% to 4% of primary shoulder arthroplasties are infected.[17,18] Because the referred population exclusively consisted of patients with a failed arthroplasty, the prevalence was suspected to be even higher. With an infection rate of 38% (11 of 29), our data confirm this. The interim results led to a running discussion of the justification to continue the study among the nuclear medicine department and the orthopedic surgeons, especially owing to the substantial patient and economic resources needed to perform the scans as a screening procedure. The conclusion was that even in this high-prevalence cohort, the WBC/BM SPECT CT was not cost-effective. As a consequence, enrollment stopped on May 15, 2015, before the planned number of patients was reached.

Definition of infection

Our primary definition of infection was set to be at least 3 positive tissue cultures with the same bacteria or if a sinus tract communicating to the prosthesis was present. This definition was based on publications from Kamme-Lindberg[11] and the Oxford group.[1] This more rigorous definition than is typically used in published reports was chosen to reduce the risk of erroneously diagnosing skin contamination from especially *C acnes* as PJI.

To evaluate whether increased diagnostic power could be obtained by adding advanced radionuclide imaging and shoulder-specific signs of infection (stiffness and unexplained pain) as minor signs of infection (Fig. 1), we modified the existing MSIS definition of PJI. In this modified MSIS definition, the primary signs remain the same as in the original MSIS definition, but the number of secondary signs needed to diagnose infection were reduced to 2, compared with 4 in the original MSIS definition. The latter was done to reduce the risk of underdiagnosing low-virulent PJI. Because immediate histology of perioperative specimens was presently not available to us, we choose to omit such findings as secondary signs of infection.

Labeling and imaging protocol of WBC/BM SPECT CT

^{111}Indium-labelled WBC imaging is a 2-day procedure. On the first day, 50 mL whole blood is extracted from the patient, and leucocytes

MSIS 2013 criteria	Modified MSIS criteria (modifications marked with underscore)
PJI Positive if at least one primary sign exists OR at least three secondary signs exist.	**PJI Positive if at least one primary sign exists OR at least <u>two</u> secondary signs exist.**
Primary signs 1)There is a sinus tract communicating with the prosthesis; OR 2)A pathogen is isolated by culture from at least two separate tissue or fluid samples obtained from the affected prosthetic joint; OR	**Primary signs** 1)There is a sinus tract communicating with the prosthesis; OR 2)A pathogen is isolated by culture from at least three separate tissue or fluid samples obtained from the affected prosthetic joint; OR
Secondary signs a) Elevated serum C-reactive protein (CRP) AND erythrocyte sedimentation rate (ESR) b) Elevated synovial fluid white blood cell (WBC) count OR ++change on leukocyte esterase test strip c) Elevated synovial fluid polymorphonuclear neutrophil percentage (PMN%) d) Positive histological analysis of periprosthetic tissue e) A single positive culture	**Secondary signs** a) <u>Elevated serum C-reactive protein (CRP) OR erythrocyte sedimentation rate (ESR) OR WBC</u> b) A single positive culture c) <u>Positive radionucleide imaging</u> d) <u>Unexplained nightly pain OR excessive joint stiffness</u>

Figure 1 Musculoskeletal Infection Society (*MSIS*) periprosthetic joint infection (*PJI*) definition and modified MSIS PJI definition.

are labeled with 30 MBq ^{111}indium-tropolone and then reinjected. Images are acquired approximately 24 hours later.

At the second day, 400 MBq freshly prepared ^{99m}Tc-sulfur colloid is injected intravenously to depict the bone marrow. Patients are scanned within 1 hour after the injection, using a gamma-camera with simultaneous dual-isotope imaging. A Symbia (T16) scanner (Siemens Healthcare Diagnostics, Tarrytown, NY, USA) equipped with medium-energy collimators were used with a 15% window centered on 140 keV, a 15% window centered on 174 keV, and a 15% window centered on 247 keV in a 128×128 matrix. Planar imaging of thorax and shoulders in anterior and posterior projection in a 128×128 matrix with 20 minutes per view were performed. To locate pathologic foci in 3D, a SPECT with low-dose CT (50 mAs) was performed in a 64×64 matrix with 32 views of 40 seconds per view. The same collimator and energy windows were used performing both planar and tomographic imaging.

Image analysis

A senior nuclear radiologist reviewed all scans. A positive scan was defined as any mismatch activity between the BM scan and WBC scan in close relation to the arthroplasty, regardless of intensity or location, providing a positive or negative answer. Examples of positive WBC/BM SPECT CT are shown in Fig. 2.

Microbiologic culture protocol

Specimens were cultured on 5% blood agar plates (SSI, Copenhagen, Denmark), anaerobic agar plates (SSI), BLL chromagar orientation medium (Becton Dickinson, Heidelberg, Germany), and inoculated in semisolid agar + pepsis blood + thioglycollate (SSI) and serum broth (SSI).

Agar plates were inspected for growth up to 4 days. Semisolid agar and serum broth was visually inspected for signs of growth on day 4 and day 14. Matrix-assisted laser desorption/ionization-time of flight mass spectrometry (Bruker, Bremen, Germany) was used for identification of strains.

Study population

During the 1-year inclusion period, 71 patients were referred to the center with a failed arthroplasty, all of whom completed the first visit. WBC/BM SPECT CT was performed on 49 patients. A detailed description of patient enrollment is shown in the flowchart (Fig. 3). Nine patients were at the surgeon's discretion assessed not to be candidates for surgery. This assessment was based on comorbidities, expected relief of symptoms, and patient's acceptance of the course of rehabilitation after surgery. In all cases, the surgeon had no suspicion of infection.

Statistical analysis

The sample size for a reasonable estimation of sensitivity and specificity was calculated using binominal proportions. The microbiologically confirmed infection rate in the revised population was estimated to be 25% and hypothesizing a WBC/BM SPECT CT sensitivity and specificity of 0.9. Based on these assumptions and including 50 patients, we expected an estimation of sensitivity ± 0.19 and specificity ± 0.09, which we considered acceptable.

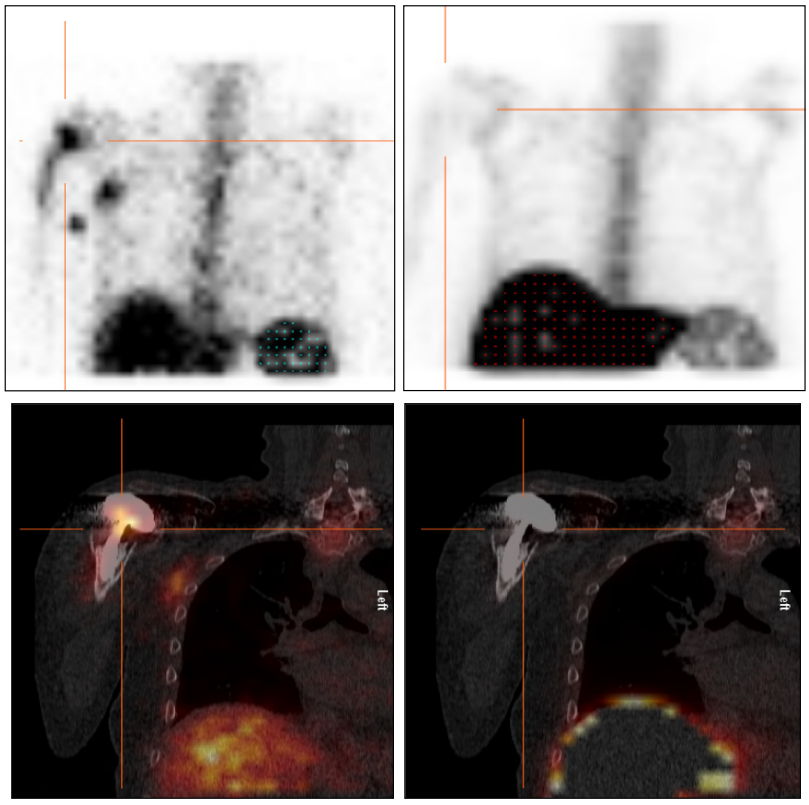


Figure 2 Example of a positive WBC/BM SPECT CT scan.

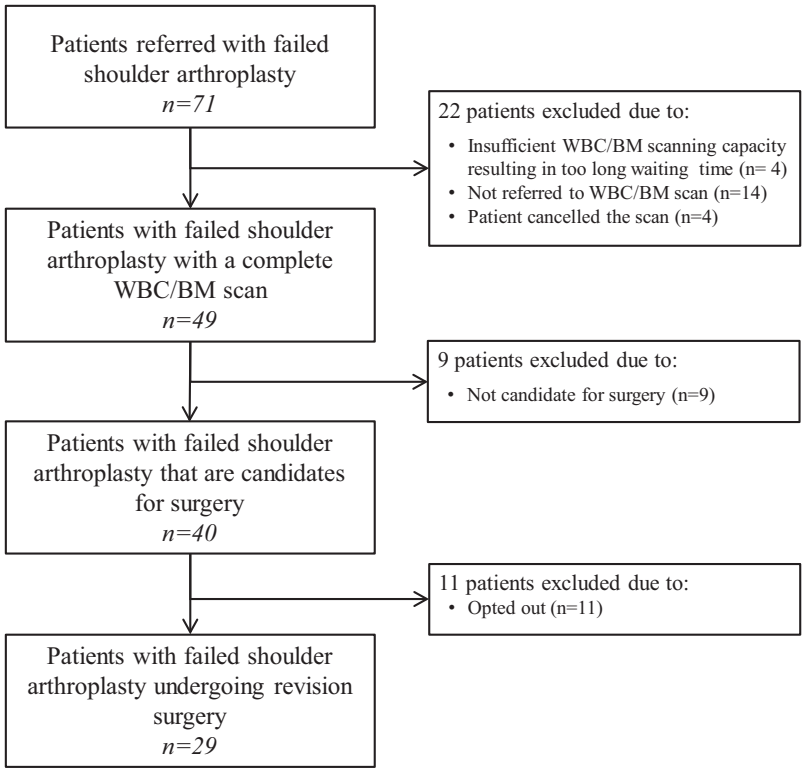


Figure 3 Flowchart of patient inclusion. *WCB/BM*, labeled leukocyte and technetium 99m sulfur colloid bone marrow.

Sensitivity, specificity, accuracy, and positive and negative predictive value (PPV and NPV, respectively) of WBC/BM SPECT CT were calculated using the microbial diagnosis as the true infection status. The same values were calculated and compared for the MSIS and modified MSIS definitions. All results are reported as fractions with 95% confidence intervals (CI). Data were analyzed using STATA 15 software (StataCorp, College Station, TX, USA).

Results

The demographics of the referred cohort and study population are described in Table I. No statistical differences between the dropout population and study cohort were found when testing for sex, mean age, or prosthesis age. No patient was being treated with antibiotics at the time of referral. A summary of the preoperative test results can be found in Table II.

Of the 49 performed WBC/BM SPECT CT studies, 2 (4%) were positive, and 47 (96%) were negative. One positive scan showed increased activity along the stem and in several lymph nodes in the axilla. The other positive scan showed intense increased activity in the shoulder joint and around the proximal part of humerus. The 2 positive WBC/BM SPECT CT studies were also culture positive.

Activity was recorded in 3 patients with negative scans, but not in areas related to the shoulder prosthesis. None of the 3 underwent operations.

Table I Demographics

Variable	Referred (n = 71)	Dropped out (n = 22)	WBC/BM SPECT CT (n = 49)	Study cohort (n = 29)
Age, mean (range), yr	65 (29-88)	67 (54-88)*	64 (29-88)	63 (29-84)
Sex, No.				
Female	33	8*	25	10
Male	38	14*	24	19
Prosthesis age, median (range), yr	3 (0-29)	5.5 (0-13)*	4 (0-13)	5 (0-13)
Revision rate		0.68 (15/22)*	0.59 (29/49)	1.00 (29/29)
Infection rate		0.2 (3/15)*	0.22 (11/49)	0.38 (11/29)

WBC/BM, labeled leukocyte and technetium 99m sulfur colloid bone marrow; *SPECT*, single-photon emission computed tomography; *CT*, computed tomography. Data are presented as median (range) or as number of patients.

* No significant difference between dropout and study cohort ($P > .05$).

Table II Results of preoperative diagnostic tests

Variable	Referred (n = 71)	Dropped out (n = 22)	Opted out and excluded (n = 20)	Study cohort (n = 29)
Indication for primary arthroplasty				
Arthrosis	25	8	8	13
Fracture/fracture sequelae	33	8	8	14
Rheumatoid arthritis	4	1	2	1
Other	9	5	2	1
Arthroplasty type				
Hemi	45 (cemented 20)	13 (cemented 8)	12 (cemented 8)	15 (cemented 13)
Total shoulder replacement	8 (cemented 3)	4 (cemented 3)	4 (cemented 3)	3 (cemented 1)
Resurfacing (hemi)	14 (uncemented)	3 (uncemented)	3(uncemented)	10 (uncemented)
Reverse	4 (cemented 4)	2 (cemented 2)	1 (cemented 1)	1 (cemented 1)
Radiologic findings				
Normal	10	2	4	4
Surgical error*	12	3	3	6
Loosening of components or osteolysis	10	4	1	5
Other†	39	13	12	14
Serologic findings				
≥1 elevated infectious marker§	21	8	8	5

* Surgical error consisted of malpositioning of implant, oversized implant.
† Other consisted of superior migration, resorption of tuberculi, glenoid attrition, dislocation, bent humoral implant.
§ C-reactive protein, white blood cell count, erythrocyte sedimentation rate.

Biopsy samples were obtained in all 29 surgically treated patients according to the protocol. In 10 patients (34%), infection was diagnosed with least 3 positive cultures. One patient had a fistula, and 2 of 5 cultures were positive with coagulase-negative staphylococci but a negative WBC/BM. This patient was classified as infected because of the fistula. Infection was ruled out in 18 patients (66%; Table III). Of the noninfected patients, all-negative cultures were found in 12 patients, and a growth pattern with less than 3 positive cultures with the same microbe or mixed flora on several plates was judged to be contamination was found in 6 patients.

The shortest interval between the latest operation to the arthroplasty and our revision was 5 months; thus, all infection can be categorized as chronic infection. The predominant organism isolated was *C acnes* found in 6 patients, followed by coagulase-negative staphylococci in 4 patients and *Peptostreptococcus* found in 1 patient. An overview of the culture results can be found in Supplementary Table S1.

Table III Summarized results of WBC/BM scan and infections status

Variable	Infected	Not infected	Total
WBC/BM SPECT CT			
Negative	9	18	27
Positive	2	0	2
Total	11	18	29

WBC/BM, labeled leukocyte and technetium 99m sulfur colloid bone marrow; *SPECT*, single-photon emission computed tomography; *CT*, computed tomography.

During surgery, only 1 of the 2 patients with a positive WBC/BM SPECT CT showed clinical signs of infection such as intra-articular purulence, membranes, stem loosening, or bone loss. For the remaining patients with positive WBC/BM SPECT CT or positive microbiological diagnosis, the surgeon did not find findings suspicious for infection preoperatively (fever, sinus, erythema, elevated blood tests) or during surgery.

The overall diagnostic values of WBC/BM SPECT CT in relation to cultures and MSIS criteria are summarized in Table IV.

Discussion

The use of radionuclide imaging to screen for low-virulent infection in the shoulder has not previously been described. A review by Yue and Tang[23] reported that WBC/BM SPECT imaging diagnosed lower extremity PJI with a high specificity and sensitivity, but large variations in estimates of diagnostic performance between studies exist. Whether a high diagnostic power can be achieved in shoulder arthroplasties is a valid question, especially keeping in mind that chronic low-grade infections (eg, caused by *C acnes*) are the most frequent cause of PJI of the shoulder.

In our study, WBC/BM SPECT CT showed a sensitivity 0.18, specificity 1.00, PPV of 1.00, NPV of 0.67, and an accuracy of 0.69 in detecting a chronic infected shoulder arthroplasty. This is in contrast to Love et al[12] who found an accuracy of 0.95, a sensitivity of 1.00, specificity of 0.91, PPV of 0.89, and NPV of 1.00 using WBC/BM scintigraphy to detect lower limb PJI in a population of both postoperative and chronic

Table IV Diagnostic performance of WBC/BM SPECT CT based on cultures, MSIS, and modified MSIS definition of periprosthetic joint infection

Variable	WBC/BM imaging (95% CI)	Modified MSIS criteria (95% CI)	MSIS criteria (95% CI)
Sensitivity	0.18 (0.00-0.41)	0.20 (0.00-0.45)	0.15 (0.0-0.35)
Specificity	1.00 (1.00-1.00)	1.00 (1.00-1.00)	1.00 (1.00-1.00)
PPV	1.00 (1.00-1.00)	1.00 (1.00-1.00)	1.00 (0.00-0.00)
NPV	0.67 (0.49-0.84)	0.70 (0.53-0.88)	0.59 (0.41-0.78)
Accuracy	0.69 (0.52-0.86)	0.72 (0.56-0.89)	0.62 (0.44-0.80)

WBC/BM, labeled leukocyte and technetium 99m sulfur colloid bone marrow; *SPECT*, single-photon emission computed tomography; *CT*, computed tomography; *MSIS*, Musculoskeletal Infection Society; CI, confidence interval; *PPV*, positive predictive value; *NPV*, negative predictive value.

infections. Their study was performed on 59 hip and knee arthroplasties, and infection was defined as growth in least 1 biopsy sample or >5 neutrophils per high-power field in a synovial sample. The difference in results could partly be explained by the population and the definition of PJI.

If an arthroplasty is infected postoperatively, an aggressive organism like *Staphylococcus aureus* is often the pathogen. This type of infection causes severe soft tissue reactions compared with the more indolent course a chronic infection can have. As a consequence, leucocyte aggregation is marked and expected to be detected more reliably using WBC/BM. Our study group consisted of chronic infections only. Most infections in the shoulder are caused by *C acnes*, and this bacterium was not identified in the hip or knee infections. Studies of *C acnes* have shown that it can escape the immune system by surviving inside macrophages and resisting phagocytosis.[22] Grosso et al[7] examined the aggregation of leucocytes in frozen sections from shoulder revisions. They concluded that fewer leucocytes were present in frozen sections from *C acnes* infections compared with other infections. This could indicate that *C acnes* might not trigger the immune system adequately to be detected by the WBC/BM SPECT CT. In consequence, this would explain the decreased sensitivity in our cohort compared with the study by Love et al.[12]

To minimize the risk of overdiagnosing PJI, our definition requires identification of the same bacteria in at least 3 of 5 separate tissue cultures. However, it is possible that 1 or 2 positive cultures represent a true infection because bacterial load may not be evenly spread throughout the joint. Thus, our definition of 3 or more positive samples could potentially underestimate the infection rate and sensitivity and overestimate the NPV in this study. Nevertheless, because no difference in WBC/BM SPECT performance could be detected using our definitions compared with the MSIS criteria requiring only 2 positive cultures to define infection, an underestimation of infection in our study seems unlikely.

Furthermore, some organisms, such as *C acnes*, exist in 2 states, biofilm and planktonic. The biofilm is found on surfaces of an implant and the planktonic form in tissue and joint fluid. It is accepted that the planktonic form can be detected with regular tissue cultures, but sonication of the extracted implant is often needed to dislodge the bacteria from the biofilm.[20] Sonication was not available to us during the study period, and biopsy samples from the implant surface were obtained by scraping any material off the implant. This could potentially have reduced the number of infected patients in our study group, resulting in even lower sensitivity than observed. However, a recent report by Grosso et al[9] found no benefits of sonication vs. standard intraoperative cultures in diagnosing shoulder PJI.

A limitation of any study investigating PJI diagnostics is the lack of a gold standard criteria for shoulder PJI.[4] Any test of a diagnostic modality can never outperform the chosen gold standard against which the modality is compared. The MSIS adopted in 2013 is a consensus criteria of PJI primarily based on lower limb arthroplasties.[15] Because numerous authors argue that excessive stiffness and nightly pain of a shoulder without any apparent explanation can be a sign of PJI, we modified the MSIS criteria to incorporate both symptoms as secondary criteria and lowered the number of secondary symptoms needed to fulfill the criteria of infection.

Using our modified MSIS criteria of infection compared with the original MSIS criteria, we could not detect a difference in diagnostic power definition (Table V). Perioperative histology was not available to us during the study period; furthermore, despite ultrasound guidance, aspiration in nearly all cases were dry taps. As a consequence, the MSIS criteria can only be positive when a primary sign is present. Because MSIS requires only 2 separate biopsy samples as a primary sign to indicate infection, and our biopsy-based definition requires 3 or more samples to indication infection, the overlap between the 2 definitions is considerable. All 3 definitions exhibit the same limitation, the difficulty of making a precise preoperative diagnosis.

All but one of our infections can be classified as unexpected positive cultures. This is a common finding in shoulder revision surgery, and the clinical implications of this is still uncertain.[3] Because *C acnes* is an abundant commensal around the shoulder girdle, this bacterium is a highly possible contaminant introduced during the revision procedure.[16] If the latter is the case, the WBC/BM SPECT CT would show significantly different results, with specificity and sensitivity near 100%. We believe that using the principle of at least 3 positive cultures from 5 separate biopsy samples as a marker of

true infection substantially lowers the risk of contamination being diagnosed as an infection.

Another limitation of this study is the low number of patients included, which is attributed to 2 factors. The first is a high dropout rate. Twenty-two patients (36% of the patients) were never referred to WBC/BM SPECT CT. This dropout was partly a consequence of far more patients than expected being referred to our center, causing an unacceptably long wait to perform the assay. A potential selection bias could have been introduced if surgeons omitted referral to WBC/BM SPECT CT of patients with low suspicion of infection. An analysis of the patients who dropped out showed no difference in demographics, revision, or infection rate. Thus, we do not believe an influential bias was introduced.

Furthermore, 11 patients who completed the preoperative diagnostics and were offered revision surgery choose to opt out. Four patients stated that due to recent non-shoulder-related surgery, they had no surplus mental and physical resources to undergo major surgery and rehabilitation. For the rest of the group, no specific reason for refusal of surgery was recorded. However, this relatively high number can partly be explained by the long rehabilitation and the sometimes limited gain in function and pain reduction after revision surgery.

Another factor for the high dropout could be attributed to the Danish health care system. In Denmark, treatment is fully paid by the government. As a consequence, a patient who opted out could, without any implications, return to the hospital for surgery at any given time.

The second limitation is the premature closure of the study after the analysis of the results after 1 year, as described in Materials and Methods. If the study had continued, the most likely effect would be more precise estimates of specificity, NPV, and accuracy.

Conclusion

Our study found a low NPV, a high PPV, and a poor accuracy using WBC/BM SPECT CT in screening a failed shoulder arthroplasty for chronic infection. Despite the premature shutdown of the study and the limitations of patient recruitment, the results of WBC/BM SPECT CT in this study are significantly inferior to similar studies of lower limb arthroplasties.

Because of the high risk of an inconclusive result, WBC/BM SPECT CT cannot currently be recommended as a method of screening a failed shoulder arthroplasty for PJI. Thus, future research in functional imaging of shoulder arthroplasties should focus on modalities that can detect chronic low-grade infections or even target specific molecules produced by the most frequent disease-causing bacteria. Furthermore, studies investigating preoperative diagnostic tools should incorporate histologic analysis of biopsy samples.

Disclaimer

Thomas Falstie-Jensen received a grant from Aarhus University covering salary during his PhD training. Aarhus University was not in any way involved in data collection, data analysis, or the preparation of or editing of the manuscript.

The authors, their immediate families, and any research foundation with which they are affiliated did not receive any financial payments or other benefits from any commercial entity related to the subject of this article.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jse.2018.10.024.

References

1. Atkins BL, Athanasou N, Deeks JJ, Crook DW, Simpson H, Peto TE, et al. Prospective evaluation of criteria for microbiological diagnosis of prosthetic-joint infection at revision arthroplasty. The OSIRIS Collaborative Study Group. J Clin Microbiol 1998;36:2932-9.
2. Brolin TJ, Hackett DJ, Abboud JA, Hsu JE, Namdari S. Routine cultures for seemingly aseptic revision shoulder arthroplasty: are they necessary? J Shoulder Elbow Surg 2017;26:2060-6. http://dx.doi.org/10.1016/j.jse.2017.07.006
3. Foruria AM, Fox TJ, Sperling JW, Cofield RH. Clinical meaning of unexpected positive cultures (UPC) in revision shoulder arthroplasty. J Shoulder Elbow Surg 2013;22:620-7. http://dx.doi.org/10.1016/j.jse.2012.07.017
4. Frangiamore SJ, Saleh A, Grosso MJ, Alolabi B, Bauer TW, Iannotti JP, et al. Early versus late culture growth of Propionibacterium acnes in revision shoulder arthroplasty. J Bone Joint Surg Am 2015;97:1149-58. http://dx.doi.org/10.2106/JBJS.N.00881
5. Frangiamore SJ, Saleh A, Grosso MJ, Kovac MF, Higuera CA, Iannotti JP, et al. α-Defensin as a predictor of periprosthetic shoulder infection. J Shoulder Elbow Surg 2015;24:1021-7. http://dx.doi.org/10.1016/j.jse.2014.12.021
6. Gemmel F, Van den Wyngaert H, Love C, Welling MM, Gemmel P, Palestro CJ. Prosthetic joint infections: radionuclide state-of-the-art imaging. Eur J Nucl Med Mol Imaging 2012;39:892-909. http://dx.doi.org/10.1007/s00259-012-2062-7
7. Grosso MJ, Frangiamore SJ, Ricchetti ET, Bauer TW, Iannotti JP. Sensitivity of frozen section histology for identifying Propionibacterium acnes infections in revision shoulder arthroplasty. J Bone Joint Surg Am 2014;96:442-7. http://dx.doi.org/10.2106/JBJS.M.00258
8. Grosso MJ, Frangiamore SJ, Saleh A, Kovac MF, Hayashi R, Ricchetti ET, et al. Poor utility of serum interleukin-6 levels to predict indolent periprosthetic shoulder infections. J Shoulder Elbow Surg 2014;23:1277-81. http://dx.doi.org/10.1016/j.jse.2013.12.023
9. Grosso MJ, Frangiamore SJ, Yakubek G, Bauer TW, Iannotti JP, Ricchetti ET. Performance of implant sonication culture for the diagnosis of periprosthetic shoulder infection. J Shoulder Elbow Surg 2018;27:211-6. http://dx.doi.org/10.1016/j.jse.2017.08.008
10. Hackett DJ Jr, Crosby LA. Evaluation and treatment of the infected shoulder arthroplasty. Bull Hosp Jt Dis (2013) 2013;71(Suppl. 2):88-93.

11. Kamme C, Lindberg L. Aerobic and anaerobic bacteria in deep infections after total hip arthroplasty: differential diagnosis between infectious and non-infectious loosening. Clin Orthop Relat Res 1981;(154):201-7.
12. Love C, Marwin SE, Tomas MB, Krauss ES, Tronco GG, Bhargava KK, et al. Diagnosing infection in the failed joint replacement: a comparison of coincidence detection 18F-FDG and 111In-labeled leukocyte/99mTc-sulfur colloid marrow imaging. J Nucl Med 2004;45:1864-71.
13. Mook WR, Garrigues GE. Diagnosis and management of periprosthetic shoulder infections. J Bone Joint Surg Am 2014;96:956-65. http://dx.doi.org/10.2106/JBJS.M.00402
14. Palestro CJ, Love C, Tronco GG, Tomas MB, Rini JN. Combined labeled leukocyte and technetium 99m sulfur colloid bone marrow imaging for diagnosing musculoskeletal infection. Radiographics 2006;26:859-70. http://dx.doi.org/10.1148/rg.263055139
15. Parvizi J, Gehrke T, International Consensus Group on Periprosthetic Joint I. Definition of periprosthetic joint infection. J Arthroplasty 2014;29:1331. http://dx.doi.org/10.1016/j.arth.2014.03.009
16. Patel A, Calfee RP, Plante M, Fischer SA, Green A. Propionibacterium acnes colonization of the human shoulder. J Shoulder Elbow Surg 2009;18:897-902. http://dx.doi.org/10.1016/j.jse.2009.01.023
17. Singh JA, Sperling JW, Schleck C, Harmsen W, Cofield RH. Periprosthetic infections after shoulder hemiarthroplasty. J Shoulder Elbow Surg 2012;21:1304-9. http://dx.doi.org/10.1016/j.jse.2011.08.067
18. Singh JA, Sperling JW, Schleck C, Harmsen WS, Cofield RH. Periprosthetic infections after total shoulder arthroplasty: a 33-year perspective. J Shoulder Elbow Surg 2012;21:1534-41. http://dx.doi.org/10.1016/j.jse.2012.01.006
19. Tashjian RZ, Granger EK, Zhang Y. Utility of prerevision tissue biopsy sample to predict revision shoulder arthroplasty culture results in at-risk patients. J Shoulder Elbow Surg 2017;26:197-203. http://dx.doi.org/10.1016/j.jse.2016.07.019
20. Trampuz A, Piper KE, Jacobson MJ, Hanssen AD, Unni KK, Osmon DR, et al. Sonication of removed hip and knee prostheses for diagnosis of infection. N Engl J Med 2007;357:654-63. http://dx.doi.org/10.1056/NEJMoa061588
21. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453-7. http://dx.doi.org/10.1016/S0140-6736(07)61602-X
22. Webster GF, Leyden JJ, Musson RA, Douglas SD. Susceptibility of Propionibacterium acnes to killing and degradation by human neutrophils and monocytes in vitro. Infect Immun 1985;49:116-21.
23. Yue B, Tang T. The use of nuclear imaging for the diagnosis of periprosthetic infection after knee and hip arthroplasties. Nucl Med Commun 2015;36:305-11. http://dx.doi.org/10.1097/MNM.0000000000000266

PAPER II

^{18}F FDG-PET/CT has poor diagnostic accuracy in diagnosing shoulder PJI

European Journal of Nuclear Medicine and Molecular Imaging
https://doi.org/10.1007/s00259-019-04381-w

ORIGINAL ARTICLE

18F FDG-PET/CT has poor diagnostic accuracy in diagnosing shoulder PJI

Thomas Falstie-Jensen[1] · J. Lange[2] · H. Daugaard[3] · M. H. Vendelbo[4,5] · A. K. Sørensen[6] · B. Zerahn[7] · J. Ovesen[1] · K. Søballe[1] · L. C. Gormsen[4] · on behalf of the ROSA study-group

Received: 14 February 2019 / Accepted: 28 May 2019

Abstract

Purpose Chronic low-grade periprosthetic joint infection (PJI) of a shoulder replacement can be challenging to diagnose. 18F-FDG PET/CT is suggested as a modality to diagnose lower-limb PJI, but no studies on shoulder replacements exist. The aim of this study was therefore to determine the diagnostic accuracy of 18F-FDG PET/CT in diagnosing chronic PJI of the shoulder.

Methods Patients evaluated for a failed shoulder replacement during a 3-year period were prospectively included in the study. All patients underwent pre-operative 18F-FDG PET/CT, and were evaluated for signs of infection by three independent reviewers using shoulder-specific criteria. Interrater-agreement was calculated between the reviewers. If the patient had revision surgery, biopsy specimens were obtained and cultured with bacterial growth in the cultures serving as gold standard of infection.

Results A total of 86 patients were included in the study. Nine patients were 18F-FDG PET/CT positive for infection, with only three true positive. Using the gold standard, infection was diagnosed after revision surgery in 22 cases. All infections were chronic and caused by low-virulent microbes. The sensitivity of 18F-FDG PET/CT was 0.14 95% CI (0.03–0.36), specificity 0.91 95% CI (0.81–0.97), positive predictive value was 0.40 95% CI (0.15–0.71) and negative predictive value 0.71 95% CI (0.67–0.75). The inter-observer agreement was 0.56 (Fleiss' kappa), indicating moderate agreement of the visual FDG-PET evaluation using the shoulder-specific criteria.

Conclusion 18F-FDG PET/CT has poor diagnostic accuracy in diagnosing low-grade PJI of the shoulder. 18F-FDG PET/CT cannot be recommended as a part of the routine preoperative workup to diagnose low-grade infection of a shoulder replacement.

Keywords FDG-PET · PJI · Infection · Shoulder · Periprosthetic joint infection

Introduction

The incidence of shoulder joint replacement surgery is expected to increase considerably due to increased life-expectancy and physical demands of senior citizens. Consequently, the absolute number of patients with complications, such as periprosthetic joint infection (PJI), is likely to increase. PJI of the shoulder is a rare, but critical condition. It can lead to prolonged morbidity, additional operations, and a reduction in functional outcome [1, 2].

This article is part of the Topical Collection on Orthopaedics

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00259-019-04381-w) contains supplementary material, which is available to authorized users.

✉ Thomas Falstie-Jensen
thomfals@rm.dk

[1] Orthopedic Department, Shoulder and elbow section, Aarhus University Hospital, Palle Juul Jensens Boulevard 99, 8200 Aarhus, Denmark

[2] Orthopedic Department, Horsens Regional Hospital, Horsens, Denmark

[3] Orthopedic Department, Gentofte Hospital, Gentofte, Denmark

[4] Department of Nuclear Medicine and PET-Center, Aarhus University Hospital, Aarhus, Denmark

[5] Department of Biomedicine, Aarhus University, Aarhus, Denmark

[6] Orthopedic Department, Herlev Hospital, Herlev, Denmark

[7] Department of Clinical Physiology, Herlev Hospital, Herlev, Denmark

Published online: 11 July 2019

PJI incidence is reported to range from 1% to up to 5% after primary total shoulder joint replacement [1, 3]. However, studies investigating routine cultures from aseptic revisions or removal of presumably non-infected hardware from the shoulder girdle report a positive culture-rate of up to 29% [4–6]. This high rate of unexpected positive cultures raises a concern of underdiagnosed low-grade shoulder PJI, since multiple positive cultures are regarded as a strong indicator of infection [7, 8].

PJIs are a challenge to diagnose clinically, because chronic low-grade infection often mimics symptoms of aseptic loosening [9]. Traditionally, the preoperative diagnostic workup consists of blood tests, synovial markers, aspiration, and conventional imaging, but these tests are neither specific nor sensitive for chronic PJI [10, 11]. So far, no single test or composite variables have been able to diagnose low-grade PJI preoperatively with high sensitivity, including radionuclide imaging, which is occasionally used to diagnose PJI. A recently published consensus paper investigated the use of advanced imaging in diagnosing PJI, and concluded that level 1 and 2 evidence is sparse, making solid recommendations difficult to give [12]. Currently, white blood cell/bone marrow (WBC/BM) SPECT CT imaging is considered the optimal radionuclide imaging technique in lower limb arthroplasties [13, 14]. However, it is debatable whether this is a favorable modality for shoulder PJIs, as evidenced by the poor results of our own recently published study [15] . In addition, WBC/BM SPEC CT is a complex and time-consuming procedure with a considerable radiation burden. Consequently, simpler functional imaging techniques are highly warranted.

An attractive alternative is ^{18}F-FDG PET/CT (FDG-PET), which has proven effective in diagnosing orthopedic and vascular graft infections [16]. First, the spatial resolution of FDG-PET outperforms gamma camera scintigraphy, theoretically improving the possibility of detecting the exact anatomical location of any infectious foci in cases with several implant components, such as total shoulder replacements [17]. Second, FDG-PET has several practical advantages, including far simpler patient and scanner logistics. Third, FDG-PET can be performed with the use of considerably less ionizing radiation than WBC/BM scintigraphy. Fourth, the diagnostic accuracy of FDG-PET to diagnose hip PJI has been shown in some studies to be high [18]. However, to our knowledge, no studies have examined the diagnostic performance of FDG-PET on failed shoulder replacements.

The aim of this study was therefore to assess the diagnostic performance of ^{18}F-FDG PET/CT in diagnosing chronic low-grade infection of failed shoulder replacements by measures of sensitivity, specificity, positive predictive value (PPV), and negative predicted value (NPV).

Material and methods

Study design

The study was a prospective, nationwide, cohort study conducted on consecutive patients with a failed shoulder replacement during the period 1st October 2015 to 30th September 2017.

In Denmark, all 5.7 million citizens have free and universal access to health care. Two public orthopedic departments are by law appointed as the only departments allowed to revise shoulder joint replacements. Patients with a failed shoulder arthroplasty are (mandatorily) referred to these two departments by general practitioners or other orthopedic departments.

The study was approved by the Local Ethical Scientific Committee (ref. no. 1–10–72-229-15) and the Danish Data Protection Agency (ref. no 1–16–02-567-13).

Study population

During the study period, all patients ($n = 292$) referred to the two departments with a failed shoulder replacement were informed about the study. Patients were included if a written consent was obtained, excluding patients with failure due to acute fracture, traumatic dislocation of the arthroplasty, or prior ipsilateral revision surgery due to chronic PJI. After enrolment, patients completed a standardized preoperative workup consisting of joint aspiration, radiographic assessment,

PJI positive if at least one primary sign exists OR at least two secondary signs exist.

Primary signs

1)There is a sinus tract communicating with the prosthesis

2)A pathogen is isolated by culture from at least three separate tissue or fluid samples obtained from the affected prosthetic joint

OR

Secondary signs

a) Elevated serum C-reactive protein (CRP) OR erythrocyte sedimentation rate OR White blood cell count
b) A single positive culture
c) Positive PET/CT imaging
d) Unexplained nightly pain OR excessive joint stiffness

Fig. 1 Modified MSIS PJI definition (mMSIS)

serum inflammatory markers (erythrocyte sedimentation rate, white blood cell count, and C-reactive protein) and FDG-PET. After completion of the FDG-PET scan, an outpatient visit was conducted, in which all diagnostic tests were reviewed and a tentative diagnosis of infection was set based on the modified Musculoskeletal Infection Society definition (MSIS) shown in Fig. 1. Revision surgery was recommended based on three conditions: i) if FDG-PET strongly indicated infection, ii) if the biochemistry or physical assessment indicated infection, or iii) if signs of component loosening was found, or obvious mechanical dysfunction was observed on conventional imaging. Patients willing to accept limitations of shoulder joint movement and level of pain, as well as patients in whom surgical treatment was not recommended, were omitted from the study ($n = 57$). A flow-chart of patient inclusion is presented in Fig. 2.

The remaining 86 patients underwent revision and comprised the diagnostic accuracy study cohort. Thus, all patients had a pre-operative FDG PET, and cultures of tissue-specimens obtained during the revision served as clinical diagnosis of infection. Baseline patient demographics are presented in Table 1.

Definition of true infection status

In order to assess the diagnostic accuracy of FDG-PET, a diagnostic gold standard in PJI is required. However, no such standard exists [19]. This study defined PJI as growth of the same organism in cultures from at least three of five separate biopsy specimens obtained during revision. The specimens were obtained with clean utensils from areas showing signs of infection, according to the method described by Kamme-Linberg et al. [20]. If no signs of infection were obvious, specimens were obtained from five random isolated intraarticular spots. All biopsy specimens were cultured for 4 days for aerobic bacteria and a further 10 days for anaerobic bacteria, the latter to detect slow-growing bacteria such as Cutibacterium acnes (C. acnes, formerly known as Proprionibacterium acnes). Any therapy with antibiotics was halted 2 weeks prior to revision surgery, and perioperative antibiotics were withheld until all biopsies were obtained. Both steps were implemented to reduce the risk of false negative results owing to iatrogenic suppression of bacterial growth.

Image acquisitions

PET/CT scans were performed on a Siemens Biograph 64 PET/CT (Erlangen, Germany). All patients fasted at least 6 h before the procedure. Bedside PET fasting blood glucose > 11 mmol/l was not accepted. In both centers, imaging was performed 60 min after intravenous administration of the radiotracer. An initial low-dose CT without contrast enhancement (50 mAs) was performed, in order to correct for photon attenuation and to co-localize FDG uptake and anatomical structures. At Center One, the FDG PET was performed from

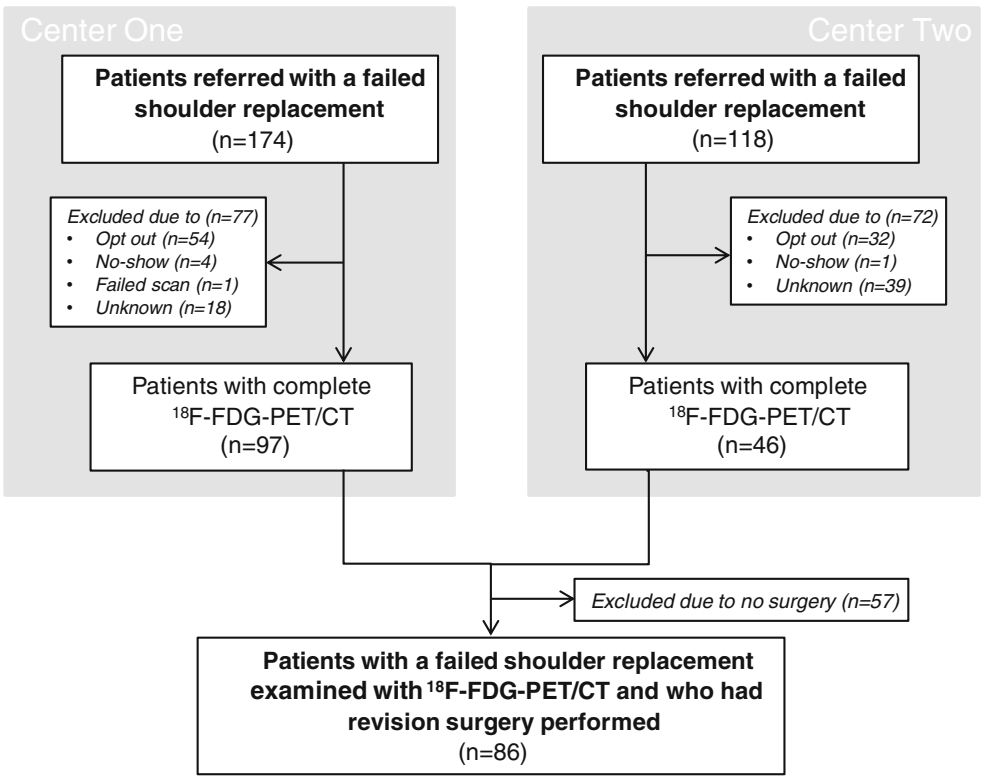


Fig. 2 Patient inclusion

Table 1 Demographics

		Referred	No consent	Excluded	Study group
No. of patients		292	149	57	86
Mean age (range)		68 (34–89)	69 (45–87)	68 (34–89)	67 (42–85)
Gender: *n* (%)	Female	182 (62%)	91 (61%)	28 (49%) *	63 (73%)
	Male	110 (38%)	58 (39%)	29 (51%) *	23 (27%)
Prosthesis age (mean)		4.2 years	4.6 years	3.8 years	4.5 years
Revision rate: % (*n*)		47% (137)	30% (45)	0% (0)	100% (86)
Infection rate % (*n*)		27% (38)	33% (16)	NA	26% (22)

*$p < 0.05$. Statistical difference compared to study group

elbow to basis cranii (4 MBq ^{18}F-FDG kg^{-1}; 2 min per bed position in three-dimensional mode) and at Center Two, from the mid-thigh to the skull (5 MBq 18F-FDG kg^{-1}; 3 min per bed position in three-dimensional mode). Both centers adhered to international guidelines [21].

Reconstruction of attenuation-corrected images was done using visually comparable, ordered subset expectation maximisation algorithms with point-spread function (PSF) (Siemens Biograph: Four iterations, 21 subsets, 3-mm Gaussian post-processing filter, matrix size 336 × 336). The estimated dose of radiation per patient was 9.5 mSv at Center One and 12.8 mSv at Center Two.

Image analysis

All images were reviewed using first attenuation-corrected images and in equivocal cases subsequently non-attenuation corrected images.

First, a senior ortho-nuclear consultant blinded for results of the preoperative work-up made an initial assessment of FDG-PET scans. Based on the consultants' experience from other cases of orthopedic or implant infections, an initial dichotomized diagnosis (infection/no infection) was set based on pattern of activity. This initial diagnosis, referred to as "best practice", was used due to the lack of a standardized algorithm for evaluating PJI in shoulders with FDG-PET.

Second, we developed a set of criteria to diagnose infection by the FDG-PET, hereafter referred to as "shoulder-specific assessment". This shoulder-specific assessment was constructed using the knowledge of initial observed patterns of activity compared with the true infection status, rate-of-loosening of the stem and cavitas components in shoulder replacements, and the criteria published by Reinartz in 2005 proposed for FDG-PET evaluation of PJI in hip arthroplasties [22]. The shoulder-specific assessment dictates that FDG-PET was positive for infection if at least one of the following patterns of activity were observed: i) increased FDG uptake in soft tissue adjoining the joint cavity, ii) increased FDG uptake along the humoral stem, or iii) increased uptake in regional lymph nodes near the affected shoulder. Examples of such patterns of activity are shown in Fig. 3.

Third, all scans were reassessed using the shoulder-specific assessment by three consultants blinded for the previous FDG-PET results. Subsequently, a new dichotomous FDG-PET diagnosis was made based on majority decision, e.g., if one reviewer judged the scan negative and two reviewers positive, the new status was positive.

Statistical analysis

Sensitivity, specificity, PPV, and NPV of FDG-PET were calculated using the microbial diagnosis as true infectious status (reference standard). All results are reported as fractions with 95% confidence intervals (CI).

Interobserver agreement was evaluated using kappa statistics for multiple raters per patient, a method put forward by Fleiss in 1971 [23]. A kappa value of zero indicates agreement as expected by chance. Kappa values were graded as poor (0.00–0.20), fair (0.21–0.40), moderate (0.41–60), good (0.61–0.80) and excellent > 0.81 [24].

Data was analyzed using STATA 15, StataCorp, College Station, TX, USA.

Results

Cultures of biopsy specimens from revisions found that 22 patients were infected. Table 2 shows pre- and perioperative observation for these patients. The most frequent isolated bacterium was Cutibacterium acnes in 17 cases, followed by *Staphylococcus epidermidis* in two cases, and one case of infection caused by each of the following; Staphylococcus hominis, Staphylococcus capitis, and unclassified coagulase-negative staphylococcus. We have enclosed a full overview of the culture results in all the infected cases and examples of different classification criteria for infection in Supplementary file 1.

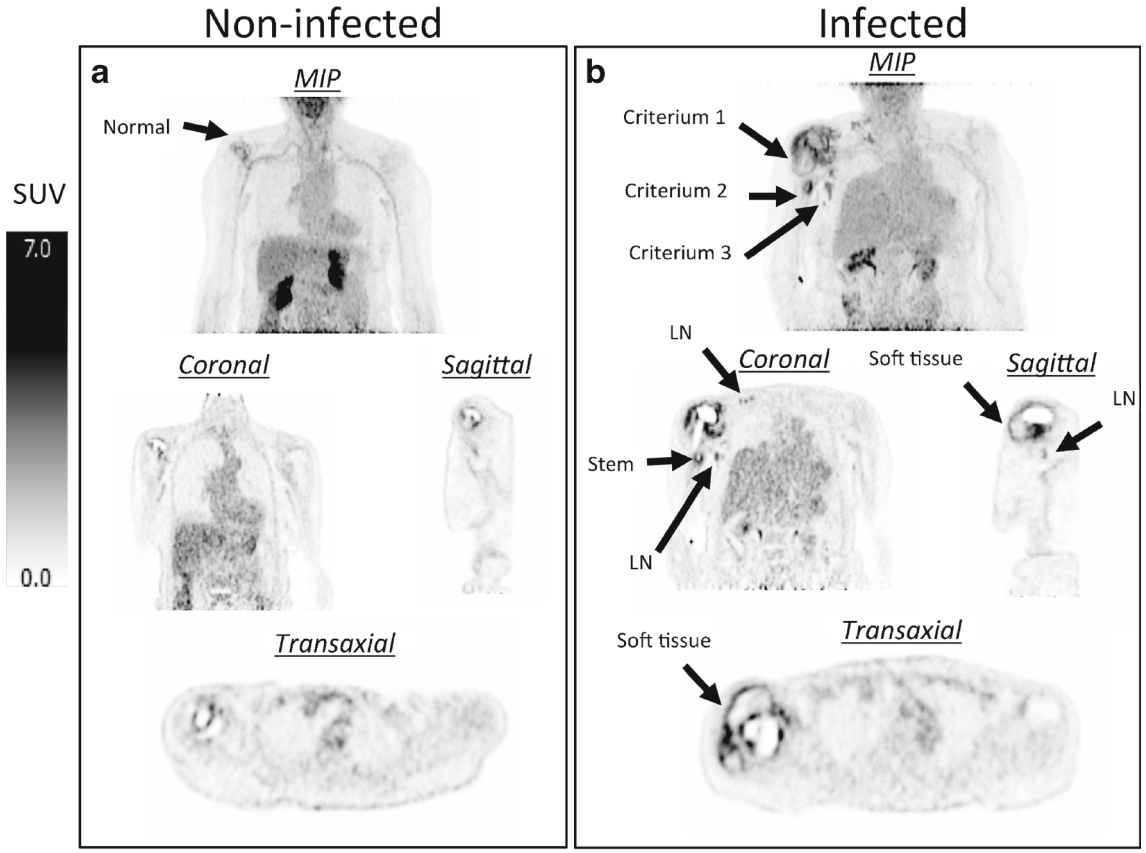


Fig. 3 Examples of criteria 1 to 3 of shoulder-specific FDG-PET assessment; *SUV* Standard Uptake Value, *LN* Lymph node, *MIP* Maximal Intensity Projection; Criterion 1 increased FDG uptake in soft tissue adjoining the joint cavity; Criterion 2 increased FDG uptake along the humoral stem; Criterion 3 increased uptake in regional lymph nodes near the affected shoulder

The mean time from last arthroplasty-related surgery to FDG-PET imaging was 4.6 years for the infected group and 4.4 years for the non-infected group ($p > 0.05$).

Using the shoulder-specific assessment, nine patients were diagnosed as infected. Three scans were true positive, six false positive, 19 false negative, and 58 true negative. Representative examples of scans are shown in Fig. 4. This results in a sensitivity of 0.14 (95% CI: 0.03–0.36) and specificity of 0.91 (95% CI: 0.81–0.97). The PPV and NPV was 0.40 (95% CI: 0.15–0.71) and 0.71 (95% CI: 0.67–0-75). Results are summarized in Tables 3, 4, and 5. The results stratified by center can be found in Table 7 (online supplementary file 2).

The overall agreement of infection diagnosis based on the shoulder-specific assessment among the three reviewers was 0.56 (Fleiss' kappa) indicating moderate inter-observer agreement. The shoulder-specific assessment contained three patterns of activity which each defined infection. When investigating the sub-criteria in the shoulder-specific assessment, the inter-observer agreement of pathologic tracer uptake in soft tissue adjoining the shoulder joint was 0.34 (fair); agreement of pathologic uptake along the stem was 0.61 (good), and for regional lymph nodes 0.57(moderate).

Discussion

The main finding of our study is that FDG-PET performs poorly in diagnosing chronic low-grade PJI of the shoulder, even with introduction of joint-specific patterns of FDG uptake previously demonstrated to perform well in lower limb prosthetic infections. In addition, the overall agreement between the three reviewers of the FDG-PET scans using the shoulder-specific assessment was only moderate, and agreement on soft-tissue tracer uptake near the joint (criterion 1) was even worse. This is hardly surprising, since prosthesis-near FDG activity varies significantly even in non-infected and well-fixed implants [25]. However, we also observed only mediocre agreement of such distinct patterns of activity as uptake along the humoral stem (criterion 2) and uptake in regional lymph nodes (criterion 3), further underscoring the difficulty in assessing the images. Such poor results suggest that any alterations in peri-prosthetic FDG uptake caused by low-grade infections are discrete, and that visual discrimination between pathological and normal FDG uptake is virtually impossible.

Table 2 Details on serum markers and perioperative observations in all infected patients

Patient number	C-reactive protein (mg/l, < 8)	White blood cell count (10^9/l, < 10.0)	Sedimentation rate (mm/h, < 30)	Majority FDG-PET diagnose	Perioperative suspicion of infection
1	1.3	6.2	11	No infection	None
2	12.6	5.17	12	No infection	None
3	22	5	22	No infection	None
4	4.5	6.84	27	No infection	None
5	2.7	9.42	1	No infection	None
6	0.5	5.59	7	No infection	None
7	4	10	16	No infection	None
8	19	9.7	34	Infection	Infection[a]
9	0.5	11.8	4	No infection	None
10	0.5	9.5	2	No infection	None
11	0.5	7.1	6	No infection	None
12	0.5	3.6	7	Infection	Infection[b]
13	0.5	7.8	16	No infection	None
14	0.5	6.4	6	No infection	None
15	0.5.	9.7	23	No infection	None
16	N/A	N/A	N/A	No infection	None
17	0.5	7.5	8	No infection	None
18	0.5	5.4	13	No infection	None
19	0.5	5.7	11	No infection	None
20	44	6.7	56	Infection	Infection[c]
21	0.5	11.2	5	No infection	None
22	0	N/A	2	No infection	None

[a] membranes and cloudy fluid, [b] membranes and pus, [c] membranes, cloudy fluid, and pus

Mixed results have been reported on FDG-PET's performance in detecting PJI of the lower limb, and the usefulness of FDG-PET in this setting is still debatable [26]. Verberne et al. reported in a review from 2016 that FDG-PET could predict PJI of hip replacements with a sensitivity and specificity of 83% and 91% [27]. FDG-PET also appears to perform poorer in knee compared to hip replacements. Verberne et al. reported in a meta-analysis from 2017 a specificity of 84% and a sensitivity of 70% and Mayer-Wagner et al. a sensitivity of 14% and specificity of 89% [28, 29]. However, the latter was based on a very small amount of material. Our diagnostic performance is markedly inferior to any study on failed hip replacements, which may partly be attributed to the definition of infection and to different microbiology. Whereas our reference standard was positive cultures in at least three of five specimens, other studies have defined infection as one positive biopsy or a combination of histopathology and microbiology [27, 30, 31]. Our more rigorous definition of infection was chosen to reduce the risk of misclassifying perioperative contamination as infection [8]. Despite the obvious lack of consensus, we have chosen to treat patients as infected if three or more separate biopsy specimens turn positive for the same bacteria, until clear guidelines on how to interpret cultures with low virulent microbes is presented.

Another factor with considerable impact on the diagnostic FDG-PET performance is the criteria applied to evaluate the pattern of pathological activity [32]. A set of criteria for failed hip replacements were developed by Reinartz et al. in 2005 [22]. It defines infection as increased FDG uptake in the periprosthetic tissue, whereas increased uptake along the stem is considered indicative of loosening. However, differentiation between aseptic and septic loosening was not possible. The inability to differentiate the cause of loosening is probably related to the polyethylene debris from wear of the prosthesis components. The inflammatory cells seen in areas with polyethylene wear or bone remodeling will cause increased FDG uptake, just as activated leucocytes in areas of infection would [28]. Since load bearing, risk of aseptic stem-loosening, and microbiology are different in hips compared to shoulders, application of these criteria without modification to our settings is questionable. In the shoulder, aseptic loosening of the humeral component is rare, and as a consequence the risk of false-positive scans from aseptic stem-loosening seem limited [33]. On the glenoid side, component loosening is more frequent, and signs of loosening, e.g., radiolucent lines on conventional imaging are not uncommon [34]. In our cohort, ten

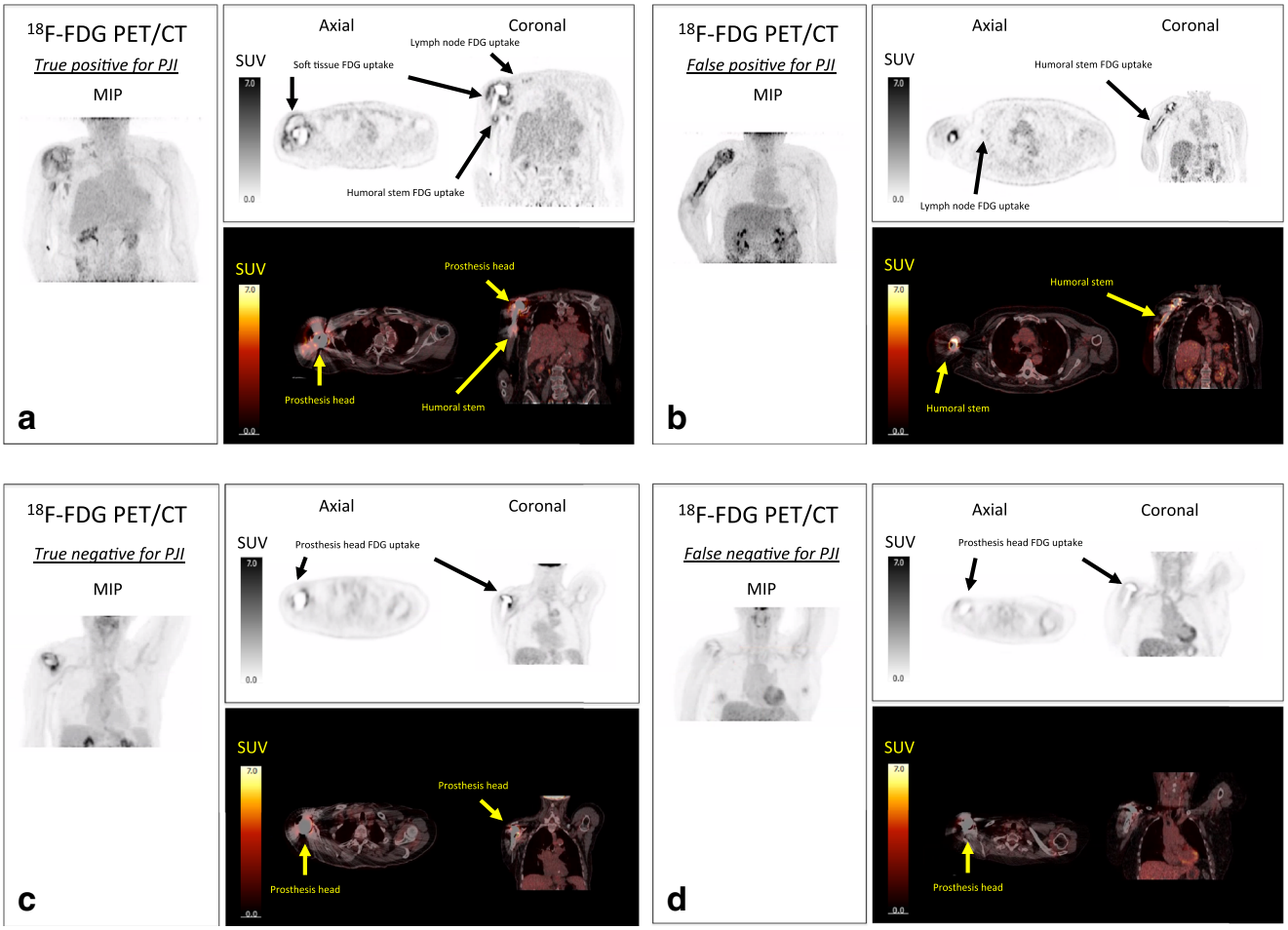


Fig. 4 Shoulder-specific assessment on representative examples of FDG-PET images; All images are attenuation corrected; a True positive scan. +Lymph node activity, + Humoral stem activity, + Soft tissue activity, positive cultures; b False positive scan. +Lymph node activity, + Humoral stem activity, - Soft tissue activity, no growth; c True negative scan. - Lymph node activity, - Humoral stem activity, - Soft tissue activity, no growth in cultures; d False negative scan. -Lymph node activity, - Humoral stem activity, - Soft tissue activity, positive

glenoid components were found to be loose; two of these shoulders were infected, but none showed increased activity on the PET scan. Thus, omitting evaluation of activity near the glenoid component in our shoulder-specific assessment does not decrease the probability of diagnosing infection.

Table 3 FDG-PET diagnosis using shoulder-specific-assessment[a] vs. true infection

	FDG-PET diagnosis		
Infection diagnosis based on biopsies	Negative	Positive	
No infection	58	6	64
Infection	19	3	22
	77	9	86

[a] Majority diagnose

Table 4 Diagnostic performance of FDG-PET using shoulder-specific assessment[a]

Sensitivity	0.14 95%CI: 0.03–0.35
Specificity	0.91 95%CI: 0.81–0.97
PPV	0.38 95%CI: 0.15–0.70
NPV	0.71 95%CI: 0.67–0.75

[a] Majority diagnose

Table 5 Serum markers and mMSIS status compared to FDG-PET and true infection diagnosis

	Infection		No injection	
	True positive[a] ($n = 3$)	False negative[a] ($n = 19$)	True negative[a] ($n = 58$)	False positive[a] ($n = 6$)
Serum markers[b]				
C-reactive protein (mg/l)[c]	19 (< 0.6–44)	< 0.6 (< 0.6–22)	1.2 (< 0.6–16.9)	9 (4.8–59)
Sedimentation rate (mm/h)	34 (7–56)	8 (1–27)	10 (2–67)	38.5 (5–54)
White blood cell count (10^9/l)	6.7 (3.6–9.7)	8.9 (4.7–12.9)	6.9 (3.7–10.7)	7.3 (5–11.8)
Infection status measured by modified MSIS criterion				
mMSIS status with postoperative information	3 positives	18 positives	53 negatives 5 positives	1 negative 5 positives
mMSIS with only preoperative information	2 positives 1 negative	18 negatives	57 negatives 1 positive	1 negative 5 positives

[a] FDG-PET diagnose measured by majority diagnose of Shoulder-Specific assessment, [b] median (range), [c] lowest detection level 0.6 mg/l

This study has several strengths. The most important is the short time-frame and prospective nature of inclusion, which ensured unaltered diagnostic modalities in the form of PET/CT scanner systems, reconstruction protocols, and microbiology cultures. Furthermore, despite shoulder revisions being less common that hip and knee revisions, the nationwide approach resulted in a high number of included patients. Last, the FDG-PET scans were performed on all patients as a screening procedure. This reduced the risk of selection bias in the study group, since no selection of patients occurred before inclusion.

A limitation of this study is the presence of culture-negative infections. This entity is thought to make up approximately 20% of revised joint replacements, and encompass cases in which no microbes can be identified despite strong clinical suspicion of infection [35]. Two of our six false-positive cases were culture-negative infections, since both patients were revised with a two-staged approach due to strong clinical suspicion of infection; but no growth was demonstrated. These culture-negative infections further illustrate the difficulties of choosing a gold standard of PJI to reference the diagnostic test. Sample images of the six false-positive cases and relevant clinical information can be found in Supplementary file 3.

The findings of this study suggest that FDG-PET has a limited place in the preoperative infectious work-up of shoulder PJI. However, due to the high prevalence of C. acnes infection of shoulders these results should be extrapolated with caution to cases with other orthopedic implants. Low virulent PJI have shown to produce less leucocyte migration to the infected joint, compared to more aggressive infections with, for example, *Staphylococcus aureus* [36–38]. Since the majority of infections in our cohort were caused by low virulent microbes, these two factors together could offer a possible explanation for the difficulty of detecting enhanced tracer uptake in infected patients. Despite this, we were able to identify two C. acnes and one *Staphylococcus epidermidis* infections.

Conclusion

Failure of a shoulder arthroplasty due to chronic low-grade infection poses a significant diagnostic challenge with an array of modalities used in the preoperative workup. In this study, we have presented data showing that FDG-PET has poor diagnostic accuracy in diagnosing chronic infection of shoulder joint replacement. Despite use of shoulder-specific criteria for evaluating the FDG-PET scans, the diagnostic performance was low. Our results lead us to conclude that FDG-PET should not constitute a routine part, if any, in the preoperative investigations of failed shoulder arthroplasties.

Acknowledgements Radiologic Department, Aarhus University Hospital, Aarhus, Denmark

Authors' contributions TFJ conceived the study, collected data, analyzed data, and drafted the manuscript. JL participated in conceiving the study, and revised the manuscript. HD revised the manuscript. MHV reviewed images. BZ reviewed images and revised the manuscript. JOV collected data and revised the manuscript. KS participated in conceiving the study, and revised the final manuscript. AKS collected data and revised the manuscript. LCG participated in conceiving the study, reviewed images, and revised the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest Author T Falstie-Jensen declares that he has no conflict of interest. Author J Lange declares that he has no conflict of interest. Author H Daugaard declares that he has no conflict of interest. Author MH Vendelbo declares that he has no conflict of interest. AKB Sørensen declares that she has no conflict of interest. Author B Zerahn declares that he has no conflict of interest. Author J Ovesen declares that she has no conflict of interest. Author K Søballe declares that he has no conflict of interest. Author LC Gormsen declares that he has no conflict of interest.

Corresponding author (T Falstie-Jensen) has received a grant from Aarhus University (no grant no.) covering part of the fee for the PhD

training program. Furthermore, corresponding author has received a research grant from A. P. Møllers og Hustru Chastine Mc-Kinney Møllers Fond til almene Formaal (no grant no.) covering study expenses related to the PhD. The two institutions have not in any way been involved in data collection, data analysis, preparation of, or editing of the manuscript.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (ref. no. 1–10–72-229-15) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Padegimas EM, Maltenfort M, Ramsey ML, Williams GR, Parvizi J, Namdari S. Periprosthetic shoulder infection in the United States: incidence and economic burden. J Shoulder Elb Surg. 2015;24(5):741–6. https://doi.org/10.1016/j.jse.2014.11.044.
2. Nelson GN, Davis DE, Namdari S. Outcomes in the treatment of periprosthetic joint infection after shoulder arthroplasty: a systematic review. J Should Elb Surg. 2016;25(8);1337–45. https://doi.org/10.1016/j.jse.2015.11.064
3. Florschutz AV, Lane PD, Crosby LA. Infection after primary anatomic versus primary reverse total shoulder arthroplasty. J Shoulder Elb Surg. 2015;24(8):1296–301. https://doi.org/10.1016/j.jse.2014.12.036.
4. Kelly JD 2nd, Hobgood ER. Positive culture rate in revision shoulder arthroplasty. Clin Orthop Relat Res. 2009;467(9):2343–8. https://doi.org/10.1007/s11999-009-0875-x.
5. Foruria AM, Fox TJ, Sperling JW, Cofield RH. Clinical meaning of unexpected positive cultures (UPC) in revision shoulder arthroplasty. J Shoulder Elb Surg. 2013;22(5):620–7. https://doi.org/10.1016/j.jse.2012.07.017.
6. Both A, Klatte TO, Lubke A, Buttner H, Hartel MJ, Grossterlinden LG, et al. Growth of Cutibacterium acnes is common on osteosynthesis material of the shoulder in patients without signs of infection. Acta Orthop. 2018;89(5):580–4. https://doi.org/10.1080/17453674.2018.1489095.
7. Padegimas EM, Lawrence C, Narzikul AC, Zmistowski BM, Abboud JA, Williams GR, et al. Future surgery after revision shoulder arthroplasty: the impact of unexpected positive cultures. J Shoulder Elb Surg. 2017;26(6):975–81.
8. Kheir MM, Tan TL, Ackerman CT, Modi R, Foltz C, Parvizi J. Culturing periprosthetic joint infection: number of samples, growth duration, and organisms. J Arthroplast. 2018. https://doi.org/10.1016/j.arth.2018.06.018.
9. Mook WR, Garrigues GE. Diagnosis and management of periprosthetic shoulder infections. J Bone Joint Surg Am. 2014;96(11):956–65. https://doi.org/10.2106/JBJS.M.00402.
10. Topolski MS, Chin PY, Sperling JW, Cofield RH. Revision shoulder arthroplasty with positive intraoperative cultures: the value of preoperative studies and intraoperative histology. J Shoulder Elbow Surg. 2006;15(4):402–6.
11. Updegrove GF, Armstrong AD, Kim HM. Preoperative and intraoperative infection workup in apparently aseptic revision shoulder arthroplasty. J Shoulder Elb Surg. 2015;24(3):491–500. https://doi.org/10.1016/j.jse.2014.10.005.
12. Signore A, Sconfienza LM, Borens O, Glaudemans A, Casar-Pullicino V, Trampuz A, et al. Consensus document for the diagnosis of prosthetic joint infections: a joint paper by the EANM, EBJIS, and ESR (with ESCMID endorsement). Eur J Nucl Med Mol Imaging. 2019;46(4):971–88https://doi.org/10.1007/s00259-019-4263-9.
13. Gemmel F, Van den Wyngaert H, Love C, Welling MM, Gemmel P, Palestro CJ. Prosthetic joint infections: radionuclide state-of-the-art imaging. Eur J Nucl Med Mol Imaging. 2012;39(5):892–909. https://doi.org/10.1007/s00259-012-2062-7.
14. Love C, Marwin SE, Tomas MB, Krauss ES, Tronco GG, Bhargava KK, et al. Diagnosing infection in the failed joint replacement: a comparison of coincidence detection 18F-FDG and 111In-labeled leukocyte/99mTc-sulfur colloid marrow imaging. J Nucl Med. 2004;45(11):1864–71.
15. Falstie-Jensen T, Daugaard H, Soballe K, Ovesen J, Arveschoug AK, Lange J, et al. Labeled white blood cell/bone marrow single-photon emission computed tomography with computed tomography fails in diagnosing chronic periprosthetic shoulder joint infection. J Shoulder Elb Surg. 2019;28(6):1040–8. https://doi.org/10.1016/j.jse.2018.10.024.
16. Basu S, Chryssikos T, Moghadam-Kia S, Zhuang H, Torigian DA, Alavi A. Positron emission tomography as a diagnostic tool in infection: present role and future possibilities. Semin Nucl Med. 2009;39(1):36–51. https://doi.org/10.1053/j.semnuclmed.2008.08.004.
17. Chryssikos T, Parvizi J, Ghanem E, Newberg A, Zhuang H, Alavi A. FDG-PET imaging can diagnose periprosthetic infection of the hip. Clin Orthop Relat Res. 2008;466(6):1338–42. https://doi.org/10.1007/s11999-008-0237-0.
18. Kwee TC, Kwee RM, Alavi A. FDG-PET for diagnosing prosthetic joint infection: systematic review and metaanalysis. Eur J Nucl Med Mol Imaging. 2008;35(11):2122–32. https://doi.org/10.1007/s00259-008-0887-x.
19. Frangiamore SJ, Saleh A, Kovac MF, Grosso MJ, Zhang X, Bauer TW, et al. Synovial fluid interleukin-6 as a predictor of periprosthetic shoulder infection. J Bone Joint Surg Am. 2015;97(1):63–70. https://doi.org/10.2106/JBJS.N.00104.
20. Kamme C, Lindberg L. Aerobic and anaerobic bacteria in deep infections after total hip arthroplasty: differential diagnosis between infectious and non-infectious loosening. Clin Orthop Relat Res. 1981;154:201–7.
21. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42(2):328–54. https://doi.org/10.1007/s00259-014-2961-x.
22. Reinartz P, Mumme T, Hermanns B, Cremerius U, Wirtz DC, Schaefer WM, et al. Radionuclide imaging of the painful hip arthroplasty: positron-emission tomography versus triple-phase bone scanning. J Bone Joint Surg Br. 2005;87(4):465–70. https://doi.org/10.1302/0301-620X.87B4.14954.
23. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378–82. https://doi.org/10.1037/h0031619.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.
25. Gelderman SJ, Jutte PC, Boellaard R, Ploegmakers JJW, Vallez Garcia D, Kampinga GA, et al. (18)F-FDG-PET uptake in non-infected total hip prostheses. Acta Orthop. 2018;89(6):634–9. https://doi.org/10.1080/17453674.2018.1525931.
26. Love C, Palestro CJ. Nuclear medicine imaging of bone infections. Clin Radiol. 2016;71(7):632–46. https://doi.org/10.1016/j.crad.2016.01.003.
27. Verberne SJ, Raijmakers PG, Temmerman OP. The accuracy of imaging techniques in the assessment of periprosthetic hip infection: a systematic review and meta-analysis. J Bone Joint Surg Am. 2016;98(19):1638–45. https://doi.org/10.2106/JBJS.15.00898.
28. Mayer-Wagner S, Mayer W, Maegerlein S, Linke R, Jansson V, Muller PE. Use of 18F-FDG-PET in the diagnosis of endoprosthetic

loosening of knee and hip implants. Arch Orthop Trauma Surg. 2010;130(10):1231–8. https://doi.org/10.1007/s00402-009-1000-z.
29. Verberne SJ, Sonnega RJ, Temmerman OP, Raijmakers PG. What is the accuracy of nuclear imaging in the assessment of periprosthetic knee infection? A meta-analysis. Clin Orthop Relat Res. 2017;475(5):1395–410. https://doi.org/10.1007/s11999-016-5218-0.
30. Aksoy SY, Asa S, Ozhan M, Ocak M, Sager MS, Erkan ME, et al. FDG and FDG-labelled leucocyte PET/CT in the imaging of prosthetic joint infection. Eur J Nucl Med Mol Imaging. 2014;41(3):556–64. https://doi.org/10.1007/s00259-013-2597-2.
31. Delank KS, Schmidt M, Michael JW, Dietlein M, Schicha H, Eysel P. The implications of 18F-FDG PET for the diagnosis of endoprosthetic loosening and infection in hip and knee arthroplasty: results from a prospective, blinded study. BMC Musculoskelet Disord. 2006;7:20. https://doi.org/10.1186/1471-2474-7-20.
32. Verberne SJ, Temmerman OPP, Vuong BH, Raijmakers PG. Fluorodeoxyglucose positron emission tomography imaging for diagnosing periprosthetic hip infection: the importance of diagnostic criteria. Int Orthop. 2018;42(9):2025–34. https://doi.org/10.1007/s00264-018-3931-x.
33. Keener JD, Chalmers PN, Yamaguchi K. The humeral implant in shoulder arthroplasty. J Am Acad Orthop Surg. 2017;25(6):427–38. https://doi.org/10.5435/JAAOS-D-15-00682.
34. Hsu JE, Hackett DJ, Jr., Vo KV, Matsen FA, 3rd. What can be learned from an analysis of 215 glenoid component failures? J Shoulder Elb Surg. 2018;27(3):478–486. https://doi.org/10.1016/j.jse.2017.09.029.
35. Tan TL, Kheir MM, Shohat N, Tan DD, Kheir M, Chen C, et al. Culture-negative periprosthetic joint infection: an update on what to expect. JB JS Open Access. 2018;3(3):e0060. https://doi.org/10.2106/JBJS.OA.17.00060.
36. Nilsdotter-Augustinsson A, Briheim G, Herder A, Ljunghusen O, Wahlstrom O, Ohman L. Inflammatory response in 85 patients with loosened hip prostheses: a prospective study comparing inflammatory markers in patients with aseptic and septic prosthetic loosening. Acta Orthop. 2007;78(5):629–39. https://doi.org/10.1080/17453670710014329.
37. Grosso MJ, Frangiamore SJ, Ricchetti ET, Bauer TW, Iannotti JP. Sensitivity of frozen section histology for identifying Propionibacterium acnes infections in revision shoulder arthroplasty. J Bone Joint Surg Am. 2014;96(6):442–7. https://doi.org/10.2106/JBJS.M.00258.
38. Pajarinen J, Cenni E, Savarino L, Gomez-Barrena E, Tamaki Y, Takagi M, et al. Profile of toll-like receptor-positive cells in septic and aseptic loosening of total hip arthroplasty implants. J Biomed Mater Res A. 2010;94(1):84–92. https://doi.org/10.1002/jbm.a.32674.

^{18}F FDG-PET/CT has poor diagnostic accuracy in diagnosing shoulder PJI

Supplementary file 1

Table 6 showing results of cultures and infection status according to different infection definitions

						Infection status	Infection status	Infection status	Infection status	
Patient #	Biopsy 1	Biopsy 2	Biopsy 3	Biopsy 4	Biopsy 5	by culture	by MSIS	by Frangiamore et al.	by shoulder MSIS	Score
1	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	5
2	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
3	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
5	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
6	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	8
9	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	5
10	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
11	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
13	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
15	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
19	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	3
21	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
22	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
24	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
26	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	3
27	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	3
28	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
29	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
31	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
35	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
36	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
37	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
40	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	7
41	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
42	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
43	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
47	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
48	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
49	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
54	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
55	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
56	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
57	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
58	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
59	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
60	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
62	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
63	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	3
64	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
65	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
66	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	4
68	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
69	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
72	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
78	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
81	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	3
82	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	3
84	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
86	No growth	No growth	No growth	No growth	No growth	No infection	No infection	No evidence for infection	Unlikely infection	1
7	CNS	No growth	No growth	No growth	No growth	No infection	No infection	Probable Contaminant	Unlikely infection	0
17	C. acnes	No growth	No growth	No growth	No growth	No infection	No infection	Probable Contaminant	Unlikely infection	0
18	C. acnes	No growth	No growth	No growth	No growth	No infection	No infection	Probable Contaminant	Unlikely infection	0
52	C. acnes	No growth	No growth	No growth	No growth	No infection	No infection	Probable Contaminant	Unlikely infection	0
73	S. Saph	No growth	No growth	No growth	No growth	No infection	No infection	Probable Contaminant	Unlikely infection	0
76	C. acnes	No growth	No growth	No growth	No growth	No infection	No infection	Probable Contaminant	Unlikely infection	0
32	C. acnes	No growth	No growth	No growth	No growth	No infection	No infection	Possible infection	Probable infection	8
61	C. acnes	No growth	No growth	No growth	No growth	No infection	No infection	Possible infection	Unlikely infection	2
23	C. acnes	CNS	CNS	No growth	No growth	No infection	Definite infection	Possible infection	Unlikely infection	3
25	CNS	CNS	Bacillus spp.	No growth	No growth	No infection	Definite infection	Possible infection	Possible infection	3
30	CNS	CNS	No growth	No growth	No growth	No infection	Definite infection	Possible infection	Possible infection	5
34	C. acnes	C. acnes	S. capitis	S. capitis	No growth	No infection	Definite infection	Possible infection	Possible infection	3
67	S. waneri	Gordonia spec	No growth	No growth	No growth	No infection	Definite infection	Possible infection	Unlikely infection	0
71	S. epidermidis	S. epidermidis	No growth	No growth	No growth	No infection	Definite infection	Possible infection	Probable infection	8
83	C. acnes	C. acnes	No growth	No growth	No growth	No infection	Definite infection	Definite infection	Probable infection	12
4	S. capitis	S. capitis	S. capitis	No growth	No growth	Definite infection	Definite infection	Possible infection	Possible infection	3
8	CNS	CNS	CNS	CNS	CNS	Definite infection	Definite infection	Possible infection	Possible infection	5
12	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	5
14	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
16	C. acnes	C. acnes	C. acnes	No growth	No growth	Definite infection	Definite infection	Possible infection	Possible infection	3
20	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
33	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
39	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
44	C. acnes	C. acnes	C. acnes	S. epidermidis	S. epidermidis	Definite infection	Definite infection	Possible infection	Possible infection	5
45	C. acnes	C. acnes	C. acnes	C. acnes	No growth	Definite infection	Definite infection	Possible infection	Possible infection	3
46	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Probable infection	6
50	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
51	C. acnes	C. acnes	C. acnes	C. acnes	B. subtillis	Definite infection	Definite infection	Possible infection	Probable infection	6
53	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
70	C. acnes	C. acnes	C. acnes	C. acnes	S. capitis	Definite infection	Definite infection	Possible infection	Possible infection	3
74	C. acnes	C. acnes	C. acnes	Micrococcus	No growth	Definite infection	Definite infection	Possible infection	Probable infection	6
75	C. acnes	C. acnes	C. acnes	C. acnes	C. acnes	Definite infection	Definite infection	Possible infection	Possible infection	3
77	S. hominis	S. hominis	S. hominis	S. hominis	No growth	Definite infection	Definite infection	Possible infection	Possible infection	3
80	C. acnes	C. acnes	C. acnes	No growth	No growth	Definite infection	Definite infection	Possible infection	Possible infection	3
85	S. epidermidis	S. epidermidis	S. epidermidis	No growth	No growth	Definite infection	Definite infection	Possible infection	Possible infection	3
38	C. acnes	C. acnes	C. acnes	Staph. pasteur	No growth	Definite infection	Definite infection	Definite infection	Probable infection	12
79	S. epidermidis	S. epidermidis	S. epidermidis	No growth	No growth	Definite infection	Definite infection	Definite infection	Probable infection	9

B. subtilis=Bacillus subtilis, C. Acnes=Cutibacterium Acnes, CNS=Coagulase Negative Staphylococcus, S. capitis=staphyloccus capitis, S. epidermidis=Staphylococcus epidermidis, S. hominis=Staphylococcus hominis, S. Saph= Staphylococcus Saphropyticus, S. waneri=Staphylococcus Waneri

^{18}F FDG-PET/CT has poor diagnostic accuracy in diagnosing shoulder PJI

Description of different infection definitions

MSIS 2013 Criteria	MSIS shoulder criteria	Periprosthetic Shoulder Infection Criteria Category by Frangiamore et al.
PIJ positive if at least on primary sign exists **OR** at least three secondary signs exist. **Primary signs** A sinus tract communicating with the prosthesis is present. A pathogen is isolated by culture from at least two separate tissue or fluid samples obtained from the affected joint **Secondary signs** Elevated serum C-reactive protein AND erythrocyte sedimentation rate Elevated synovial fluid white blood cell count OR ++ changes on leucocyte esterase test strip Elevated synovial fluid polymorphnuclear neutrophil percentage Positive histological analysis of periprosthetic tissue A single positive culture from biopsy specimen	Meeting one of the following Major criteria is diagnostic of **Definite** periprosthetic shoulder infection: **Major Criteria** A sinus tract communicating with the prosthesis is present. Gross intra-articular pus Two positive cultures with phenotypically-identical virulent organisms **Minor Criteria Weight** Unexpected wound drainage — Weight Single positive tissue culture (virulent organism) — 4 Single positive tissue culture (low-virulence organism) — 3 Second positive tissue culture (identical low-virulence organism) — 1 Humeral loosening — 3 Positive frozen section (5 PMN in at least 5 high-power fields) — 3 Positive pre-operative aspirate culture (low or high-virulence) — 3 Elevated synovial neutrophil percentage (>80%)* — 3 Elevated Synovial WBC count (>3,000 cells / µL)* — 2 Elevated ESR (>30 mm/hr)* — 2 Elevated CRP (>10 mg/L)* — 2 Elevated synovial alpha-defensin — 2 Cloudy fluid — 2 *beyond six weeks from recent surgery ESR=erythrocyte sedimentation rate; CRP=C-reactive protein; PMN=polymorphoneuclear leukocyte; WBC=white blood cell; µL=microliter **EVALUATION SCORING** Weighted values for all positive tests performed as part of the diagnostic evaluation of a failed shoulder arthroplasty are summed (Table 1). **EVALUATION SCORING** Weighted values for all positive tests performed as part of the diagnostic evaluation of a failed shoulder arthroplasty are summed (Table 1). • 6 or greater with identified organism = **probable PJI** • 6 or greater *without* identified organism = **possible PJI** • 6 or less with single positive culture virulent organism = **possible PJI** • two positive cultures low-virulence organism = **possible PJI** • negative cultures or only single positive culture for low virulent organism = **PJI unlikely**	**Definite infection** At least one positive preoperative or intraoperative finding of infection and multiple positive intraoperative cultures with the same organism OR One positive preoperative (aspirate) culture and one positive intraoperative culture with the same organism **Probable infection** At least one positive preoperative or intraoperative finding of infection and one positive intraoperative culture **OR** No preoperative or intraoperative findings of infection and multiple positive intraoperative cultures with the same organism. **Probable contaminant** No preoperative or intraoperative findings of infection and one positive intraoperative culture No preoperative or intraoperative findings of infection and no positive intraoperative cultures *Preoperative or intraoperative findings of infection: preoperative clinical signs (swelling, sinus tract, redness, drainage), positive ESR or CRP value, intraoperative gross findings (purulent drainage, necrosis), positive intraoperative frozen section.

Supplementary file 2:

Results of PET-Diagnose stratified by criterion and center:

Using the shoulder-specific criterium altered the diagnose in 28 patients of ALL scans performed.

After applying the shoulder-specific-assessment, the FDG-PET diagnose changed in 23 patients at Center One, corresponding to 8 changed from negative to positive and 15 from positive to negative. At Center Two, five patients changed FDG-PET diagnose after applying shoulder-specific-assessment, corresponding to two from negative to positive and three from positive to negative.

Table 7: FDG-PET Diagnose using Intial assessment vs. Shoulder-specific-assessment[a]

	Shoulder-specific-assessment		
Best practice	Negative	Positive	
Negative	50	2	52
Positive	26	8	34
	76	10	86

[a]majority diagnose

The diagnostic performance of the FDG-PET varied between the centers, both according to best-practice and the Shoulder-specific-criterium. As can be seen from table below confidence-intervals overlap. Hence, no statistical-significant difference between the centers were observed excluding sensitivity and PPV for Best-practice.

Table 8: Diagnostic performance of PET-CT stratified by center and criterion applied

	Best-practice (95% CI)		Shoulder-specific-criterium (95%CI)	
	Center no. 1	Center no. 2	Center no. 1	Center no. 2
Sensitivity	0.33 (0.09-0.57)	0.0 (0.0-0.0)	0.42 (0.14-0.70)	0.16 (0.00-0.46)
Specificity	0.72 (0.58-0.87)	0.96 (0.88-1.0)	0.74 (0.61-0.87)	0.96 (0.88-1.0)
PPV	0.33 (0.10-0.57)	0 (0.00-0.00)	0.31 (0.08-0.54)	0.50 (0.00-1.00)
NPV	0.67 (0.43-0.91)	1.00 (1.00-1.00)	0.82 (0.70-0.94)	0.83 (0.69-0.97)
Accuracy	0.61 (0.47-0.74)	0.80 (0.66-0.94)	0.67 (0.55-0.80)	0.81 (0.67-0.95)

Unexpected positive cultures after revision shoulder arthroplasty

-does it impact outcome

PAPER III

Unexpected positive cultures after revision shoulder arthroplasty -does it impact outcome?

T Falstie-Jensen M.D.[1], J Lange M.D., PhD[4], H Daugaard M.D., Phd[3], AKB Sørensen, M.D., PhD[3], J Ovesen M.D., Phd[1], K Soeballe M.D., DMse[1] on behalf of the ROSA study group

ROSA study group: BS Olsen, M.D., PhD[3], LC Gormsen M.D., PhD[2], B Zerahn M.D.[5], HVS Johanssen M.D.[1], B Elmengaard M.D., PhD[7], TM Thillemann M.D., PhD[1], L Bolvig M.D.[6]

[1] Orthopedic department, Aarhus University Hospital, Denmark [2] Nuclear medicine department, Aarhus University Hospital, Denmark [3] Orthopedic department, Herlev-Gentofte Hospital, Denmark, Denmark. [4] Orthopedic department, Horsens Regional Hospital, Denmark. [5] Department of Clinical Physiology, Herlev Hospital, [6] Radiologic department, Aarhus University Hospital, Denmark, [7]Regional hospital Silkeborg, Denmark

Corresponding author:

Thomas Falstie-Jensen[1], Orthopedic department, Shoulder and elbow section, Aarhus University Hospital, Palle Juul Jensens Boulevard 99, 8200 Aarhus N, Denmark

e-mail: thomfals@rm.dk

Abstract

Background

Numerous studies have confirmed the high rate of unexpected positive cultures (UPC) after presumably aseptic revisions of shoulder arthroplasties but the impact on outcome is still unclear. The purpose of this prospective study was to compare the patient reported outcome of standard revisions with and without emergence of UPC.

Methods

During a 3-year period we included all patients who were revised for other reasons than suspicion of infection from two orthopedic centers. Five biopsies were obtained from every revision and UPC was defined as growth with the same bacteria in at least 3 biopsy-specimens. Patients were assessed with Oxford Shoulder Score (OSS) and range of motion.

Results

126 patients were included with a median follow-up of 28 months (range 26-32). UPC emerged after 28 revisions (24%) with Cutibacterium acnes accounting for 64% (18/28). At baseline, OSS was 20 in both the culture negative and the UPC group. At follow-up, OSS was 37 in the culture negative group and 35 in the UPC group ($p>0.05$). Similarly, forward elevation increased 46 and 34 degrees in the two groups and external rotation was unchanged ($p>0.05$). Subgroup-analyses stratified by different implant designs showed equal patterns with no statistical differences in outcome.

Conclusion

We could not demonstrate that patients who later developed UPC presented with a lower OSS score at baseline compared to patients who did not develop UPC. Neither did we find a difference in short term outcome after a standard revision regardless of emergence of UPC.

Introduction

Periprosthetic joint infection (PJI) can be detrimental to a well-functioning joint replacement. Any patients undergoing joint replacement surgery are at risk; either of acute infection with sudden onset or a chronic infection with more insidious symptoms. Whichever condition, the treatment can be challenging and often the patient will suffer from permanent disability caused by reduced function and pain.

The incidence of shoulder PJI is estimated to range from 1-5% after primary surgery[1]. However, during the last decade there has been an increased focus on slow-growing and low-virulent bacteria as cause of PJI, thus the incidence might be underestimated. This group of bacteria have triggered the enigma of frequent isolation of bacteria after a clinically non-infected shoulder revision; or even primary surgery[19, 33]. Especially the presence of Cutibacterium acnes (C acnes) has proven a challenge to interpret[22]. Studies have shown that C acnes grown unexpectedly in up to 29% of biopsy specimens obtained during removal of metal hardware or joint replacements; even with no clinical or paraclinical suspicion of infection. [9, 18, 24]. This situation with no pre- or perioperative suspicion of infection, yet growth in cultures, has led to the designation *Unexpected Positive Cultures* (UPC). The interpretation and recommended treatment of revised joint replacements with UPC are still unclear together with the impact on functional outcome [2, 9, 27]. Several studies have examined UPC in relation to revision of shoulder arthroplasties[3, 8-11, 14, 18, 20, 21, 24, 26, 32]. They are in general characterized by being retrospective and evaluate a hard end-point such as re-revision or re-infection. Endpoints such as re-revision and ROM are important aspects to evaluate; but they do not incorporate the patient's perception of the result of the operation. Patient Reported Outcome Measures (PROM) can shed a light on subjective aspects such as pain and changes important to daily function. However, to our knowledge no prospective studies have looked into the aspect of whether UPC in shoulder revision influences patient reported outcome (PRO) scores.

Thus, the purpose of this prospective study was to assess if emergence of UPC impacted the PRO score and outcome after revision of failed shoulder replacements.

Method

We conducted a prospective, nationwide, cohort study from April 2014 to October 2017 including all patients treated with a standard revision of a shoulder replacement. At the time-of-study, the Danish Nation Board of Health had appointed the two participating centers as the only hospitals in Denmark to perform shoulder revisions.

All patients referred with a symptomatic shoulder replacement underwent a standardized outpatient diagnostic program (which included: serology, x-ray, physical examination and aspiration if deemed

relevant by the treating surgeon). A case-by-case shared-decision between the patient and the treating surgeon was made in relation to a planned revision and to the extent of a such.

Information was on symptoms, co-morbidities and ROM were uniformly collected by the surgeons during the outpatient program at the two participating centers. Furthermore, patients completed a paper version of the Danish translation of the Oxford Shoulder Score (OSS). After revision, patients were followed at least two years. At follow-up patients were assessed with x-ray and ROM, and completed the OSS. If a patient did not show up for the scheduled appointment, they were contacted by phone or mail, and invited to a new consultation. If she/he declined or failed to appear again, questionnaires were retrieved by mail or online using RedCap (a secure open-source system for managing research-data and online capture of survey data)[12].

Patients were excluded from the current study if they were diagnosed with PJI before the revision according to the definition forwarded by Musculoskeletal Infection Society (MSIS)[25], had another procedure than single-stage component exchange planned, did not complete two-year follow-up, had a history of ipsilateral chronic PJI or an acute periprosthetic fracture. A flow-chart describing patient enrollment are presented in Figure 1.

Thus, the study-cohort comprised of patients with a symptomatic shoulder joint replacement without suspicion of infection, and who underwent a standard revision with complete component exchange.

The study was registered according to Danish law at the Danish Data protection Agency (ref. no. 1-16-02-567-13). No formal approval was needed according to correspondence to the local research ethical committee (ref. no. 217/2013).

Surgical technique

During surgery a standardized single-stage exchange was performed at both centers. A deltopectoral approach was utilized and any scar-tissue was excised. All cement was removed and relevant soft-tissue debridement was performed followed by irrigation with 1000 ml saline mixed with 2 g Vancomycin and 240 mg Gentamycin. In general, an anatomic total shoulder replacement (TSR) was implanted in patients with intact and functional rotator-cuff and a reverse shoulder arthroplasty (RSA) was implanted in patients with absent or non-functional rotator cuff. A hemi arthroplasty (HA) was chosen in selected cases at surgeon's discretion. If a RSA design was used Gentamycin impregnated fleece was placed behind the glenosphere. If a cemented implant was used, high-viscosity methyl methacrylate cement containing Gentamycin was used. The wound was closed in layers; no drain or pain-catheters were inserted.

Sample acquisition and culturing:

Immediately after exposure of the arthroplasty, five separate biopsy-specimens in close proximity to the bone/cement interface, synovium and from areas with membranes or necrotic tissue were obtained with clean utensils[16]. Antibiotics were withheld until biopsies-specimens were obtained. All specimens

were cultured on blood Agar and Anaerobic Agar plates (SSI, Denmark), BLL Chromagar Orientation Medium (Becton Dickinson, Germany) and inoculated in Semisolid Agar+Pepsis Blood+Thioglycollate (SSI, Denmark) and Serum Broth (SSI, Denmark). Agar plates were inspected daily for growth up to four days. Semisolid Agar and Serum Broth were visually inspected for signs of growth only on day four and day fourteen to decrease risk of contamination. Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight mass spectrometry (MALDI-TOF, Bruker, Germany) was used for identification of bacteria strains.

Definition of culture status

Cultures were defined as positive when growth with identical bacteria was detected in least three of the five biopsy specimens. If growth from less than tree biopsy specimens were present, the cultures were designated negative. Likewise, if more than three biopsy specimens showed growth but with different bacteria; it was considered contamination and cultures defined as negative.

Hence, according to the MSIS definition patients were diagnosed with PJI if cultures were positive. However, due to the purpose of this study and the preoperative lack of suspicion of PJI, we designated these patients as having UPC. If cultures were negative, patients were designated "culture-negative".

Antibiotic treatment

Preoperative antibiotics were withheld until biopsy-specimens were obtained after which an intravenous (IV) dose of either Cefuroxim 1,5 g or 1g Dicloxacillin in combination with 2 MIE penicillin was administered.

Post-operative IV-antibiotics were administered the first 24-hours, and patients were typically discharged after the last administration. Patients were recommended antibiotic treatment covering at least C acnes either oral or IV at surgeon's discretion until final culture-result was available. If cultures showed growth at readings up to the 4th post-operative day, two weeks IV treatment followed by four weeks oral treatment was initiated. If cultures were positive at the reading on the 14th postoperative day, an additional 4-week oral treatment was recommended. Both IV and oral antibiotic treatment was based on the antibiogram of the cultured microorganism and recommendations of the local microbiology department. Antibiotic treatment was discontinued if cultures were negative the 14th postoperative day.

Outcome measures

The Oxford Shoulder Score (OSS) is validated to assess outcome after shoulder surgery due to degenerative disease[7]. Furthermore, it has shown to be sensitive, responsive, translated to several languages and easy to use for both patient and researcher[5]. It consists of 12 items; 4 representing a dimension of pain and 8 a dimension of function. One item can be awarded zero to 4 points and the total OSS score range from 0 to 48 with higher score representing better outcome. A validated Danish version of OSS obtained from Oxford University Innovation, Oxford, UK was utilized.

We measured the range of forward elevation and external rotation using a manual goniometer with 10 degrees intervals. Patients were asked to forward elevate the shoulder in the scapular plane to the maximal pain-tolerable height with a free scapula, the elbow in full extension (range 0-180°). External rotation was measured with a free scapula, the elbow flexed 90° and placed to the side and rotated to the maximal pain-tolerable position. Zero degrees are represented by the sagittal plane (range -90°-90°).

Statistics

OSS questionnaires were checked for completeness. If more than one field in an item was ticked off, the lowest scoring response was adopted. One or two missing values were handled as proposed by Dawson et al. Hence, missing values were imputed by a mean of the patient's responses[7]. If more than two questions were missing, the questionnaire was omitted and the patient was excluded from the study. Perspective of pain was analyzed using four items in the OSS; number 1, 8, 11 and 12 which all describe different aspects of pain[6]; the sum of the four items represented the pain score.

Cutoff values for categories of outcome e.g. excellent, good or poor has not been established for the OSS. Consequently, we report the effect size (ES) to assess the magnitude of change between the groups[17]. An ES of 0.2, 0.5 and 0.8 is generally accepted as a small, medium and large change.

Further, potential issues with responsiveness of the OSS were evaluated by estimating the occurrence of ceiling- or floor effect. Ceiling- and floor effect were defined as 15% or more patients scoring the highest or lowest achievable score.

To determine the outcome of revision, the difference from baseline to 2-year follow-up in OSS, external rotation and elevation were assessed. Mean-difference was calculated as mean of differences between baseline and 2-year follow-up.

To evaluate if UPC affected the patients' preoperative and postoperative status, the mean OSS, mean elevation and mean external rotation were compare at each time-point.

Normality was checked by plotting data and using Shapiro-Wilks test. Parametric data are presented with means and 95% CI. Non-parametric data are presented as medians with 25-75% interquartile range (IQR).

Means at same timepoints were compared between groups with two-way analysis-of-variance (ANOVA) when applicable. Levene's test was used for testing equality of variances prior to ANOVA testing. Post-hoc analyses were performed if main- or interaction effect were demonstrated. However, OSS at follow-up (in contrast to OSS at baseline) were not normal-distributed; hereby not allowing analysis with ANOVA. Hence, comparisons of OSS at follow-up were conducted pairwise with Mann-Whitney two sample test knowing the risk of type 1-error was increased due to multiple testing. Means of parametric

data were compared using unpaired t-test; Mann-Whitney two-sample test were used for non-parametric data.

Binomial data was compared with Chi^2-test. In all statistical tests level of significance was set to 0.05. Data was analyzed using STATA version 15, StataCorp LLC, Texas, USA.

Study population

During the study period, 394 patients were referred with an unsatisfactory result after a shoulder replacement. Two-hundred sixty-eight (68%) of the referred patients were excluded; 11 (3%) due to previous treatment for chronic PJI, non-operative approach was chosen for 141 (36%) patients and 57 (14%) received other operative treatment than single-stage component exchange, 24 (6%) patients had missing information in their case-report forms, five (1%) patients did not return for follow-up visit despite several reminders, and 11 (3%) died from causes unrelated to the shoulder-surgery before completing two-year follow-up. Thus, the study-cohort comprised of 126 patients undergoing a standard single-stage revision (Figure 1, flow-chart of patient selection).

The median follow-up was 28 months IQR (26-32). The study population consisted of 61% (77/126) females and 39% (49/126) males with a mean age of 67 years (95%CI, 65-68). The median age of the revised implant was 4 years IQR (1-6). The demographics of the study-cohort patients are presented in table I.

Main findings

Unexpected positive cultures were present in 28 patients (23%). The three most frequent isolated bacteria were C acnes in 64% (18/28), Staphylococcus epidermidis in 11% (3/28) patients and Coagulase negative staphylococcus in 7% (2/28) patients. In one case the cultures were reported positive in the case report form without specification of microorganism in either the medical record or the case report form. A graph of frequency and type of all cultured bacteria can be found in Figure 2. Both distribution of sex and indication of primary surgery differed significantly between the culture-negative and UPC groups. More men had UPC and more patients with arthritis were found in the UPC groups. Differences between the culture-negative and UPC groups regarding mean age, mean age of implant, type of implant and mean of inflammatory markers were not statistically significant

The OSS increased 14 (95%CI 11-16) points in the culture negative group and 13 (95%CI 8-17) points in the UPC group (Table II). Similarly, significant increase in OSS were found in subgroups regardless of culture-status or implant type. Effect-sizes were all >1.2 which indicates a large change in outcome in this cohort. The largest ES was found for the TSR patients. The mean forward elevation increased significant regardless of culture-result for patients receiving a RSA and for culture-negative TSR patients (table III). Although not statistically significant, an increase in forward elevation were likewise found for the TSR patients with UPC and all HA patients (table III). None of the groups showed a significant change in external rotation (table III).

The mean pain score increased (higher score=less pain) from 4 (95%CI, 3-5) to 10 (95%CI, 9-11) in the culture-negative group and from 4 (95%CI, 4-5) to 10 (95%CI, 8-12) in UPC group. Comparable increases were found all groups regardless of culture-status and implant type. A Forrest plot showing mean pain score and 95%CI illustrate this increased pain score across all patient groups (Figure 3).

To assess if UPC had an impact in the level of OSS or ROM a comparison at baseline and 2-year follow-up were conducted. The culture-negative group had a mean baseline OSS of 20 (95%CI, 18-22) compared to the UPC-group with a mean OSS of 20 (95%CI, 17-23). At 2-year follow-up, the culture-negative group had a median OSS of 37 (IQR, 26-44) compared to 35 (IQR, 24-45) in the UPC group ($p>0.05$). Similarly, at baseline the culture-negative group had a mean elevation of 79° (95%CI, 69-88) and mean external rotation of 32° (95%CI, 27-36) compared to the UPC group with 80° (95%CI, 63-98) mean elevation and 38° (95%CI, 31-44) mean external rotation in. At 2-year follow-up the culture-negative group had a mean forward elevation of 125° (95%CI, 116-134) and 34° (95%CI, 29-40) external rotation. The corresponding values in the UPC group were 114° (95%CI, 94-135) and 38° (95%CI, 28-47). Hence, no significant differences in OSS or ROM were found between the groups at baseline or 2-year follow-up. Similar analyses stratified by design of implanted prosthesis showed a comparable pattern, with non-statically significant differences between most parameters (table II & III). However, at 2-year follow-up the median OSS was significantly higher in the UPC HA group compared to the culture-negative group. At both baseline and 2-year follow-up the UPC HA group had significant better elevation and external rotation compared to the culture-negative group (Table III).

No patients scored lowest achievable OSS score and 9 (7%) patients scored the highest achievable score. Consequently, no ceiling or floor effect were detected.

Discussion

This study shows that patients in general gain function and experience decreased level of pain after a standardized single-stage revision regardless of emergence of UPC. Likewise, an increase in mean forward elevation was demonstrated in all groups; though not all subgroups reached statistical significance. In contrast, external rotation remained unchanged. Only two parameters were affected negatively by presence of UPC; the post-operative forward elevation in the RSA group and the preoperative OSS in the TSR group. To our best judgement, the data does not demonstrate a pattern in OSS or ROM which could suggest a negative impact of UPC, neither pre- nor postoperative.

Our data shows that minor differences can be found between the culture-negative and UPC groups at baseline and 2 years after revision. We did observe a higher OSS and external rotation in the HA UPC group compared to the culture-negative group at 2-years follow-up. However, this is more likely due to the small sample size ($n<5$) than a true positive effect of UPC. All implant groups showed increase in mean OSS from baseline to two-year follow-up. The largest improvement of OSS was found in the TSR group (23 points) which could be expected since the majority of these patients were uncomplicated revisions of resurfacing hemiarthroplasties due to glenoid attrition and with intact rotator cuff. It was more surprising, that UPC TSR patients reported the highest gain in OSS. Since no logic argument

explains this, it could be attributed to small sample size. Alternatively, the finding could support the hypothesis of Jacobsen et al., who demonstrated that implants can provide a niche for a non-pathologic microbiome and consequently UPC could reflect this situation [15].

A minimal clinically important difference (MCID) for OSS has not been established for shoulder replacements, but have been estimated between 5 and 17 in cohorts of patients treated for subacromial impingement syndrome or rotator cuff tear[4]. An approximation can be calculated as a half standard deviation of the change[7]. In our study-cohort this corresponds to a MCID of approximately 6 points. Similarly, the MCID for pain can be estimated to 2. Thus, the reduction of pain and the overall functional gain are both clinically relevant to all the patients. However, it should be noted that the vast majority of patients experience some degree of chronic pain after revision surgery. This is important to communicate to the patient when the indication and expectations to the results of a revision are discussed.

Although, we did not detect a ceiling-effect 7% (9/126) of the patients reached the maximum OSS score. It implies that the OSS cannot detect further clinical improvement for these patients. This is surprisingly high share of patients compared to empiric observations from the outpatient clinic. Ceiling effect is not unknown to the OSS and The National Joint Registry of England, Wales, Northern Ireland and Isle of Man reported in 2016 (NJR 2016) a pronounced ceiling-effect after primary joint replacement[23]. Interestingly, they reported preoperative median OSS of 16 and a median mean increase in 18 points after elective primary shoulder replacements. Despite our cohort selectively incorporate revision cases our results with a mean baseline of 20 and increase of 14 points are surprisingly similar to the NRJ 2016 data. However, it should be noted that in our material 11% (14/126) patients reported a lower OSS postoperative than preoperative in comparison to 8% after primary surgery in the NRJ 2016 data.

In our cohort, we saw a significant increase in both OSS and mean forward elevation regardless of culture result in most groups. Despite shorter follow-up, our results are in line with Hsu et al. who reported on a group of 55 patients with UPC of C acnes only[14]. They concluded that no difference in patient reported outcome of pain and function could be demonstrated after four-years follow up.

Studies reporting outcome of aseptic revisions present results concordant to ours. Hartel et al reported outcome measured by OSS in 19 patients revised to a TSR or RSA with a mean of 41 months follow-up[13]. The TSR group had a postoperative OSS score of 42 versus 28 for the RSA group, outcomes which is almost identical to our results. Sheth et al reported on 28 patients who underwent conversion of HA to a TSR due to progressive glenoid arthrosis with a mean 5.1-year follow-up[30]. Although no preoperative score was available, the mean American Shoulder and Elbow Surgeon score (ASES score 0-100 points) was 78 and patients reported a satisfactory result in 81% of the cases. Sevelda et al reported in 2018 results of 14 one-stage revisions with a mean follow-up of 5 years (range 2-10 years) and a preoperative identified microorganism (9 C acnes and 5 Staph. epidermidis)[29]. An overall improvement of mean Constant score was found with a trend towards TSR doing better than RSA. The mean forward elevation at last follow-up was 99° and 82° for TSR and RSA, respectively. Our results are in line with this and the comparison to one-stage revisions seem fair, since a majority of our patients

were treated with implant removal, meticulous cement removal prior to new implantation followed by at least two weeks treatment with antibiotics. Ortmaier et al retrospectively investigated the clinical outcome of 50 patients with a HA revised to a RSA. Significant improvement in Constant score from 18.5 to 49.3 and forward elevation from 47° to 98° were reported. The patients investigated included those treated for PJI and periprosthetic fracture. Our results show similar improvements in mean forward elevation and OSS regardless presence of UPC.

Our study has some limitations. First, we do not have a control group. Nearly 50% of the patients received two weeks of antibiotics and 90% of the culture-positive patients were treated for six weeks with the majority receiving either Penicillin or Amoxicillin (due to C acnes susceptibility of these drugs). Consequently, our UPC patients were de facto treated with one-stage PJI revisions. However, the efficacy of this oral therapy might be questioned. The majority of the UPC were caused by biofilm forming bacteria and few patients received Rifampicin, which is generally recommended in bone and joint infections due to the increased penetration of biofilm and bone-penetration[28].

Second, this study only reports on short-term outcome. Despite two patients were treated for subsequent infection (both were culture-negative at the first revision) it does not necessarily mean they were the only patients suffering from a reinfection or new low-grade PJI. Surgeons performing revision surgery are often too familiar with a moderate outcome, and acceptance of some degree of pain and limited ROM may lead to hesitation of further re-revisions. Since low-grade PJI often requires tissue biopsies obtained arthroscopically, open surgery or even a new revision to be diagnosed, it could be argued that some surgeon-delay probably will occur[31].

Third, stratification lead to small subgroups. Consequently, confidence intervals are wide and care should be taken when interpreting the result. Point estimates e.g. regarding mean difference may indicate a change above the MCID, however, when CIs are considered the change may also correspond to a change not clinically relevant in some groups. Nevertheless, having this in mind, we find it informative for the reader to assess trends in outcome of the different groups.

Fourth, we did try to analyze the limited information we had on the 39 patients excluded due to missing information in the case report form. A comparison of sex, age, distribution of primary implant design and UPC-rate to the study cohort showed significant differences in distribution of primary implant design and UPC-rate compared to the study cohort. A higher UPC rate (42%) was found among the excluded patients and more resurfacing, TSR and HA and less RSA were present in the study cohort compared to the excluded patients. We do not have an explanation for this difference but it could indicate a potential selection bias.

Conclusion

In conclusion, emergence of UPC does not impact short term outcome of a standard revision. Patients can expect an increase in function and a significant reduction of pain, a gain in elevation and unchanged external rotation after revision.

Since we, in practice, performed a one-stage PJI revision and treated our patients with antibiotics postoperatively the data cannot answer the enigma if UPC represents true infection, colonization or

contamination. Combining evidence presented in this study and by other authors, we find a future prospective study with a true control-group consisting of patients with UPC receiving no antibiotic treatment both ethical and feasible.

Reference

1 Bonnevialle N, Dauzeres F, Toulemonde J, Elia F, Laffosse JM, Mansat P. Periprosthetic shoulder infection: an overview. EFORT Open Rev 2017;2:104-109. 10.1302/2058-5241.2.160023

2 Both A, Klatte TO, Lubke A, Buttner H, Hartel MJ, Grossterlinden LG et al. Growth of Cutibacterium acnes is common on osteosynthesis material of the shoulder in patients without signs of infection. Acta Orthop 2018;89:580-584. 10.1080/17453674.2018.1489095

3 Cheung EV, Sperling JW, Cofield RH. Revision shoulder arthroplasty for glenoid component loosening. J Shoulder Elbow Surg 2008;17:371-375. 10.1016/j.jse.2007.09.003

4 Christiansen DH, Frost P, Falla D, Haahr JP, Frich LH, Svendsen SW. Responsiveness and Minimal Clinically Important Change: A Comparison Between 2 Shoulder Outcome Measures. J Orthop Sports Phys Ther 2015;45:620-625. 10.2519/jospt.2015.5760

5 Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. The Journal of bone and joint surgeryBritish volume 1996;78:593-600.

6 Dawson J, Harris KK, Doll H, Fitzpatrick R, 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;7:195-203. 10.2147/PROM.S115488

7 Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. Arch Orthop Trauma Surg 2009;129:119-123. 10.1007/s00402-007-0549-7

8 Dodson CC, Craig EV, Cordasco FA, Dines DM, Dines JS, Dicarlo E et al. Propionibacterium acnes infection after shoulder arthroplasty: a diagnostic challenge. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2010;19:303-307. 10.1016/j.jse.2009.07.065 [doi]

9 Foruria AM, Fox TJ, Sperling JW, Cofield RH. Clinical meaning of unexpected positive cultures (UPC) in revision shoulder arthroplasty. Journal of shoulder and elbow surgery 2013;22:620-627. 10.1016/j.jse.2012.07.017

10 Frangiamore SJ, Saleh A, Grosso MJ, Alolabi B, Bauer TW, Iannotti JP et al. Early Versus Late Culture Growth of Propionibacterium acnes in Revision Shoulder Arthroplasty. The Journal of bone and joint surgery American volume 2015;97:1149-1158. 10.2106/JBJS.N.00881

11 Grosso MJ, Sabesan VJ, Ho JC, Ricchetti ET, Iannotti JP. Reinfection rates after 1-stage revision shoulder arthroplasty for patients with unexpected positive intraoperative cultures. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2012;21:754-758. 10.1016/j.jse.2011.08.052; 10.1016/j.jse.2011.08.052

12 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377-381. 10.1016/j.jbi.2008.08.010

13 Hartel BP, Alta TD, Sewnath ME, Willems WJ. Difference in clinical outcome between total shoulder arthroplasty and reverse shoulder arthroplasty used in hemiarthroplasty revision surgery. Int J Shoulder Surg 2015;9:69-73. 10.4103/0973-6042.161426

14 Hsu JE, Gorbaty JD, Whitney IJ, Matsen FA, 3rd. Single-Stage Revision Is Effective for Failed Shoulder Arthroplasty with Positive Cultures for Propionibacterium. The Journal of bone and joint surgeryAmerican volume 2016;98:2047-2051. 10.2106/JBJS.16.00149 [doi]

15 Jakobsen TH, Eickhardt SR, Gheorghe AG, Stenqvist C, Sonderholm M, Stavnsberg C et al. Implants induce a new niche for microbiomes. APMIS 2018;126:685-692. 10.1111/apm.12862

16 Kamme C, Lindberg L. Aerobic and anaerobic bacteria in deep infections after total hip arthroplasty: differential diagnosis between infectious and non-infectious loosening. Clinical orthopaedics and related research 1981;(154):201-207.

17 Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:S178-189.

18 Kelly JD, 2nd, Hobgood ER. Positive culture rate in revision shoulder arthroplasty. Clin Orthop Relat Res 2009;467:2343-2348. 10.1007/s11999-009-0875-x

19 Levy O, Iyer S, Atoun E, Peter N, Hous N, Cash D et al. Propionibacterium acnes: an underestimated etiology in the pathogenesis of osteoarthritis? Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2013;22:505-511. 10.1016/j.jse.2012.07.007; 10.1016/j.jse.2012.07.007

20 Lorenzetti AJ, Wongworawat MD, Jobe CM, Phipatanakul WP. Cyanoacrylate microbial sealant may reduce the prevalence of positive cultures in revision shoulder arthroplasty. Clin Orthop Relat Res 2013;471:3225-3229. 10.1007/s11999-013-2854-5

21 Lucas RM, Hsu JE, Whitney IJ, Wasserburger J, Matsen FA, 3rd. Loose glenoid components in revision shoulder arthroplasty: is there an association with positive cultures? J Shoulder Elbow Surg 2016;25:1371-1375. 10.1016/j.jse.2015.12.026

22 Moroder P, Trampuz A, Scheibel M. Propionibacterium: We Found It, Now We Have to Deal with It: Commentary on an article by Jason E. Hsu, MD, et al.: "Single-Stage Revision Is Effective for Failed Shoulder Arthroplasty with Positive Cultures for Propionibacterium". The Journal of bone and joint surgeryAmerican volume 2016;98:e112. 10.2106/JBJS.16.00838 [doi]

23 National Joint Registry for England W, Northern Ireland and Isle of Man. 13th Annual Report. 2016.

24 Padegimas EM, Lawrence C, Narzikul AC, Zmistowski BM, Abboud JA, Williams GR et al. Future surgery after revision shoulder arthroplasty: the impact of unexpected positive cultures. Journal of shoulder and elbow surgery 2017;26:975-981. S1058-2746(16)30568-7 [pii]

25 Parvizi J, Zmistowski B, Berbari EF, Bauer TW, Springer BD, Della Valle CJ et al. New definition for periprosthetic joint infection: from the Workgroup of the Musculoskeletal Infection Society.

Clinical orthopaedics and related research 2011;469:2992-2994. 10.1007/s11999-011-2102-9 [doi]

26 Pottinger P, Butler-Wu S, Neradilek MB, Merritt A, Bertelsen A, Jette JL et al. Prognostic factors for bacterial cultures positive for Propionibacterium acnes and other organisms in a large series of revision shoulder arthroplasties performed for stiffness, pain, or loosening. J Bone Joint Surg Am 2012;94:2075-2083. 10.2106/JBJS.K.00861

27 Renz N, Mudrovcic S, Perka C, Trampuz A. Orthopedic implant-associated infections caused by Cutibacterium spp. - A remaining diagnostic challenge. PLoS One 2018;13:e0202639. 10.1371/journal.pone.0202639

28 Seaton RA, Ritchie ND, Robb F, Stewart L, White B, Vallance C. From 'OPAT' to 'COpAT': implications of the OVIVA study for ambulatory management of bone and joint infection. J Antimicrob Chemother 2019. 10.1093/jac/dkz122

29 Sevelda F, Fink B. One-stage exchange of septic shoulder arthroplasty following a standardized treatment algorithm. J Shoulder Elbow Surg 2018. 10.1016/j.jse.2018.06.004

30 Sheth MM, Sholder D, Abboud J, Lazarus MD, Ramsey ML, Williams GR et al. Revision of failed hemiarthroplasty for painful glenoid arthrosis to anatomic total shoulder arthroplasty. J Shoulder Elbow Surg 2018;27:1884-1890. 10.1016/j.jse.2018.03.018

31 Tashjian RZ, Granger EK, Zhang Y. Utility of prerevision tissue biopsy sample to predict revision shoulder arthroplasty culture results in at-risk patients. Journal of shoulder and elbow surgery 2017;26:197-203. 10.1016/j.jse.2016.07.019

32 Topolski MS, Chin PY, Sperling JW, Cofield RH. Revision shoulder arthroplasty with positive intraoperative cultures: the value of preoperative studies and intraoperative histology. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2006;15:402-406. S1058-2746(05)00311-3 [pii]

33 Wong JC, Schoch BS, Lee BK, Sholder D, Nicholson T, Namdari S et al. Culture positivity in primary total shoulder arthroplasty. J Shoulder Elbow Surg 2018;27:1422-1428. 10.1016/j.jse.2018.05.024

Table I: Baseline demographics				
		Study-cohort	Patients with negative cultures	Patients with UPC
No. of patients		126	98 (77%)	28 (23%)
Mean Age (years, range)		67 (29-85)	67 (29-85)	68 (44-81)
Sex	Female	77 (61%)	70 (71%)	7 (25%) *
	Male	49 (39%)	28 (29%)	21 (75%) *
Prosthesis age (years, IQR)		3 (1-6)	4 (1-7)	5 (1-6)
Primary arthroplasty				
Hemiarthroplasty, with stem		66 (52%)	51 (52%)	15 (54%)
Hemiarthroplasty, resurfacing		47 (37%)	36 (37%)	11 (39%)
Total Shoulder replacement		9 (7%)	8 (8%)	1 (4%)
Reverse prosthesis		4 (3%)	3 (3%)	1 (4%)
Indication for primary arthroplasty				
Fracture/fracture seq.		40 (32%)	31 (32 %)	9 (32 %)*
Arthrosis		67 (53%)	49 (50%)	18 (64 %)*
Rheumatoid arthritis		6 (5%)	6 (6%)	0 (0 %)*
Other/unknown		13 (10%)	12 (12%)	1 (4 %)*
C-reactive protein		3.3	3.5	2.7
(mg/l, reference <8)		(95% CI:2.5-4.1)	(95% CI:2.5-4.4)	(95% CI:0.9-4.4)
White blood cell count		7.3	7.3	7.2
(10^9/l, reference <10.0)		(95% CI:7.0-7.6)	(95% CI:7.0-7.6)	(95% CI:6.3-8.1)
Erythrocyte sedimentation ratio		12	13	9
(mm/hour, reference <30)		(95% CI:10-14)	(95% CI:10-16)	(95% CI:6-11)

*=statistical difference compared to culture negative group (p>0.05)

Table II: Outcome after revision measured by OSS

			Oxford Shoulder Score					
	Infection status	Patients n=126	Baseline		2-year follow-up***	Mean difference	Effect size	SD Baseline
All patients			**Mean (95% CI)**	**Median (IQR)**	**Median (IQR)**	**(95% CI)**		
	Negative cultures	98	20 (18-22)	22 (13-26)	37 (26-44)	14 (11-16) **	1.7	8.14
	UPC	28	20 (17-23)	22 (16-24)	35 (24-45)	13 (8-17) **	1.8	7.44
Patients by final joint implant								
Reverse	Negative cultures	67	18 (16-20)	17 (12-24)	29 (23-34)	15 (11-18) **	1.9	7.81
	UPC	18	19 (15-23)	21 (15-25)	26 (21-40)	9 (3-15) **	1.2	7.73
TSR	Negative cultures	21	27 (25-29) x	28 (24-31)	31 (43-35)	11 (8-15) **	2.3	4.73
	UPC	4	20 (14-25) x	20 (17-23)	45 (37-48)	23 (11-34) **	6.6	3.51
Hemi	Negative cultures	10	18 (12-23)	16 (16-25)	29 (23-34) x	12 (3-21) **	1.6	7.71
	UPC	6	24 (16-32)	24 (21-26)	45 (38-46) x	17 (3-30) **	2.2	7.78

x =significant difference between culture negative and UPC group, **=significant difference between baseline and 2-year follow up

***Due to non-parametric distribution of data, scores are only presented as median and IQR (25^{th} and 75^{th} percentile)

Table III: Outcome after revision measured by elevation and external rotation

			Mean forward elevation				Mean external rotation			
		Patients n=103	Baseline	2-year follow-up	Difference	Patients n=74	Baseline	2-year follow-up	Mean difference
All patients			**Mean (95% CI)**	**Mean (95% CI)**	**Mean (95% CI)**		**Mean (95% CI)**	**Mean (95% CI)**	**(95% CI)**
	Negative cultures	79	79 (69-88)	125 (116-134)	46 (37-55)	59	32 (27-36)	34 (29-40)	3 (-3-8)
	UPC	24	80 (63-98)	114 (94-135)	34 (16-52)	15	38 (31-44)	38 (28-47)	1 (-10-12)
Patients by final joint implant									
Reverse	Negative cultures	50	66 (58-75)	124 (114-134)	58 (48-68)	36	30 (24-38)	30 (23-36)	-1 (-7-5)
	UPC	14	56 (36-76)	101 (73-128)	45 (20-69)	7	38 (28-47)	41 (21-60)	3 (-20-26)
TSR	Negative cultures	19	121 (101-141)	148 (131-164)	27 (4-48)	18	36 (30-42)	45 (36-56)	10 (-3-23)
	UPC	4	128 (106-139)	144 (116-199)	16 (-35-68)	4	33 (16-50)	28 (6-50)	-5 (-29-20)
Hemi	Negative cultures	10	61 (32-89)	85 (60-110)	25 (0-49)	5	25 (12-37)*	26 (9-43)*	1 (-24-26)
	UPC	6	74 (53-96)	122 (63-181)	25 (-32-82)	4	43 (26-59)*	45 (45-45)*	2.5 (-14-19)

*=significant difference between culture negative and UPC group

Figure 1: flowchart of patient inclusion

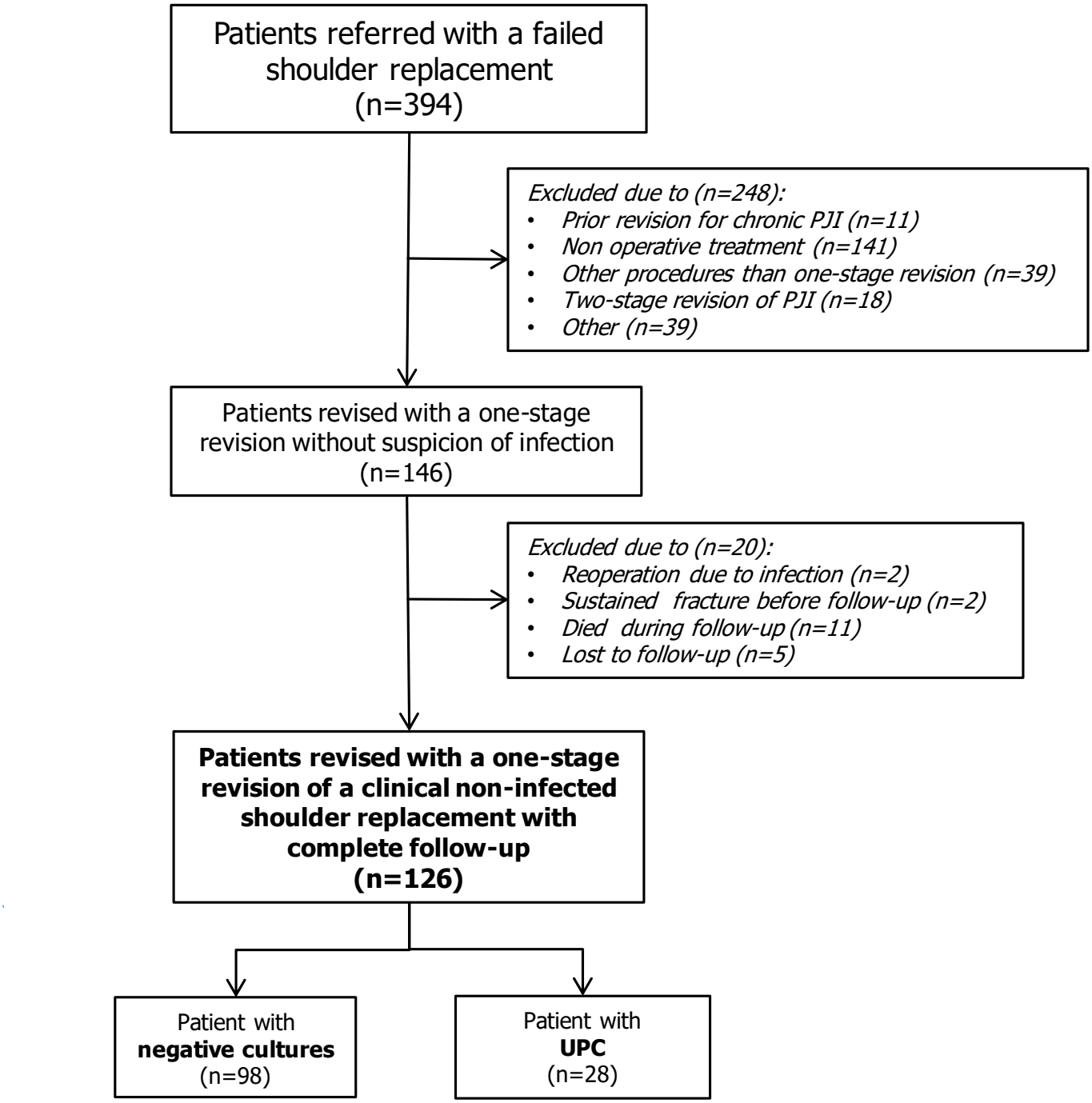


Figure 2: Isolated bacteria from revision

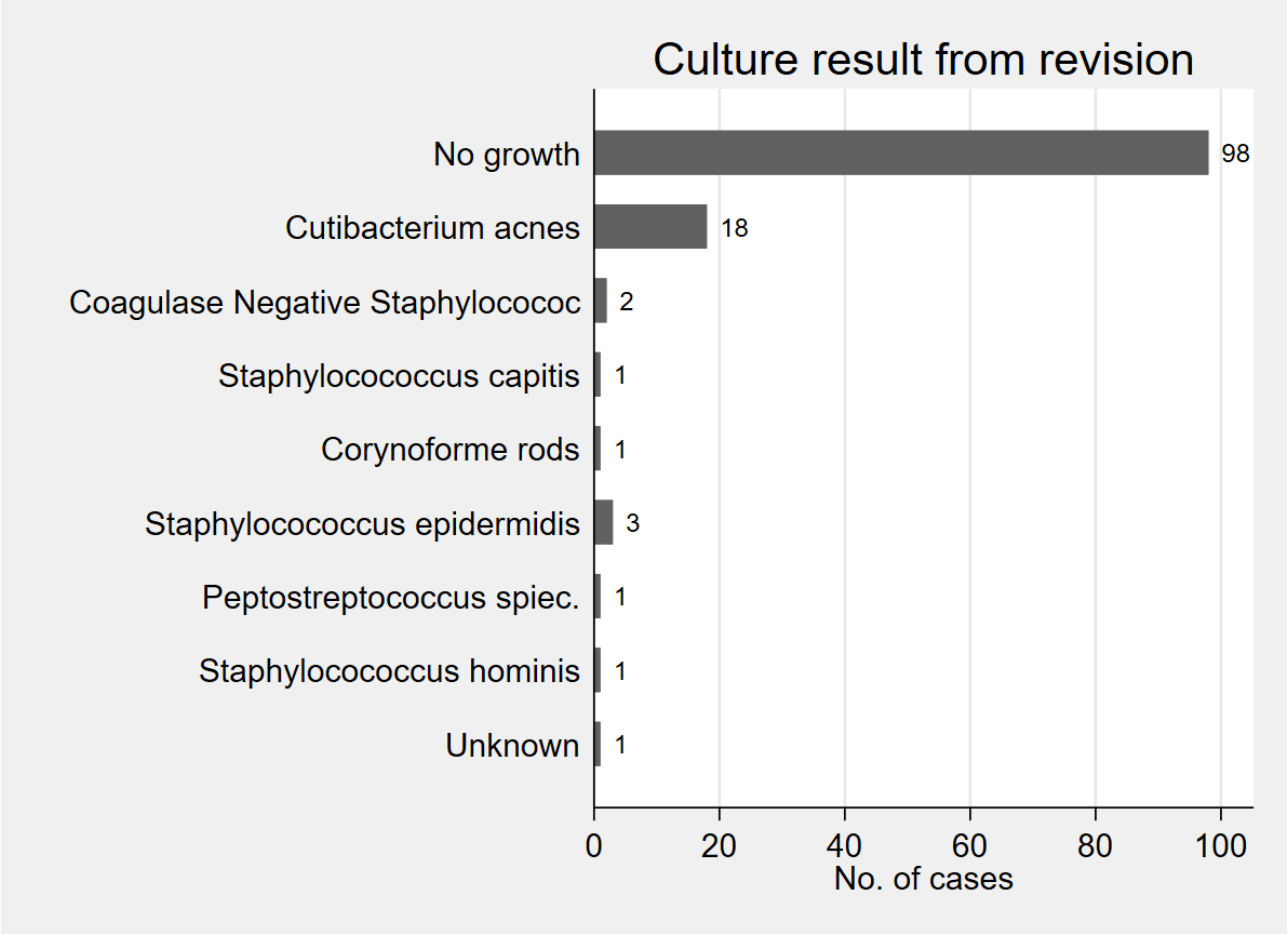


Figure 3: Forrest plot of pain-score stratified by culture-result and design of final implant

Higher score=less pain, whiskers=95% Confidence interval.