WASHINGTON STATE HEALTH CARE AUTHORITY

Hip Surgery Procedures for Treatment of Femoroacetabular Impingement Syndrome

Health Technology Assessment

Friday, August 26, 2011

Health Technology Assessment Program
676 Woodland Square Loop SE
P.O. Box 42712
Olympia, WA 98504-2712
http://www.hta.hca.wa.gov

Hip Surgery Procedures for Treatment of Femoroacetabular Impingement Syndrome

Provided by:

Spectrum Research, Inc.

Prepared by:

Joseph R. Dettori, PhD, MPH
Robin Hashimoto, PhD
Erika Brodt, BS

This technology assessment report is based on research conducted by a contracted technology assessment center, with updates as contracted by the Washington State Health Care Authority. This report is an independent assessment of the technology question(s) described based on accepted methodological principles. The findings and conclusions contained herein are those of the investigators and authors who are responsible for the content. These findings and conclusions may not necessarily represent the views of the HCA/Agency and thus, no statement in this report shall be construed as an official position or policy of the HCA/Agency.

The information in this assessment is intended to assist health care decision makers, clinicians, patients and policy makers in making sound evidence-based decisions that may improve the quality and cost-effectiveness of health care services. Information in this report is not a substitute for sound clinical judgment. Those making decisions regarding the provision of health care services should consider this report in a manner similar to any other medical reference, integrating the information with all other pertinent information to make decisions within the context of individual patient circumstances and resource availability.

TABLE OF CONTENTS

TABLES

FIGURES

APPENDICES

EXECUTIVE SUMMARY

Introduction

Femoroacetabular impingement (FAI) syndrome is a recently recognized diagnosis in primarily younger individuals where relatively minor abnormalities in the joint (orientation or morphology) are thought to cause friction/impingement and pain. It is theorized that FAI starts the breakdown of cartilage, leading to osteoarthritis. There are two types of FAI: cam impingement (non-spherical femoral head or abnormality at the head-neck junction) and pincer impingement (deep or retroverted acetabulum resulting in overcoverage of the femoral head). Proponents believe that surgical correction of the impinging deformities will alleviate the symptoms and retard the progression of OA degeneration.

Surgery to correct FAI includes arthroscopy, open dislocation of the hip and arthroscopy combined with a mini-open approach. The purpose of the surgery is to remove abnormal outgrowths of bone and damaged cartilage, and to reshape the femoral neck to ensure that there is sufficient clearance between the rim of the acetabulum and the neck of the femur.

The causes of hip pain, the natural history of FAI and its relationship to osteoarthritis are unclear, and the case definition and selection criterion of patients for this procedure is uncertain. Furthermore, questions remain about the efficacy and effectiveness, safety and cost effectiveness of hip surgery for FAI. Therefore, this health technology assessment set out to answer the following key questions:

Key question 1

Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?

Key question 2

What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Has clinically meaningful improvement in outcomes been defined for FAI?

Key question 3

What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI? Including consideration of short-term and long-term:

- Need of or time to total hip arthroplasty
- Development or progression of osteoarthritis
- Impact on function, pain, range of motion, quality of life, activities of daily living and return to work

- Other reported measures

Key question 4
What is the evidence of the safety of hip surgery for FAI compared with no surgery? Including consideration of:

- Revision/re-operation rates
- Adverse events type and frequency (peri-operative, fractures, nerve damage, mortality, other major morbidity)

Key question 5
What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations? Including consideration of:

- Gender
- Age
- Psychological or psychosocial comorbidities
- Baseline functional status: e.g. type of deformity, extent of osteoarthritis or cartilage damage
- Other patient characteristics or evidence-based patient selection criteria
- Provider type, setting or other provider characteristics
- Payer/beneficiary type: including worker's compensation, Medicaid, state employees

Key question 6
What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI? Including consideration of:

- Costs (direct and indirect) and cost effectiveness
- Short-term and long-term

Methods for evaluation comparative effectiveness

We conducted a formal, structured systematic search of the peer-reviewed literature across a number of databases in addition to searches of pertinent databases related to clinical guidelines and previously performed assessments. Pertinent studies were critically appraised using our Level of Evidence (LoE) system which evaluates the methodological quality based on study design as well as factors which may bias studies. An overall Strength of Evidence combines the LoE with consideration of the number of studies and the consistency of the findings to describe an overall confidence regarding the stability of estimates as further research is available. Included economic studies were also formally appraised based on criteria for quality of economic studies and pertinent epidemiological precepts.

We selected articles to summarize based on the inclusion and exclusion criteria in the following table:

Study Component	Inclusion	Exclusion
Population	• Patients undergoing treatment for FAI	• Congenital hip dysplasia, slipped capital femoral epiphysis, Legg-Calve-Perthes
Intervention	• Operative treatment for FAI (open, arthroscopic, or combination)	
Comparator	• Nonoperative care (activity modification, NSAIDs, injections, etc)	
Outcomes	Short-term: • Functional outcome (patient- and clinician-reported hip scores) • Pain • Range of motion • Return to work • Complications/adverse events (safety) • Reoperation (safety) Long-term: • Conversion to THA • Function • Pain • Range of motion	• Non-clinical outcomes
Study Design	• Prospective studies listing inclusion criteria, and reliability/validity studies for KQ1 • Reliability/validity studies for KQ2 • Comparative studies and if need be, case series for questions 3-5 and case reports for safety. • Formal economic studies for question 6	• Case reports (except for KQ 4, safety) • Non-clinical studies
Publication	• Studies published in English in peer reviewed journals, published HTAs or publically available FDA reports • Full formal economic analyses (e.g. cost-utility studies) published in English in a HTA or in a peer-reviewed journal published after those represented in previous HTAs	• Abstracts, editorials, letters • Duplicate publications of the same study which do not report on different outcomes • Single reports from multicenter trials • Studies reporting on the technical aspects of FAI surgery • White papers • Narrative reviews • Articles identified as preliminary reports when results are published in later versions • Incomplete economic evaluations such as costing studies

Results

KEY QUESTION 1
Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?

To answer this key question, we first sought to identify and compare the inclusion criteria from all prospective studies evaluating the effectiveness of treatment for FAI. Inclusion and exclusion criteria of a clinical trial define the population of interest, in this case, those thought to have FAI. Secondly, we looked for studies that assessed the validity of the "diagnosis" of FAI using the patients' symptoms, clinical exam and imaging results either in combination or individually. For validity, we included only those studies that used visual inspection at the time of surgery as the reference standard for comparison against the test. Lastly, we searched for studies whose purpose was to test the reliability of common clinical tests (e.g. impingement test) or imaging exams (e.g. alpha angle) believed to be important criteria for diagnosing FAI.

Only four studies were identified as prospective studies that listed study inclusion and exclusion criteria. Pain and a positive impingement test are two inclusion criteria specified in three of the four studies. All four studies included a positive impingement test. All four included a positive imaging study to confirm the diagnosis. The α-angle was used in three of the studies to diagnose cam FAI: >50º in two studies and >55º in the other. One study listed range of motion or limited hip motion as an inclusion criterion, but did not state the criteria of what defines "limited" motion.

Two studies attempted to assess validity: one assessed the clinical exam against the diagnosis of FAI, and one evaluated the impingement test and the α-angle, separately. Six experienced orthopedic surgeons made a diagnosis of labral tear, FAI or capsular laxity in eight patients with musculoskeletal hip-related pathologies from clinical exam alone, and these results were compared with a final diagnoses made at the time of surgery. The diagnoses obtained from the clinical exam had only a 65% agreement compared with that made from surgical inspection. The impingement test and α-angle measurement via MR arthrography had a sensitivity, specificity, positive and negative predictive value (PPV, NPV) of 77%, 87%, 86%, 79% for impingement, and 39%, 70%, 55% and 54% for the α-angle measurement.

One study assessed the reliability of the impingement test, and several evaluated the presence or absence of one or more individual imaging test. The interobserver reliability between one therapist and one surgeon for the impingement test was moderate ($\kappa = 0.58$). The α-angle showed moderate to high interobserver reliability in several studies. Other imaging tests

assessing abnormalities of the femur (head-neck offset, pistol-grip deformity, focal prominence, head sphericity, flattening of the femoral head) and acetabulum (crossover sign, posterior wall sign, ischial spine sign, excessive acetabular coverage, acetabular depth, acetabular inclination, pelvic rotation) had variable degrees of reliability, but none were tested for diagnostic validity.

KEY QUESTION 2
What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Has clinically meaningful improvement in outcomes been defined in FAI?

We identified seven hip outcome measures commonly used in the FAI patient population, but only two were evaluated for validity and/or reliability in an FAI population: the Hip Outcome Score, German version (HOS-D) and the modified Western Ontario and McMaster Universities Arthritis Index (M-WOMAC). Neither of these measures was adequately tested for validity or reliability in the FAI population. One instrument, the Nonarthritic Hip Score (NAHS) was validated in a young hip-pain patient population; however, its reliability was inadequately tested.

The minimal clinically important difference (MCID) was defined in only one measure, the HOS-D, and found to be 9 points for the ADL subscale and 6 points for the sports subscale in FAI patients.

KEY QUESTION 3
What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI?

We found no randomized controlled trials (RCTs) comparing surgery with conservative care for FAI or comparing different surgical treatments for FAI.

We identified one study that retrospectively compared conservatively treated patients versus those receiving FAI surgery versus patients having a total hip arthroplasty in the short-term (<5 year follow-up). In addition we found four comparative studies which investigated the effectiveness of various surgical treatments for FAI: labral debridement versus labral refixation (two studies) and osteoplasty versus no osteoplasty (two studies). The first study poorly describes the selection of patients so that it was not possible to tell how the treatment and control groups were obtained. The last four studies use historical controls. There is no evidence that one specific treatment resulted in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty).

In addition, 27 case series were found that reported on clinical outcomes following treatment for FAI in non- or recreational athletes. All report improvement in pain, patient-reported and

clinician-reported hip outcome scores, patient satisfaction and return to normal activities following FAI surgery. However, whether this improvement is a result of the surgery, or the postoperative rehabilitation, or the change in activity subsequent to the surgery or placebo is not known.

Approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years.

There are no long-term (≥10 years) data available to assess long-term effectiveness of FAI surgery. There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.

KEY QUESTION 4
What is the evidence of the safety of hip surgery for FAI compared with no surgery?

Six comparative studies, 31 case-series and three case-reports were found that reported complications following surgical treatment for FAI in non- or recreational athletes. Altogether, 20 studies reported on arthroscopy, ten on open dislocation and seven on the mini-open procedure.

Reoperation for reasons other than a conversion to a total hip arthroplasty occurred 3.8% in patients undergoing arthroscopy, 4.4% in those receiving open dislocation and 8.7% in patients following a mini-open procedure. There was only one reported head-neck fracture (<0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion. Heterotopic ossification occurred in 2% to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation.

Neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature. Three case-reports described an occurrence of extravasation of fluid into the abdomen/chest during arthroscopic treatment of FAI. In one case, the fluid extravasation resulted in an intra-abdominal compartment syndrome that presented as cardiopulmonary arrest.

KEY QUESTION 5
What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations?

No studies were found comparing the differential effectiveness of surgery versus nonsurgical care in FAI patients. However, five studies were identified that looked at outcomes following surgical treatment for FAI in two subpopulations, those with varying degrees of osteoarthritis as assessed by the Tönnis grade and patients with varying degrees of chondral damage assessed during surgery.

Outcomes following FAI surgery were consistently worse in patients with greater preoperative osteoarthritis compared with those with less osteoarthritis. In one study, the relative risk of a conversion to total hip arthroplasty (THA) in those with preoperative Tönnis grade 2–3 was 58 (95% CI: 8, 424) compared with Tönnis grade 0-1. There was no reported difference in outcomes in patients with varying degrees of chondral damage assessed during surgery. No data from other subpopulations were found.

KEY QUESTION 6
What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI?

No cost effectiveness, cost utility or costing studies were found on FAI surgery.

Summary

Key Question 1: Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?

	Strength of evidence	Conclusions/Comments
Case definition	VERY LOW	• The most consistent case definition of FAI (cam or mixed) as defined by inclusion/exclusion criteria in prospective studies of treatment effectiveness includes hip/groin pain, positive clinical impingement test, and an α-angle >50-55º • There is no evidence that the diagnosis of FAI can be obtained from clinical exam in one small study. One clinical test, the impingement sign, had a positive and negative predictive value of 86% and 79% in one study where the prevalence of FAI was 50%; however, in another study, the reliability of the impingement sign was only moderate. • Even though the α-angle showed moderate to high interobserver reliability in several studies, it had poor diagnostic value in identifying FAI. Other imaging tests assessing abnormalities of the femur and acetabulum had variable degrees of reliability, but no others were tested for diagnostic validity.

Key Question 2: What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Has clinically meaningful improvement in outcomes been defined for FAI?

	Strength of evidence	Conclusions/Comments
Hip osteoarthritis (Tönnis classification)	VERY LOW	• The Tönnis classification is often used to determine the extent of osteoarthritis in the hip. There were no studies found that assessed its validity. Reliability was tested in only one study and intra- and interobserver reliability in that study was moderate.
Patient- and clinician-reported outcome measures	VERY LOW	• Seven hip outcomes measures were used commonly in FAI patients. Three have undergone psychometric analysis in FAI (HOS-D, M-WOMAC) or young hip-pain (HOS, NAHS) patient populations. • Only one (NAHS) of the three instruments was adequately tested for validity, and it was performed in a young hip-pain patient population. • Reliability was inadequately tested for all three instruments. • The MCID was defined to be 9 points for the ADL subscale and 6 points for the sports subscale of the HOS-D in FAI patients. The MCID has not been defined for any other outcome measures in FAI or young hip-pain patients.

Key Question 3: What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI?

	Strength of evidence	Conclusions/Comments
Efficacy	NO EVIDENCE	• There are no data available to assess the short- or long-term efficacy of FAI surgery compared with no surgery
Effectiveness short-term	VERY LOW	• There is no evidence that one specific treatment resulted in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty). • Several case series report improvement in pain, patient reported and clinician reported hip outcome scores, patient satisfaction and return to normal activities following FAI surgery. However, whether this improvement is a result of the surgery, or the postoperative rehabilitation, or the change in activity subsequent to the surgery or placebo is not known. • Approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years.
Effectiveness long-term	NO EVIDENCE	• There are no data available to assess long-term effectiveness of FAI surgery compared with no surgery. • There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.

Key Question 4: What is the evidence of the safety of hip surgery for FAI compared with no surgery?

	Strength of evidence	Conclusions/Comments
Safety	LOW	• The risk of reoperation (other than conversion to THA) occurred in 4% (arthroscopy and open dislocation) and 9% of the patients (mini-open). • There was only one reported head-neck fracture (0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion. • Heterotopic ossification occurred in 2% to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation. • Neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature.

Key Question 5: What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations?

	Strength of evidence	Conclusions/Comments
Differential efficacy, effectiveness or safety	VERY LOW	• We found no studies comparing the differential efficacy, effectiveness or safety of surgery versus nonsurgical care in FAI patients. • Outcomes following FAI surgery were consistently worse in patients with greater preoperative osteoarthritis compared with those with less osteoarthritis. • There was no reported difference in outcomes in patients with varying degrees of chondral damage assessed during surgery. • No data from other subpopulations were found.

Key Question 6: What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI?

	Strength of evidence	Conclusions/Comments
Cost-effectiveness	NO EVIDENCE	There were no cost-effectiveness, cost utility or costing studies found on FAI surgery.

1. Appraisal

1.1. Rationale

Femoroacetabular impingement (FAI) syndrome is a recently recognized diagnosis in primarily younger individuals where relatively minor abnormalities in the joint (orientation or morphology) are thought to cause friction/impingement and pain. It is theorized that FAI starts the breakdown of cartilage, leading to osteoarthritis. There are two types of FAI: cam impingement (most common in young athletic males) and pincer impingement (most common in middle-aged women). Proponents believe that surgical correction of the impinging deformities will alleviate the symptoms and retard the progression of OA degeneration.

Hip surgery is an invasive procedure to correct FAI using either an open surgery or arthroscopic approach. The surgeon cuts off abnormal outgrowths of bone, removes damaged cartilage, and reshapes the femoral neck to ensure that there is sufficient clearance between the rim of the joint socket and the neck of the femur. After corrective surgery, avoidance of weight bearing for several weeks to months and rehabilitation is required.

The causes of hip pain, the natural history of FAI and its relationship to osteoarthritis are unclear; THE case definition and selection criterion of patients for this procedure is uncertain.
Significant questions remain about the efficacy and effectiveness, safety and cost effectiveness of hip surgery for FAI. Effectiveness questions particularly center on whether the potential beneficial outcomes of long-term pain and functional improvement, and prevention of a total hip replacement due to osteoarthritis deterioration occur with surgical intervention. With respect to safety, it is important to understand the risks of the intervention, and how often complications arise.

1.2. Key Questions

Key questions are developed by the Washington State Health Technology Assessment Program.

When used in patients with Femoroacetabular Impingement (FAI):

Key Question 1:
Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?

Key Question 2:
What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Has clinically meaningful improvement in outcomes been defined for FAI?

Key Question 3:
What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI? Including consideration of short-term and long-term:

- Impact on function, pain, range of motion, quality of life, activities of daily living and return to work
- Development or progression of osteoarthritis
- Need of or time to total hip arthroplasty ("continuing" or "subsequent intervention" that is not THA would go in the safety section)
- Other reported measures

Key Question 4:
What is the evidence of the safety of hip surgery for FAI compared with no surgery? Including consideration of:

- Adverse events type and frequency (peri-operative, fractures, nerve damage, mortality, other major morbidity)
- Revision/re-operation rates

Key Question 5:
What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations? Including consideration of:

- Gender
- Age
- Psychological or psychosocial comorbidities
- Baseline functional status: e.g. type of deformity, extent of osteoarthritis or cartilage damage
- Other patient characteristics or evidence based patient selection criteria
- Provider type, setting or other provider characteristics
- Payer/ beneficiary type: including worker's compensation, Medicaid, state employees

Key Question 6:
What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI?

- Costs (direct and indirect) and cost-effectiveness
- Short-term and long-term

1.3. Key considerations highlighted by clinical experts:

1.3.1. What is the prevalence of FAI? How frequent is it?

The prevalence of FAI has only been recently studied. We identified five studies assessing the prevalence based on the presence of morphological characteristics as found on imaging. The estimates vary depending on the criteria used to determine whether the morphological feature associated with FAI was present.

Laborie et al conducted a prospective population-based study of healthy young adults to determine the prevalence of radiographic findings thought to be associated with FAI.[74] The study population included a one year cohort born in a single Norwegian hospital in 1989. The initial cohort consisted of 5068 newborns that participated in an earlier randomized controlled trial assessing ultrasound screening on the primary diagnosis, management, and prevalence of late cases of developmental dysplasia of the hip (DDH). Of that initial cohort, 1062 were not available as a result of death (n=61), emigration abroad (n=256) or living out of catchment area at the time of birth (n=745). Of the 4006 remaining individuals, 2081 (52%) agreed to participate in the study. Twenty one were excluded due to suboptimal radiographs or uncertainty of pregnancy leaving 2060 individuals as study participants. Each participant underwent a standardized weight-bearing anteroposterior (AP) and supine frog-leg radiographic evaluation. The radiographer underwent specific training for the examination and ensured correct posture during the exam. The radiographs were read by an experienced pediatric musculoskeletal radiologist to assess both cam- and pincer-type findings as follows:

- Cam-type findings: pistol-grip deformity, focal prominence of the femoral neck, flattening of the lateral aspect of the femoral head.
- Pincer-type findings: posterior wall sign, excessive acetabular coverage, crossover sign.
- Fibrocystic changes (FCC) were also noted.

The results of this study showed that radiographic findings suggestive of FAI are common in a population of healthy young adults, especially males. The prevalence of one or more findings for cam-type deformities was 35% for males and 10% for females; for pincer-type deformities, 34% for males and 17% for females, Table 1. This study found little overlap between cam- and pincer-type findings. This supports the views of Cobb et al. who reported that cam hips are shallower than normal hips, and that normal hips are shallower than pincer hips.[29] They concluded that cam and pincer hips are distinct pathoanatomic entities.

Table 1. The prevalence of cam-and pincer-type findings in a healthy young population.

	Male (n=868) %	Female (n=1192) %
Cam-type deformities		
Pistol-grip	21.5	3.3
Focal prominence	10.3	2.6
Flattening lateral head	14.4	6.2
Cam-type (≥1 finding)	*35.0*	*10.2*
Pincer-type deformities		
Posterior wall sign	23.4	11.0
Excessive acetabular coverage	14.6	4.9
Crossover sign	51.4	45.5
Pincer-type (≥1 finding)	*34.3*	*16.6*
Fibrocystic changes	5.8	1.6

Reichenbach et al. recently assessed the prevalence of cam-type deformity in a population-base,d cross-sectional study of young males undergoing conscription for the Swiss army.[115] A random sample of 244 asymptomatic males (mean age, 19.9 years) received a clinical exam to assess hip range of motion and an MRI of the hip to determine the presence of cam-type deformity. From the MRI, the head-neck offset was determined using a "semiquantitative" scoring system that ranged from 0-3 where grades 2 and 3 were defined as definite or severe deformity with an established decrease. One hundred seventy nine subjects (73%) showed some MRI evidence of a cam-type deformity (grades 1, 2 or 3) while 24% had deformity grades 2 or 3. The prevalence estimates increased with decreasing internal rotation; 48% among those with reduced internal rotation of <30º, 21% among those with internal rotation between 30º and 40º, and 13% in those with internal rotation ≥40º. The authors concluded that cam-type deformity can be found on MRI in every fourth young asymptomatic male and in every second male with decreased internal rotation.

Hack et al. studied the prevalence of cam-type deformity in 200 asymptomatic volunteers (46% males; mean age, 29 years, range, 21 to 51 years) from among hospital and medical school workers at one institution in Canada.[45] Subjects had no prior hip surgery or childhood hip problems. Each underwent a MRI of both hips. α-angles were measured, and those with angles >50.5º were considered to have a cam-type deformity. In this population, 14% of the volunteers had a cam-type deformity in at least one hip. The prevalence of cam-type deformity was higher in males compared with females, 25% versus 5% and higher in those with ≤20º of internal rotation compared with those >20º,

25% versus 5%. The authors conclude that the prevalence of cam-type FAI deformity is higher in males and in individuals with less hip internal rotation.

Gosvig et al. conducted a population-based study to determine the prevalence of osseous malformations associated with FAI and to assess whether pain in the groin was associated with the malformations.[43] Data were collected from the Copenhagen Osteoarthritis Substudy (n = 4,151), a subgroup of a larger Copenhagen City Heart Study, a longitudinal survey of an adult, predominantly white cohort from the county of Østerbro in Copenhagen. The investigators excluded individuals with hip replacement surgery, Legg-Calvé-Perthes disease, childhood hip disease, rheumatoid arthritis, unreadable radiographs, or radiographs demonstrating excessive rotation leaving 3,620 individuals making up the study cohort. The mean age for the cohort was 60.5 years (range, 21 to 90 years), and there were nearly twice as many women as men (63% vs. 37%). Each individual received standing radiographs with feet pointing straight forward and legs in slight abduction. From these, the investigators recorded the presence of acetabular dysplasia (center-edge angle of Wiberg ≤20º), a deep acetabular socket (center-edge angle ≥45º), pistol-grip deformity (triangular index ≥0º), and osteoarthritis (joint space width ≤2mm). The prevalence of acetabular dysplasia was 4.3% in men and 3.6% in women; a deep acetabular socket, 15.2% of men and 19.4% of women; a pistol-grip malformation, 19.6% of men and 5.2% of women, the combination of pistol-grip deformity and deep acetabular socket, 2.9% of men 0.9% of women. There was no statistically significant increased prevalence of groin pain in individuals with radiographic abnormalities compared with those without abnormalities ($P = .13$). A deep acetabular socket and a pistol-grip deformity were risk factors for the development of osteoarthritis, adjusted risk ratios 2.4 (95% CI, 2.0, 2.9) and 2.2 (95% CI, 1.7, 2.8), respectively. The authors concluded that a deep acetabular socket and a pistol-grip deformity were common radiographic finds, and that these findings were associated with hip osteoarthritis.

Kang et al. used CT scans of asymptomatic hip patients (n=50 patients, 100 hips; 46% male; age range, 15 to 40 years) that received a scan for abdominal trauma or non-specific abdominal pain to determine the prevalence of bony abnormalities thought to predispose one to FAI.[66] Patients were excluded if they had current hip symptoms, history of congenital hip dysplasia, Perthés disease, slipped capital femoral epiphysis, hip fracture, hip surgery, and any arthropathy that could cause secondary alterations of the hip joint. Raw data from the abdominal CT scan were reformatted using bone algorithm into several different planes. From the reformatted images, the following assessments were made: acetabular retroversion (acetabular version angle of <15º), acetabular crossover sign, center edge angle (coxa profunda defined as center edge angle >40º), α-angle (abnormal defined as >55º) and head-neck offset (abnormal defined as <8mm).

Fourteen percent of the individuals had acetabular retroversion (the prevalence in males, 57%; the prevalence in females, 4%), 20% had a positive crossover sign, 16% had coxa profunda, 10% had abnormal α-angles, and 12% had abnormal head-neck offset. Thirty nine percent of all individuals had at least one of the above findings; the prevalence among males was 48% and females, 31%. Sphericity of the femoral head (abnormal if the femoral head projected outside a circular template on the femoral head) was also evaluated. Twenty six percent had an aspherical femoral head at the anterior head-neck junction, 42% at the lateral head-neck junction, and 59% at the anterolateral head-neck junction. Seventy-four percent had an aspherical femoral head in a t least one of the above planes. This study demonstrated substantial prevalence of bony characteristics thought to be associated with FAI in asymptomatic individuals.

Weir et al. examined the prevalence of radiographic signs of FAI in athletes with long-standing adductor-related groin pain.[135] Athletes who had groin pain for at least 2 months, pain located at the proximal attachment of the adductor muscles, pain with resistive adduction and pain during or after sporting activities were included. Those with pain above the conjoint tendon, with symptoms of prostatitis or urinary tract infection, or with low back pain were excluded. Thirty four patients (68 hips) made up the study population and were assessed by a single physician for iliopsoas length, hip range of motion, and hip impingement. A separate physician evaluated radiographs for three signs of cam-type FAI (pistol-grip deformity, centrum-collum-diaphyseal angle, the femoral head neck index of Heymann and Herndon) and five signs of pincer-type FAI (coxa profunda, protrusion acetabuli, the lateral center edge angle, the acetabular index and the crossover sign). Ten patients had bilateral adductor pain and 24 had unilateral pain resulting in 44 painful adductor-related hips and 24 non-painful adductor-related hips. There were 128 abnormal radiological signs observed in the 68 hips. The prevalence of having one or more FAI signs was 94% with only 4 hips (6%) without any signs of FAI. There was no association between the number of radiological signs and a positive anterior hip impingement test nor hip range of motion. The authors concluded that radiological signs of FAI are frequent in patients presenting with long-standing adductor-related groin pain. Furthermore, radiographic signs of FAI are often present with a negative impingement test, and the impingement test may not be specific for femoroacetabular impingement. Finally, the authors suggest that clear diagnostic criteria for FAI are needed.

1.3.2. What is the natural history of FAI?

The conceptual model of FAI suggests that there are morphological abnormalities of the proximal femur and or acetabulum resulting in abnormal contact at the end range of motion, particularly in flexion, internal rotation and adduction. Initially the hip is

asymptomatic but continued contact through excessive motion results in pain, chondral lesions, labral tears and progressive hip osteoarthritis.[39,70,106,127] However, the data to support this hypothesis is lacking. For example, a recent study by Hartofilakidis et al. investigated the association between morphological abnormalities suggestive of FAI and the development of osteoarthritis in asymptomatic hips.[48] The investigators identified 205 patients that received total hip arthroplasty (THA) in one hip but had an asymptomatic contralateral hip at the time of the joint replacement. Each had radiographs of the asymptomatic hip. From among the 205 patients, 96 were ≤65 years of age (mean age, 49 years), had no radiological evidence of osteoarthritis in the asymptomatic hip, and had one or more of the following morphological features associated with FAI: pistol-grip deformity, increased α-angle (>68º for men and >50º for women), coxa vara (neck-shaft angle <125º), acetabular inclination (≤0º), center-edge angle (≥35º), crossover sign, posterior wall sign or anterior rim prominence. Each radiograph was read independently by two authors (the crossover sign by three). The senior author was one of the reviewers of the radiographs. It was not clear how differences of opinions were resolved except for the crossover sign, which was resolved by consensus. Patients were followed for a mean of 18.5 years (range, 10 to 40 years). Radiographs at final follow-up were assessed by the senior author for the presence of early osteoarthritis defined as any subtle indication of joint space narrowing and/or the presence of marginal osteophytes on the femoral head. At final follow-up the prevalence of osteoarthritis was only 17.7%. The prevalence rate of osteoarthritis among those who did not have a morphological feature associated with FAI was not reported. Nevertheless, from this study it appears that a substantial proportion of hips with morphological features associated with FAI may not develop radiographic evidence of osteoarthritis in the long-term.

A another study by Bardakos et al investigated the effect of radiological parameters on the progression of osteoarthritis in patients under 55 years of age with a history of symptomatic idiopathic hip arthritis, Tönnis grade 1 or 2.[4] Those that had AP radiographs available with at least 10 year follow-up were selected for review. From among these records, patients with radiographs that suggested cam-type FAI as defined by the presence of a discernible reduced superior offset at the femoral head neck junction were included. After excluding patients with hip dysplasia, inflammatory arthritis, osteonecrosis of the femoral head, trauma and inadequate radiographic quality, 43 patients (43 hips) made up the study population. From the radiographs a single observer measured the α-angle, the neck-shaft angle, the center-edge angle of Wiberg, the medial proximal femoral angle, the Tönnis angle, the crossover sign, the posterior wall sign, coxa profunda and protrusion acetabuli. The outcome was progression of the initial osteoarthritis as measured by the Tönnis grade. Over the 10-plus year period, twenty eight (65%) showed evidence of osteoarthritis progression. The prevalence of progression

was similar between hips with initial Tönnis grade 1 or grade 2 osteoarthritis. Comparison of the hips with and without progression of arthritis revealed a significant difference in the mean medial proximal femoral angle (81 degrees vs. 87 degrees, p = 0.004) and the presence of the posterior wall sign (39% vs. 7%, p = 0.02), but not in the other radiographic measurements typically associated with cam FAI. The authors concluded that mild to moderate osteoarthritis in hips with a pistol-grip deformity will not progress rapidly in all patients. In one-third, progression will take more than ten years to manifest, if ever. While individual geometry of the proximal femur and acetabulum partly influences this phenomenon, a hip with cam impingement is not always destined for end-stage arthritic degeneration.

1.4. Washington State utilization and cost data

Femoroacetabular Impingement (FAI) Disclaimer:

There is no specific procedure code for FAI surgery. It is commonly billed under procedure codes that are described as "unlisted" procedures for either open or arthroscopic surgery. State utilization data also shows other hip surgery codes were paid on the same surgery date as these unlisted codes. Specific codes are identified in the list at the end of the Agency Experience section.

Because hip surgery for FAI has no specific procedure code and FAI itself has no specific diagnosis code, providing relevant data to support the HTCC decision was challenging. Note that due to the difficulty of identifying FAI cases, and differences in data capture and claims structures, each agency used a strategy to identify FAI claims that was best suited to their data environment. Differences in strategy are noted in report headings.

The following tables that list "potential" FAI surgeries are based on procedure codes that could have been billed for this surgery during 2007 – 2010, and though we know that some of the cases represent actual FAI surgeries, many of them may not.

PEB Data
HCA was able to definitely identify the following cases of Hip Surgery for FAI via the appeals database; however, data was compiled using normal state utilization sources.

PEB – Day of Surgery Costs per FAI Claim

Year	Member Number	Paid Per Surgery Date
2008	**1**	$4,103
	2	$4,103
	3	$14,533
	4	$4,103
	5	$3,899
	6	$6,900
2009	**7**	$11,222
	8	$11,696
2010	**9**	$5,307
	10	$8,982
	11a	$9,448
	11b	$11,174
Grand Total		$95,470

PEB – Payments and Member Counts by Diagnosis Code on FAI Claims

Diagnosis code Description	Member Ct	Total Paid
ARTICULAR CARTILAGE DISORDER, PELVIC REGION AND THIGH	9	$32,522
CHONDROMALACIA	1	$8,982
ENTHESOPATHY OF HIP REGION	1	$10,039
OTHER JOINT DERANGEMENT, NOT ELSEWHERE CLASSIFIED, PELVIC REGION AND THIGH	8	$28,080
OTHER SYNOVITIS AND TENOSYNOVITIS	1	$4,601
SPRAIN AND STRAIN OF OTHER SPECIFIED SITES OF HIP AND THIGH	1	$11,246
Grand Total		$95,470

PEB – FAI Claims, Mbrs by Gender & Age

Gender/Age Group	Member Count
F	
19-35	3
36-50	5
51-65	1
Total	**9**
M	
19-35	1
51-65	1
Total	**2**
Total	**11**

HCA was able to identify the following cases as potential FAI surgeries based on CPT codes and diagnoses:

PEB Potential FAI Hip Surgery, 2007-2010***

Open Hip Surgery (CPT 27299)	**2007**	**2008**	**2009**	**2010**	**Grand Total**
Members	1	8	6	3	17**
Procedure Counts	1	8	6	3	18
Procedure only Amount Paid	$219	$7,500	$23,173	$1,054	$31,946
Day of Surgery Amt Paid	$267	$28,796	$139,635	$6,338	$175,036
Average Amt Paid per Day of Surgery*		$4,114	$6,451	$2,924	$4,500
Arthroscopic Hip Surgery (CPT 29999)	**2007**	**2008**	**2009**	**2010**	**Grand Total**
Members	7	13	10	16	44**
Procedure Counts	8	14	11	17	50
Procedure only Amount Paid	$33,400	$22,667	$12,487	$14,170	$82,724
Day of Surgery Amt Paid	$61,359	$67,857	$86,180	$112,800	$328,196
Average Amt Paid per Day of Surgery*	$7,670	$4,902	$9,230	$6,635	$6,829
All Hip Surgeries, 27299, 29999	**2007**	**2008**	**2009**	**2010**	**Grand Total**
Members	8	21	16	19	61**
Procedure Counts	9	22	17	20	68
Procedure only Amount Paid	$33,619	$30,167	$35,660	$15,224	$114,670
Day of Surgery Amt Paid	$61,626	$96,653	$225,815	$119,138	$503,232
Average Amt Paid per Day of Surgery*	$7,670	$4,626	$8,535	$6,245	$6,355

**In order to calculate a representative average, two procedures were excluded as outliers (more than 4 standard deviations above the mean), and secondary payer claims were excluded.*
*** Overall Member counts are not the sum of annual member counts, but a separate count of members over 4 years.*
****Over all years, without outliers and secondary claims, the maximum day of surgery cost was $12,496, the median day of surgery cost was $5781, and average day of surgery cost was $6355 with a standard deviation of $3251.*

DSHS Data

HCA was not able to identify any clear cases of FAI in Medicaid data. The following table shows all Hip Surgeries with correct coding for FAI conditions during the 2007-2010 time frame; however, these codes are also correctly used for other conditions.

DSHS - DSHS Potential FAI Hip Surgery, 2007-2010***

Open Hip Surgery (CPT 27299)	2007	2008	2009	2010	Grand Total
Members	4	14	22	29	67**
Procedure Counts	4	15	22	31	72
Procedure only Amount Paid	$5,306	$10,219	$20,953	$39,100	$75,577
Claim Amt Paid	$6,044	$12,845	$44,722	$62,458	$126,070
Average Amt Paid per Claim*	$1,511	$856	$2,033	$1,643	$1,590
Arthroscopic Hip Surgery (CPT 29999)	**2007**	**2008**	**2009**	**2010**	**Grand Total**
Members	11	17	22	26	75**
Procedure Counts	11	17	22	27	77
Procedure only Amount Paid	$10,055	$12,794	$20,006	$16,244	$59,099
Claim Amt Paid	$21,869	$30,825	$50,272	$68,896	$171,862
Average Amt Paid per Claim*	$1,988	$1,813	$2,285	$2,552	$2,232
All Hip Surgeries, 27299, 29999	**2007**	**2008**	**2009**	**2010**	**Grand Total**
Members	15	31	44	55	142**
Procedure Counts	15	32	44	58	149
Procedure only Amount Paid	$15,361	$23,013	$40,959	$55,344	$134,677
Claim Amt Paid	$27,914	$43,670	$94,995	$131,354	$297,932
Average Amt Paid per Claim*	$1,861	$1,365	$2,159	$2,074	$1,924

**In order to calculate a representative average, one procedure was excluded as an outlier (more than 4 standard deviations above the mean).*
***Overall Member counts are not the sum of annual member counts, but a separate count of members over 4 years.*
****Over all years, excluding the outlier, the maximum claim paid was $8,181, the median payment was $1,599 and the average payment was $1,924 with a standard deviation of $1647.*

L&I Data

Clinical chart review was used to identify a usual billing pattern for arthroscopic hip surgery where FAI was noted in the chart. Charges were then accumulated for similar billing combinations; however these codes are also correctly used for other conditions.

L&I – L&I Potential FAI Hip Surgery, 2007-2010

Calendar Year	# of Claims	Total Paid*	Minimum Allowed	Maximum Allowed	Average/ Claim
2007	11	$166,204	$6,380	$25,922	$15,109
2008	24	$345,206	$6,669	$29,285	$14,384
2009	23	$388,364	$8,450	$33,721	$16,885
2010	35	$553,039	$6,502	$29,182	$15,801
Total	**93**	**$1,452,813**	**$6,380**	**$33,721**	**$15,622**

**All other paid charges for the date of surgery were included for this report. Bills for interpreter services, vocational services and claimant travel were excluded.*

All Agency Hip Reconstruction Data, 2005-2008

The following tables are provided as context for overall hip procedures, especially relevant since hip surgery for FAI may prevent or delay progress to these procedures. Note that the utilization data time frame in these tables differs from the time frame in previous report sections.

All Agencies:

Count of Procedures by Year, 2005-2008

UMP, L&I, & Medicaid

ICD-9 Procedure Codes	2005	2006	2007	2008	Total
00.85 (total hip resurfacing)	0	3	20	22	**45**
00.86 (resurfacing, femoral head)	0	1	2	2	**5**
00.87 (resurfacing, acetabulum)	0	0	0	0	**0**
81.51 (total hip replacement)	432	471	487	614	**2004**
81.52 (partial hip replacement)	108	100	82	102	**392**
Total	**540**	**575**	**591**	**740**	**2446**

All Agencies:

Amount Paid* by Procedure by Year, 2005-2008

UMP, L&I, & Medicaid

ICD-9 Procedure Codes	2005	2006	2007	2008	Total
00.85 (total hip resurfacing)	$0	$69,406	$404,120	$454,032	**$927,558**
00.86 (resurfacing, femoral head)	$0	$19,991	$36,344	$60,457	**$116,792**
00.87 (resurfacing, acetabulum)	$0	$0	$0	$0	**$0**
81.51 (total hip replacement)	$5,639,160	$6,378,458	$6,389,632	$9,036,877	**$27,444,126**
81.52 (partial hip replacement)	$1,264,504	$940,592	$957,011	$1,246,261	**$4,408,368**
Total	**$6,903,663**	**$7,408,447**	**$7,787,107**	**$10,797,626**	**$32,896,844**

Related Medical Codes			
Code Type	**Codes**	**Short Description**	**Additional Info**
ICD9 Diagnosis	ICD-9		**Not limited to:**
	718.05	Articular cartilage disorder, pelvic region	Expected Diagnosis
	718.45	Contracture of joint, pelvic region and thigh	Expected Diagnosis
	718.65	Unspecified intrapelvic protrusion acetabulum, pelvic region and thigh	Expected Diagnosis
	718.85	Other joint derangement, not elsewhere classified	Expected Diagnosis
	718.95	Unspecified derangement of joint	Expected Diagnosis
	719.45	Pain in joint, pelvic region and thigh	Expected Diagnosis
	719.55	Stiffness of joint, not elsewhere classified, pelvic region and thigh	Expected Diagnosis
	719.7	Difficulty in walking	Expected Diagnosis
	719.85	Other specified disorders of join, pelvic region and thigh	Expected Diagnosis
	719.95	Unspecified disorder of joint, pelvic region and thigh	Expected Diagnosis
	736.30	Acquired deformities of hip, unspecified deformity	Expected Diagnosis
	736.39	Acquired deformities of hip, other	Expected Diagnosis
Comorbidity	ICD-9		
	715-715.9	Osteoarthritis	Comorbidity
Treatments	CPT		
	27036	Capsulectomy or capsulotomy, hip, with or without excision of heterotopic bone, with release of hip flexor muscles	Hip surgery additionally billed code
	27299	Unlisted procedure, pelvis or hip joint [when specified as open procedure for femoroacetabular impingement syndrome]	Hip surgery selection
	29862	Arthroscopy, hip, with debridement/shaving or articular cartilage (chondroplasty), abrasion athroplasty, and/or resection of labrum.	Hip surgery additionally billed code

	29863	Arthroscopy, hip, with synovectomy	Hip surgery additionally billed code
	29914*	Arthroscopy, hip, surgical; with femoroplasty (i.e., treatment of cam lesion)	Future research
	29915*	Arthroscopy, hip, surgical; with acetabuloplasty (i.e., treatment of pincer lesion)	Future research
	29916*	Arthroscopy, hip, surgical; with labral repair	Future research
	29999	Unlisted procedure, arthroscopy	Hip surgery selection
Progression to replacement	CPT		
	27120	Acetabulum Replacement	Hip Replacement progression code
	27122	Resection femoral head	Hip Replacement progression code
	27125	Partial hip replacement	Hip Replacement progression code
	27130	Total hip replacement	Hip Replacement progression code

**New codes for 2011, not in use during study*

2. Background

2.1. History of Femoroacetabular Impingement as a Diagnosis

In the early 1960s, hip damage as a result of femoroacetabular contact was reported as a consequence of childhood disease, particularly slipped capital femoral epiphysis.[20,55] In 1974, Stulberg et al noted that subtle anatomic abnormalities of the hip, particularly a decreased head-neck offset of the proximal femur, was associated with early development of osteoarthritis.[129] Ganz et al in the early 1990s described six cases of femoral neck-acetabular impingement following fracture and malunion of the femoral neck.[37] Ganz subsequently in 2001 described a technique for surgical dislocation of the hip that allowed direct observation of the joint.[38] Following this description, Ganz and colleagues proposed FAI as a mechanism for the development of early osteoarthritis for nondysplastic hips based on in situ inspection of the damage pattern of over 600 surgical hip dislocations.[39] From that point forward, a significant number of other articles on the treatment, prognosis and diagnosis of FAI have been published, as well as numerous reviews.

2.2. Mechanism of Femoroacetabular Impingement

The proposed mechanism of FAI is one where abnormal contact occurs between the proximal femur and acetabulum during the end range of hip motion, particularly flexion and internal rotation. This abnormal contact is believed to be due to morphologic abnormalities of the acetabulum or proximal femur (or both) resulting in labrum tears, chondral lesions, and progressive osteoarthritis.[39,61,106]

2.3. Classification of Femoroacetabular Impingement

Two distinct types of FAI have been proposed by Ganz and coworkers[39] depending on where the abnormal morphology occurs; abnormal morphology of the femur is termed “cam impingement” and of the acetabulum, “pincer impingement”.

Cam-type impingement is associated with a non-spherical femoral head or an abnormality at the head-neck junction.[39,67,77,81] The malshaped proximal femur has the effect of increasing the radius of the femoral head, leading to abnormal contact with the acetabulum at the end of hip motion. This type of impingement has also been associated with slipped capital femoral epiphysis, Leg-Calvé-Perthes disease, osteonecrosis and post-traumatic deformities of the femur.[67,81,93,127]

Pincer-type FAI is characterized by a functionally deep or retroverted acetabulum resulting in overcoverage of the femoral head. [39,67,77,81,127] This overcoverage may be a relative anterior overcoverage, as seen in retroverted acetabuli, focal anterior overcoverage or a global acetabular overcoverage, often the result of coxa profunda or protusio acetabuli.[39,67,93]

While FAI has been classified into these two types, a mixed-type impingement, with characteristics of both cam and pincer-type FAI, has also been described.[16,17,48,58,109,111] However, at least one study that purposed to evaluate the acetabulum in those with a diagnosis of cam or pincer FAI reported that cam hips were slightly shallower than normal whereas pincer hips were deeper.[29] They concluded that cam and pincer hips are distinct pathoanatomic entities.

2.4. Treatment (surgical and nonsurgical)

2.4.1. Non-operative treatment

A variety of non-operative approaches have been used to treat FAI including non-steroidal anti-inflammatory medications, core strengthening, physical therapy, steroid injections, activity modification and pelvic postural retraining.[32,39,61,62,67,81] It has been hypothesized that while conservative measures may be employed to alleviate symptoms, these treatments will not address the underlying pathomechanics and as a result, will ultimately lead to progressive degeneration of the hip and development of osteoarthritis.[39,67,81] Some authors have suggested that physical therapy may in fact aggravate impingement symptoms in some cases.[61,106]

2.4.2. Operative Treatment

The fundamental goals of surgical intervention in the treatment of FAI, regardless of the type of impingement or the surgical technique used, are to correct the underlying morphologic abnormalities of the femur and/or acetabulum and address the pathologic changes present in the labrum and articular cartilage in order to improve hip range of motion and alleviate areas of abnormal contact.[39,61,67,81,106]

- *Cam Impingement*
  In cam-type impingement, the goals of surgery are to remove any asphericity of the femoral head and improve the head-neck offset. Debridement of bony abnormalities of the head and femoral osteotomy at the level of the head and neck, base of neck and intertrochanteric level, have all been used to correct underlying morphologic abnormalities and restore the head-neck offset in femoral causes of FAI.[39,67,81,93]

- *Pincer Impingement*
  In pincer-type impingement, surgery is aimed at reducing the prominence of the acetabular rim, debriding the degenerative labral tissue and reattaching normal labral tissue.[39,67,81] Periacetabular osteotomy has also been performed in cases of severe acetabular retroversion.[81,93] In dealing with labral abnormalities, it has been suggested that resecting the labrum should be avoided if at all possible. However, if the acetabular rim needs to be resected the labrum can be taken down as part of the approach and surgically refixed.[39,81]

Three surgical approaches are commonly used to accomplish the goals of surgical intervention; an open approach, arthroscopy or arthroscopy with a limited open approach (mini-open).

- *Open Approach*
  Ganz et. al. described an open procedure for the treatment of FAI in which a lateral hip incision is made, followed by trochanteric osteotomy and Z-shaped capsulotomy. The hip is dislocated anteriorly, allowing for a full 360˚ view of the femoral head and acetabulum.[38] Since this original description, other open approaches have been used based on the nature of the underlying abnormal morphology present. In one case series, the trochanteric slide exposure was utilized when there was extensive posterior-inferior acetabular impingement and the iliofemoral exposure was performed when isolated anterior FAI was present.[93] The open approaches allow for adequate debridement of aspherical portions of the femoral head and the acetabular rim as well as providing excellent exposure to inspect articular surfaces.[39,93] Femoral osteotomies at the level of the head and neck, base of neck and intertrochanteric osteotomies can also be performed when an open approach is utilized. It has been hypothesized that complex bony abnormalities are better treated with an open approach, compared to arthroscopic or arthroscopic assisted procedures.[81] However, the disadvantages of this procedure include a relatively long rehabilitation time due to trochanteric osteotomy and a potential impairment of hip proprioception due to capsulotomy and resection of the ligamentum teres.[25,82]

- *Arthroscopy*
  Hip arthroscopy is a minimally invasive procedure that has gained favor in treating cam, pincer or mixed-type FAI.[67,77,128] A meta-analysis grading the current indications for hip arthroscopy noted that FAI has become one of the most accepted indications for hip arthroscopy and assigned it a level B rating, indicating fair evidence exists to support hip arthroscopy for the treatment of FAI.[128] Similar to treatment through an open approach, the goals of treating FAI with arthroscopy are

correcting underlying structural deformities of the femur in cam-type impingement and reduction of overcoverage of the acetabulum in pincer-type FAI. This is accomplished through a minimally invasive approach, utilizing 2 to 3 ports and, unlike the open approach, does not involve surgical dislocation of the hip. Although this procedure is less invasive than an open procedure, there are limitations to what can be done arthroscopically. Posterior based lesions can be challenging to treat, difficulty assessing the true depth of bony resection may lead to over or under resection, resecting a retroverted acetabulum is technically difficult, and it is difficult to treat chondral lesions.[67,77] Protrusio has also been cited as being difficult to treat arthroscopically given the difficulty in performing dynamic assessment of hip motion intraoperatively.[77]

- *Arthroscopy with Limited Open Approach*
  Combining arthroscopy with a mini-open approach allows for the treatment of focal cam impingement and addressing labral and chondral lesions with an improved exposure compared to arthroscopy alone.[47,80,117] Similar to arthroscopy, the inability to address posterior based lesions is a known limitation when utilizing this approach.[67]

2.5. Indications and Contraindications

Continued pain despite treatment with a conservative approach in patients diagnosed with FAI, and not having severe osteoarthritis, has been reported as an indication for surgical intervention.[67,77,79] The duration of the trial of conservative treatment varies from as little as 6 weeks[79] up to 6 months.[67]

Several authors have cited patient selection as an important factor in outcomes of surgery for FAI, particularly the difficulty of achieving a successful outcome in patients with advanced osteoarthritis prior to surgery.[57,80,101,117] One author found that patients with greater than 50% joint space narrowing, predominance of aching pain at rest and bipolar grade 4 lesions on MRI had universally poor outcomes following surgery for FAI.[79]

2.6. Potential complications/harms of FAI surgery

It has been suggested that complications/harms for FAI surgery can be grouped into major, moderate and minor categories.[27] Potential major complications include avascular necrosis, femoral head-neck fracture, loss of fixation requiring revision, deep infection, symptomatic or significant limitation of hip motion due to heterotopic ossification, neurovascular injury,

and symptomatic venous thromboembolism. Potential moderate complications include symptomatic hardware, with or without removal. Potential minor complications include asymptomatic heterotopic ossification, superficial infection, and urinary tract infection.

2.7. Clinical Guidelines

2.7.1. National Guideline Clearinghouse

No clinical guidelines related to surgery (open or arthroscopic) for FAI syndrome were found.

2.7.2. National Institute for Health and Clinical Excellence

Arthroscopy

The National Institute for Health and Clinical Excellence (NICE), (which provides guidance on health technologies and clinical practice for the National Health Service in England and Wales) concluded in 2007 that current evidence on the efficacy and safety of both arthroscopic surgery for the treatment of FAI syndrome "does not appear adequate for these procedures to be used without special arrangements for consent and for audit or research"; further publications of safety and efficacy outcomes will be needed. NICE stated that only surgeons with specialist expertise in arthroscopic hip surgery should perform this procedure for FAI and that the natural history of FAI syndrome and the selection of patients for this procedure are uncertain; further research on these issues will be useful.[97]

In July 2011, NICE published an updated report on arthroscopy for FAI syndrome in the form a rapid review of the medical literature and specialist opinion. The review is based on approximately 1126 patients from three non-randomized controlled trials, five case-series, and one case-report. Several short-comings in the available literature were addressed such as overall poor study quality, limited prospective data collection in case-series, variability of outcome assessment scales used and lack of validation of these scales, heterogeneity in treatments making comparison between studies difficult, and descriptions of hip impingement pathology/lesions not well defined in all studies. The specialists' concluded that "there is no proof yet that this procedure is efficacious, but the technique may have a place in preventing the development of osteoarthritis of the hip in some patients". They also stated that use of this procedure will become more widespread, but should remain with the confines of the specialist dealing with hip disorders in young adults.[98]

Open Dislocation

NICE published an updated guidance report on open surgery for FAI in July 2011 stating that "current evidence on the efficacy of open femoro-acetabular surgery for hip impingement syndrome is adequate in terms of symptom relief in the short and medium term. With regard to safety, there are well recognized complications. Therefore this procedure may be used

provided that normal arrangements are in place for clinical governance, consent and audit with local review of outcomes."[99]

2.8. Previous Systematic Reviews/Technology Assessments

Bedi et al conducted a systematic review of the literature up to May 2008 on labral tears and FAI.[10] Of the 19 articles included all but one (a level III) were retrospective case-series with short-term follow-up. The studies suggested that 65% to 85% of patients who receive open surgical dislocation with labral debridement and 67% to 100% of patients who receive arthroscopic treatment of labral tears will be satisfied with their outcome at a mean of 3 years after surgery. All series reported an increased incidence of failure among patients with substantial pre-existing osteoarthritis. The authors concluded that the quality of literature reporting outcomes of surgical intervention for labral tears and FAI is limited.

A systematic evidence review prepared by the Health Care Insurance Board of the Netherlands found no randomized or prospective comparative studies of surgery for FAI syndrome. The review noted that evidence consists largely of retrospective case-series, which are heterogeneous in terms of patient populations, treatment and outcomes. The Health Board of the Netherland concluded that there is insufficient evidence of the effectiveness of surgical treatment of FAI syndrome, in terms of reduction in pain and improvements in function, and reduction in osteoarthritis progression, compared with standard therapy.[54]

Clohisy and colleagues performed a systematic review of the literature between 1950 and 2009 for all studies reporting on surgical treatment of FAI. Studies with clinical outcome data and minimum 2-year follow-up were analyzed. A total of 11 studies met criteria for inclusion -- 9 were Level IV and 2 were Level III. Reduced pain and improvement in hip function were reported in all studies. Conversion to total hip arthroplasty was reported in 0% to 26% of cases. Major complications occurred in 0% to 18% of the procedures. The authors noted that limitations in the literature are substantial and primarily result from the limited number of published studies, the heterogeneous study methods and surgical techniques used in the included studies. Current evidence regarding FAI surgery is primarily of poor quality and suggests the various surgical techniques are associated with pain relief and improved function in 68% to 96% of patients over short-term follow-up. The authors also stated that long-term follow-up is needed to determine survivorship and impact on osteoarthritis progression and natural history.[27]

Only one previously conducted technology assessment was found, a Hayes Brief published in 2010 investigating arthroscopic hip surgery for FAI syndrome.[53] No assessments were located evaluating open surgery for FAI. Table 2 summarizes the previous assessment.

Table 2. Overview of previous health technology assessments on surgery for FAI syndrome

Assessment (year)	Lit search dates	Evidence base available*†	Critical Appraisal‡	Comments	Primary Conclusions
Hayes Brief (2010)	2005 to January 2010	• 2 nonrandomized comparative studies (N = 123 hips, mean f/u 1.6–2 years, %f/u NR) • 5 case-series (N = 652, mean f/u 10 months–2.3 years, %f/u NR)	Not described	Hayes Rating C: Potential but unproven benefit. Some positive published evidence regarding safety and/or efficacy support use of the technology for the cited application(s), but a beneficial impact on health outcomes has not been proven because data are sparse and the level of evidence is low, or data are inconsistent or conflicting.	• None of the available studies compared arthroscopic surgery with open surgery or involved more than 2.5 years of evaluation of the results of surgery; therefore, additional studies are needed to determine whether the short-term benefits of arthroscopic hip surgery are offset by worse long-term results. • Efficacy: insufficient/no evidence • Safety: minor safety issues • Intended patient population: somewhat defined • Patient-centered outcomes: insufficient/no evidence
Hayes update (2011)	December 2009 through 2011	• 2 SRs • 2 prospective studies • 3 retrospective studies • 2 clinical studies • 2 case-series • 4 case-reports • 3 reviews	NA	• Efficacy: unchanged • Safety: additional information available • Patient selection criteria: unchanged • Long-term follow-up: up to 60 months	• No change from initial conclusions in 2010 brief

NA: not available; NR: not reported; SR: systematic review.
*Percent follow-ups are weighted based on sample size, and were calculated using the N reported in the assessment. Percent follow-ups were not given for all RCTs or case series. Mean time to follow-up is reported here.
†N reflects numbers before loss to follow-up.
‡Critical appraisal refers to formal evaluation of individual study quality using criteria such as the Jadad or GRADE methods of scoring and the determination of overall strength of evidence.

2.9. Medicare and Representative Private Insurer Coverage Policies

Currently there are no national or local coverage determinations or policies for The Centers for Medicare and Medicaid Services (CMS) regarding the surgical treatment of FAI syndrome. Coverage policies are consistent (apart from Aetna) for surgical treatment, either

open or arthroscopic, of FAI syndrome for selected bell-weather payers. The payers will provide coverage for surgical intervention as long as certain patient conditions are met. Table 3 provides an overview of policy decisions.

- Medicare
  The Centers for Medicare and Medicaid Services (CMS) have no policies or national or local coverage determinations currently.

- Aetna[1]
  Aetna considers surgery, either arthroscopic or open, for the treatment of FAI syndrome experimental and investigational, citing insufficient evidence in the peer-reviewed literature to support its effectiveness.

- Blue Cross/Blue Shield[13]
  Open or arthroscopic treatment of FAI may be medically necessary when ALL of the following criteria are met:
  - Age
    - Adolescent patients should be skeletally mature with documented closure of growth plates
    - Adult patients should be young enough to be considered inappropriate candidates for THA or other reconstructive hip surgery (e.g., < 55 years of age)
  - Symptoms
    - Moderate-to-severe hip pain that is worsened by flexion activities (e.g. squatting or prolonged sitting) that significantly limits activities
    - Unresponsive to conservative therapy for ≥ 3 months or conservative therapy is contraindicated (e.g., history of falls due to mechanical instability of hip joint). Conservative therapy for FAI should include:
      - Activity modification including avoidance of hip stretching activities
      - Restriction of athletics pursuits
      - Avoidance of symptomatic motion
    - Positive impingement sign on clinical examination (i.e., pain elicited with 90° of flexion and internal rotation and adduction of the femur
  - Imaging
    - Morphology indicative of cam-type or pincer-type FAI, e.g., pistol-grip deformity, femoral head-neck offset with an alpha angle greater than 50 degrees, a positive wall sign, acetabular retroversion (overcoverage with crossover sign), coxa profunda or protrusion, or damage of the acetabular rim
    - High probability of a causal association between the FAI morphology and damage, e.g., a pistol-grip deformity with a tear of the acetabular labrum and articular cartilage damage in the anterosuperior quadrant

- No evidence of advanced osteoarthritis, defined as Tonnis grade II or III, or joint space of less than 2 mm, except when there is mechanical instability
- No evidence of severe (Outerbridge grade IV) chondral damage

- Cigna[22]
  Covers open or arthroscopic hip surgery, including labral repair with or without grafting, for FAI syndrome as medically necessary when ALL of the following criteria are met:
  - Moderate-to-severe persistent hip or groin pain that limits activity and is worsened by flexion activities (e.g., squatting or prolonged sitting)
  - Pain unresponsive to medical management (e.g., restricted activity, nonsteroidal anti-inflammatory drugs)
  - Positive impingement sign (i.e., sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation)
  - Radiographic confirmation of FAI (e.g., pistol-grip deformity, alpha angle greater than 50 degrees, coxa profunda, and/or acetabular retroversion)
  - Absence of ALL of the following:
  - Tönnis grade 2 osteoarthritis (i.e., small cysts in femoral head or acetabulum, increasing narrowing of joint space, moderate loss of sphericity of femoral head)
  - Tönnis grade 3 osteoarthritis (i.e., large cysts, severe narrowing or obliteration of joint space, severe deformity of femoral head, avascular necrosis)
  - Outerbridge grade III cartilage damage (i.e., fissuring to the level of subchondral bone in an area with a diameter more than 1.5 centimeters)
  - Outerbridge grade IV cartilage damage (i.e., exposed subchondral bone head)

- Harvard Pilgrim[49]
  Arthroscopic hip surgery for FAI is covered when ALL the following criteria have been met:
  - Pain unresponsive to medical management
  - Positive impingement sign with sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation
  - Imaging studies confirming FAI
  - No evidence of advanced OA (Tonnis grade II or III) and the absence of severe chondral damage

  No policy on open hip surgery for FAI syndrome was found.

- UnitedHealthcare[132]
  Surgical treatment, both arthroscopic and open, for FAI is medically necessary and covered. Best surgical outcomes are achieved in patients who have ALL of the following:
  - Pain unresponsive to medical management (e.g., restricted activity, nonsteroidal anti-inflammatory drugs);

- Moderate-to-severe persistent hip or groin pain that limits activity and is worsened by flexion activities (e.g., squatting or prolonged sitting);
- Positive impingement sign (i.e., sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation);
- Radiographic confirmation of FAI (e.g., pistol-grip deformity, alpha angle greater than 50 degrees, coax profunda, and/or acetabular retroversion);
- Do not have advanced osteoarthritis (i.e., Tönnis grade 2 or 3) and/or severe cartilage damage (i.e., Outerbridge grade III or IV)

Table 3.Overview of payer technology assessments and policies for hip surgery procedures for FAI.

Payer (year)	Lit search dates	Evidence base available*†	Policy	Rationale/comments
Centers for Medicare and Medicaid Services (CMS)	NA	• NA	No national or local coverage determination	• NA
Aetna (2011)	NR	• No RCTs or prospective cohorts found • 2 guidance documents (NICE 2007a, 2007b) • 3 SRs (N = 30 studies for 2 SRs, number of studies NR for third SR; mostly case-series; %f/u and f/u NR) • 1 cohort (N = 60 hips; %f/u NR; f/u 2 years) • 8 case-series (N = 744 hips; %f/u NR; f/u 9 months–5 years)	Surgery, open or arthroscopic, for the treatment of FAI syndrome is experimental and investigational	• There is currently insufficient evidence to support the effectiveness of surgery (open or arthroscopic) for the treatment of FAI syndrome • There is a lack of evidence that surgical intervention slows the rate of progression to OA of the hip in these patients • Long-term follow-up is needed

BCBS of MA‡ (2011)	NR	• NR (but reference listed)	Open or arthroscopic treatment of FAI may be medically necessary when all of the following criteria are met:
			• Age
			  • Adolescent patients should be skeletally mature with documented closure of growth plates
			  • Adult patients should be young enough to be considered inappropriate candidates for THA or other reconstructive hip surgery (e.g., < 55 years of age)
			• Symptoms
			  • Moderate-to-severe hip pain that is worsened by flexion activities (e.g. squatting or prolonged sitting) that significantly limits activities
			  • Unresponsive to conservative therapy for ≥ 3 months or conservative therapy is contraindicated (e.g., history of falls due to mechanical instability of hip joint).
				• Not stated

Payer (year)	Lit search dates	Evidence base available*†	Policy	Rationale/comments
			• Conservative therapy for FAI should include: • Activity modification including avoidance of hip stretching activities • Restriction of athletics pursuits • Avoidance of symptomatic motion • Positive impingement sign on clinical examination (i.e., pain elicited with 90° of flexion and internal rotation and adduction of the femur • Imaging • Morphology indicative of cam-type or pincer-type FAI, e.g., pistol-grip deformity, femoral head-neck offset with an alpha angle greater than 50 degrees, a positive wall sign, acetabular retroversion (overcoverage with crossover sign), coxa profunda or protrusion, or damage of the acetabular rim • High probability of a causal association between the FAI morphology and damage, e.g., a pistol-grip deformity with a tear of the acetabular labrum and articular cartilage damage in the anterosuperior quadrant • No evidence of advanced osteoarthritis, defined as Tonnis grade II or III, or joint space of less than 2 mm, except when there is mechanical instability • No evidence of severe (Outerbridge grade IV) chondral damage Surgical treatment of FAI is considered investigational for all other indications	
Anthem BCBS (2011)	NR	• No RCTs or prospective comparative studies found • 7 case-series (N = 646, %f/u NR, 10 months–2.4 years)	Surgical treatment of femoroacetabular impingement syndrome (FAIS) is considered medically necessary when all of the following criteria have been met: • Individual exhibits symptoms of FAIS, including hip pain (primarily in the groin) that interferes with activities of daily living; Radiographs confirm diagnosis of FAIS, with evidence of cam impingement (alpha angle greater than 50 degrees), pincer impingement (acetabular retroversion or coxa profunda), or both • Individual has failed conservative therapy for a duration of at least 6 months, including: • Activity modification, with restriction of athletic pursuits, if any, that include avoidance of symptomatic movements; and • Treatment with NSAIDs or acetaminophen • Etiology of hip pain is confirmed by relief after injection of local anesthetic into the joint, and there is no other explanation for	• Expert opinion and uncontrolled case series suggest that for relatively young active people for whom no other options exist, surgical treatment for FAI syndrome significantly improves quality of life and pain symptoms, and may help avoid the development of hip osteoarthritis later on.

Payer (year)	Lit search dates	Evidence base available*†	Policy	Rationale/comments
			pain • Individual has minimal degenerative changes of the hip joint (Tönnis grade 1 or less) • Individual has had no prior hip surgery or arthroscopy.	

Cigna (2011)	NR	• SRs, case-series, retrospective reviews with up to 12 years f/u
		• 2 guidance documents (NICE 2007a, 2007b)
			Covers open or arthroscopic hip surgery, including labral repair with or without grafting, for FAI syndrome as medically necessary when ALL of the following criteria are met:
			• moderate-to-severe persistent hip or groin pain that limits activity and is worsened by flexion activities (e.g., squatting or prolonged sitting)
			• pain unresponsive to medical management (e.g., restricted activity, nonsteroidal anti-inflammatory drugs)
			• positive impingement sign (i.e., sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation)
			• radiographic confirmation of FAI (e.g., pistol-grip deformity, alpha angle greater than 50 degrees, coxa profunda, and/or acetabular retroversion)
			• absence of ALL of the following:
			  • Tönnis grade 2 osteoarthritis (i.e., small cysts in femoral head or acetabulum, increasing narrowing of joint space, moderate loss of sphericity of femoral head)
			  • Tönnis grade 3 osteoarthritis (i.e., large cysts, severe narrowing or obliteration of joint space, severe deformity of femoral head, avascular necrosis)
			  • Outerbridge grade III cartilage damage (i.e., fissuring to the level of subchondral bone in an area with a diameter more than 1.5 centimeters)
			  • Outerbridge grade IV cartilage damage (i.e., exposed subchondral bone head)
			
			Labral repair, with or without grafting, during surgical treatment of FAI is not covered because it is considered experimental, investigational, or unproven.
				• Evidence in the published peer-reviewed scientific literature supports open and arthroscopic hip surgery, including labral repair with or without grafting, as safe and effective for the treatment of femoroacetabular impingement (FAI) syndrome in a carefully selected subset of patients.

Regence (2010)	NR	• No RCTs found
		• 2 guidance documents (NICE 2007a, 2007b)
		• 1 SR (N = 19 studies; mainly case-series)
		• 1 controlled cohort (N = 71, %f/u NR, 1 year f/u)
		• 11 case-series (N = 711, %f/u NR, 10 months–4 years f/u)
			Open or arthroscopic treatment of FAI may be medically necessary when all of the following criteria are met:
			• Age
			  • Adolescent patients should be skeletally mature with documented closure of growth plates
			  • Adult patients should be young enough to be considered inappropriate candidates for THA or other reconstructive hip surgery (e.g., < 55 years of age)
			• Symptoms
				• Treatment of FAI is most effective in younger patients without osteoarthritis (Tonnis grade 0 or I) or severe cartilage damage.
				• There is a high association between FAI pathology and idiopathic osteoarthritis, but this may represent a small proportion of the total cases of hip osteoarthritis.
				• It is not known whether

Payer (year)	Lit search dates	Evidence base available*†	Policy	Rationale/comments

			• Moderate-to-severe hip pain that is worsened by flexion activities (e.g. squatting or prolonged sitting) that significantly limits activities
			• Unresponsive to conservative therapy for ≥ 3 months or conservative therapy is contraindicated (e.g., history of falls due to mechanical instability of hip joint). Conservative therapy for FAI should include:
			  • Activity modification including avoidance of hip stretching activities
			  • Restriction of athletics pursuits
			  • Avoidance of symptomatic motion
			• Positive impingement sign on clinical examination (i.e., pain elicited with 90° of flexion and internal rotation and adduction of the femur
			• Imaging
			  • Morphology indicative of cam-type or pincer-type FAI, e.g., pistol-grip deformity, femoral head-neck offset with an alpha angle greater than 50 degrees, a positive wall sign, acetabular retroversion (overcoverage with crossover sign), coxa profunda or protrusion, or damage of the acetabular rim
			  • High probability of a causal association between the FAI morphology and damage, e.g., a pistol-grip deformity with a tear of the acetabular labrum and articular cartilage damage in the anterosuperior quadrant
			  • No evidence of advanced osteoarthritis, defined as Tonnis grade II or III, or joint space of less than 2 mm, except when there is mechanical instability
			  • No evidence of severe (Outerbridge grade IV) chondral damage
			
			Surgical treatment of FAI is considered investigational for all other indications
				arthroscopic or open approaches result in better net health outcome when patients are matched for severity of FAI morphology and articular cartilage damage.
				• The evidence is insufficient to permit conclusions concerning the effect of surgical procedure on the development of osteoarthritis. Therefore, treatment of FAI morphology in the absence of symptoms is considered investigational.

Harvard Pilgrim (2011)	NR	• No RCTs found • 1 HTA (Hayes 2010) • Guidance document (NICE 2007b) • 1 SR (N = 19 studies) • 8 case-series (N = 735, %f/u NR, 10 months–2.3 years f/u)	Arthroscopic hip surgery for FAI is covered when ALL the following criteria have been met: • Pain unresponsive to medical management • Positive impingement sign with sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation • Imaging studies confirming FAI • No evidence of advanced OA (Tonnis grade II or III) and the absence of severe chondral damage	• There is increasing evidence in clinical literature to support arthroscopy for FAI. The best results are seen in younger patients with little to no osteoarthritis and the absence of severe chondral damage. Studies show significant improvement in quality of life and pain symptoms in patients.
United Healthcare	NR	• No RCTs or prospective comparative studies found	Surgical treatment, both arthroscopic and open, for FAI is medically necessary and covered. Best surgical outcomes are achieved in patients who have ALL of the following:	• Clinical evidence supporting the surgical treatment of femoroacetabular impingement (FAI) syndrome

Payer (year)	Lit search dates	Evidence base available*†	Policy	Rationale/comments
		• 1 HTA (Hayes 2010) • 2 guidance documents (NICE 2007a, 2007b) • 1 cohort (N = 75, f/u NR) • 15 case-series (N = 1086, %f/u NR, 6 months–2.7 years f/u)	• pain unresponsive to medical management (e.g., restricted activity, nonsteroidal anti-inflammatory drugs); • moderate-to-severe persistent hip or groin pain that limits activity and is worsened by flexion activities (e.g., squatting or prolonged sitting); • positive impingement sign (i.e., sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation); • radiographic confirmation of FAI (e.g., pistol-grip deformity, alpha angle greater than 50 degrees, coxa profunda, and/or acetabular retroversion); • do not have advanced osteoarthritis (i.e., Tönnis grade 2 or 3) and/or severe cartilage damage (i.e., Outerbridge grade III or IV)	is limited and conflicting. UnitedHealthcare will continue to review clinical evidence surrounding the surgical treatment of FAI and may modify this conclusion at later date based upon the evolution of the published clinical evidence
Washington State Payers				

LifeWise Health Plan of Washington (2011)	through April 2010	• No RCTs or prospective comparative studies found
		• 2 guidance documents (NICE 2007a, 2007b)
		• 1 SR (N = 19 studies; mainly case-series)
		• 1 controlled cohort (N = 71, %f/u NR, 1 year f/u)
		• 11 case-series (N = 711, %f/u NR, 10 months–4 years f/u)
			Open or arthroscopic treatment of FAI may be medically necessary when all of the following criteria are met:
			• Age
			  • Adolescent patients should be skeletally mature with documented closure of growth plates
			  • Adult patients should be young enough to be considered inappropriate candidates for THA or other reconstructive hip surgery (e.g., < 55 years of age)
			• Symptoms
			  • Moderate-to-severe hip pain that is worsened by flexion activities (e.g. squatting or prolonged sitting) that significantly limits activities
			  • Unresponsive to conservative therapy for ≥ 3 months or conservative therapy is contraindicated (e.g., history of falls due to mechanical instability of hip joint).
			Conservative therapy for FAI should include:
			  • Activity modification including avoidance of hip stretching activities
			  • Restriction of athletics pursuits
			  • Avoidance of symptomatic motion
			  • Positive impingement sign on clinical examination (i.e., pain elicited with 90° of flexion and internal rotation and adduction of the femur
			• Imaging
			  • Morphology indicative of cam-type or pincer-type FAI, e.g., pistol-grip deformity, femoral head-neck offset with an alpha angle greater than 50 degrees, a positive wall sign, acetabular retroversion (overcoverage with crossover sign), coxa profunda or protrusion, or damage of the acetabular rim
				• Treatment of FAI is most effective in younger patients without osteoarthritis (Tonnis grade 0 or I) or severe cartilage damage.
				• There is a high association between FAI pathology and idiopathic osteoarthritis, but this may represent a small proportion of the total cases of hip osteoarthritis.
				• It is not known whether arthroscopic or open approaches result in better net health outcome when patients are matched for severity of FAI morphology and articular cartilage damage.
				• The evidence is insufficient to permit conclusions concerning the effect of surgical procedure on the development of osteoarthritis. Therefore, treatment of FAI morphology in the absence of symptoms is considered investigational

Payer (year)	Lit search dates	Evidence base available*†	Policy	Rationale/comments
			• High probability of a causal association between the FAI morphology and damage, e.g., a pistol-grip deformity with a tear of the acetabular labrum and articular cartilage damage in the anterosuperior quadrant • No evidence of advanced osteoarthritis, defined as Tonnis grade II or III, or joint space of less than 2 mm, except when there is mechanical instability • No evidence of severe (Outerbridge grade IV) chondral damage Surgical treatment of FAI is considered investigational for all other indications If femoroacetabular impingement (FAI) morphology is identified, patients should be advised not to play aggressive sports. No more frequent than annual follow-up with magnetic resonance (MR) arthrography may be indicated for FAI morphology to evaluate cartilage changes before damage becomes severe. Treatment of FAI should be restricted to centers experienced in treating this condition and staffed by surgeons adequately trained in techniques addressing FAI. Because of the differing benefits and risks of open and arthroscopic approaches, patients should make an informed choice between the procedures.	

BCBS: BlueCrossBlueShield; FAI: femoroacetabular syndrome; NA: not applicable; NR: not reported.
*Formal critical appraisals were not reported in any of the payer HTAs except Group Health. Percent follow-ups were not given for RCTs or case series. Mean time to follow-up is reported here.
†N reflects numbers before loss to follow-up.
‡Based on BCBSA policy 7.01.118, issued 4/09.

3. The Evidence

3.1. Methods of the Systematic Literature Review

3.1.1. Inclusion/exclusion

Inclusion and exclusion criteria are summarized in Table 4.

- *Population.*

  KQ1, 3-6: We included studies of patients undergoing operative or nonoperative treatment for FAI. Studies that included a heterogeneous population of FAI and other hip conditions that did not report results specific for FAI were excluded.

  KQ2: We included studies of patients undergoing operative treatment for FAI or studies of young patients receiving operative treatment for unspecified hip pain.

- *Intervention.*

  KQ1-6: Open dislocation, arthroscopic techniques and arthroscopically assisted mini-open techniques to correct FAI were included.

- *Comparator.*

  KQ1, 2: Not applicable.

  KQ3-6: Any nonoperative care was used as a comparison with any operative care. In addition, we included studies that used one type of surgery as a comparison to another type of surgery for FAI (e.g., open dislocation compared with arthroscopy for FAI). .

- *Outcomes.*

  KQ1, 2: We included studies that assessed reliability and validity of diagnostic criteria or outcomes measures.

  KQ3, 5, 6: We included studies that reported on at least one of the following outcomes: physical function/disability, pain, range of motion, return to work, development or progression of osteoarthritis, quality of life, activities of daily living, return to work, need for continuing and/or subsequent intervention (e.g., conversion to total hip arthroplasty).

  KQ4: Studies that reported on at least one of the following outcomes were included: perioperative adverse events or complications, revision surgery, heterotopic ossification, trochanteric nonunion, failure of labral refixation.

- *Study design.*

  KQ1: We included prospective studies listing inclusion criteria, and reliability/validity studies.
  KQ2: Reliability/validity studies.
  KQ3-5: We included comparative studies and case series.
  KQ6: Formal economic studies.

Table 4. Summary of inclusion and exclusion criteria

Study Component	Inclusion	Exclusion
Population	• Patients undergoing treatment for FAI	• Congenital hip dysplasia, slipped capital femoral epiphysis, Legg-Calve-Perthes
Intervention	• Operative treatment for FAI (open, arthroscopic, or combination)	
Comparator	• Nonoperative care (activity modification, NSAIDs, injections, etc)	
Outcomes	Short-term: • Functional outcome (patient- and clinician-reported hip scores) • Pain • Range of motion • Return to work • Complications/adverse events (safety) • Reoperation (safety) Long-term: • Conversion to THA • Function • Pain • Range of motion	• Non-clinical outcomes
Study Design	• Prospective studies listing inclusion criteria, and reliability/validity studies for KQ1 • Reliability/validity studies for KQ2 • Comparative studies and if need be, case series for questions 3-5. • Formal economic studies for question 6	• Case reports • Non-clinical studies
Publication	• Studies published in English in peer reviewed journals, published HTAs or publically available FDA reports • Full formal economic analyses (e.g. cost-utility studies) published in English in a HTAs or in a peer-reviewed journal published after those represented in previous HTAs	• Abstracts, editorials, letters • Duplicate publications of the same study which do not report on different outcomes • Single reports from multicenter trials • Studies reporting on the technical aspects of FAI surgery • White papers • Narrative reviews • Articles identified as preliminary reports when results are published in later versions • Incomplete economic evaluations such as costing studies

3.1.2. Data sources and search strategy

The clinical studies included in this report were identified using the algorithm shown in Appendix A. The search took place in four stages. The first stage of the study selection process consisted of a comprehensive literature search using electronic means and hand searching. We then screened all possible relevant articles using titles and abstracts in stage

two. This was done by two individuals independently. Those articles that met a set of *a priori* retrieval criteria based on the criteria above were included. Any disagreement between screeners that were unresolved resulted in the article being included for the next stage. Stage three involved retrieval of the full text articles remaining. The final stage of the study selection algorithm consisted of the selection of those studies using a set of a priori inclusion criteria, again, by two independent investigators. Those articles selected form the evidence base for this report.

Electronic databases searched included PubMed, EMBASE, CINAHL, ClinicalTrials.gov, CRISP, HSTAT, *The Cochrane Library*, EconLIT, PsychINFO, AHRQ, and INAHTA for eligible studies, including health technology assessments (HTAs), systematic reviews, primary studies and FDA reports. Reference lists of all eligible studies were also searched. The search strategies used for PubMed and EMBASE, are shown in Appendix B. Figure 1 shows a flow chart of the results of all searches for included primary studies. Articles excluded at full-text review are listed in Appendix C.

3.1.3. Data extraction

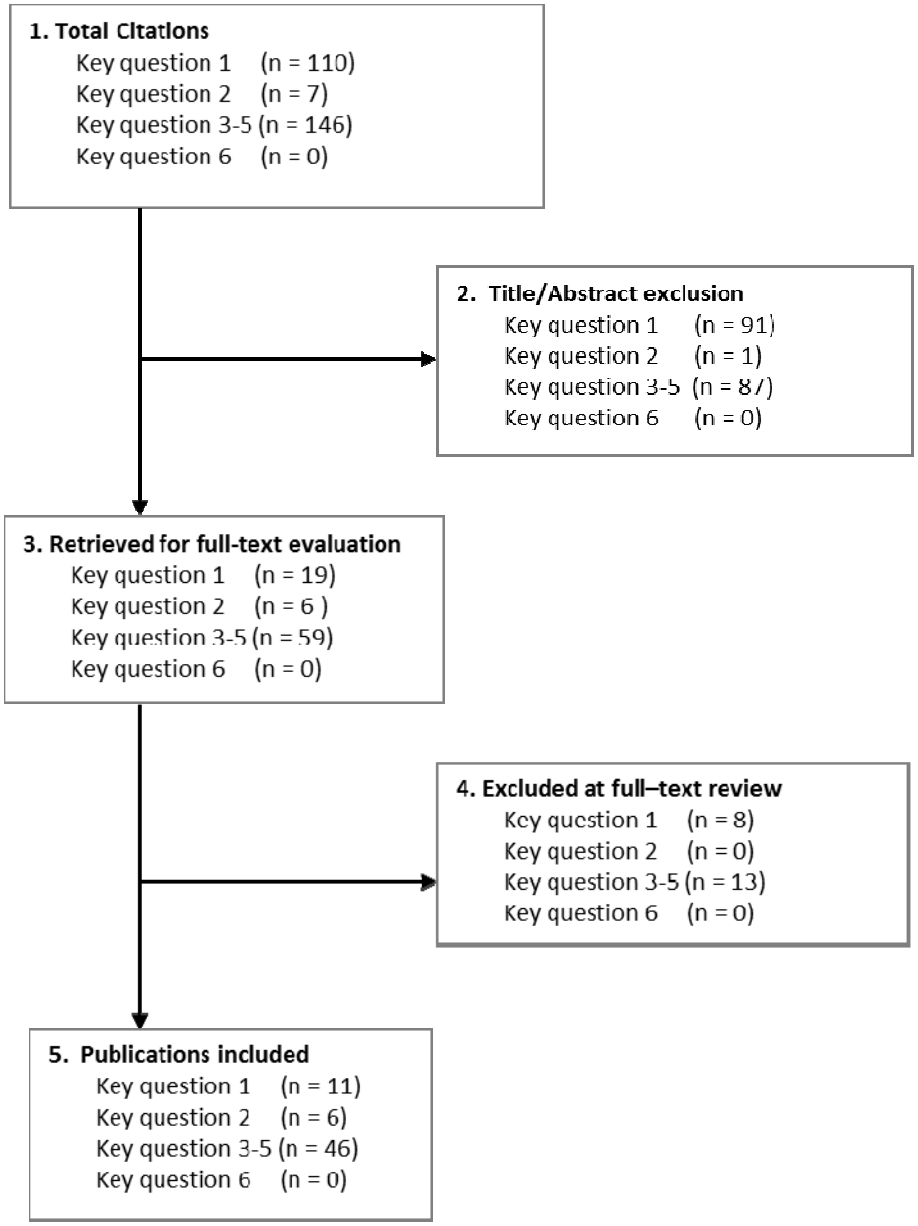
Reviewers extracted the following data from the included clinical studies: study population characteristics, study type, patient demographics, study interventions, follow-up time, study outcomes, complications/adverse events (reoperation, femoral neck fracture, avascular necrosis, trochanteric nonunion, heterotopic ossification, avascular necrosis, osteonecrosis, death, infection, deep vein thrombosis, pulmonary embolism, neurovascular injury, etc.). An attempt was made to reconcile conflicting information among multiple reports presenting the same data.

3.1.4. Study quality assessment: Level of evidence (LoE) evaluation

The method used by Spectrum Research, Inc. (SRI) for assessing the quality of evidence of individual studies as well as the overall quality of evidence incorporates aspects of the rating scheme developed by the Oxford Centre for Evidence-based Medicine, precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group, and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).

Details of the Level of Evidence (LoE) methodology are found in Appendix D. Each clinical/human study chosen for inclusion was given a LoE rating based on the quality criteria listed in Appendix D. Standardized abstraction guidelines were used to determine the LoE for each study included in this assessment.

Figure 1. Flow chart showing results of literature search



3.2. Quality of Literature Available

Quality of studies retained.

We initially found 263 citations using the search strategy in Appendix B.

For Key Question 1 we identified three prospective case series that identified inclusion and exclusion criteria for study entry, and two validity and 10 reliability studies.

For Key Question 2 we found two studies assessing reliability of hip outcome measures in an FAI population and four studies assessing reliability in a young population undergoing hip arthroscopy for pain.

For Key Questions 3 and 4 on effectiveness and safety we found a total of 36 case-series, all level of evidence (LoE) IV and six cohort studies (all LoE III). For effectiveness, 32 case series and five cohorts were included; for safety, 34 and six, respectively. Three case reports were found that addressed safety and were also included.

To address outcomes following FAI surgery in special populations (Key Question 5), we included three retrospective cohort studies (LoE III).

Tables summarizing the level of evidence can be found in APPENDIX E.

4. Results

4.1. Key Question 1:

Is there a consistent or agreed upon case definition for FAI? What is the evidence of validity and reliability of these case definitions?

4.1.1. Background

Ganz et al first described the clinical presentation and radiographic findings of FAI based on their series of over 600 surgical dislocations of the hip.[39] Since then, there have been a significant number of publications describing various clinical and imaging criteria for FAI. Although the specific criteria are not universally agreed upon, in general, most case series report that patients have pain at or around the groin associated with prolonged sitting, walking or athletic activities[61,106,120,140]; have loss of range of motion of the hip, primarily in flexion, internal rotation and adduction[7,12,64,73,118]; have a positive impingement test[24,77,85,123]; and have one or more imaging findings suggestive of morphological abnormalities. A number of morphological abnormalities have been described and include:

1. Loss of normal concavity of the anterosuperior region of the femoral head-neck junction[6,103]
2. Focal bony prominence to the femoral neck[74]
3. Flattening of the lateral aspect of the femoral head[74]
4. Asphericity of the femoral head[6,26,79]
5. Large alpha angle[6,28,113,126]
6. Reduced head-neck offset[6,14,23,64,83,96]

7. Pistol-grip deformity[4,43,74,130]
8. Crossover sign (COS)[42,65,66,68,69,74]
9. Posterior wall sign[4,74,123,136]
10. Excessive acetabular coverage[60,61,74,79]

4.1.2. Strategy to answer the question

To answer this key question, we first sought to identify and compare the inclusion criteria from all prospective studies evaluating the effectiveness of treatment for FAI. Inclusion and exclusion criteria of a clinical trial define the population of interest, in this case, those thought to have FAI. Secondly, we looked for studies that assessed the validity of the "diagnosis" of FAI using the patients' symptoms, clinical exam and imaging results either in combination or individually. For validity, we included only those studies that used visual inspection at the time of surgery as the reference standard for comparison against the test. Lastly, we search for studies whose purpose was to test the reliability of common clinical tests (e.g. impingement test) or imaging exams (e.g. alpha angle) believed to be important criteria for diagnosing FAI.

4.1.3. Inclusion criteria from prospective studies

We identified seven reports that appeared to be prospective. Three do not explicitly describe inclusion criteria for the study.[32,42,62,112] In the remaining four studies, one was in a population of professional male hockey players[110], mean age of 27 years (range 18 to 36), average duration of symptoms was 19 months (range, 1.5 to 99); and three were in a population of young, mostly male patients, mean age of 34 years (range 17 to 54)[34], 41 years (range 17 to 66)[58], and 42 years (range, 18 to 67).[126] Pain and a positive impingement test are two inclusion criteria specified in three of the four studies, Table 5. The study by Horisberger et al[58] states that the diagnosis was by clinical exam but did not specify the exam components. All four studies included a positive impingement test. All four included a positive imaging study to confirm the diagnosis. The α-angle was used in three of the studies to diagnose cam FAI: >50º in two studies[58,126] and >55º in the other.[110] One study listed range of motion or limited hip motion as an inclusion criterion, but did not state the criteria of what defines "limited" motion.[58]

Table 5. Inclusion and exclusion criteria in prospective studies assessing treatment for FAI

	Horisberger et al 2010	Philippon et al 2010	Fletcher et al 2011	Stahelin et al 2008
INCLUSION CRITERIA				
Pain	"symptomatic"	yes, preventing hockey play	yes, mechanical pain in inguinal fold, buttocks or trochanter area	no
Length of pain	no	no	≥6 months	no
Failed non-operative treatment	no	no	yes	no
Positive impingement	yes	yes	yes	yes
Type of FAI	cam or mixed	cam, pincer or mixed	not stated	cam only
Positive imaging sign	yes, osseous bump, α-angle >50º	cam: abnormal bone characteristics at the head-neck junction AND α-angle >55º pincer: coxa profunda OR coxa protrusion OR acetabular retroversion	yes, but unspecified	yes, α-angle >50º
Limited hip motion	yes (IR, flex, adduction), though specific criteria defining limited motion not described	no	no	
EXCLUSION CRITERIA				
Osteoarthritis	Tönnis grade III	no	Tönnis grade II or III	Tönnis grade III
Previous surgery	yes	no	no	yes
Precedent trauma	yes	no	no	no
Other		Retired hockey players Non English speakers		coxa profunda, protrusio coxae, or a crossover sign

4.1.4. Validity studies

Two studies attempted to assess validity: one assessed the clinical exam against the diagnosis of FAI,[87] and one evaluated the impingement test and the α-angle, separately.[85]

Diagnosis of FAI using the clinical exam

Martin et al recently assessed the ability of six experienced orthopedic surgeons to agree on a diagnosis of labral tear, FAI or capsular laxity in eight patients with musculoskeletal hip-related pathologies from clinical exam alone.[87] Patients were a small group all deemed in need of surgery as determined by a treating physician who had additional access to radiographs and magnetic resonance (MR) arthrography. Each surgeon examined the

patients as they would in their own practice. The final diagnoses were made at the time of surgery. Overall, the diagnoses obtained by the experienced surgeons from the clinical exam had only a 65% agreement with that made from surgical inspection, Table 6. The authors concluded that the agreement with the correct diagnosis was low, and that clinical examination techniques used for diagnosing hip pathologies need to be improved and standardized.

Diagnosis of FAI using the impingement test or the α-angle

Lohan et al performed a retrospective study to determine the ability of the impingement test or the α-angle to predict a diagnosis of cam-FAI.[85] Seventy-eight patients identified through chart review had surgical (open or arthroscopic) and test data available and therefore made up the study population. Those with prior hip surgery, post-traumatic deformity, Legg-Calve-Perthes disease, osteonecrosis, advanced osteoarthritis, slipped capital femoral epiphysis or hip dysplasia were excluded. Patients were designated to have a cam-type FAI if they received a femoral head-neck junction bony osteochondroplasty/arthroscopic femoral debridement (n = 39). All other patients receiving other interventions were designated as not having FAI (e.g., labral or articular cartilage debridement or repair, joint wash-out).

The diagnostic accuracy (i.e. validity) of the impingement test as identified from the medical chart was evaluated against the surgical diagnosis.[85] The impingement test is a clinical test performed with the patient supine and the hip passively flexed to 90º, internally rotated and adducted. A positive test reproduces the symptoms. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the impingement test were 76.9%, 87.2%, 85.7% and 79.1%, respectively, Table 6. The predictive values (if the persons tests positive or negative, what is the probability that he or she has or does not have the condition) is dependent on the prevalence of the disease, and the prevalence in this study may not represent the prevalence in other settings. The authors concluded that the impingement test proved to be more valuable than imaging tests in suggesting the presence or absence of cam-FAI.

The α-angle was also assessed for its diagnostic ability.[85] The α-angle is a measure first described by Nötzli et al in an attempt to quantify abnormalities of the anterior femoral head-neck junction.[103] It is a measurement of the asphericity of the femoral head-neck junction and is defined as the intersection of the following two lines at the center of the femoral head: (1) line drawn along the femoral neck so that it passes through the center of the femoral head, and (2) line drawn from the center of the femoral head to the point where the femoral head sphericity ends. Larger angles represent diminished concavity at the junction suggestive of cam-type FAI. Three different observers measured the α-angle from MR arthrograms. The

mean sensitivity, specificity, PPV and NPV among the 3 observers was 39.3%, 70.1%, 54.7% and 53.5%, Table 6. The authors concluded that the α-angle performed poorly and was statistically of no value in suggesting the presence or absence of cam-FAI.

Table 6. Summary of the validity of clinical tests and imaging findings commonly described in diagnosing FAI compared with findings at surgery.

	Sensitivity	Specificity	Additional measurements
Clinical exam (Martin et al[87])			% agreement: 65% (6 surgeons)
Impingement test (Lohan et al[85])	76.9%	87.2%	PPV: 85.7% NPV: 79.1%
α-angle MR arthrography (Lohan et al[85])	Observer #1: 35.8% Observer #2: 43.5% Observer #3:38.5% Mean: 39.3%	Observer #1: 69.2% Observer #2: 79.5% Observer #3:61.5% Mean: 70.1%	PPV: Observer #1: 46.1% Observer #2: 68.0% Observer #3:50.0% Mean: 54.7 NPV: Observer #1: 51.9% Observer #2: 58.5% Observer #3: 50.0% Mean: 53.5%

FAI: femoroacetabular impingement; MR: magnetic resonance; NPV: negative predictive value; PPV: positive predictive value

4.1.5. Reliability studies

We identified 10 studies that purposed to assess reliability (intraobserver and/or interobserver) of various tests. One study assessed a clinical test (impingement sign), one an imaging diagnosis of FAI and eight evaluated the presence or absence of one or more individual imaging tests. These are summarized in Table 7.

The results of reliability studies are presented in terms of a kappa statistic (κ, for nominal or ordinal data) or the intraclass correlation coefficient (ICC, for continuous data). Each of these statistics has variations depending on certain assumptions of the data. In the studies represented below, the details of assumptions were not disclosed; rather, the studies reported only the coefficients. While there are some who argue against interpreting the coefficients using a standard "rule of thumb",[15] for the sake of this report we describe the strength of agreement following the standards as proposed by Landis and Koch[76] for the kappa and apply this equally to the ICC[35] as follows:

≤0 = poor
.01–.20 = slight
.21–.40 = fair

.41– .60 = moderate
.61–.80 = substantial
.81–1 = almost perfect

Clinical signs

Impingement test

One study reported the interobserver reliability of the impingement test between a physical therapist and orthopedic surgeon in 68 patients with hip pain who were referred to a specialty clinic for evaluation for surgery.[91] The therapist and surgeon had extensive training in conducting the test over a 2 month period with approximately 25 patients. The interobserver reliability in this study was only moderate ($\kappa = 0.58$, 95% confidence interval, 0.29 to 0.87). The intraobserver reliability was not examined.

Diagnosis

Imaging diagnosis

One study evaluated the reliability of the imaging diagnosis using six observers with tests performed six weeks apart in a group of patients with normal hips (n = 25), FAI (n = 25), or acetabular dysplasia (n = 27).[23]

- First, the reliability of making one of four diagnoses (normal, FAI, dysplasia, or combined FAI + dysplasia) was assessed: the intraobserver reliability was substantial ($\kappa = 0.61$; 95% confidence interval, 0.56 to 0.67), however, only one of the six readers had "good or excellent reliability" (not defined); the interobserver reliability was also substantial ($\kappa = 0.80$).
- Next, the reliability of making one of three diagnoses (normal, FAI or combined FAI + dysplasia, or dysplasia) was evaluated in an attempt to discern the ability of the readers to diagnosis whether FAI was present or absent (even it was present in combination with features of hip dysplasia). Both the intraobserver and interobserver reliabilities were moderate (($\kappa = 0.56$, 95% confidence interval, 0.48 to 0.65) (with no readers having good or excellent reliability) and ($\kappa = 0.46$), respectively).

Imaging measurements

α- angle

The reliability of the α-angle measurement was assessed by four studies[6,26,72,85] using x-ray radiographs in three studies[6,26,72] and magnetic resonance (MR) arthrography in one study.[85] Overall, measurements of the α-angle had moderate to high intraobserver reliability (range, 0.60 to 0.98) and moderate to high interobserver (range, 0.52 to 0.95) reliability, though the

estimates varied by study, by imaging view and by imaging modality (radiograph, MRI), Table 7.

- The α-angle as measured on plain radiographs from the:
  - Anteroposterior (AP) view had moderate to high intraobserver reliability (range, 0.60 to 0.88) and high interobserver reliability (range, 0.85 to 0.95) as reported by two studies.[6,26]
  - Cross table lateral view had substantial to high intraobserver (range, 0.63 to 0.95) and moderate to high interobserver reliability (range, 0.52 to 0.85) as measured by the same two small studies.[6,26]
  - Dunn view had high intraobserver and interobserver reliability (0.98 and 0.90, respectively) as reported by one study.[6]
  - Frog leg view had high intraobserver reliability (range, 0.88 to 0.98) and moderate to high interobserver reliability (range, 0.78 to 0.83) as reported by two studies.[26,72]
- The α-angle as measured on MR arthrograms had moderately high intraobserver agreement as reported by one study, which employed three different observers (range of intraobserver agreement, 81 to 88%).[85]

Head-neck offset

The femoral head-neck offset was assessed in two studies [23,26] and is classified in relation to the posterior femoral head-neck junction based on the gross appearance of the curvature radius at each location. Specifically, symmetric concavity was defined when the anterior and posterior concavities appeared generally symmetric. Moderate decreased head-neck offset was defined when the concavity at the anterior head-neck junction had a radius of curvature greater than that of the posterior head-neck junction. Prominence was defined as when the anterior head-neck junction had a convexity, as opposed to a concavity.

- Clohisy et al (2007) reported substantial intraobserver (range between radiographic views, 0.63 to 0.74) and moderate interobserver (range between views, 0.52 to 0.53) reliability for measuring the head-neck offset.[26]
- Clohisy et al (2009) reported the reliability of classifying the head-neck offset as symmetric, moderately decreased, or prominent (as described above).[23] There was fair to moderate intraobserver reliability (range, 0.30 to 0.55) and slight to fair interobserver reliability (range, 0.19 to 0.24) between six observers.

Pistol-grip deformity

Pistol-grip deformity is present when there is substantial loss of the waisting as well as flattening of the normal concavity of the head-neck junction.[74] Laborie et al (2011) reported

substantial intraobserver (range of two observers, 0.65 to 0.78, with ≥ 3 months between tests) and high interobserver reliabilities (0.84 between two observers) of recording the presence or absence of a pistol-grip deformity on either the AP or frog-leg view (both views available).[74]

Focal prominence

Laborie et al (2011) reported substantial intraobserver reliability (range of two observers, 0.65 to 0.77) and high interobserver reliability (0.84) of reporting a positive or negative focal prominence, which was defined as a convex bump at the femoral neck.[74]

Head sphericity

The sphericity of the femoral head is typically assessed using a circular template that corresponds to the size of the femoral head. If the femoral head-neck junction extends outside the circular template (one study used a threshold of at least 2 mm[23]) and extends in a convex shape at the base of the neck, the femoral head is considered to be aspherical.[23,26] Reliability was evaluated by three studies[23,26,72] using measurements taken on plain radiographs. Overall, assessment of sphericity of the femoral head had moderate to high intraobserver (range, 0.55 to 0.99) and interobserver (range, 0.46 to 0.82) reliability.

- Sphericity of the femoral head as measured on plain radiographs from the:
    - Anteroposterior (AP) view had moderate to high intraobserver (range, 0.55 to 0.98) and interobserver reliability (range, 0.46 to 0.78) as reported by three studies (N = 77–350).[23,26,74]
    - Crosstable lateral view had moderate to high intraobserver reliability (range, 0.55 to 0.98) and moderate interobserver reliability (range, 0.41 to 0.66) as reported by two studies.[23,26]
    - Frog leg view had moderate to high intraobserver reliability (range, 0.57 to 0.99).[23,26,72] Two studies reported interobserver reliability, which ranged from moderate to high (range, 0.44 to 0.82).[23,26] As might be expected, the study with lower interobserver reliability employed six observers while the study with high interobserver reliability used only two observers.

Alternatively, one study reported asphericity as flattening of the lateral aspect of femoral head. Intraobserver reliability was moderate to substantial (range, 0.55 to 0.77) for two observers; interobserver reliability was substantial ($\kappa = 0.76$).[74]

Crossover sign

The crossover sign is a measure of acetabular retroversion. A hip is considered to be positive for the crossover sign when the posterior rim crosses the line that represents the anterior rim prior to the lateral edge of the weight bearing zone.[116] One study referred to the crossover sign as "acetabular version".[23] Four studies assessed the reliability of the crossover sign using plain radiographs taken in the AP view. The intraobserver reliability of this radiographic evaluation ranged from fair to high among 15 observers total (range, 0.33 to 1.00), and the interobserver variation was also fair to high (range, 0.39 to 0.82).[23,65,69,74]

Posterior wall sign

The posterior wall sign is a measure of insufficient posterior femoral head coverage and is considered positive when the visible outline of the edge of the posterior wall of the acetabulum descends medial to the center of the femoral head.[116] Reliability was evaluated by two studies.[69,74] The intraobserver reliability ranged from moderate to high among seven different observers (range, 0.48 to 0.95), while the interobserver was substantial to high (range, 0.63 to 0.83).

Ischial spine sign

The ischial spine sign is classified as positive when the ischial spine projects into the pelvic cavity.[65] Reliability was examined by two studies[65,69], one of which first defined this sign and examined two aspects of it: the total length of the ischial spine extending into the pelvic cavity (PRIS 1), and the total length of the ischial spine (PRIS 2[65]). The observers in the other study only evaluated whether this sign was positive or negative. Intraobserver reliability among seven observers ranged widely from fair to high (range, 0.38 to 0.92), and interobserver reliability was moderate to high (range, 0.54 to 0.91).

Acetabular coverage

Extensive acetabular coverage is indicative of pincer-type impingement, and is considered to be present if the lateral acetabular rim extends inferiorly and/or laterally.[74] Reliability of assessing the presence or absence of extensive acetabular coverage was examined by one study on plain AP radiographs.[74] The intraobserver reliability was moderate to substantial for two observers (κ range, 0.49 to 0.71) while the interobserver reliability was substantial (κ = 0.75).

Acetabular depth, acetabular inclination, and pelvic rotation

Other less commonly used radiographic measurements were assessed for reliability (but not validity), including acetabular depth[23], acetabular inclination[23], pelvic rotation[23] (see Table 7 for details).

SUMMARY

Consistent or agreed upon case definition

- The most consistent case definition of FAI (cam or mixed) as defined by inclusion/exclusion criteria in prospective studies of treatment effectiveness included the following:
    - hip/groin pain
    - positive clinical impingement test
    - α-angle >50-55º

Evidence of validity and reliability

- There is no evidence that the diagnosis of FAI can be obtained from clinical exam. One clinical test, the impingement sign, had a positive and negative predictive value of 86% and 79% in one study where the prevalence of FAI was 50%; however, in another study, the reliability of the impingement sign was only moderate.

- Even though the α-angle showed moderate to high interobserver reliability in several studies, it had poor diagnostic value in identifying FAI. Other imaging tests assessing abnormalities of the femur and acetabulum had variable degrees of reliability, but no others were tested for diagnostic validity.

Table 7. Summary of reliability coefficients for clinical tests and imaging commonly described in diagnosing FAI

		Intraobserver reliability coefficient (95% CI)	Interobserver reliability coefficient (95% CI)	Test condition (# observers, time between test and retest)	Patient population (normal, FAI, dysplasia) (number of patients)
Clinical exam					
Impingement test	Martin[91]		κ = 0.58 (0.29, 0.87)	2 observers, 1 hour	Hip pain (n = 68)
Imaging diagnosis					
"FAI", "dysplasia", "combined FAI + dysplasia", or "normal"	Clohisy,[23]	κ = 0.61 (0.56, 0.67)	κ = 0.80	6 observers, 6 weeks	FAI (n = 25); normal (n = 25) dysplasia (n = 27)
"FAI OR "combined FAI + dysplasia", "dysplasia only", or "normal"	Clohisy,[23]	κ = 0.56 (0.48, 0.65)	κ = 0.46	6 observers, 6 weeks	FAI (n = 25); normal (n = 25) dysplasia (n = 27)
α-angle					
AP view	Barton [6] Clohisy[26]	ICC*: 0.88 (0.75, 0.95)† ICC = 0.60	ICC: 0.95 (0.91, 0.97)† ICC = 0.85	2 observers, 4 weeks 2 observers, 2 weeks	FAI (N = 68) FAI (n = 61); normal (n = 24)
Crosstable lateral	Barton [6] Clohisy[26]	ICC*: 0.95 (0.88, 0.98)† ICC = 0.63	ICC*: 0.85 (0.75, 0.91)† ICC = 0.52		
Dunn view	Barton [6]	ICC*: 0.98 (0.95, 0.99)†	ICC*: 0.90 (0.83, 0.94)†		
Frog leg view	Clohisy[26] Konan[72]	ICC = 0.98 κ = 0.88 (0.71, 0.90)	ICC = 0.78 ICC = 0.83 (0.69, 0.89)‡		
MR arthrography	Lohan[85]	observer #1 agreement = 81% observer #2 agreement = 85% observer #3 agreement = 88%		3 observers, 2 weeks	Acetabular labral or cartilage abnormalities (N = 78)
Head-neck offset					
AP view	Clohisy[26] Clohisy,[23]	ICC = 0.63 κ = 0.43 (0.36, 0.50)	ICC = 0.53 κ = 0.24	2 observers, 2 weeks 6 observers, 6 weeks	FAI (n = 61); normal (n = 24) FAI (n = 25); normal (n = 25) dysplasia (n = 27)
Crosstable view	Clohisy,[23] Clohisy [26]	κ = 0.30 (0.23, 0.37) ICC = 0.73	κ = 0.22 ICC = 0.52		
Frog leg lateral	Clohisy,[23]	κ = 0.55 (0.48, 0.62)	κ = 0.19		

		Intraobserver reliability coefficient (95% CI)	Interobserver reliability coefficient (95% CI)	Test condition (# observers, time between test and retest)	Patient population (normal, FAI, dysplasia) (number of patients)
view	Clohisy[26]	ICC = 0.74	ICC = 0.52		
Pistol-grip deformity					
Positive on AP and/or frog leg view	Laborie [74]	observer #1 κ = 0.65 observer #2 κ = 0.78	κ = 0.65	2 observers, 3+ months	Primarily normal (N = 350)
Focal prominence					
Positive on AP and/or frog leg view	Laborie[74]	observer #1 κ = 0.65 observer #2 κ = 0.77	κ = 0.84		
Head sphericity					
AP view	Clohisy,[23]	κ = 0.56 (0.48, 0.63)	κ = 0.46	6 observers, 6 weeks	FAI (n = 25); normal (n = 25) dysplasia (n = 27)
	Clohisy [26]	ICC = 0.98	ICC = 0.78	2 observers, 2 weeks	FAI (n = 61); normal (n = 24)
Crosstable lateral	Clohisy,[23]	κ = 0.55 (0.45, 0.64)	κ = 0.41		
	Clohisy [26]	ICC = 0.98	ICC = 0.66		
Frog leg view	Clohisy,[23]	κ = 0.60 (0.52, 0.68)	κ = 0.44		
	Konan[72]	κ = 0.57 (0.48, 0.71)		4 weeks	FAI (N = 32)
	Clohisy [26]	ICC = 0.99	ICC = 0.82		
Flattening of the femoral head					
AP view	Laborie[74]	observer #1 κ = 0.55 observer #2 κ = 0.77	κ = 0.76	2 observers, 3+ months	Primarily normal (N = 350)
Crossover sign					
AP view	Clohisy,[23]	κ = 0.46 (0.37, 0.55)	κ = 0.39	6 observers, 6 weeks	FAI (n = 25); normal (n = 25) dysplasia (n = 27)
	Kappe[69]	observer #1 κ = 0.33 observer #2 κ = 0.85 observer #3 κ = 0.68 observer #4 κ = 0.70 observer #5 κ = 1.00	κ = 0.51	5 observers, 6 weeks	Non-arthritic hip pain (N = 20)
	Kalberer[65]	Reliability: 0.83	Reliability: 0.65	2 observers, 1+ week	Mixed population (normal, FAI dysplasia, Perthes) (N = 26)
	Laborie[74]	observer #1 κ = 0.59	κ = 0.82	2 observers, 3+ months	Primarily normal (N = 350)

		Intraobserver reliability coefficient (95% CI)	Interobserver reliability coefficient (95% CI)	Test condition (# observers, time between test and retest)	Patient population (normal, FAI, dysplasia) (number of patients)
		observer #2 κ = 0.80			
Posterior wall sign					
AP view	Kappe[69]	observer #1 κ = 0.75 observer #2 κ = 0.89 observer #3 κ = 0.59 observer #4 κ = 0.48 observer #5 κ = 0.95	κ = 0.63	5 observers, 6 weeks	Non-arthritic hip pain (N = 20)
	Laborie[74]	observer #1 κ = 0.55 observer #2 κ = 0.73	κ = 0.83	2 observers, 3+ months	Primarily normal (N = 350)
Ischial spine sign					
AP view	Kappe[69]	observer #1 κ = 0.58 observer #2 κ = 0.68 observer #3 κ = 0.60 observer #4 κ = 0.38 observer #5 κ = 0.90	κ = 0.54	5 observers, 6 weeks	Non-arthritic hip pain (N = 20)
AP view (PRIS 1)§	Kalberer[65]	Reliability: 0.92§	Reliability: 0.91§	2 observers, 1+ week	Mixed population (normal, FAI dysplasia, Perthes) (N = 26)
AP view (PRIS 2)**	Kalberer[65]	Reliability: 0.87**	Reliability: 0.77**		
Excessive acetabular coverage					
AP view	Laborie[74]	observer #1 κ = 0.49 observer #2 κ = 0.71	κ = 0.75	2 observers, 3+ months	Primarily normal (N = 350)
Acetabular depth††					
AP view	Clohisy,[23]	κ = 0.61 (0.54, 0.68)	κ = 0.39	6 observers, 6 weeks	FAI (n = 25); normal (n = 25) dysplasia (n = 27)
Acetabular inclination‡‡					
AP view	Clohisy,[23]	κ = 0.73 (0.68, 0.79)	κ = 0.64	6 observers, 6 weeks	FAI (n = 25); normal (n = 25) dysplasia (n = 27)
Pelvic rotation					
AP view	Clohisy,[23]	κ = 0.57 (0.50, 0.65)	κ = 0.21	6 observers, 6 weeks	FAI (n = 25);

	Intraobserver reliability coefficient (95% CI)	Interobserver reliability coefficient (95% CI)	Test condition (# observers, time between test and retest)	Patient population (normal, FAI, dysplasia) (number of patients)
				normal (n = 25) dysplasia (n = 27)

AP: anteroposterior; FAI: femoroacetabular impingement; ICC: intraclass correlation coefficient, κ: kappa; MR: magnetic resonance
* These data were reported as "reliability" in the study, but actually refer to ICC (personal correspondence with one of the authors).
† % CI not specified.
‡ A subset of 10 radiographs were evaluated for interobserver reliability.
§ Measurement of the prominence of the ischial spine: specifically, the amount of the ischial spine that could be seen extending medially into the pelvis inlet.
** Measurement of the prominence of the ischial spine: specifically, the total width of the ischial spine that could be seen (even if its radiopaque shadow fell behind the pelvic brim).
†† Acetabular depth was assessed by determining the relationship of the floor of the fossa acetabuli and the femoral head in relation to the ilioischial line; hips classified as "profunda" (floor of fossa acetabuli tohced or fell medially to the ilioischial line) or "protrusio" (medial edge of the femoral head fell medially to the ilioischial line) were considered to be at risk of pincer-type impingment.
‡‡ Acetabular inclination was labeled as normal, increased, or decreased based on the degree of the Tönnis angle; hips with a decreased Tönnis angle are considered to be at risk for pincer-type FAI while those with an increased angle are thought to be at risk for structural instability.

4.2. Key question 2:

What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Has clinically meaningful improvement in outcomes been defined in FAI?

The goals of FAI surgery include preventing or delaying osteoarthritis (OA) of the hip and total hip arthroplasty in the long-term, and improving function and restoring activity in the short-term. With respect to osteoarthritis of the hip, this outcome is often determined by radiography, most commonly using the Tönnis grading system. The Tönnis classification system has four grades:

- Grade 0: No signs of OA
- Grade 1: Increased sclerosis, slight joint space narrowing, no or slight loss of head sphericity
- Grade 2: Small cysts, moderate joint space narrowing, moderate loss of head sphericity
- Grade 3: Large cysts, severe joint space narrowing, severe deformity of the head, or evidence of necrosis.

With respect to identifying improved function and restoration of activity, patient- and clinician reported functional outcomes measures are often employed.

4.2.1. Long-term outcome, osteoarthritis

The Tönnis classification

We found no study that sought to validate the Tönnis classification for hip arthritis. We identified one study, Clohisy et al 2009[23], that evaluated the intraobserver and interobserver reliability of the Tönnis classification in a series of 77 patients with diagnoses of cam, pincer, or combined FAI (n = 25); acetabular dysplasia (n = 27), or normal hips (n = 25). Six orthopedic surgeons independently evaluated preoperative plain radiographs of all patients based on written definitions of the Tönnis classification system. The observers were blinded to patients' characteristics as well as to their signs and symptoms. The time between test and retest was 6 weeks during which the observers were instructed to not discuss their findings. At the retest, each observer reassessed the radiographs, which were scrambled in order and numbering. Four different radiographic views were available: anteroposterior (AP) pelvis, crosstable lateral, frog-leg lateral and false profile views. The combined intraobserver reliability (kappa value) was 0.60 (95% CI, 0.54, 0.66), and the interobserver reliability was 0.59. Further, only three of the six observers had a $\kappa \geq 0.50$. Therefore, good reproducibility was not been demonstrated for the Tönnis grading system of osteoarthritis.

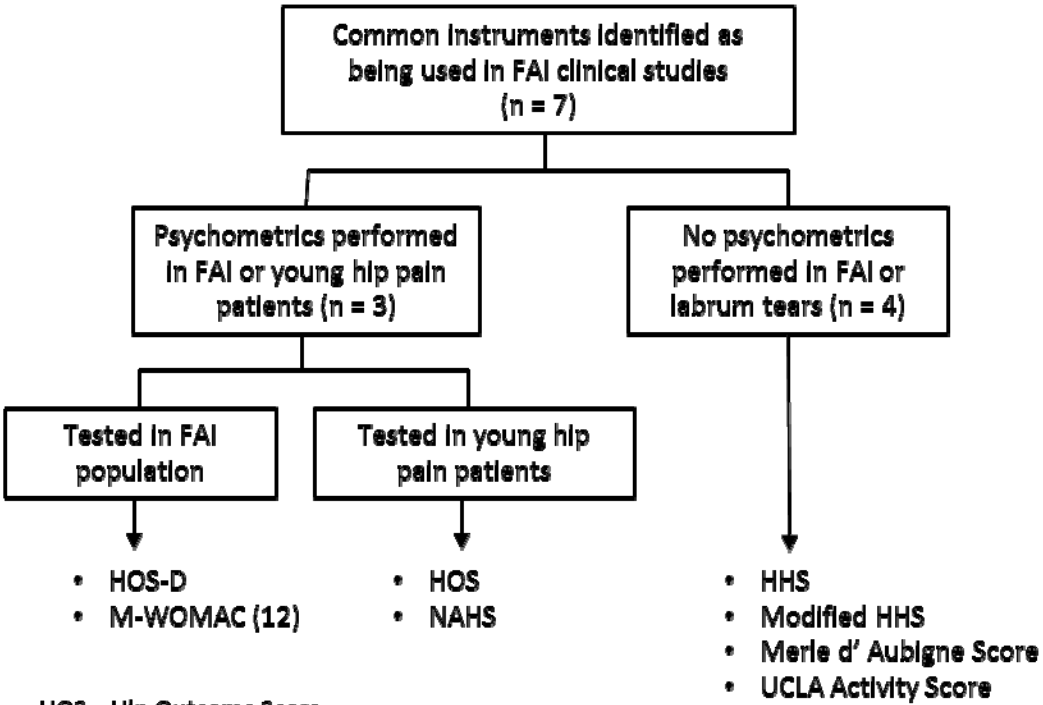
4.2.2. Short-term outcomes, patient- and clinician-reported functional outcomes

We identified seven commonly used functional hip outcome measures used in the FAI patient population, Figure 2:

1. Hip Outcome Score (HOS)/ German version (HOS-D)
2. 12-item modified Western Ontario and McMasters Universities Osteoarthritis Index (M-WOMAC (12))
3. Nonarthritic Hip Score (NAHS)
4. Harris Hip Score (HHS)
5. M-HHS (modified HHS)
6. Merle d'Aubigne Score (MA)
7. UCLA Activity Score

Of these, two outcome measures have been tested for validity in FAI patients: HOS-D and M-WOMAC (12). In addition, two outcome measures have been validated in young hip-pain population: HOS and NAHS. The latter instruments are described below and additionally summarized in Table 8. Thus, a total of three instruments have been validated in either FAI or young hip-pain patients: HOS/HOS-D, M-WOMAC, and NAHS (Figure 2)

Figure 2. Functional outcomes measures commonly reported in studies on FAI patients.



We evaluated the HOS/HOS-D, M-WOMAC, and NAHS based on the following quality criteria:[102,131]

- **Validity.** Validity evaluates whether an outcome instrument measures what it was intended to measure.[102] We evaluated three aspects of validity:

  - *Content validity* evaluates whether the outcomes of interest are comprehensively represented by the questions in the instrument.[102,131] We gave the studies credit if there was a clear description of each of the following: the aim of the outcome measure, the target population, the concepts being assessed, and the method by which the items were selected. In addition, the population of interest (and either investigators or experts) should have been involved in item selection.[131]

  - *Criterion validity* refers to whether the scores relate to a "gold standard" on the same theme[102,131]; for credit, we looked for a correlation with the gold standard of at least 0.70.[131]

  - *Construct validity* evaluates whether scores relate to other measures in accordance with specific hypotheses that are theoretically derived. The instrument of interest and another related outcome measure may have convergent (high correlation if they measure similar concepts) or divergent (low correlation if they measure different concepts) validity with one another.[102,131] For credit, specific hypotheses need to be stated, and 75% or more of the results should be consistent with these hypotheses as tested in at least 50 patients.[131]

- **Reliability** evaluates the extent to which repeated measurements in stable patients (test-retest) yield similar responses.[102] There are two aspects of reliability:

  - *Internal consistency* assesses whether the items in the questionnaire are correlated, in that they evaluate the same concept.[131] Questions should correlate highly with one another and with the overall (sub)scale score.[102] For credit, factor analysis should be performed on a minimum of 100 patients to determine whether the construct is uni- or multidimensional; Cronbach's alpha should range from 0.70 to 0.95 for each subscale, which is an indication of good internal consistency.[131]

  - *Reproducibility* measures whether patients can be differentiated from each other in spite of measurement error (relative measurement error).[102,131] For credit, the ICC (intraclass correlation coefficient) or weighted Kappa coefficient should be ≥ 0.70 when measured in at least 50 patients. The Pearson correlation coefficient is

not an adequate measurement of reliability, as it does not account for systematic differences.[131]

- **Responsiveness** assesses whether a questionnaire is able to detect clinically important changes over time (i.e., the score changes with the status of the patient).[102,131] For credit, ***one*** of the following should be demonstrated:
  1. SDC < MIC[131], where:
     - SDC (smallest detectable change) = 1.96 x √2 x SEM (standard error of measurement); thus the SDC is the smallest intraperson change in score that should be interpreted as "real" change, or change greater than measurement error.[131]
     - MIC (minimal important change) is defined as "the smallest difference in score in the domain of interest which patients perceive as beneficial and would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management".[131] MIC may also be written as MCID (minimal clinically important difference).

  2. MIC should be outside the limits of agreement (LOA)[131]
     LOA = mean change in scores of repeated measurements ± 1.96 x standard deviation of the changes[131]
  3. RR (responsiveness ratio) > 1.96[131]
  4. AUC ≥ 0.70[131]
     AUC (area under the ROC (receiver operating characteristics) curve) measures whether a questionnaire is able to differentiate between patients who have and have not changed, as measured by some other criteria (usually the patient's own perception of change)[131]

- **Floor or ceiling effects** are absent if the lowest or highest possible score, respectively, was reached by less than 15% of patients. Credit is given if no floor or ceiling effects are found in a sample size of 50 patients or more.[131]

- **MCID** (minimal clinically important difference, see MIC under responsiveness above for definition) assesses whether the authors reported the MCID for the questionnaire based on comparisons with patient-reported evaluation of overall outcome (i.e., function).

A summary of the quality assessment of the HOS, NAHS, and 12-item modified WOMAC may be found in Appendix F, and detailed information is provided below.

Table 8. Quality assessment of outcome measures evaluated in the FAI, labral tear, and hip arthroscopy populations.

Instrument	Validity			Reliability				
	Content validity	Criterion validity	Construct validity	Internal consistency	Reproducibility	Floor/ceiling	Responsiveness	MCID
FAI population								
HOS-D[94]	NR	NR	+	+	+/-	+	NR	NR
12-item modified WOMAC[119]	+/-	NR	+/-	+/-	NR	NR	NR	NR
Young hip-pain population								
HOS[88-90]	–	NR	+	+	+/-	NR	+	+
NAHS[21]	+	+	+	+/-	+/-	+	NR	NR

Table adapted from Lodhia et al. (2011)[84] and Terwee et al. (2007)[131]

"+" indicates criteria were met, "+/-" indicates the quality assessment was inadequate or indeterminate, "-" indicates the criteria were not met; NR indicates the quality assessment was not reported or performed.

HIP OUTCOME SCORE (HOS)

Martin et al. (2006)[88] created the Hip Outcome Score (HOS) as an instrument that could be used to assess outcomes in patients with acetabular tears. Like FAI patients, this population functions across a spectrum of abilities. The HOS is summarized in Table 9. Briefly, it consists of two subscales, activities of daily living (ADL) (19 items, 2 items not scored) and sports (9 items). Each item is graded on a 5-point Likert scale (and also includes an unscored option of "non-applicable" if something other than their hip pathology affects the patient's answer). The ADL and sports subscales are graded separately: the total score is divided by the maximum possible score (based on the number of questions answered), and the result multiplied by 100 to obtain a percentage. A higher the score represents better function. In addition, the questionnaire asks the patients to rate their current level of function during their usual ADL and sports activities on a scale from 0 to 100 for each, with 100 indicating the patients' level of function prior to hip problems. The patients are also asked to indicate their overall current level of function as "normal", "nearly normal", "abnormal", and "severely abnormal".

The HOS has been evaluated by four different studies and in different patient subsets, including FAI patients by Naal et al. (2011)[94]:

The HOS was initially validated by Martin et al. (2006)[88] in a prospective cohort of 507 patients with a primary diagnosis of labral tear; all patients were under the care of a single orthopedic surgeon. The HOS questionnaire was filled out at a regularly scheduled appointment. The mean age of these patients was 38 ± 13 years (range, 13 to 66 years), and 45.8% were male. The mean duration of symptoms was 3.4 ± 5 years (range, 11 days to 29

years). Half of the patients (52%) had undergone arthroscopic surgery for their labral tear at a mean of 6.7 months (range, 2 days to 3.86 years) prior.[88]

In order to demonstrate validity of the HOS as a longer term outcome measure in patients who had undergone hip arthroscopy, Martin et al. (2007)[89] published a second prospective cohort study of patients who underwent hip arthroscopy at least two years prior to completing the HOS questionnaire; 107 of the 337 patients (34%) who were mailed the questionnaire returned the completed form and met the inclusion criteria. Males comprised 48% of this patient set, and the mean age was 42 ± 14 years (range, 14 to 79 years). Arthroscopy had been performed a mean of 3.1 ± 0.49 years ago (range, 2 to 4.6 years).

In 2008, Martin et al. published a third study,[90] a retrospective cohort study of prospectively collected data. Patients who underwent hip arthroscopy by the senior author and who were prospectively part of a larger ongoing study with preoperative and postoperative records available to review were evaluated for inclusion. After review of the HOS questionnaires, only those patients who fit into a predefined "change" or "stable" group were included. To meet the criteria for the "change" group, patients had to categorize their postoperative improvement (compared with their pre-injury function) as "much improved" or "somewhat improved" and rate their level of function as "normal" or "nearly normal"; 108 patients were in the "change" group. Patients who qualified for the "stable" group (n = 18) rated their postoperative improvement as "unchanged" or "abnormal" and assessed their level of function as "abnormal" or "severely abnormal". A total of 126 patients were included in the study and had a mean age of 41 ± 16 years (range, 13 to 80 years), and 47% were male. Patients completed their follow-up HOS questionnaires at a mean of 7 months ± 96 days (range, 55 to 420 days) post-operation.

Finally, a German language version of the HOS (HOS-D) was validation-tested in 85 preoperative FAI patients in a recent prospective study by Naal et al. (2011).[94] Consecutive FAI patients who were scheduled for surgery were included; no other limitations were placed regarding inclusion and exclusion criteria. Patients were at an average age of 33 ± 12 years, and 58% were male. The HOS-D questionnaires were mailed to patients two to three weeks prior to the scheduled surgery and returned at admission. In addition, a subset of 33 patients volunteered to complete the questionnaire again upon admission for use for the reliability assessment.

Quality assessment of the HOS/HOS-D:

- *Content validity* was not demonstrated for the HOS, as patients were not involved in item selection during the development of the outcome measure.[88]

- The purpose of the HOS of the outcome instrument is to evaluate treatment outcomes in acetabular patients who function across a wide range of ability levels.[88]
- The HOS was designed to evaluate function. It reports the overall level of patient-perceived function as well as separate scores that distinguish function in activities of daily living (ADL) from that in sports-related activities. The questions adequately evaluate these different domains and include activities that cover a wide range of abilities. For example, the ADL subscale evaluates a patient's ability to sit for 15 minutes as well as the ease with which they are able to perform heavy work (pushing/pulling, climbing, carrying).[88]
- Items were selected based on the input from physicians and physical therapists that treat patients with musculoskeletal hip disorders.[88] However, the target population was not involved in the item selection, which is necessary for demonstration of content validity.[94]

- *Criterion validity* was not tested (no information found).

- *Construct validity* was demonstrated in three different studies.[88,89,94]
  - Convergent validity was demonstrated in the initial cohort of 507 labral tear patients[88] by evaluating the association between the HOS and SF-36 physical function subscale and SF-36 physical component summary score. The Pearson correlation coefficients between the ADL subscale and the two SF-36 outcomes were 0.76 and 0.74, respectively; that between the sports subscale and the respective SF-36 physical function scores were 0.72 and 0.68.
  - Convergent validity of the HOS was subsequently tested in 107 patients who had undergone hip arthroscopy at least two years prior.[89] The Pearson correlation coefficients between the HOS ADL subscale and SF-36 physical function outcomes described above were 0.86 and 0.80, respectively; those between the HOS sports subscale and the same SF-36 outcomes were 0.84 and 0.81, respectively.
  - Divergent validity was assessed in the 507 labral tear patients[88] by examining the correlation between the HOS and the SF-36 mental health subscale, and SF-36 mental component summary score. Pearson correlation coefficients between the ADL subscale and SF-36 scores were 0.27 and 0.18, respectively; between the sports subscale and SF-36 scores, the values were 0.23 and 0.1, respectively.
  - Divergent validity was additionally tested in those who had received hip arthroscopic surgery at least two years before completion of the questionnaire (n = 107).[89] Between the HOS ADL subscale and the SF-36 mental health outcomes (as described above), the Pearson correlation coefficients were 0.41 and 0.17,

respectively, and were 0.43 and 0.18 between the HOS sports subscale and the same SF-36 mental health outcome scores.

- Validity was also evaluated among patients two or more years post arthroscopy (n = 107)[89] based on reported differences in their activity level ("normal" (23%), "near normal" (42%), "abnormal" (24%), and "severely abnormal" (6%) (question not answered by 5%)). One-way ANOVA and post hoc comparison demonstrated that there were significant differences between each of the groups in the ADL and sports subscales ($P < .0005$ for each), with better patient-perceived activity level correlating with higher scores. Similarly, there were statistically meaningful differences between the scores from those patients who reported an excellent/good surgical outcome (80%) compared with those who reported a fair/poor outcome (20%) for both the ADL and sports subscales ($P < .0005$ for each). Finally, patients above the median age of 44.2 years had higher scores for both of the HOS subscales compared with those below the median age ($P < .0005$ for each). Detailed scores for each of these groups are located in Appendix F.

- The German language version of the HOS demonstrated convergent and divergent validity in 85 pre-surgical patients. Convergent validity was tested by correlating HOS-D scores with several different outcome measures, including: UCLA activity scale; WOMAC total score; WOMAC pain, function, stiffness domains; Oxford Hip Score, and SF-12 physical component scale. Divergent validity was evaluated by comparing the HOS-D scores with the SF-12 mental component scale. In all cases and for both the ADL and sports subscales, convergent ($P < .001$) and divergent ($P > .45$) validity was shown. Data are reported in Appendix F.

- *Reliability*
  - *Internal consistency* was demonstrated for the HOS/HOS-D.[88,94]

    1. Martin et al. (2006)[88] performed exploratory factor analysis using PRELIS software. For both the ADL and sports subscales, data from the 430 (85%) and 343 (68%) patients, respectively, with no missing responses were compared with that from the other patients in the study (i.e., those with missing responses). No differences were found for the subscale for age, length of symptoms, time since surgery (for those who had undergone surgery), and current function rating, with a α value set at 0.005 (due to the large number of comparisons). A significant difference was found for gender for the sports, but not the ADL, subscale ($P < .0005$), and the authors suggested that this was due to the lower ratio of females to males in the group of patients with no

missing data compared to those patients with one or two missing responses. Factor analysis showed that the items in the ADL subscale loaded on two items (items 3 (putting on socks and shoes) and 11 (sitting), which were changed to unscored items. The remaining 17-item ADL subscale loaded on one factor that accounted for 68% of the variance. The 9-item sports subscale loaded on one factor, which accounted for 80.3% of the variance. The 17-item ADL and 9-item sports subscales were used for subsequent analyses.

2. Martin et al. (2006)[88] calculated Cronbach's coefficient to be 0.96 and 0.95 for the 17-item ADL and 9-item sports subscales, respectively, indicating high correlations among the items in the scale as tested in labral tear patients.

3. Naal et al. (2011)[94] reported high Cronbach coefficient alpha values for both the ADL (0.95) and sports (0.91) subscales of the HOS-D in 85 preoperative FAI patients, indicating good internal consistency in FAI patients.

- *Reproducibility* was not adequately demonstrated for the HOS/HOS-D, as both studies[90,94] tested for reliability in less than 50 patients.

1. Martin et al. (2008)[90] assessed reproducibility in the "stable" group of patients (n = 18), however, it has been recommended that a group of at least 50 patients are needed to demonstrate test-retest reliability.[131] These 18 "stable" patients filled out the HOS questionnaire preoperatively and again at a mean of 7 months post-arthroscopy and rated their level of activity (compared with their preoperative functioning) as "unchanged" or "abnormal" and assessed their level of function as "abnormal" or "severely abnormal". Test-retest reliability was assessed in this group of patients by using their preoperative compared with their postoperative HOS scores. The ICC (intraclass correlation coefficient) values for the ADL and sports subscales were 0.98 and 0.92, respectively, with MDC (minimal detectable change) values of ± 3 for both subscales.

2. Naal et al. (2011) evaluated reproducibility in a subset of 33 FAI patients, who filled out the HOS-D questionnaire twice preoperatively at a median interval of 10 days. The ICC values were high for both the ADL and sports subscales, at 0.94 (95% CI, 0.98 to 0.97) and 0.89 (95% CI, 0.80 to 0.95), respectively.

- *Floor/ceiling* effects of the HOS were tested by Naal et al. (2011) in 85 pre-surgical FAI patients[94]; there were no floor or ceiling effects when the MDC was not considered.

    - When the minimal detectable difference (MDC) was taken into account, the ADL subscale showed no floor effect but had a ceiling effect in 21% of patients. The sports subscale had both floor and ceiling effects in 18% and 12% of patients, respectively.[94]
    - However, when the MDC was not considered and only the lowest (0 points) and highest (100 points) possible scores were evaluated, there were no floor or ceiling effects for either subscale (0%), the ceiling effects were low for the ADL (2%) and sports (1%) subscales.[94] Terwee et al. (2007) consider floor and ceiling effects to be absent when less than 15% of a sample of 50 or more patients is affected.

- *Responsiveness* was demonstrated for the HOS by Martin et al. (2008)[90] using three different methods. The mean difference in ADL subscale scores in the "change" (n = 108) group was 22.4 ± 18 (range, –57 to 76) points and in the "stable" (n = 18) group was 3.7 ± 7.3 (range, –24 to 100) points. The 2-way ANOVA (analysis of variance) was significant ($P < .0005$). Similarly, the mean difference in the sports subscale scores in the "change" and "stable" groups were 34.5 ± 26.2 (range, –24 to 100) and –3.7 ± 13.3 (range, –25 to 30) with a significant ANOVA ($P < .0005$). The effect sizes for the ADL and sports subscales were calculated to be 1.2 and 1.5, respectively, which are considered to be large effect sizes. Finally, the AUCs (area under the curve) were calculated for the ADL and sports subscales to be 0.88 (95% CI, 0.80 to 0.95) and 0.90 (95% CI, 0.83 to 0.97), respectively; an AUC of more than 0.70 is considered to be an adequate indicator of responsiveness.[131]

- *MCID* was defined by Martin et al. (2008) in a population of 126 patients at a mean of 7 months post-arthroscopy. A minimal clinically important difference (MCID) of 9 points was associated with a sensitivity of 0.82 and a specificity of 0.89 for the ADL subscale. For the sports subscale, a MCID of 6 points was associated with a sensitivity of 0.85 and a specificity of 0.87.[90]

MODIFIED 12-ITEM WESTERN ONTARIO AND MCMASTERS UNIVERSITIES OSTEOARTHRITIS INDEX (WOMAC)

Rothenfluh et al. (2008)[119] conducted a prospective evaluation of the internal construct validity of the WOMAC osteoarthritis index in patients with FAI or osteoarthritis (OA) of the hip. Patients were included in the FAI group on the basis of a positive impingement test,

radiographic imaging showing decreased offset, femoral head asphericity, or deep retroverted acetabulum) and MRI images, and the absence of radiographic signs of OA of the hip. Those in the OA group had a positive diagnosis by clinical exam and radiography with a Tönnis grade of more than 1. Patients who had undergone prior hip surgery or had comorbidities that affected ambulation or caused pain were excluded from the study. The German version of the full WOMAC was distributed to 200 patients at their first office visit, and 157 questionnaires were completed. Of these, 100 were filled out by FAI patients and with mean age of 31.7 ± 9.7 years (45% male); 57 were completed by OA patients with a mean age of 60.3 ± 11.7 years (49% male). An age- and gender-matched set of 200 patients without hip pain was also included for some analyses; these patients had a mean age of 32.6 ± 5.6 years (49.5% male) and served as a control for the FAI group. The authors noted that a high percentage of FAI and healthy control patients participated in sporting activities (83% and 85%, respectively) compared with OA patients (35%).

Items on the questionnaire given to the patients were scored on a 7-point Likert scale, as opposed to the 5-point scale on the original WOMAC. This was done in an attempt to increase the resolution of the scores. However, the authors found that the 7-point scale showed threshold disordering and thus needed to rescore all patient answers back to a 5-point scale (using RUMM2020 software).[119]

Rasch analysis was performed on the WOMAC in order to determine whether this questionnaire had internal construct validity in FAI and OA patients.[119] The authors first found that the original (rescored, German version of the) total WOMAC had misfit to the Rasch model for both FAI and OA patients, as well as both patient groups combined. The authors concluded that a total WOMAC score is multidimensional in both these patient sets, and thus its use is not supported. Detailed scores can be found in Supplementary Table X. Subsequent analysis was done separately for the pain and function subscales. Furthermore, the data of FAI and OA patients were combined into a FAI + OA patient group since patients with FAI and OA responded similarly.

Rothenfluh et al. (2008)[119] next used a progressive, step-by-step process to generate a modified 12-item WOMAC that was unidimensional (i.e., fit the Rasch model) in the combined FAI + OA patient group (n = 157). Briefly, first the pain and function subscales were analyzed separately for adequate fit. One item was removed from the pain subscale, item 3 (pain sitting/lying); in addition, item 4 (pain walking flat) was rescored from a 5-point to a 4-point scale. Six items were removed from the function subscale, items 7 (lying in bed), 20 (bending), 21 (putting on socks), 22 (taking off socks), 24 (heavy chores), and 25 (light chores). The stiffness subscale was also eliminated from the modified WOMAC, as it only contains two items and this subscale showed misfit to the Rasch model. Once the two reduced pain and function subscales were recombined into a 15-item scale, three additional

items showed misfit to the model and were eliminated: pain item 1 (night pain) and function items 11 (getting out of bed) and 18 (getting on/off toilet). Additional details regarding the removal of items to improve overall fit of the questionnaire to the Rasch model may be found in Appendix F.

An analysis of the resulting modified 12-item WOMAC was then performed in the FAI patient group (n = 100); some analyses were also performed on the OA group (n = 57) and the healthy patient (FAI control) group (n = 200). The following quality assessment evaluates these results.

Quality assessment of the modified 12-item WOMAC:

- *Content validity* was not adequately demonstrated for the 12-item WOMAC.[119]
  - The purpose of the modified WOMAC is to evaluate patients with FAI and OA of the hip.
  - Overall, the outcome measure is designed to evaluate function, and does so by evaluating both pain and physical function. As described above, this questionnaire was derived from the original WOMAC, and the remaining questions seem to evaluate these different domains in a manner appropriate for FAI and OA patients.
  - The WOMAC was selected as a starting point for generating an outcome measure with good internal construct validity as it is frequently used to evaluate functional outcomes. Items were selected for good fit to the Rasch model of unidimensionality as described above. However, patients received the original WOMAC and were not involved in the selection of items to be included in the modified version of this outcome measure.[119]

- *Criterion validity* was not evaluated (no information found).[119]

- *Construct validity* was inadequate. Although differences in overall scores were shown for patients with different diagnoses (FAI patients compared with a matched normal population, FAI compared with OA patients), no formal hypothesis was stated for the expected differences between these groups, which could lead to bias.[119]
  - The overall mean scores were lower (indicating less disability) in the normal population (FAI control) (n = 200) compared with FAI patients (n = 100) (0.39 ± 2.90 versus 8.32 ± 7.32, respectively; $t = -8.5269$; $P < .001$). The effect size (r) was 0.71, which is greater than 0.5 and thus indicates a large difference between populations.
  - The overall scores of the OA patients (n = 57) were substantially higher (indicating greater disability) than those of the FAI population (n = 100) (16.23 ± 8.04 versus 8.32 ± 7.32, respectively; $t = -7.7034$; $P < .001$).

- *Reliability*
  - *Internal consistency* was inadequate. Although the Rasch model demonstrated unidimensionality, credit was not given, as Cronbach's coefficient alpha was not calculated.[119]

    1. Factor analysis was performed as described above so that all 12 items in the modified WOMAC demonstrated unidimensionality; 157 patients were assessed. The final mean item fit residual was –0.111 ± 1.045; the mean person fit residual was –0.318 ± 1.286, the chi square interaction value was 25.534 ($P = 0.377$). Furthermore, a PSI (person separation index) of 0.93 suggests high internal consistency reliability for individual patients. T-tests were performed as a last test of unidimensionality and were significant at 6.62%.[119]

    2. Cronbach's coefficient alpha was not calculated.[119]

  - *Reproducibility* was not tested (no information found) .[119]

- *Floor/ceiling* effects were not evaluated (no information found) .[119]

- *Responsiveness* was not evaluated (no information found) .[119]

- *MCID* was not defined (no information found) .[119]

NONARTHRITIC HIP SCORE (NAHS)

The Nonarthritic Hip Score (NAHS) was developed by Christensen et al. (2003)[21] in order to provide a scoring system that evaluated hip pain and function specifically in young (20 to 40 years of age) and active patients with hip pain and without a clear radiographic diagnosis. This patient-reported outcome measure was derived in part from the WOMAC index and evaluates function with four domains: pain (5 items, 20 points), physical function (5 items, 20 points), mechanical symptoms (4 items, 16 points), and level of activity (6 items, 24 points). The final score is obtained by multiplying the total points by 1.25. Total scores range from 0 to 100, with the maximum score (100 points) indicating normal hip function. A summary of the NAHS may be found in Appendix F.

Christensen et al. (2003)[21] prospectively evaluated the properties of the NAHS in consecutive patients with hip pain of at least six months' duration that had not responded to non-steroidal anti-inflammatory drugs, therapy, or injections. The questionnaire was filled out at (and/or

before) the first office visit. Internal consistency and validity was evaluated in 48 consecutive patients who ranged in age from 16 to 45 years (mean age: 33 years); 40% of these patients were male (n = 19). All 48 patients completed the questionnaire, and the mean (preoperative) NAHS score was 56.0 ± 18.1 (range, 12.5 to 92.5). Reproducibility was assessed in an additional 17 consecutive patients (11 females and six males) with a mean age of 32 years (range was not reported). The mean time between the test and retest was 5.5 days (range, 1 to 16 days).

Quality assessment of the NAHS:

- *Content validity* was demonstrated for the NAHS.[21]
    - As stated above, the purpose of the outcome instrument was to provide a highly sensitive scoring system that is able to discriminate between high levels of activity in the target population of young (20 to 40 years of age), active patients with hip pain and no obvious radiographic diagnosis.[21]
    - Overall, the outcome measure is designed to evaluate function, and does so by evaluating four different domains: pain, physical function, mechanical symptoms, and level of activity. The questions adequately evaluate these different domains. For example, the questions that address activity evaluate the patient's ability to participate in high and low demand sports, walking and jogging for exercise, and the ability to perform heavy and light household duties.[21]
    - Items were selected based on the input from patients, surgeons, physical therapists, and epidemiologists. All items that evaluate pain and function were taken from the WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index). Preliminary testing was conducted in patients with a range of educational levels as well as in health professionals and did not affect the content of the questionnaire.

- *Criterion validity* was demonstrated for the NAHS, which was shown to have excellent correlation with the HHS ("gold standard").[21]
    - The Pearson correlation coefficient was 0.82 between the NAHS and the Harris Hip Score (HHS), both of which were completed preoperatively. Pearson correlation coefficients between the domains of the NAHS and the HHS were also calculated: pain: r = 0.73; physical function: r = 0.73; mechanical symptoms: r = 0.61; activity level: r = 0.76. The authors chose the HHS as a "gold standard" against which the NAHS was measured and noted that although the HHS is a well-accepted measure of hip function, it has not been validated in this patient population.

- *Construct validity* was demonstrated for the NAHS, which, according to the authors, had good correlation with the SF-12.[21]
  - The SF-12 (Short Form-12) is a widely accepted measure of global health status. The Pearson correlation coefficient between the NAHS and (what is presumably the overall score for) the SF-12 was 0.59; the Pearson correlation coefficient was 0.37 and 0.51 between the NAHS and the physical and emotional subscales of the SF-12, respectively.[21]

- *Reliability*
  - *Internal consistency* was inadequate.

    1. Factor analysis was not performed.[21]
    2. Cronbach's coefficient alpha was calculated for each of the four domains on 48 consecutive patients, and ranged from 0.69 to 0.92 (pain: $\alpha = 0.87$; function: $\alpha = 0.85$; mechanical symptom: $\alpha = 0.69$; activity level: $\alpha = 0.92$).[21] For credit to be given factor analysis should be performed on at least 100 patients and Cronbach's alpha should range from 0.70 to 0.95.[94]

  - *Reproducibility* was inadequately evaluated in an additional subset of 17 patients; the mean time between the test and retest was 5.5 days (range, 1–16 days).[21] The authors assessed reliability by calculating the Pearson correlation coefficient, which was 0.96 for the NAHS as a whole. The Pearson correlation coefficient was also measured for each of the four domains: pain, 0.92; physical function, 0.92; mechanical symptom, 0.87; and activity level, 0.95. While these numbers suggest good agreement between test results in stable patients over time, the Pearson correlation coefficient does not account for systematic differences and is not an adequate measure of reliability.[94] In addition, at least 50 patients need to be evaluated in order to demonstrate reliability for the NAHS.[94]

- *Floor/ceiling* measurements were not reached, as no patients scored the highest or lowest possible scores (scores ranged from 12.5 to 92.5).[21] Ideally, 50 patients would be tested[94]; we gave credit as 48 patients were evaluated.

- *Responsiveness* was not tested (no information found).[21]

- *Interpretability* was not evaluated (no information found).

SUMMARY:

- The Tönnis classification is often used to determine the extent of osteoarthritis in the hip. There were no studies found that assessed its validity. Reliability was tested in only one study and intra- and interobserver reliability in that study was moderate.
- Three patient- reported outcomes measures commonly used in FAI patients have undergone psychometric analysis in FAI (HOS-D, M-WOMAC) or young hip-pain (HOS, NAHS) patient populations.

  – Validity: only one (NAHS) of the three instruments was adequately tested for validity, and it was performed in a young hip-pain patient population. Content validity was inadequate for the other two (HOS, M-WOMAC) instruments primarily because patients were not involved in item selection; criterion validity was not tested for the same two instruments. Construct validity was demonstrated for both the HOS/HOS-D and the NAHS, but was inadequately tested for the M-WOMAC as no hypothesis was made as to expected differences in scores between patient groups.

  – Reliability was inadequately tested for all three outcome measures. While good internal consistency was shown in both the FAI and young hip-pain patient populations for the HOS/HOS-D, reproducibility of this instrument was inadequate as too few patients were tested.

  – The MCID was defined to be 9 points for the ADL subscale and 6 points for the sports subscale of the HOS-D in FAI patients. The MCID has not been defined for any other outcome measures in FAI or young hip-pain patients.

4.3. Key Question 3

What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI? Including consideration of short-term and long-term:

- Need of or time to total hip arthroplasty
- Development or progression of osteoarthritis
- Impact on function, pain, range of motion, quality of life, activities of daily living and return to work
- Other reported measures

4.3.1. Efficacy

We considered randomized controlled trials as providing evidence on efficacy. We found no randomized controlled trials (RCTs) comparing surgery with conservative care for FAI or comparing different surgical treatments for FAI.

4.3.2. Effectiveness

We identified one study that retrospectively compared conservatively treated patients versus those receiving FAI surgery versus patients having a total hip arthroplasty.[62] In addition we found four comparative studies which investigated the effectiveness of various surgical treatments for FAI: labral debridement versus labral refixation (two studies)[33,78] and osteoplasty versus no osteoplasty (two studies).[3,100] The first study by Jager et al poorly describes the selection of patients so that it was not possible to tell how the treatment and control groups were obtained. The last four studies use historical controls. The results of these studies should be taken with caution. The fact that these studies (1) are retrospective cohorts mostly using historical controls, (2) did not clearly account for all excluded patients, and (3) only included patients who completed follow-up or who had complete clinical and radiographic data creates the potential for selection, performance and attrition bias. Selection bias is an inherent problem with cohort studies since systematic differences arise from self-selection or physician-directed selection of treatments. In these cases, selecting patients for inclusion based on the completeness of the data in one's database is likely to produce a subset of patients that are different than patients not in the database but who received the treatment of interest. Performance bias in these studies is a real possibility due to the use of historical controls. For example, differences in the level and competency of care may exist between historical controls and those treated with more current and improved surgical methods or by surgeons who have acquired more experience over time. Finally, attrition bias can result when those who do not return for final follow-up are systematically different from those who remain in the study, thus changing the overall group characteristic in a way that is unable to be controlled or accounted for.

In addition to the cohort studies, we identified 32 case-series reporting clinical outcomes following various treatments for FAI. We report the findings stratified by (1) population (non- or recreational athletes and competitive/professional athletes) since these two populations may be substantially different, and (2) surgical procedure (arthroscopy, open dislocation and arthroscopy augmented with a mini-open procedure). In non- or recreational athletes, 14 studies reported on the effect of arthroscopy, seven on open dislocation, five on the mini-open approach, and one on conservative care. In competitive or professional athletes, four studies reported on the effect of arthroscopy, one on open dislocation, and no studies reported either mini-open or conservative care in this patient population. Details on each study can be found in Appendix G. In the report we provide a summary table using pooled estimates. For categorical data, we calculate pooled risks (%) and 95% confidence intervals weighted by sample size. For continuous data, we provide pooled percent mean differences comparing follow-up with baseline scores. We were unable to calculate a standardized effect measure or a confidence interval as many studies did not include standard deviations. We did not include in the pooled estimates those studies that were limited to adolescents[111] or patients 60 years or older[63], or those that did not provide preoperative scores for continuous data.[80,109] However, the individual data for these studies are presented in the text and can be found in Appendix G.

Conservative versus Arthroscopic or Open Surgery versus THA

Jager et al 2004 compared 17 patients (22 hips) with FAI who underwent three different treatments: nonoperative care with physiotherapy and anti-inflammatory cyclooxygenase-2 (COX-2) inhibitors, arthroscopy or open dislocation, and total hip replacement (THA).[62] There were nine patients (10 hips) in the nonoperative group, six patients (eight hips) in the FAI surgery group, and two patients (four hips) in the THA group. The authors gave no indication of how these patients were selected or how many patients overall may have been eligible for the study; they simply stated that radiographic findings of osseous bump deformities on the anterolateral head-neck junction were found in all patients along with typical symptoms of FAI. They did, however, admit that the treatment received was based according to clinical and radiographic findings and MRI, thus acknowledging the potential of confounding by indication. Those with moderate clinical symptoms but morphological signs of degenerative destruction of the hip joints underwent nonoperative treatment. Those with labral defects but only minor cartilage destruction on MRI underwent FAI surgery. The two patients who received THA did so as a result of having severe signs of osteoarthritis on radiographs. The two patients who received THA were the oldest (average age 49.5 years), followed by the conservatively treated group (34.5 years) and the arthroscopy/open dislocation group (27.3 years). Overall, males comprised 76.5% of the population (not described by group) and all patients had cam-type impingement. Average follow-up periods were 1.4 years for the conservatively treated patients, 1.8 years for patients treated by arthroscopy or open dislocation, and 2.2 years in patients who underwent THA. There is no

mention of independent or blind assessment of outcomes. This study is a retrospective cohort and does not provide any information regarding the patient selection process or loss to follow-up. There was no description of baseline characteristics apart from the mean age of patients. With respect to age, there were potentially important differences in ages of the patients among the three treatment groups. Only pain and return to work/sports are reported at final follow-up, with patients in the conservative group showing the poorest results overall: none were pain free at final follow-up compared with 100% of the patients in both surgical groups. Only 67% had returned to their previous work or sports level again compared with 100% of the patients in both surgical groups, Table 9. It is difficult to draw any conclusions from this study as the patient groups compared with clearly different in many characteristics.

Labral debridement vs. Labral refixation

Two studies compared labral debridement (historical controls) with labral refixation for the treatment of FAI.[33,78] Collectively, there were 54 patients (61 hips) and 69 patients (74 hips) in each group, respectively. Mean ages were 30.4 years (range, 16–57) and 28.8 years (range, 16–56) and males comprised 67.6% and 62.9% of the patients. The results of these two studies should be taken with caution. The fact that these studies are retrospective cohorts using historical controls, did not clearly account for all excluded patients and only included patients who completed follow-up creates the potential for selection, performance and attrition bias. Selection bias is an inherent problem with cohort studies since systematic differences arise from self-selection or physician-directed selection of treatments. Performance bias in these studies is a real possibility due to the use of historical controls. For example, differences in the level and competency of care may exist between historical controls and those treated with more current and improved surgical methods or by surgeons who have acquired more experience over time. Finally, attrition bias can result when those who do not return for final follow-up are systematically different from those who remain in the study, thus changing the overall group characteristic in a way that is unable to be controlled or accounted for.

Larson et al 2009 reviewed patients who underwent arthroscopic labral debridement during a period before the development of labral repair techniques and identified those who would have fulfilled the current criteria for labral refixation (a relatively healthy portion of the labrum was available for refixation without complex tearing, intralabral ossification, or calcification).[78] All patients included in the study had magnetic resonance imaging, plain radiographs, and detailed operative notes; however, it is unclear whether some patients were excluded from the study because they did not have this documentation. The remaining patients were compared with a cohort who underwent current arthroscopic labral refixation. Inclusion criteria were radiographic and intraoperative findings consistent with pincer or combined cam-pincer FAI, minimal to no radiographic degenerative changes, and a minimum of 1 year follow-up. During the time periods of interest, 34 patients (36 hips) that

underwent labral debridement and 37 patients (39 hips) that underwent labral refixation met the inclusion criteria. The debridement group tended to be slightly older and to have a greater proportion of males compared with the fixation group (mean age 31 years (range, 16–57) vs. 27 years (range, 16–56) and 73.5% male vs. 62.2% male). Mean follow-up was 1.8 years (range, 0.5–3.0) and 1.4 years (range, 0.5–2.0), respectively. Preoperative diagnosis was combined cam-pincer impingement in 83.3% and 84.6% of hips in the debridement and refixation groups, respectively, and isolated pincer impingement in 16.7% and 15.4%. Preoperative Tönnis grade 0 was in 72.2% and 76.9% of hips, grade 1 in 22.2% and 20.5% of hips and grade 2 in 5.6% and 2.6%, respectively. Intraoperative, the degree of acetabular chondral damage (Outerbridge classification) was as follows in the debridement and refixation groups, respectively: grade 1, 2.8% versus 5.1%; grade 2, 8.3% versus 2.6%; grade 3, 52.8% versus 46.2%; and grade 4, 33.3% versus 35.9%. Corresponding findings for femoral chondral damage were: 0% versus 2.6%; 13.9% versus 12.8%; 8.3 versus 0%; and 0% versus 0%. Following surgery, patients who underwent simple debridement were allowed to bear weight as tolerated with crutches as needed. Patients undergoing labral refixation were restricted to toe-touch weight bearing for 2 weeks with range of motion encouraged but avoiding extremes of external rotation. Outcomes for both groups were prospectively measured with the modified Harris Hip Score, Short Form 12, and visual analog scale (VAS) for pain preoperatively and postoperatively at 6 weeks, 3 months, 6 months, and 1 year. There was no mention of independent or blind assessment of outcomes. Failure, defined as a modified Harris Hip Score less than 70, subsequent debridement of a hip that had undergone labral fixation or conversion to THA occurred more often in the debridement compared with the refixation group: 11.1% and 7.7%, respectively. Conversion to THA was reported in none of the hips that underwent labral debridement and in one hip (2.6%) that had labral refixation; however, this patient had a 2.5 cm full-thickness acetabular chondral defect at the time of arthroplasty. Radiographic OA progression, measured by mean change in Tönnis grade pre- to postoperatively, was 0.3 in the debridement group and 0.25 in refixation group. Mean change in modified Harris Hip Scores from preoperative to 1 year was greater in the refixation group compared with the debridement group, 31.3 versus 23.9, translating to a percent improvement of 49.7% versus 36.8% (Table 9), and more excellent/good scores (MHHS > 80) were obtained in the refixation group compared to the debridement group: 89.7% (n = 35) versus 66.7% (n = 24). Both groups had similar percent improvement in pain at 1 year: 83.9% in the refixation group and 81.3% in the debridement group (decrease of 5.2 points on the VAS (0–10) for both treatments), Table 9. Percent improvement in SF-12 scores was higher in the refixation group, 40.3% versus 32.8%, Table 9.

Espinosa et al 2006 reviewed a cohort of 141 consecutive patients who underwent surgical dislocation of the hip for FAI and compared the first 20 patients (25 hips) who underwent labral debridement and acetabular rim resection with the next 32 patients (35 hips) who

underwent labral refixation after rim resection.[33] Inclusion criteria were complete preoperative and postoperative clinical scores and radiographic documentation. Exclusion criteria included open growth plates, age greater than 40 years, previous hip surgery, and professional or semi-professional athletes. Twenty-two percent of patients did not meet the inclusion criteria and 34% were excluded due to incomplete medical records, resulting in the remaining 52 patients. Average age of the entire population was 30 years (range, 20–40) and 63.5% were male. Whether the two groups were similar demographically was unable to be determined since they were not described separately. The groups did have similar preoperative clinical hip scores, pain score, and radiographic grades of osteoarthritis; however, the debridement group had more restricted internal-external rotation preoperatively compared with the refixation group (34º vs. 43º). FAI type was not reported. Postoperative rehabilitation was not mentioned. Clinical exams were performed preoperatively and at 1 and 2 years postoperatively by independent observers. Loss to follow-up was unable to be determined. Radiographic OA progression, measured by mean change in Tönnis grade pre- to postoperatively, was 0.7 in the debridement group and 0.3 in refixation group. The Merle d'Aubigné Hip Score was used to measure hip function. The debridement group showed less improvement at 2 years compared with the refixation group with percent change from pre- to postoperative scores of 25.0% versus 41.7%, respectively, Table 9. Furthermore, overall excellent/good results were obtained in fewer patients: 76.0% (n = 19) versus 94.3% (n = 33). Both groups showed significant pain improvement (Merle d'Aubigné pain score) at 2 years following treatment; however, the refixation group showed a decrease in pain over baseline scores of the debridement group, 4.1 points versus 2.6 points, respectively, Table 9. Internal-external rotation improved by 1 degree in the debridement group compared with 6 degrees in the refixation group for total range of motion of 35 and 49 degrees, respectively. Failure and conversion to THA were not reported.

No Osteoplasty versus Osteoplasty

Bardakos et al 2008 retrospectively reviewed patients who had undergone hip arthroplasty for labral pathology secondary to FAI and compared those who had not undergone osteoplasty with those who had undergone osteoplasty.[3] The first group represented an earlier time period in which femoral osteoplasty was not part of the standard treatment for FAI (historical control group). Inclusion criteria were cam impingement with the presence of an obvious "pistol-grip" deformity on plain anteroposterior radiographs or a clearly reduced anterior head-neck offset on a cross-table lateral view, and a minimum of 1 year follow-up. Exclusion criteria included the presence of pincer impingement, a history of hip fracture, previous hip surgery, hip dysplasia, osteonecrosis, sepsis, rheumatoid arthritis, Perthes' disease and osteoarthritis greater than Tönnis grade 2. It is unclear how many patients overall were eligible for the study. There were 47 patients (47 hips) in the no osteoplasty group and 24 patients (24 hips) in the osteoplasty group. Patients in the no osteoplasty group were slightly older and there were fewer males compared with the osteoplasty group: 35

years (27–46) versus 33 years (27–41) and 49% vs. 58%, respectively. The no osteoplasty group had slightly worse function preoperatively compared with the osteoplasty group: mean modified Harris Hip Score (MHHS) of 55 (37–72) versus 59 (52–64), respectively. All patients were enrolled in a formal physiotherapy program following surgery which included gait training, proprioception and range of movement (ROM) exercises. Muscle strengthening was allowed only after ROM had been recovered and full weight-bearing had been initiated. No high-impact activities were allowed for eight to 12 weeks. Clinical follow-up was obtained at 6 weeks, 6 months, and 1 year and annually thereafter, either by a follow-up visit or by a telephone interview conducted by a research coordinator. There is no mention as to whether this coordinator was blinded. Only preoperative and 1 year scores were used for this study. Mean change in MHHS from preoperative to 1 year was similar between groups, 22 in the no osteoplasty group and 24 in the osteoplasty group translating to percent mean changes of 40% and 40.7%, respectively, Table 9. However, more excellent/good scores (MHHS > 80) were obtained in the osteoplasty group compared to the no osteoplasty group: 83.3% (n = 20) versus 59.6% (n = 28), $P = 0.06$.

Nepple et al 2009 compared patients with isolated cam impingement who underwent arthroscopic partial labral resection and chondroplasty without osteoplasty with those who had an augmentation of this procedure with limited open (mini-open) osteochondroplasty.[100] Initially, isolated cam FAI was treated with arthroscopic partial labral resection and chondroplasty without addressing the bony structural abnormality, and it wasn't until later that osteochondroplasty became a standard addition to the treatment. A retrospective review of 221 consecutive arthroscopic procedures for FAI identified 48 patients (48 hips) with isolated cam FAI. Of these, 23 patients (23 hips) underwent partial labral resection and chondroplasty without osteoplasty (historical controls) and 25 (25 hips) underwent the modified combined surgery (plus osteochondroplasty). Exclusion criteria included less than 1 year of follow-up, pure pincer impingement, or combined cam-pincer impingement. The patients in the arthroscopic chondroplasty group were slightly older and there were fewer males compared with the osteochondroplasty group: mean 37 years versus 33 years and 52% versus 68%. The mean length of follow-up was slightly longer for the chondroplasty group (2.3 years) than osteochondroplasty group (1.7 years). Preoperative osteoarthritis Tönnis grades in the chondroplasty versus the osteochondroplasty group were grade 0 in 65.2% versus 72.0% of hips, grade 1 in 30.4% versus 28.0%, and grade 2 in 4.3% versus 0%, respectively. Acetabular chondral lesions grades 3 or 4 were present in 56.5% versus 84.0% of hips with chondroplasty and osteochondroplasty, respectively. Femoral chondral lesions grades 3 or 4 were present in 34.8% and 20% of hips, respectively. The average preoperative modified Harris Hip Score was 59.8 for the chondroplasty group and 64.5 for the osteochondroplasty group. Postoperative clinical outcomes were collected at 1 year, 2 year and longest follow-up (range, 1.0–3.9) and there was no mention of independent or blind assessment. Failure was defined as a modified Harris Hip Score of < 70 points or the need

for additional surgery. In the chondroplasty group, failure was reported in 21.7% of patients and conversion to THA was necessary in 8.7% compared with no patient in the osteochondroplasty group having either outcome, Table 9. Mean change in modified Harris Hip Score at last follow-up was 23.3 in the chondroplasty group and 26 points in the osteochondroplasty group, translating to similar percent improvements: 38.9% and 40.3%, respectively, Table 9. However, more excellent/good scores (MHHS > 80) were obtained in the osteochondroplasty group compared with the chondroplasty group: 92.0% (n = 23) versus 73.9% (n = 17).

Table 9. Comparative effectiveness of different surgical procedures for the treatment of FAI.

Author	Demographics	Outcome	Results			Mean Follow-up years (range)
			No. of cases (%)			
			Debridement n=34	Refixation n=37		
Larson 2009	Mean age:29 years (16–57) % male: 68 pincer (16%); mixed (84%)	Failure* Conversion to THA	4 (11.1) 0 (0)	3 (7.7) 1 (2.6)†		1.6 (1.0–3.0)
			Mean Change score (% change)			
		Tönnis score	0.30 (200)	0.25 (100)		
		Modified Harris Hip Score	23.9 (36.8)	31.3 (49.7)		
		VAS (0-10)	5.2 (81.3)	5.2 (83.9)		
Espinosa 2006	Mean age: 30 years (20–40) % male: 63	Tönnis score Merle d'Aubigné Pain score	0.7 (117) 2.6 (186)	0.3 (60) 4.1 (273)		2.0 (NR)
			No Osteoplasty n=47	Osteoplasty n=24		
Bardakos 2008‡	Mean age:29 years 34 (27–46) % male: 5 cam (100%)	Modified Harris Hip Score	22 (40.0)	24 (40.7)		>1
			No Osteoplasty n=23	Osteoplasty n=25		
Nepple 2009§	Mean age: 35 years % male: 60 cam (100%)	Failure‡ Conversion to THA Modified Harris Hip Score	5 (21.7) 2 (8.7) 23.3 (38.9)	0 (0) 0 (0) 26 (40.3)		2.0 (1.0–4.0)
			No. of cases (%)			
			Nonoperative n=9	Open FAI n=6	THA n=2	
Jager 2004	Mean age: 34 years % male: 77 cam (100%)	Pain free Return to work/sports	0 (0) 6 (67)	6 (100) 6 (100)	2 (100) 2 (100)	1.8 (NR)

THA: Total hip arthroplasty; FAI: femoroacetabular impingement; NR = not reported.
*Failure definition: Modified Harris Hip Score < 70, subsequent debridement of a hip that had undergone labral refixation, or conversion to THA.
† In a patient with a 2.5-cm full-thickness acetabular chondral defect at the time of arthroscopy.
‡Failure definition: Modified Harris Hip Score < 70 or need for additional surgery

Case-series in non- or recreational athletes

Twenty seven case series were found that reported on clinical outcomes following treatment for FAI in non- or recreational athletes; 14 reporting results following arthroscopy[17,19,34,41,42,51,57,58,60,63,96,109,111,126] , seven following open dislocation[8,9,44,93,107,123,139], five after an open mini procedure[28,47,80,83,118] and one following conservative care.[80] The studies included a total of 1567 patients (1718 hips). The vast majority of patients had either cam or mixed-type FAI. Mean follow-up ranged from 0.5 to 4.9 years (6 months to 6.7 years) across the studies.

Failure, Conversion to THA, and OA Progression (Table 10)

Eight studies reported failure rates; two studies reported failure rates following arthroscopy (N=131 hips)[60,109], five following open dislocation (N=204 hips)[8,9,93,107,123] and one after a mini-open procedure (N=100 hips).[80] Overall failure was defined differently across studies but generally included conversion to total hip arthroplasty (THA), and/or progression of osteoarthritis, and/or a worsening of pain or function (Appendix G). The risk of failure was 8.4% (95% CI: 4.7, 14.4) in those receiving arthroscopy, 12.3% (95 CI: 8.4, 17.5) in those receiving open dislocation, and 11.0% (95% CI: 6.3, 18.6) in patients following a mini-open procedure.

Conversion to THA was reported in 18 studies; eight studies reporting results following arthroscopy (N=775 hips)[19,34,42,57,58,60,96,109], five studies after open dislocation (N=204 hips)[8,9,93,107,123] and five studies after a mini-open procedure (N=226 hips).[22,38,66,69,98] Conversion to THA following arthroscopy occurred in 4.9% (95% CI: 3.6%, 6.7%) of hips treated with arthroscopy, 12.3% (95 CI: 8.4, 17.5) in those receiving open dislocation, and 6.2% (95% CI: 3.7, 10.1) in patients following a mini-open procedure. One study looking at arthroscopic treatment included only patients 60 years and older.[63] In that population, 17.5% converted to THA at a mean follow-up of 2.5 years.

Radiographic OA progression was reported in eight studies, three following arthroscopy (N=168 hips)[38,39,54], two studies following open dislocation (N=115 hips)[9,107], and three studies after a mini-open procedure (N=157 hips).[28,80,83] The proportion of patients with progression of osteoarthritis during short-term follow-up of 0.5 to 3 years was 2.4 % (95% CI: 0.9, 5.9) in those treated with arthroscopy, 23.5% (95% CI: 16.7, 32.0) among patients having open dislocation and 8.9% (95% CI: 55.4, 14.4) in patients receiving a mini-open procedure.

Hip Scores (Table 10)

Seven patient-reported and two clinician-based hip scores were used to report function in 13 studies. Most of the patient-reported scores were in populations receiving arthroscopy.

The *Hip Outcome Score (HOS) – ADLs and Sport* –was reported by two studies following arthroscopy. One study was comprised of 112 patients (112 hips), mean age 40.6 years (38-44), and 44.6% male.[109] The second study was conducted in 16 adolescents (mean age 15 years, range 11–16 years, 12.5% male) who participated in athletics and who underwent arthroscopy for idiopathic FAI.[111] FAI type was predominately cam-pincer combined in both studies and mean follow-up was 2.3 and 1.4 years, respectively. Mean change in HOS ADLs was 17.8 points for the older group and 36.0 points for the adolescents, translating to a mean percent improvement from baseline of 25.4% and 62.1%. Corresponding numbers for the HOS Sport were 26 and 56 points and 60.5% and 170%.

The change in the *Nonarthritic Hip Score (NAHS)* was reported by seven studies following arthroscopy (N= 328 patients)[15,31,47,51,52,63,117], one study of a mini-open procedure[28] and the lone series on conservative treatment.[32] Mean age was 38 years in six of the arthroscopy studies,[17,34,51,57,58,109,126] and one study was conducted specifically in older patients with a mean age of 65 years (range, 65–82).[63] Mean follow-up ranged from 0.5 to 3.0 years across the seven studies. The pooled percent mean improvement was 44.2% among patients treated with arthroscopy, 20.1% in those with a mini-open procedure, and 26.4% among those receiving conservative care (physical therapy, non-steroidal anti-inflammatory medication, and activities avoiding extremes of motion). The proportion of patients who achieved very good/good outcomes based on the NAHS were reported by only two arthroscopy studies and were 54.5% and 20%.[57,126]

The change in the *Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)* was reported in three studies following arthroscopy (N= 80 patients) with mean age of 35 years (range, 17–64) and males comprising 57.9% to 86.8% of patients[34,41,60], and one study following open dislocation (mean age, 40.5 years).[8] The majority of patients had cam or mixed-type FAI. Mean follow-up ranged from a minimum of 1 year to 5.1 years. The pooled percent mean improvement was 30% in those receiving arthroscopy and 33% in those receiving open dislocation.

The change in the *Modified Harris Hip Score (mHHS)* was reported by four studies following arthroscopy (N=250 patients); and one study using the mini-open approach. FAI type was primarily cam and mixed. In the arthroscopy studies, two were in populations outside the typical populations for FAI surgery; one in adolescent population (mean age 15 years)[111] and one in an older population (mean age 65 years).[63] Mean change from preoperative to postoperative in mHHS was highest in the adolescent population, and lowest in the oldest group, a percent mean improvement of 63.6% and 31.7%, respectively. The remaining two arthroscopy studies had a pooled percent mean improvement of 38%.[51,109] Those receiving a mini-open approach had a pooled percent mean improvement of 35.3%.

Clinician-reported hip scores
The *Merle d'Aubigné* was reported by two studies following arthroscopy (N=61 patients)[34,41], four studies following open dislocation (N=110 patients)[9,44,93,123] and one study following an mini-open surgery (N=32 patients).[118] Mean follow-ups ranged from 1 to 4.7 years. The pooled percent mean improvement in the Merle d'Aubigné score was 24.6% in the arthroscopy studies, 25% in the open dislocation studies and 22.5% in the mini-open study.

The percent mean improvement (pooled) in the *Harris Hip Score (HHS)* was 30% in four studies of arthroscopy (N=386 patients)[19,34,96], 34% in two studies of open dislocation (N=108 patients)[107,139], and 19% in one study of a mini-open procedure (N=14 patients).[83]

Pain, Quality of Life, Patient Satisfaction, and Return to Normal Activities (Appendix G)
Pain was reported following arthroscopy by four studies (N=183 patients)[17,57,58,126] and following conservative care in one study (N=37 patients)[32] using a visual analog scale (0–10). The majority of patients had cam FAI, followed closely by mixed-type impingement. Mean follow-up ranged from 0.5 to 3 years. The pooled percent mean improvement was 73% in those undergoing arthroscopy and 67% in those undergoing conservative care.

Patient Satisfaction following arthroscopy was reported in six studies. Four reported the proportion of patients who were satisfied/very satisfied with their outcome following arthroscopic surgery.[17,41,42,63] There were a total of 241 patients (242 hips), mean ages ranged from 31 to 65 years (16–82), and males ranged from 65% to 87% of the populations. Cam and mixed were the predominant types of FAI in these patients. Mean follow-up ranged from 1.0 to 2.5 years. Overall, 80.9% of patients were satisfied or very satisfied with their outcome with proportions in individual studies ranging from 77.3% to 94.3%. Two studies reported satisfaction on a scale on a scale of 1-10, with 10 being very satisfied. One study was in 16 adolescents (17 hips) who participated in sports, mean age 15 years, and 12.5% male.[111] The other study had scores available in only 90 patients (81%) of their patients (total population: mean age 40.6 years and 45% male).[109] Satisfaction scores in both studies were 9 out of 10.

Return to normal activities/work following arthroscopy was reported in four studies with a total of 199 patients (200 hips).[17,63,109,111] Mean ages ranged from 15 to 65 years (range, 11–82), and males ranged from 12.5% to 77%. Mean follow-up ranged from 1.4 to 2.5 years. Overall, 71.9% of the patients returned to normal activities or their previous job. Across the four studies, proportions ranged from 68.9% to 100%, with the greatest numbers seen in both the youngest (100%)[111] and oldest (90.9%)[63] groups of patients.

Range of Motion (ROM) (Appendix G)

ROM was reported in four studies following arthroscopic treatment for FAI (N=200 hips)[17,57,58,126], one study following open dislocation (N=29 hips)[123], one study following mini-open surgery (N=16 hips)[83] and one study following conservative management (N=37 hips).[32] Mean internal rotation improved by 9º, 10º, 5º and 0.6º, respectively. Flexion also increased over baseline motion following surgical treatment by 2.8º following arthroscopy, 7º following open dislocation and 15.9º following the mini-open procedure. Flexion in the in those treated conservatively decreased by 7º.

Case-series in competitive athletes

Four case-series were found that reported on clinical outcomes following arthroscopy for FAI in competitive or professional athletes.[108,110,121,124] There were a total of 103 patients (106 hips) with a mean age of 27.7 years (range, 16–61) and 97.1% were male. The majority of patients had cam or mixed-type FAI. Mean follow-up ranged from 1.6 to 2.0 years (6 months to 5.5 years) across all studies.

Only one study reported functional outcomes in professional male athletes following open dislocation for FAI.[95] There were 22 athletes (30 hips), including 14 (64%) ice hockey players, four (18%) floorball players (called "unihockey" in Switzerland, comparable to street hockey but played indoors), three (14%) soccer players, one (5%) table tennis player. Mean age of the patients was 19.7 years (range, 16–25) and all had either cam or mixed-type impingement. Mean follow-up was 3.8 years (range, 1.0–6.6).

Hip Outcomes (Table 10)

The *Nonarthritic Hip Score (NHS)* following arthroscopy was reported by one study conducted in Australian Football League players.[124] There were 24 male athletes (27 hips) with a mean age of 22 years (range, 16–29). The majority had isolated cam-type FAI (81.5%) followed by pincer and mixed. Mean follow-up was 1.8 years. Preoperative mean NHS scores were high, 81 points, which is to be expected in a physically fit population such as this. Mean percent improvement in scores was 18.5% (mean change of 15 points).

The *Modified Harris Hip Score (mHHS)* following arthroscopy was reported by two studies with 52 patients (55 hips). Twenty-eight patients were professional football players in Australia[124] and 24 were professional hockey players in the National Hockey League of the United States.[110] Mean age of all the athletes was 24.7 years (16–37) and all were male. Most patients had isolated cam-type impingement, followed by mixed and pincer. Mean follow-up for both studies was 2 years (range, 0.5–5). Mean percent improvement in mHHS from preoperative to final follow-up was 35.7% in the hockey players and 11.6% in the

football players (mean change in scores, 25 and 10 points, respectively). The football player had higher preoperative mHHS scores than the hockey players, 86 versus 70 points, possible explaining the smaller percent improvement seen at follow-up.

The *Hip Outcome Score (HOS) – ADLs and Sport* (score range, 0-100 for both subscales), and the *University of California Los Angeles (UCLA) Activity Score* (score range, 1–10) were used to assess patients following open dislocation. Only postoperative scores were reported and are as follows, for the HOS ADLs, HOS Sport, and UCLA score, respectively, 94.5 ± 9.3, 89.1 ± 16.0, and 9.8 ± 0.8. For the HOS, higher scores represent a higher level of physical function and for the UCLA score, 10 represents regular participation in impact sports.

Quality of Life, Patient Satisfaction, and Return to Sports (Table 10)

Patient satisfaction following arthroscopy was reported in two studies with 52 patients (55 hips), 28 professional football players in Australia[124] and 24 professional hockey players in the National Hockey League of the United States.[110] Mean age of all the athletes was 24.7 years (range, 16–37) and all were male. Most patients had isolated cam-type impingement, followed by mixed and pincer. Mean follow-up for both studies was 2 years (range, 0.5–5). All of the football players reported they were satisfied or very satisfied. On a scale of 1 to 10, the hockey players reported a median of 10, indicating they were very satisfied.

Return to sports following arthroscopy was reported by three studies. There were 75 patients (78 hips), mean age 27.9 years and 96% were male. Most patients had cam or mixed FAI. In one study the majority of patients played professional hockey (53%), followed by golf (13%), football (11%), soccer (7%), and various other professional sports (16%).[108] In a second study, all six patients were professional soccer players in England[121] and in the last all patients played football in Australia.[124] Mean follow-up ranged from 1.6 to 1.8 years. Overall, 80.8% of athletes return to their previous level of sports. Across the three studies, rates ranged from 77.8% to 95.9%, with the highest being in the Australia football players.

One study of open dislocation recorded mean postoperative scores for the SF-12 Physical and the SF-12 Mental were 51.1 ± 8 and 54.3 ± 731, respectively.[95] Preoperative scores were not reported. Overall, 81.8% patients reported that they were very satisfied or satisfied with the results and 95.5% had returned to their previous level of sports competition at final follow-up.

Table 10. Summary of outcomes in case-series following surgical treatment for FAI in non- or recreational athletes

Outcome	Arthroscopy				Open Dislocation				Mini-open			
	No. Studies	No. hips	Mean age years (range)	Risk, % (95% CI)	No. Studies	No. hips	Mean age years (range)	Risk, % (95% CI)	No. Studies	No. hips	Mean age years (range)	Risk, % (95% CI)
Failure	2[60,109]	131	39.6 (27–44)	8.4 (4.7, 14.4)	5[8,9,93,107,123]	204	31.1 (19–54)	12.3 (8.4, 17.5)	1[80]	100	33.4 (16–56)	11.0 (6.3, 18.6)
Conversion to THA	9*[19,34,42,57,58,60,77,96,109]	875	34.4 (15–66)	4.7 (3.5, 6.3)	5[8,9,93,107,123]	204	31.1 (19–54)	8.3 (5.3, 12.9)	5[28,47,80,83,118]	226	33.8 (15–56)	6.2 (3.7, 10.1)
OA progression	3[41,42,60]	168	32.5 (16–64)	2.4 (0.9, 5.9)	2[9,107]	115	29.3 (14–52)	23.5 (16.7, 32.0)	3[28,80,83]	157	33.9 (16–56)	8.9 (5.4, 14.4)
Patient Satisfaction	3†[17,41,42]	201	34.9 (16–66)	82.1 (76.2, 86.8)	1[8]	34	40.5 (19–54)	82.4 (66.5, 91.7)	1[47]	33	31 (15–47)	90.9 (76.4, 96.9)
Return to Work/Activities	2‡[17,109]	165	41.1 (17–66)	58.8 (51.2, 66.0)					1[118]	32	36.2 (23–48)	100 (89.3, 100)
	No. Studies	**No. patients**	**Mean age years (range)**	**% mean change**	**No. Studies**	**No. patients**	**Mean age years (range)**	**% mean change**	**No. Studies**	**No. patients**	**Mean age years (range)**	**% mean change**
PATIENT REPORTED												
Hip Outcome Score - ADLs	1**[109]	112	40.6 (38–44)	25.4								
Hip Outcome Score - Sport	1**[109]	112	40.6 (38–44)	60.5								
Nonarthritic Hip Score	6††[17,34,51,57,58,126]	288	37.8 (17–67)	44.2					1§[28]	41	34 (16–48)	20.1
WOMAC	3[34,41,60]	80	35.0 (17–64)	29.9	1[8]	34	40.5 (19–54)	33.0				
Modified Harris Hip Score	2‡[51,109]	194	35.9 (14–63)	38.0					2[28,47]	74	32.7 (15–48)	35.3
UCLA Activity Score					1[8]	34	40.5 (19–54)	56.3	1[28]	41	34 (16–48)	37.7
CLINICIAN BASED												
Merle d'Aubigné hip score	2[34,41]	61	35.2 (17–64)	24.6	4[9,44,93,123]	110	32.0 (14–54)	25.0	1[118]	32	36.2 (23–48)	22.5
Harris Hip Score	4[19,34,77,96]	482	33.3 (15–64)	32.2	2[107,139]	108	29.0 (14–54)	33.9	1[83]	14	37 (17–51)	19.3
PAIN AND QoL												
VAS Pain (0-10)	5[17,57,58,77,126]	300	42.0 (17–67)	73.7								
SF-12 Total	1[77]	100	34.7 (16-64)	35.0								
SF-12 Physical					1[8]	34	40.5 (19–54)	22.3				
SF-12 Mental					1[8]	34	40.5 (19–54)	10.3				

*Javed 2011 was excluded from the pooled analysis since the study was in patients aged 65 years or older.

†Javed 2011 was excluded from the pooled analysis since the study was in patients aged 65 years or older; and Philippon 2008 and 2009 excluded from the pooled analysis since they did not provide preoperative scores.

‡ Javed 2011 was excluded from the pooled analysis since the study was in patients aged 65 years or older; and Philippon 2008 was excluded from the pooled analysis since the study was in adolescents aged 16 years or younger.

§Laude 2009 was excluded from the pooled analysis since they did not provide preoperative scores.

**Philippon 2008 was excluded from the pooled analysis since the study was in adolescents aged 16 years or younger.

††Javed 2011 was excluded from the pooled analysis since the study was in patients aged 65 years or older; and Philippon 2009 was excluded from the pooled analysis since they did not provide preoperative scores.

SUMMARY

Efficacy of surgery for FAI

- There are no data available to assess the short- or long-term efficacy of FAI surgery compared with no surgery.

Effectiveness of surgery for FAI

Short-term 0-5 years:

- There is no evidence that one specific treatment resulted in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty).
- Several case series report improvement in pain, patient reported and clinician reported hip outcome scores, patient satisfaction and return to normal activities following FAI surgery. However, whether this improvement is a result of the surgery, or the postoperative rehabilitation, or the change in activity subsequent to the surgery or placebo is not known and cannot be fully determined without well conducted comparative studies.
- Approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years.

Long-term (≥10 years):

- There are no data available to assess long-term effectiveness of FAI surgery compared with no surgery.
- There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.

4.4. Key Question 4

What is the evidence of the safety of hip surgery for FAI compared with no surgery? Including consideration of:

- Adverse events type and frequency (peri-operative, fractures, nerve damage, mortality, other major morbidity)
- Revision/re-operation rates

Complications

Six comparative studies,[3,33,62,78,100,117] 31 case-series[8,9,17,19,28,31,34,41,42,44,47,51,57-60,63,77,80,83,92,93,96,107,109,111,112,118,123,126,139] and three case-reports[36,75,134] were found that reported complications following surgical treatment for FAI in non- or recreational athletes. Altogether, 20 studies reported on arthroscopy,[3,17,19,34,41,42,51,57-60,63,75,77,78,92,96,109,111,126] ten on open dislocation[8,9,31,33,44,93,107,112,123,139] and seven on mini-open.[28,47,80,83,100,117,118] One of the comparative studies[62] reported on arthroscopy and open dislocation combined as well as a group of patients treated conservatively, and one other case-series investigated the effects of conservative treatment for FAI.[32] Of the five case-series that examined surgery in competitive or professional athletes only three studies reported complications: two following arthroscopy[108,121] and one following open dislocation.[95] Details of each study can be found in Appendix H.

Surgery

Non- or recreational athletes (Table 11)

Reoperation

Reoperation risk following arthroscopy (14 studies, 1263 hips), open dislocation (7 studies, 180 hips), and mini-open (6 studies, 334 hips) were 3.8%, 4.4%, and 8.7% respectively.

Head-neck fracture

Cumulative incidences of head-neck fracture were very low in all groups: 0.2% of cases following arthroscopy (11 studies, 688 hips), 0% of cases following open dislocation (5 studies, 227 hips), and 0.6% of cases following mini-open (3 studies, 175 hips).

Avascular necrosis (AVN) and osteonecrosis (ON)

There were no incidences of either AVN or ON in any study.

Trochanteric nonunion

Trochanteric nonunion occurred in the open dislocation group only, 2.0% of cases (4/202 hips) in 5 studies.

Heterotopic Ossification (HO)

Heterotopic ossification following arthroscopy (11 studies, 1319 hips), open dislocation (5 studies, 168 hips), and mini-open (5 studies, 327 hips) was reported in 1.7%, 6.0%, and 3.4% of

cases, respectively. In one study conducted in 300 patients undergoing arthroscopy for FAI, NSAID prophylaxis was shown to significantly decrease the incidence of HO development: 33% (5/15) in those who did not receive NSAIDs versus 0% (0/285) in those who did.[114]

Deep vein thrombosis (DVT)/pulmonary embolism (PE)
Out of 13 studies that reported incidences of DVT/PE, one case (2.4%) was reported in a study involving 41 patients following mini-open surgery.

Neurological
Neurological complications occurred most often in patients undergoing mini-open procedures, 22.2% (243 hips, 5 studies), followed by arthroscopy, 1.2% (1431 hips, 15 studies), and open dislocation, 0% (227 hips, 5 studies). The majority (41/54) of neurological complications reported occurred in two studies and included femoral nerve palsy/parasthesias and transient numbness in the areas of the pudendus nerve and lateral cutaneous femoral nerve, all except four (9.8%) of which resolved spontaneously by last follow-up or did not cause any significant impairment to the patient.[47,117]

Infection
Incidences of infection were low overall: 0.3% of cases following arthroscopy (13 studies, 1148 hips), 0% of cases following open dislocation (5 studies, 139 hips), and 1.2% of cases following mini-open (3 studies, 258 hips).

Other
Other complications, including superficial tear of the labia minora, symptomatic hardware and temporal hypesthesia were reported in 1.7% of cases following arthroscopy (3 studies, 236 hips) and 12.9% of cases following open dislocation (2 studies, 85 hips). In one study, hypertrophic scar formation was reported in 32 cases (27.4%) following mini-open surgery, as well as two cases of hematoma requiring drainage.[117] In the study that combined arthroscopy and open dislocation, there was one case (16.7%) of a persisting hematoma that required surgical removal; no other perioperative complications were reported.[62] We also found three case-reports that described an occurrence of extravasation of fluid into the abdomen/chest during arthroscopic treatment of FAI.[36,75,134] Fluid extravasation has been reported in hip arthroscopic treatment for reasons other than FAI.[5,50,122] In one case, a 50 year old man had a loose body removed arthroscopically following a successful repair of an acetabular fracture.[5] The arthroscopic fluid extravasated through the fracture site under pump pressure and resulted in an intra-abdominal compartment syndrome that presented as cardiopulmonary arrest. In non-traumatic cases the irrigation fluid is thought to travel into the abdominal cavity through a retroperitoneal approach.[134]

Table 11. Summary of complications in studies reporting treatment for FAI in non- or recreational athletes.

	Arthroscopy			Open Dislocation			Mini-Open		
Complication	**No. Studies**	**N (hips)**	**Cases (%)**	**No. Studies**	**N (hips)**	**Cases (%)**	**No. Studies**	**N (hips)**	**Cases (%)**
Reoperation*	14[3,19,34,41,51,52,57,63,78,79,92,96,100,126]	1263	48 (3.8)	7[8,31,33,93,112,123,139]	180	8 (4.4)	6[28,80,83,100,117,118]	334	29 (8.7)
Head-neck fracture	11[3,34,42,58-60,63,77,92,96,126]	688	1 (0.2)	5[33,44,107,112,139]	227	0 (0)	3[28,47,80]	175	1 (0.6)
Avascular necrosis	8[3,59,60,63,92,96,111,126]	366	0 (0)	6[9,33,44,93,112,139]	173	0 (0)	2[80,117]	217	0 (0)
Osteonecrosis	5[3,34,58,77,92]	304	0 (0)	6[8,31,33,107,112,139]	227	0 (0)	3[28,47,118]	110	0 (0)
Trochanteric nonunion	4[3,34,92,126]	121	0 (0)	5[33,93,107,112,139]	202	4 (2.0)	2[47,80]	134	0 (0)
Heterotopic ossification	11[3,19,34,42,78,79,92,96,114,126]	1319	22 (1.7)	5[8,33,44,112,139]	168	10 (6.0)	5[28,47,80,118]	327	11 (3.4)
DVT/PE	8[3,34,51,52,63,92,109,126]	607	0 (0)	4[33,44,112,139]	131	0 (0)	1[28]	41	1 (2.4)
Neurological†	15[3,19,34,41,42,51,52,58,59,63,77,79,92,109,126]	1431	17 (1.2)	5[33,44,107,112,139]	227	0 (0)	5[28,47,83,118]	243	54 (22.2)
Infection‡	12[3,34,51,52,60,63,77,79,92,96,109,111,126]	1148	4 (0.3)	5[31,33,44,112,139]	139	0 (0)	3[28,80,117]	258	3 (1.2)
Other§	3[42,57,58]	236	4 (1.7)	2[8,44]	85	11 (12.9)	1[117]	117	34 (29.1)

DVT/PE: deep vein thrombosis/pulmonary embolism
*Excluding conversion to total hip arthroplasty
†Including nerve palsy, paresthesia, and neuropraxia, and other
‡Superficial versus deep infection not reported
§Including superficial tear of the labia minora, temporal hypesthesia, symptomatic hardware, hypertrophic scar formation

Competitive or Professional Athletes (Table 12)

Reoperation

Reoperation rates were reported by two studies (51 hips) investigating arthroscopy and one study (30 hips) following open dislocation. Five (9.8%) and one (3.3%) hip, respectively, had undergone or needed subsequent surgery by final follow-up.

Other

There were no reported incidences of head-neck fracture, avascular necrosis, osteonecrosis, trochanteric nonunion, heterotopic ossification, deep vein thrombosis or pulmonary embolism, neurological complications, or infection in either of the two studies reporting on arthroscopy for FAI in competitive athletes. In the study reporting on open dislocation in this population, screw removal was necessary in six cases (20%); no other complications were reported.

Nonsurgical

Non- or recreational athletes

Two studies reported complications following conservative treatment (avoidance of excessive physical activity, the use of anti-inflammatory drugs, and physiotherapy) of FAI in non- or recreational athletes. In one study, treatment failure (the need for surgical intervention) occurred in four (10.8%) of the 37 patients.[32] In the second study, persistent hip pain and dysfunction were reported in all nine non-operatively treated patients at 1.4 years follow-up.[62]

Table 12. Summary of complications in studies reporting treatment for FAI in competitive or professional athletes.

	Arthroscopy			Open Dislocation		
Complication	**No. Studies**	**N (hips)**	**Cases (%)**	**No. Studies**	**N (hips)**	**Cases (%)**
Reoperation*	2[108,121]	51	5 (9.8)	1[95]	30	1 (3.3)
Head-neck fracture	2[108,121]	51	0 (0)	- - -	- - -	- - -
Avascular necrosis	2[108,121]	51	0 (0)	- - -	- - -	- - -
Osteonecrosis	1[121]	6	0 (0)	- - -	- - -	- - -
Trochanteric nonunion	1[121]	6	0 (0)	- - -	- - -	- - -
Heterotopic ossification	1[121]	6	0 (0)	- - -	- - -	- - -
DVT/PE	1[121]	6	0 (0)	- - -	- - -	- - -
Neurological†	2[108,121]	51	0 (0)	- - -	- - -	- - -
Infection	1[121]	6	0 (0)	- - -	- - -	- - -
Other‡	- - -	- - -	- - -	1[95]	30	6 (20%)

AVN: avascular necrosis; DVT/PE: deep vein thrombosis/pulmonary embolism; HO: heterotopic ossification; ON: osteonecrosis.

*Excluding conversion to total hip arthroplasty

†Including nerve palsy, paresthesia, and neuropraxia, and other

‡Including superficial tear of the labia minora, temporal hypesthesia, and symptomatic hardware

SUMMARY

Safety of FAI surgery

- The risk of reoperation (other than conversion to THA) occurred in 4% (arthroscopy and open dislocation) and 9% of the patients (mini-open).
- There was only one reported head-neck fracture (0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion.
- Heterotopic ossification occurred in 2 to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation.
- Neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature.

4.5. Key question 5

What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations? Including consideration of:

- Gender, age, psychological or psychosocial co-morbidities
- Baseline functional status: e.g. type of deformity, extent of osteoarthritis or cartilage damage
- Other patient characteristics or evidence based patient selection criteria
- Provider type, setting or other provider characteristics
- Payer/ beneficiary type: including worker's compensation, Medicaid, state employees

Subpopulations

We found no studies comparing the differential effectiveness of surgery versus nonsurgical care in FAI patients. However, we identified five studies[17,41,52,79,126] that looked at outcomes following surgical treatment for FAI in two subpopulations, those with varying degrees of osteoarthritis as assessed by the Tönnis grade and patients with varying degrees of chondral damage assessed during surgery. We report the results from those five studies.

Varying degrees of osteoarthritis (Table 13)

Five studies[17,41,79,117,126] compared outcomes following arthroscopy for FAI in patients with and without significant preoperative osteoarthritis (OA) classified radiographically using the Tönnis grade. In general, patients with more severe OA tended to have worse outcomes overall compared with those who had little to no preoperative OA. In one study, patients with more advanced OA (Tönnis 2–3) failed four times more often (P <.001) and needed significantly more conversions to total hip arthroplasty compared with those patients who had less advanced or no preoperative hip OA (Tönnis 2–3), 35.4% versus 0.6%, respectively (RR=58; 95% CI=8, 424; P < .001).[79] Likewise, clinical hip outcomes tended to show less mean percent improvement in patients with more severe OA over 6 months to 3 years follow-up: Nonarthritic Hip Score, 22.2% versus 59.6% (P <.05)[126]; Modified Harris Hip Score, 35.0% versus 59.0% (P < .001)[79]; Merle d'Aubigné Score, 9.5% versus 16.2% (P = ns)[41] and 7.4% versus 8.3% (P = ns) [117]; WOMAC, 32.0% versus 35.1% (P = ns)[41] and 23.5% versus 41.8% (P < .001). [117] In two studies, worse postoperative VAS (0-10) pain scores were also seen in those with more severe OA compared with patients with less advanced stages of OA: 3.2 versus 5.2 (P < .05)[126] and 2.6 versus 4.5 (P = NR).[79] Less range of motion, both internal rotation and flexion, was seen postoperatively in patients with Tönnis grade 1 or 2 compared to those with grade 0 in one study, P < .05.[126] One study reported no difference in the Nonarthritic Hip Scores and Sports Frequency Scores at mean of 2 years follow-up in patients with grade 2 compared with grade 1 OA.[17]

Varying degrees of chondral damage (Table 14)

One retrospective cohort study compared outcomes following arthroscopy for FAI in patients with varying degrees of chondral damage assessed intraoperative.[52] There were no differences

between groups with respect to conversion to total hip arthroplasty, the Nonarthritic Hip Score, or the Modified Harris Hip Score at a mean of 2 years follow-up.

SUMMARY

Differential efficacy/effectiveness/safety

- We found no studies comparing the differential effectiveness or safety of surgery versus nonsurgical care in FAI patients.
- Outcomes following FAI surgery were consistently worse in patients with greater preoperative osteoarthritis compared with those with less osteoarthritis. In one study, the relative risk of a conversion to THA in those with preoperative Tönnis grade 2–3 was 58 (95% CI: 8, 424) compared with Tönnis grade 0-1.
- There was no reported difference in outcomes in patients with varying degrees of chondral damage assessed during surgery.
- No data from other subpopulations were found.

Table 13. The effect of osteoarthritis on patients receiving FAI surgery.

Outcome	No. Patients (no. hips)	Mean age	Mean pre-op score (±SD)	No. Hips with outcome (%)	RR (95% CI) or P-value	Mean Follow-up in years (range)
Failure						
Larson 2011*	Tönnis 0–1: 154 (169) Tönnis 2–3: 56 (58)	32 years 45 years	- - - - -	20 (12.0) 30 (52.0)	RR=4.4 (2.7, 71)	2.3 (1.0–5.0)
Conversion to THA						
Larson 2011	Tönnis 0–1: 154 (169) Tönnis 2–3: 56 (58)	32 years 45 years	- - - - -	1 (0.6) 20 (34.5)†	RR=58.3 (8.0, 424)	2.3 (1.0–5.0)
				Mean Change Pre-Post (% mean change)		
Nonarthritic Hip Score						
Brunner 2009	Tönnis 1: 32 (32) Tönnis 2: 13 (13)	NR	57.3 48.1	29.4 (51.3) 35.4 (73.6)	ns	2.4 (2–3.2)
Stahelin 2008	Tönnis 0: 14 (14) Tönnis 1–2: 8 (8)	NR	52 ± 20 45 ± 12	31 ± 22 (59.6) 10 ± 20 (22.2)	< .05	0.5 (NR)
Sports Frequency Score						
Brunner 2009	Tönnis 1: 32 (32) Tönnis 2: 13 (13)	NR	0.74 0.86	0.97 (131) 1.28 (149)	ns	2.4 (2–3.2)
Modified Harris Hip Score						
Larson 2011	Tönnis 0–1: 154 (169) Tönnis 2–3: 56 (58)	32 years 45 years	65.2 63.3	22.8 (35.0) 3.7 (5.9)	< .001	2.3 (1.0–5.0)
Merle d'Aubigné Score						
Gédouin, Duperron 2010	Tönnis 0: 29 (29) Tönnis 1: 7 (7)	NR	14.8 ± 1.8 13.7 ± 1.5	2.4 (16.2) 1.3 (9.5)	ns	1.3 (0.5–3.0)
Ribas 2010	Tönnis 0: NR (32) Tönnis 1: NR (61) Tönnis 2: NR (24)	NR	16.7 16.2 14.9	1.1 (6.6) 1.5 (9.3) 1.1 (7.4)	ns	3.7 (3–5.5)
WOMAC						
Gédouin, Duperron 2010	Tönnis 0: 29 (29) Tönnis 1: 7 (7)	NR	57 ± 17 50 ± 20	20 (35.1) 16 (32.0)	ns	1.3 (0.5–3.0)
Ribas 2010	Tönnis 0: NR (32) Tönnis 1: NR (61) Tönnis 2: NR (24)	NR	68.3 66.2 62.2	27.8 (40.7) 28.1 (42.4) 14.6 (23.5)	< .001	3.7 (3–5.5)
VAS pain (0–10)						
Larson 2011	Tönnis 0–1: 154 (169) Tönnis 2–3: 56 (58)	32 years 45 years	NR	4.5 (NR) 2.6 (NR)	NR	2.3 (1.0–5.0)
Stahelin 2008	Tönnis 0: 14 (14) Tönnis 1–2: 8 (8)	NR	5.8 ± 2.3 5.8 ± 1.6	5.2 ± 2.4 (89.7) 3.2 ± 2.0 (55.2)	< .05	0.5 (NR)
Quality of Life—SF-12						
Larson 2011	Tönnis 0–1: 154 (169) Tönnis 2–3: 56 (58)	32 years 45 years	NR	20.9 (NR) 4.3 (NR)	NR	2.3 (1.0–5.0)
ROM—Internal Rotation (°)						
Stahelin 2008**	Tönnis 0: 14 (14) Tönnis 1–2: 8 (8)	NR	8 ± 8.0 -1.1 ± 8.7	29 ± 11.7‡ 9.4 ± 6.8‡	< .05	0.5 (NR)
ROM—Flexion (°)						
Stahelin 2008**	Tönnis 0: 14 (14) Tönnis 1–2: 8 (8)	NR	112 ± 14.1 98 ± 11.1	132 ± 8.0‡ 106 ± 21‡	< .05	0.5 (NR)

NR: not reported; ns: not statistically significant; SD: standard deviation; SF-12: Short Form 12 Questionnaire;VAS: visual analog scale; WOMAC: Western Ontario and MacMaster University Osteoarthritis Index.
*Failure definition: Modified Harris Hip Score < 70 or conversion to THA.
†Of the 20 hips in group 2 that converted to a THA, 8 (22.2%) had mild to moderate preoperative joint space narrowing and 12 (57.1%) had advanced preoperative joint space narrowing.
‡Postoperative ROM (not pre-post difference).

Table 14. The effect of chondral damage on patients receiving FAI surgery.

Outcome	No. Patients (no. hips)	Mean age	Mean pre-op score	No. Hips with outcome (%)	P-value	Mean Follow-up years (range)
Conversion to THA						
Haviv, Singh 2010*	Grade 1: NR (35)	32 years		0 (0)	NR	1.8 (1.0–3.0)
	Grade 2: NR (83)	35 years	- - - - -	0 (0)		
	Grade 3: NR (52)	43 years		2 (1.2)		
				Mean Change Pre-Post (% mean change)		
Nonarthritic Hip Score						
Haviv, Singh 2010*	Grade 1: NR (35)	32 years	70.7	16.3 (23.1)	NR	1.8 (1.0–3.0)
	Grade 2: NR (83)	35 years	69.1	14.4 (20.8)		
	Grade 3: NR (52)	43 years	60.5	17.5 (28.9)		
Modified Harris Hip Score						
Haviv, Singh 2010*	Grade 1: NR (35)	32 years	74.1	15.7 (21.2)	NR	1.8 (1.0–3.0)
	Grade 2: NR (83)	35 years	73.4	11.3 (15.4)		
	Grade 3: NR (52)	43 years	62.3	15.1 (24.2)		

FAI: femoroacetabular impingement

*Grade 1 = small partial-thickness defect; Grade 2 = full-thickness cartilage loss with a maximum width < 30% of the distance from the acetabular edge to the fovea); Grade 3 = full-thickness cartilage loss with a maximum width > 30% of the distance from the acetabular edge to the fovea.

4.6. Key question 6

What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI? Including consideration of: :

- Costs (direct and indirect) and cost effectiveness
- Short-term and long-term

A major goal of FAI surgery is to delay or prevent hip osteoarthritis. The expected cost benefit of FAI surgery theoretically would come from the costs associated with osteoarthritis's effect (e.g., loss productivity) and treatment (e.g. total hip arthroplasty). At this review, there are no long-term data to determine the efficacy or effectiveness of FAI surgery with respect to neither delaying or preventing osteoarthritis, nor are there data comparing the rate of THA among those with symptomatic FAI who do and don't receive FAI surgery. We were unable to find any cost-effectiveness, cost utility or costing studies on this topic.

5. Summary by Key Question

Key Question 1: Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
Case definition	VERY LOW	• The most consistent case definition of FAI (cam or mixed) as defined by inclusion/exclusion criteria in prospective studies of treatment effectiveness includes hip/groin pain, positive clinical impingement test, and an α-angle >50-55º • There is no evidence that the diagnosis of FAI can be obtained from clinical exam in one small study. One clinical test, the impingement sign, had a positive and negative predictive value of 86% and 79% in one study where the prevalence of FAI was 50%; however, in another study, the reliability of the impingement sign was only moderate. • Even though the α-angle showed moderate to high interobserver reliability in several studies, it had poor diagnostic value in identifying FAI. Other imaging tests assessing abnormalities of the femur and acetabulum had variable degrees of reliability, but no others were tested for diagnostic validity.	-	-	-

Key Question 2: What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Has clinically meaningful improvement in outcomes been defined for FAI?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
Hip osteoarthritis (Tönnis classification)	VERY LOW	• The Tönnis classification is often used to determine the extent of osteoarthritis in the hip. There were no studies found that assessed its validity. Reliability was tested in only one study and intra- and interobserver reliability in that study was moderate.	-	-	-
Patient- and clinician-reported outcomes measures	VERY LOW	• Seven hip outcomes measures were used commonly in FAI patients. Three have undergone psychometric analysis in FAI (HOS-D, M-WOMAC) or young hip-pain (HOS, NAHS) patient populations. • Only one (NAHS) of the three instruments was adequately tested for validity, and it was performed in a young hip-pain patient population. Reliability was inadequately tested for all three instruments. • The MCID was defined to be 9 points for the ADL subscale and 6 points for the sports subscale of the HOS-D in FAI patients. The MCID has not been defined for any other outcome measures in FAI or young hip-pain patients.	-	-	-

Key Question 3: What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
Efficacy	**No evidence**	• There are no data available to assess the short- or long-term efficacy of FAI surgery compared with no surgery			
Effectiveness short-term	**Very low evidence**	• There is no evidence that one specific treatment resulted in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty). • Several case series report improvement in pain, patient reported and clinician reported hip outcome scores, patient satisfaction and return to normal activities following FAI surgery. However, whether this improvement is a result of the surgery, or the postoperative rehabilitation, or the change in activity subsequent to the surgery or placebo is not known. • Approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years.	-	-	+
Effectiveness long-term	**No evidence**	• There are no data available to assess long-term effectiveness of FAI surgery compared with no surgery. • There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.			

Key Question 4: What is the evidence of the safety of hip surgery for FAI compare with no surgery?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
Safety	**Low**	• The risk of reoperation (other than conversion to THA) occurred in 4% (arthroscopy and open dislocation) and 9% of the patients (mini-open). • There was only one reported head-neck fracture (0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion. • Heterotopic ossification occurred in 2 to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation. • Neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature.	-	+	+

Key Question 5: What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
Differential efficacy, effectiveness or safety	**Very low evidence**	• We found no studies comparing the differential efficacy, effectiveness or safety of surgery versus nonsurgical care in FAI patients. • Outcomes following FAI surgery were consistently worse in patients with greater preoperative osteoarthritis compared with those with less osteoarthritis. • There was no reported difference in outcomes in patients with varying degrees of chondral damage assessed during surgery. • No data from other subpopulations were found.	-	-	-

Key Question 6: What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
Cost-effectiveness	**No evidence**	There were no cost-effectiveness, cost utility or costing studies found on FAI surgery.			

References

1. Aetna. *Femoro-Acetabular Surgery for Hip Impingement.* Clinical Policy Bulletin; Number 0736. July 2011. Available at: http://www.aetna.com/cpb/medical/data/700_799/0736.html.

2. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. Bmj 2004;328:1490.

3. Bardakos NV, Vasconcelos JC, Villar RN. Early outcome of hip arthroscopy for femoroacetabular impingement: the role of femoral osteoplasty in symptomatic improvement. J Bone Joint Surg Br 2008;90:1570-5.

4. Bardakos NV, Villar RN. Predictors of progression of osteoarthritis in femoroacetabular impingement: a radiological study with a minimum of ten years follow-up. J Bone Joint Surg Br 2009;91:162-9.

5. Bartlett CS, DiFelice GS, Buly RL, Quinn TJ, Green DS, Helfet DL. Cardiac arrest as a result of intraabdominal extravasation of fluid during arthroscopic removal of a loose body from the hip joint of a patient with an acetabular fracture. J Orthop Trauma 1998;12:294-9.

6. Barton C, Salineros MJ, Rakhra KS, Beaule PE. Validity of the alpha angle measurement on plain radiographs in the evaluation of cam-type femoroacetabular impingement. Clin Orthop Relat Res 2011;469:464-9.

7. Beall DP, Sweet CF, Martin HD, et al. Imaging findings of femoroacetabular impingement syndrome. Skeletal Radiol 2005;34:691-701.

8. Beaule PE, Le Duff MJ, Zaragoza E. Quality of life following femoral head-neck osteochondroplasty for femoroacetabular impingement. J Bone Joint Surg Am 2007;89:773-9.

9. Beck M, Leunig M, Parvizi J, Boutier V, Wyss D, Ganz R. Anterior femoroacetabular impingement: part II. Midterm results of surgical treatment. Clin Orthop Relat Res 2004:67-73.

10. Bedi A, Chen N, Robertson W, Kelly BT. The management of labral tears and femoroacetabular impingement of the hip in the young, active patient. Arthroscopy 2008;24:1135-45.

11. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833-40.

12. Bizzini M, Notzli HP, Maffiuletti NA. Femoroacetabular impingement in professional ice hockey players: a case series of 5 athletes after open surgical decompression of the hip. Am J Sports Med 2007;35:1955-9.

13. BlueCross BlueShield of Massachusetts *Surgical Treatment of Femoroacetabular Impingement.* Medical Policy; Number 145. July 2011. Available at: http://www.bluecrossma.com/common/en_US/medical_policies/145%20Surgical%20Treatment%20of%20Femoroacetabular%20Impingement%20prn.pdf#page=2.

14. Bredella MA, Stoller DW. MR imaging of femoroacetabular impingement. Magn Reson Imaging Clin N Am 2005;13:653-64.

15. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. BMJ 1992;304:1491-4.

16. Brunner A, Hamers AT, Fitze M, Herzog RF. The plain beta-angle measured on radiographs in the assessment of femoroacetabular impingement. J Bone Joint Surg Br 2010;92:1203-8.

17. Brunner A, Horisberger M, Herzog RF. Sports and recreation activity of patients with femoroacetabular impingement before and after arthroscopic osteoplasty. Am J Sports Med 2009;37:917-22.

18. Byrd JW, Jones KS. Prospective analysis of hip arthroscopy with 2-year follow-up. Arthroscopy 2000;16:578-87.

19. Byrd JW, Jones KS. Arthroscopic femoroplasty in the management of cam-type femoroacetabular impingement. Clin Orthop Relat Res 2009;467:739-46.

20. Carlioz H, Pous JG, Rey JC. [Upper femoral epiphysiolysis]. Rev Chir Orthop Reparatrice Appar Mot 1968;54:387-491.

21. Christensen CP, Althausen PL, Mittleman MA, Lee JA, McCarthy JC. The nonarthritic hip score: reliable and validated. Clin Orthop Relat Res 2003:75-83.

22. Cigna. *Hip Surgery for Femoroacetabular Impingement (FAI) Syndrome.* Medical Coverage Policy; Number 0485. April 2011. Available at: http://www.cigna.com/customer_care/healthcare_professional/coverage_positions/medical/mm_0485_hip_arthroscopy_for_femoroacetabular_impingement_syndrome.pdf.

23. Clohisy JC, Carlisle JC, Trousdale R, et al. Radiographic evaluation of the hip has limited reliability. Clin Orthop Relat Res 2009;467:666-75.

24. Clohisy JC, Knaus ER, Hunt DM, Lesher JM, Harris-Hayes M, Prather H. Clinical presentation of patients with symptomatic anterior hip impingement. Clin Orthop Relat Res 2009;467:638-44.

25. Clohisy JC, McClure JT. Treatment of anterior femoroacetabular impingement with combined hip arthroscopy and limited anterior decompression. Iowa Orthop J 2005;25:164-71.

26. Clohisy JC, Nunley RM, Otto RJ, Schoenecker PL. The frog-leg lateral radiograph accurately visualized hip cam impingement abnormalities. Clin Orthop Relat Res 2007;462:115-21.

27. Clohisy JC, St John LC, Schutz AL. Surgical treatment of femoroacetabular impingement: a systematic review of the literature. Clin Orthop Relat Res 2010;468:555-64.

28. Clohisy JC, Zebala LP, Nepple JJ, Pashos G. Combined hip arthroscopy and limited open osteochondroplasty for anterior femoroacetabular impingement. J Bone Joint Surg Am 2010;92:1697-706.

29. Cobb J, Logishetty K, Davda K, Iranpour F. Cams and pincer impingement are distinct, not mixed: the acetabular pathomorphology of femoroacetabular impingement. Clin Orthop Relat Res 2010;468:2143-51.

30. D'Aubigne RM, Postel M. Function al results of hip arthroplasty with acrylic prosthesis. J Bone Joint Surg Am 1954;36-A:451-75.

31. Eijer H, Myers SR, Ganz R. Anterior femoroacetabular impingement after femoral neck fractures. J Orthop Trauma 2001;15:475-81.

32. Emara K, Samir W, Motasem el H, Ghafar KA. Conservative treatment for mild femoroacetabular impingement. J Orthop Surg (Hong Kong) 2011;19:41-5.

33. Espinosa N, Rothenfluh DA, Beck M, Ganz R, Leunig M. Treatment of femoro-acetabular impingement: preliminary results of labral refixation. J Bone Joint Surg Am 2006;88:925-35.

34. Flecher X, Dumas J, Argenson JN. Is a hip distractor useful in the arthroscopic treatment of femoroacetabular impingement? Orthop Traumatol Surg Res 2011.

35. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and Psychological Measurement 1973;33:613-9.

36. Fowler J, Owens BD. Abdominal compartment syndrome after hip arthroscopy. Arthroscopy 2010;26:128-30.

37. Ganz R, Bamert P, Hausner P, Isler B, Vrevc F. [Cervico-acetabular impingement after femoral neck fracture]. Unfallchirurg 1991;94:172-5.

38. Ganz R, Gill TJ, Gautier E, Ganz K, Krugel N, Berlemann U. Surgical dislocation of the adult hip a technique with full access to the femoral head and acetabulum without the risk of avascular necrosis. J Bone Joint Surg Br 2001;83:1119-24.

39. Ganz R, Parvizi J, Beck M, Leunig M, Notzli H, Siebenrock KA. Femoroacetabular impingement: a cause for osteoarthritis of the hip. Clin Orthop Relat Res 2003:112-20.

40. Garellick G, Malchau H, Herberts P. Specific or general health outcome measures in the evaluation of total hip replacement. A comparison between the Harris hip score and the Nottingham Health Profile. J Bone Joint Surg Br 1998;80:600-6.

41. Gedouin JE, Duperron D, Langlais F, Thomazeau H. Update to femoroacetabular impingement arthroscopic management. Orthop Traumatol Surg Res 2010;96:222-7.

42. Gedouin JE, May O, Bonin N, et al. Assessment of arthroscopic management of femoroacetabular impingement. A prospective multicenter study. Orthop Traumatol Surg Res 2010;96:S59-67.

43. Gosvig KK, Jacobsen S, Sonne-Holm S, Palm H, Troelsen A. Prevalence of malformations of the hip joint and their relationship to sex, groin pain, and risk of osteoarthritis: a population-based survey. J Bone Joint Surg Am 2010;92:1162-9.

44. Graves ML, Mast JW. Femoroacetabular impingement: do outcomes reliably improve with surgical dislocations? Clin Orthop Relat Res 2009;467:717-23.

45. Hack K, Di Primio G, Rakhra K, Beaule PE. Prevalence of cam-type femoroacetabular impingement morphology in asymptomatic volunteers. J Bone Joint Surg Am 2010;92:2436-44.

46. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. J Bone Joint Surg Am 1969;51:737-55.

47. Hartmann A, Gunther KP. Arthroscopically assisted anterior decompression for femoroacetabular impingement: technique and early clinical results. Arch Orthop Trauma Surg 2009;129:1001-9.

48. Hartofilakidis G, Bardakos NV, Babis GC, Georgiades G. An examination of the association between different morphotypes of femoroacetabular impingement in asymptomatic subjects and the development of osteoarthritis of the hip. J Bone Joint Surg Br 2011;93:580-6.

49. Harvard Pilgrim Health Care. *Hip Arthroscopy for Femoroacetabular Impingement (FAI).* Medical Policy. Effective December 2008, Revised 2011. Available at: https://www.harvardpilgrim.org/pls/portal/docs/PAGE/PROVIDERS/MEDMGMT/STATEMENTS/HIP_ARTHROSCOPIC_FAI_12.10.PDF.

50. Haupt U, Volkle D, Waldherr C, Beck M. Intra- and retroperitoneal irrigation liquid after arthroscopy of the hip joint. Arthroscopy 2008;24:966-8.

51. Haviv B, O'Donnell J. Arthroscopic treatment for symptomatic bilateral cam-type femoroacetabular impingement. Orthopedics 2010;33:874.

52. Haviv B, Singh PJ, Takla A, O'Donnell J. Arthroscopic femoral osteochondroplasty for cam lesions with isolated acetabular chondral damage. J Bone Joint Surg Br 2010;92:629-33.

53. Hayes Health Technology Brief. . Hayes, Inc. *Arthroscopic Hip Surgery for Femoroacetabular Impingment (FAI)*. March 3, 2010. Available at: http://www.hayesinc.com/hayes/products_and_services/knowledge-center/knowledge-center-highlights/full-ht-briefs-table-of-contents-a-through-i/.

54. Health Care Insurance Board/College voor zorgverzekeringen (CVZ). Femoro-acetabulaire chirurgie ter behandeling van FAI voldoet niet aan het criterium van de stand van de wetenschap en praktijk [Femoroacetabular surgery for the treatment of femoroacetabular impingement does not meet the scientific criteria for practice]. Diemen, The Netherlands: CVZ; February 25, 2010.

55. Herndon CH, Heyman CH, Bell DM. Treatment of Slipped Capital Femoral Epiphysis by Epiphyseodesis and Osteoplasty of the Femoral Neck. A Report of Further Experiences. J Bone Joint Surg Am 1963;45:999-1012.

56. Hoeksma HL, Van Den Ende CH, Ronday HK, Heering A, Breedveld FC. Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in osteoarthritis of the hip. Ann Rheum Dis 2003;62:935-8.

57. Horisberger M, Brunner A, Herzog RF. Arthroscopic treatment of femoral acetabular impingement in patients with preoperative generalized degenerative changes. Arthroscopy 2010;26:623-9.

58. Horisberger M, Brunner A, Herzog RF. Arthroscopic treatment of femoroacetabular impingement of the hip: a new technique to access the joint. Clin Orthop Relat Res 2010;468:182-90.

59. Ilizaliturri VM, Jr., Nossa-Barrera JM, Acosta-Rodriguez E, Camacho-Galindo J. Arthroscopic treatment of femoroacetabular impingement secondary to paediatric hip disorders. J Bone Joint Surg Br 2007;89:1025-30.

60. Ilizaliturri VM, Jr., Orozco-Rodriguez L, Acosta-Rodriguez E, Camacho-Galindo J. Arthroscopic treatment of cam-type femoroacetabular impingement: preliminary report at 2 years minimum follow-up. J Arthroplasty 2008;23:226-34.

61. Jaberi FM, Parvizi J. Hip pain in young adults: femoroacetabular impingement. J Arthroplasty 2007;22:37-42.

62. Jager M, Wild A, Westhoff B, Krauspe R. Femoroacetabular impingement caused by a femoral osseous head-neck bump deformity: clinical, radiological, and experimental results. J Orthop Sci 2004;9:256-63.

63. Javed A, O'Donnell JM. Arthroscopic femoral osteochondroplasty for cam femoroacetabular impingement in patients over 60 years of age. J Bone Joint Surg Br 2011;93:326-31.

64. Johnston TL, Schenker ML, Briggs KK, Philippon MJ. Relationship between offset angle alpha and hip chondral injury in femoroacetabular impingement. Arthroscopy 2008;24:669-75.

65. Kalberer F, Sierra RJ, Madan SS, Ganz R, Leunig M. Ischial spine projection into the pelvis : a new sign for acetabular retroversion. Clin Orthop Relat Res 2008;466:677-83.

66. Kang AC, Gooding AJ, Coates MH, Goh TD, Armour P, Rietveld J. Computed tomography assessment of hip joints in asymptomatic individuals in relation to femoroacetabular impingement. Am J Sports Med 2010;38:1160-5.

67. Kaplan KM, Shah MR, Youm T. Femoroacetabular impingement--diagnosis and treatment. Bull NYU Hosp Jt Dis 2010;68:70-5.

68. Kappe T, Kocak T, Neuerburg C, Lippacher S, Bieger R, Reichel H. Reliability of radiographic signs for acetabular retroversion. Int Orthop 2010.

69. Kappe T, Kocak T, Neuerburg C, Lippacher S, Bieger R, Reichel H. Reliability of radiographic signs for acetabular retroversion. Int Orthop 2011;35:817-21.

70. Kassarjian A, Brisson M, Palmer WE. Femoroacetabular impingement. Eur J Radiol 2007;63:29-35.

71. Kirmit L, Karatosun V, Unver B, Bakirhan S, Sen A, Gocen Z. The reliability of hip scoring systems for total hip arthroplasty candidates: assessment by physical therapists. Clin Rehabil 2005;19:659-61.

72. Konan S, Rayan F, Haddad FS. Is the frog lateral plain radiograph a reliable predictor of the alpha angle in femoroacetabular impingement? J Bone Joint Surg Br 2010;92:47-50.

73. Kubiak-Langer M, Tannast M, Murphy SB, Siebenrock KA, Langlotz F. Range of motion in anterior femoroacetabular impingement. Clin Orthop Relat Res 2007;458:117-24.

74. Laborie LB, Lehmann TG, Engesaeter IO, Eastwood DM, Engesaeter LB, Rosendahl K. Prevalence of Radiographic Findings Thought to Be Associated with Femoroacetabular Impingement in a Population-based Cohort of 2081 Healthy Young Adults. Radiology 2011.

75. Ladner B, Nester K, Cascio B. Abdominal fluid extravasation during hip arthroscopy. Arthroscopy 2010;26:131-5.

76. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

77. Larson CM, Giveans MR. Arthroscopic management of femoroacetabular impingement: early outcomes measures. Arthroscopy 2008;24:540-6.

78. Larson CM, Giveans MR. Arthroscopic debridement versus refixation of the acetabular labrum associated with femoroacetabular impingement. Arthroscopy 2009;25:369-76.

79. Larson CM, Giveans MR, Taylor M. Does arthroscopic FAI correction improve function with radiographic arthritis? Clin Orthop Relat Res 2011;469:1667-76.

80. Laude F, Sariali E, Nogier A. Femoroacetabular impingement treatment using arthroscopy and anterior approach. Clin Orthop Relat Res 2009;467:747-52.

81. Leunig M, Beaule PE, Ganz R. The concept of femoroacetabular impingement: current status and future perspectives. Clin Orthop Relat Res 2009;467:616-22.

82. Leunig M, Casillas MM, Hamlet M, et al. Slipped capital femoral epiphysis: early mechanical damage to the acetabular cartilage by a prominent femoral metaphysis. Acta Orthop Scand 2000;71:370-5.

83. Lincoln M, Johnston K, Muldoon M, Santore R. Combined arthroscopic and modified open approach for cam femoroacetabular impingement: a preliminary experience. Arthroscopy 2009;25:392-9.

84. Lodhia P, Slobogean GP, Noonan VK, Gilbart MK. Patient-reported outcome instruments for femoroacetabular impingement and hip labral pathology: a systematic review of the clinimetric evidence. Arthroscopy 2011;27:279-86.

85. Lohan DG, Seeger LL, Motamedi K, Hame S, Sayre J. Cam-type femoral-acetabular impingement: is the alpha angle the best MR arthrography has to offer? Skeletal Radiol 2009;38:855-62.

86. Mahomed NN, Arndt DC, McGrory BJ, Harris WH. The Harris hip score: comparison of patient self-report with surgeon assessment. J Arthroplasty 2001;16:575-80.

87. Martin RL, Kelly BT, Leunig M, et al. Reliability of clinical diagnosis in intraarticular hip diseases. Knee Surg Sports Traumatol Arthrosc 2010;18:685-90.

88. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. Arthroscopy 2006;22:1304-11.

89. Martin RL, Philippon MJ. Evidence of validity for the hip outcome score in hip arthroscopy. Arthroscopy 2007;23:822-6.

90. Martin RL, Philippon MJ. Evidence of reliability and responsiveness for the hip outcome score. Arthroscopy 2008;24:676-82.

91. Martin RL, Sekiya JK. The interrater reliability of 4 clinical tests used to assess individuals with musculoskeletal hip pain. J Orthop Sports Phys Ther 2008;38:71-7.

92. May O, Matar WY, Beaule PE. Treatment of failed arthroscopic acetabular labral debridement by femoral chondro-osteoplasty: a case series of five patients. J Bone Joint Surg Br 2007;89:595-8.

93. Murphy S, Tannast M, Kim YJ, Buly R, Millis MB. Debridement of the adult hip for femoroacetabular impingement: indications and preliminary clinical results. Clin Orthop Relat Res 2004:178-81.

94. Naal FD, Impellizzeri FM, Miozzari HH, Mannion AF, Leunig M. The German Hip Outcome Score: validation in patients undergoing surgical treatment for femoroacetabular impingement. Arthroscopy 2011;27:339-45.

95. Naal FD, Miozzari HH, Wyss TF, Notzli HP. Surgical hip dislocation for the treatment of femoroacetabular impingement in high-level athletes. Am J Sports Med 2011;39:544-50.

96. Nassif NA, Pekmezci M, Pashos G, Schoenecker PL, Clohisy JC. Osseous remodeling after femoral head-neck junction osteochondroplasty. Clin Orthop Relat Res 2010;468:511-8.

97. National Institute for Health and Clinical Excellence (NICE) *Arthroscopic femoro-acetabular surgery for hip impingement syndrome*. Interventional procedure guidance (IPG) 213. March 2007. Available at: http://guidance.nice.org.uk/IPG213/Guidance/pdf/English. .

98. National Institute for Health and Clinical Excellence (NICE) *Interventional procedure overview of arthroscopic femoro-acetabular surgery for hip impingement syndrome*. June 2011. Available at: http://guidance.nice.org.uk/IPG213/Overview/pdf/English.

99. National Institute for Health and Clinical Excellence (NICE) *Open femoro-acetabular surgery for hip impingement syndrome*. Interventional procedure guidance (IPG) 403. July 2011. Available at: http://guidance.nice.org.uk/IPG403/Guidance/pdf/English.

100. Nepple JJ, Zebala LP, Clohisy JC. Labral disease associated with femoroacetabular impingement: do we need to correct the structural deformity? J Arthroplasty 2009;24:114-9.

101. Ng VY, Arora N, Best TM, Pan X, Ellis TJ. Efficacy of surgery for femoroacetabular impingement: a systematic review. Am J Sports Med 2010;38:2337-45.

102. Norvell DC, Dettori JR, Suk M. What makes a quality outcomes instruments? In: Suk M, Hanson BP, Norvell DC, Helfet DL, eds. AO handbook of musculoskeletal outcomes measures and instruments. Stuttgart: Thieme; 2005:7-24.

103. Notzli HP, Wyss TF, Stoecklin CH, Schmid MR, Treiber K, Hodler J. The contour of the femoral head-neck junction as a predictor for the risk of anterior impingement. J Bone Joint Surg Br 2002;84:556-60.

104. Ovre S, Sandvik L, Madsen JE, Roise O. Comparison of distribution, agreement and correlation between the original and modified Merle d'Aubigne-Postel Score and the Harris Hip Score after acetabular fracture treatment: moderate agreement, high ceiling effect and excellent correlation in 450 patients. Acta Orthop 2005;76:796-802.

105. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. J Clin Epidemiol 2010;63:513-23.

106. Parvizi J, Leunig M, Ganz R. Femoroacetabular impingement. J Am Acad Orthop Surg 2007;15:561-70.

107. Peters CL, Schabel K, Anderson L, Erickson J. Open treatment of femoroacetabular impingement is associated with clinical improvement and low complication rate at short-term followup. Clin Orthop Relat Res 2010;468:504-10.

108. Philippon M, Schenker M, Briggs K, Kuppersmith D. Femoroacetabular impingement in 45 professional athletes: associated pathologies and return to sport following arthroscopic decompression. Knee Surg Sports Traumatol Arthrosc 2007;15:908-14.

109. Philippon MJ, Briggs KK, Yen YM, Kuppersmith DA. Outcomes following hip arthroscopy for femoroacetabular impingement with associated chondrolabral dysfunction: minimum two-year follow-up. J Bone Joint Surg Br 2009;91:16-23.

110. Philippon MJ, Weiss DR, Kuppersmith DA, Briggs KK, Hay CJ. Arthroscopic labral repair and treatment of femoroacetabular impingement in professional hockey players. Am J Sports Med 2010;38:99-104.

111. Philippon MJ, Yen YM, Briggs KK, Kuppersmith DA, Maxwell RB. Early outcomes after hip arthroscopy for femoroacetabular impingement in the athletic adolescent patient: a preliminary report. J Pediatr Orthop 2008;28:705-10.

112. Pierannunzii L, d'Imporzano M. Treatment of femoroacetabular impingement: a modified resection osteoplasty technique through an anterior approach. Orthopedics 2007;30:96-102.

113. Rakhra KS, Sheikh AM, Allen D, Beaule PE. Comparison of MRI alpha angle measurement planes in femoroacetabular impingement. Clin Orthop Relat Res 2009;467:660-5.

114. Randelli F, Pierannunzii L, Banci L, Ragone V, Aliprandi A, Buly R. Heterotopic ossifications after arthroscopic management of femoroacetabular impingement: the role of NSAID prophylaxis. J Orthop Traumatol 2010;11:245-50.

115. Reichenbach S, Juni P, Werlen S, et al. Prevalence of cam-type deformity on hip magnetic resonance imaging in young males: a cross-sectional study. Arthritis Care Res (Hoboken) 2010;62:1319-27.

116. Reynolds D, Lucas J, Klaue K. Retroversion of the acetabulum. A cause of hip pain. J Bone Joint Surg Br 1999;81:281-8.

117. Ribas M, Ledesma R, Cardenas C, Marin-Pena O, Toro J, Caceres E. Clinical results after anterior mini-open approach for femoroacetabular impingement in early degenerative stage. Hip Int 2010;20 (Suppl 7):36-42.

118. Ribas M, Marin-Pena OR, Regenbrecht B, De La Torre B, Vilarrubias JM. Hip osteoplasty by an anterior minimally invasive approach for active patients with femoroacetabular impingement. Hip Int 2007;17:91-8.

119. Rothenfluh DA, Reedwisch D, Muller U, Ganz R, Tennant A, Leunig M. Construct validity of a 12-item WOMAC for assessment of femoro-acetabular impingement and osteoarthritis of the hip. Osteoarthritis Cartilage 2008;16:1032-8.

120. Samora JB, Ng VY, Ellis TJ. Femoroacetabular impingement: a common cause of hip pain in young adults. Clin J Sport Med 2011;21:51-6.

121. Saw T, Villar R. Footballer's hip a report of six cases. J Bone Joint Surg Br 2004;86:655-8.

122. Sharma A, Sachdev H, Gomillion M. Abdominal compartment syndrome during hip arthroscopy. Anaesthesia 2009;64:567-9.

123. Siebenrock KA, Schoeniger R, Ganz R. Anterior femoro-acetabular impingement due to acetabular retroversion. Treatment with periacetabular osteotomy. J Bone Joint Surg Am 2003;85-A:278-86.

124. Singh PJ, O'Donnell JM. The outcome of hip arthroscopy in Australian football league players: a review of 27 hips. Arthroscopy 2010;26:743-9.

125. Soderman P, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? Clin Orthop Relat Res 2001:189-97.

126. Stahelin L, Stahelin T, Jolles BM, Herzog RF. Arthroscopic offset restoration in femoroacetabular cam impingement: accuracy and early clinical outcome. Arthroscopy 2008;24:51-7 e1.

127. Standaert CJ, Manner PA, Herring SA. Expert opinion and controversies in musculoskeletal and sports medicine: femoroacetabular impingement. Arch Phys Med Rehabil 2008;89:890-3.

128. Stevens MS, Legay DA, Glazebrook MA, Amirault D. The evidence for hip arthroscopy: grading the current indications. Arthroscopy 2010;26:1370-83.

129. Stulberg SD, Harris WH. Acetabular dysplasia and development of osteoarthritis of hip. In: Harris WH, ed. The Hip Proceedings of the Second Open Scientific Meeting of the Hip Society. St Louis, MO: Mosby; 1974:82-93.

130. Tanzer M, Noiseux N. Osseous abnormalities and early osteoarthritis: the role of hip impingement. Clin Orthop Relat Res 2004:170-7.

131. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.

132. United Healthcare. *Femoroacetabular Impingement Syndrome Treatment for Commercial Plans.* Medical Policy, Number: SURGERY 074.4 T1. February 2011. Available at: https://www.oxhp.com/secure/policy/femoroacetabular_impingement_syndrome_211.html.

133. Unver B, Karatosun V, Gunal I. Assessing the results of thrust plate prosthesis: a comparison of four different rating systems. Clin Rehabil 2005;19:654-8.

134. Verma M, Sekiya JK. Intrathoracic fluid extravasation after hip arthroscopy. Arthroscopy 2010;26:S90-4.

135. Weir A, de Vos RJ, Moen M, Holmich P, Tol JL. Prevalence of radiological signs of femoroacetabular impingement in patients presenting with long-standing adductor-related groin pain. Br J Sports Med 2011;45:6-9.

136. Werner CM, Copeland CE, Ruckstuhl T, et al. Radiographic markers of acetabular retroversion: correlation of the cross-over sign, ischial spine sign and posterior wall sign. Acta Orthop Belg 2010;76:166-73.

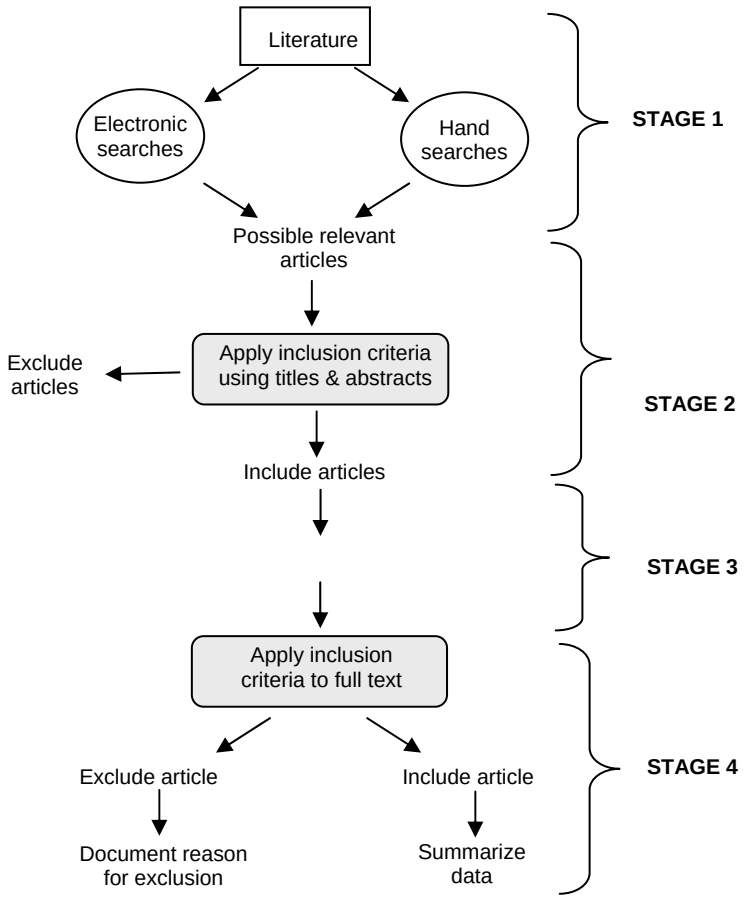
137. West S, King V, Carey TS, et.al. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center, Contract No. 290-97-0011). In: Agency for Healthcare Research and Quality, Rockville, MD; 2002.

138. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239-46.

139. Yun HH, Shon WY, Yun JY. Treatment of femoroacetabular impingement with surgical dislocation. Clin Orthop Surg 2009;1:146-54.

140. Zebala LP, Schoenecker PL, Clohisy JC. Anterior femoroacetabular impingement: a diverse disease with evolving treatment options. Iowa Orthop J 2007;27:71-81.

APPENDIX A. ALGORITHM FOR ARTICLE SELECTION



APPENDIX B. SEARCH STRATEGIES

Below is the search strategy for PubMed. Parallel strategies were used to search other electronic databases listed below. Keyword searches were conducted in the other listed resources.

Key Question 1

1.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR "femoral osteochondroplasty" OR "femoral osteoplasty"
2.	SENSITIVITY[TIAB] OR SPECIFICITY[TIAB] OR PREDICT*[TIAB] OR "Reproducibility of Results"[Mesh] OR RELIAB*[TI] OR VALID* OR INTERTEST* OR INTEROBSERV* OR INTRATEST* OR INTRAOBSERV* OR INTERRAT* OR INTRARAT* OR "Validation Studies" [Publication Type] OR "Reproducibility of Results"[Mesh]
3.	PROSPECTIV*
4.	#1 AND #2 (LIMIT ENGLISH)
5.	#1 AND #3 (LIMIT ENGLISH)

Key Question 2

6.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR "femoral osteochondroplasty" OR "femoral osteoplasty"
7.	#4 (LIMIT TO SYSTEMATIC REVIEW/METAANALYSIS)
8.	"Merle d'Aubigné" OR "HARRIS HIP SCORE" OR "Western Ontario and McMaster Universities Osteoarthritis Index" OR WOMAC OR "NON ARTHRITIC HIP SCORE" OR "NONARTHRITIC HIP SCORE" OR "HIP OUTCOME SCORE"
9.	"Reproducibility of Results"[Mesh] OR RELIAB*[TI] OR VALID* OR INTERTEST* OR INTEROBSERV* OR INTRATEST* OR INTRAOBSERV* OR INTERRAT* OR INTRARAT*) OR "Validation Studies" [Publication Type]) OR "Reproducibility of Results"[Mesh]
10.	#6 AND #7 (LIMIT ENGLISH)
11.	#6 AND #8 AND #9 (LIMIT ENGLISH)

Key Question 3, 4, 5

12.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR "femoral osteochondroplasty" OR "femoral osteoplasty"
13.	"Reoperation"[Mesh] OR "Femur Head Necrosis"[Mesh] OR "Arthroplasty, Replacement, Hip"[Mesh] OR REOPERATION REATTACHMENT OR AVN OR AVASCULAR NECROSIS OR TOTAL HIP OR TOTAL JOINT OR ARTHROPLASTY OR INFECTION* OR DEATH OR COMPLICATION* OR ADVERSE EVENT OR "Intraoperative Complications"[Mesh] OR SCIATIC* OR NERVE OR NEURO* OR FRACTURE* OR INTRAABDOM* OR CARDIAC ARREST OR THROMBO* OR EMBOL* OR INSTABILITY
14.	#9 AND #10 (LIMIT ENGLISH)

Key Question 6

15.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR

	FEMORAL) AND IMPINGMENT*) OR “femoral osteochondroplasty” OR “femoral osteoplasty”
16.	COST OR "Cost-Benefit Analysis"[Mesh])
17.	#12 AND #13 (LIMIT ENGLISH)

Electronic Database Searches

The following databases have been searched for relevant information:
Agency for Healthcare Research and Quality (AHRQ)
Cumulative Index to Nursing and Allied Health (CINAHL)
Cochrane Database of Systematic Reviews (through June 2011)
Cochrane Registry of Clinical Trials (CENTRAL) (through June 2011)
Cochrane Review Methodology Database (through June 2011)
Computer Retrieval of Information on Scientific Projects (CRISP)
Database of Reviews of Effectiveness (Cochrane Library) (through June 2011)
EMBASE (1985 through Jun 1, 2011)
PubMed (1975 through Jun 1, 2011)
Informational Network of Agencies for Health Technology Assessment (INAHTA)
NHS Economic Evaluation Database (Cochrane Library through June 2011)
HSTAT (Health Services/Technology Assessment Text)
EconLIT

Additional Economics, Clinical Guideline and Gray Literature Databases

AHRQ- Healthcare Cost and Utilization Project
Canadian Agency for Drugs and Technologies in Health
Centers for Medicare and Medicaid Services (CMS)
Food and Drug Administration (FDA)
Google
Institute for Clinical Systems Improvement (ICSI)
National Guideline Clearinghouse

APPENDIX C. EXCLUDED ARTICLES

Exclude at full-text review

Author	year	Reason for exclusion
KQ 1		
Emara	2011	no inclusion criteria stated
Gedouin, May	2010	no inclusion criteria stated
Pierannunzii	2007	no inclusion criteria stated
Stahelin	2008	no inclusion criteria stated
Jager	2004	no inclusion criteria stated
Byrd	2009	retrospective study
Gedouin,	2010	retrospective study
Philippon	2009	retrospective study
KQ 3-5		
Farjo	1999	Labral tears; not FAI
Guanche	2005	Not FAI specific
McCarthy	2003	Not FAI specific
Santori	2000	Labral tears; not FAI
Bizinni	2007	N=5, no safety discussed
Byrd	2000	Not FAI specific
Kang	2009	correlation of acetabular labral tears a with femoroacetabular
Kim	2007	OA as predictor of outcome
O'Leary	2001	Not FAI specific
Peters	2006	Same study as Peters 2010
Pollard	2010	Genetics study
Potter	2005	Labral tears; not FAI
Tanzer	2004	Labral tears; not FAI

APPENDIX D. LEVEL OF EVIDENCE DETERMINATION

Each study was rated against pre-set criteria that resulted in an evidence rating (Level of Evidence I, II, III, or IV) and presented in a table. The criteria are listed in the Tables below.

Table D1. Definition of the different levels of evidence for articles on therapy and prognosis

	Studies of Therapy		Studies of Prognosis	
Level	**Study design**	**Criteria**	**Study design**	**Criteria**
I	Good quality RCT	• Concealment • Blind or independent assessment for important outcomes • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size	Good quality cohort	• Prospective design • Patients at similar point in the course of their disease or treatment • F/U rate of 80%+ • Patients followed long enough for outcomes to occur • Controlling for extraneous prognostic factors*
II	Moderate or poor quality RCT	• Violation of any of the criteria for good quality RCT	Moderate quality cohort	• Prospective design, with violation of one of the other criteria for good quality cohort study • Retrospective design, meeting all the rest of the criteria in level I
	Good quality cohort	• Blind or independent assessment in a prospective study, or use of reliable data* in a retrospective study • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size • Controlling for possible confounding†		
III	Moderate or poor quality cohort	• Violation of any of the criteria for good quality cohort	Poor quality cohort	• Prospective design with violation of 2 or more criteria for good quality cohort, or • Retrospective design with violation of 1 or more criteria for good quality cohort
	Case-control	• Any case-control design	Case-control	• Any case-control design
IV	Case series	• Any case series design	Case series	• Any case series design

* Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Table D2. Definitions of the different levels of evidence for diagnostic test accuracy/validity studies

Level	Study type	Criteria
I	Good quality prospective study	• Broad spectrum of persons with the expected condition • Appropriate reference standard used • Adequate description of test and reference for replication • Blinded comparison of tests with appropriate reference standard • Reference standard performed independently of diagnostic test
II	Moderate quality prospective study	• Violation of any one of the criteria for a good quality prospective study (LoE I)
	Good quality retrospective study	• Broad spectrum of persons with the expected condition • Appropriate reference standard used • Adequate description of test and reference for replication • Blinded comparison of tests with appropriate reference standard • Reference standard performed independently of diagnostic test
III	Poor quality prospective study	• Violation of any two or more of the criteria for a good quality prospective study (LoE I)
	Moderate quality retrospective study	• Violation of any one of the criteria for a good quality retrospective study (LoE II)
IV	Poor quality retrospective study	• Violation of any two or more of the criteria for a good quality retrospective study (LoE II)
	Case-Control Study	

Table D3. Definitions of the different levels of evidence for reliability studies

Level	Study type	Criteria
I	Good quality study	• Broad spectrum of persons with the expected condition • Adequate description of methods for replication • Blinded performance of tests, measurements or interpretation • Second test/interpretation performed independently of the first
II	Moderate quality	• Violation of any one of the criteria for a good quality study
III	Poor quality study	• Violation of any two of the criteria
IV	Very poor quality study	• Violation of any three of the criteria

Determination of Overall Strength of Evidence

Following the assessment of the quality of each individual study included in the report, an overall "strength of evidence" for the relevant question or topic is determined. Methods for determining the overall strength of evidence are variable across the literature and are most applicable to evaluation of therapeutic studies.

SRI's method incorporates the primary domains of quality (LoE), quantity of studies and consistency of results across studies as described by AHRQ.

The following definitions are used by SRI to determine whether or not the body of evidence meets the criteria for each domain:

Domain	Definition/Criterion
Quality	• At least 80% of the studies are LoE I or II
Quantity	• There are at least three studies which are adequately powered to answer the study question
Consistency	• Study results would lead to a similar conclusion (similar values, in the same direction) in at least 70% of the studies

Based on the criteria described above, the possible scenarios that would be encountered are described below. Each scenario is ranked according to the impact that future research is likely to have on both the overall estimates of an effect and the confidence in the estimate. This ranking describes the overall "Strength of Evidence" (SoE) for the body of literature on a specific topic. The method and descriptions of overall strength are adapted from the system described by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group[2] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[105,137]

SoE	Description	Further Research Impact	Domain Criterion Met		
			Quality	Quantity	Consistency
1	**High**	Very unlikely to change confidence in effect estimate	+	+	+
2	**Moderate**	Likely to have an important impact on confidence in estimate and *may* change the estimate	+	-	+
			+	+	-
3	**Low**	Very likely to have an important impact on confidence in estimate and *likely* to change the estimate	+	-	-
			-	+	+
4	**Very Low**	Any effect estimate is uncertain	-	+	-
			-	-	+
			-	-	-

Assessment of Economic Studies

Full formal economic analyses evaluate both costs and clinical outcomes of two or more alternative interventions. The four primary types are cost minimization analysis (CMA), cost-utility analysis (CUA), cost-effectiveness analysis (CEA), and cost-benefit analyses (CBA). Each employs different methodologies, potentially complicating critical appraisal, but some common criteria can be assessed across studies.

No standard, universally accepted method of critical appraisal of economic analyses is currently in use. A number of checklists [Canadian, BMJ, AMA] are available to facilitate critique of such studies. The Quality of Health Economic Studies (QHES) instrument developed by Ofman, et al. QHES embodies the primary components relevant for critical appraisal of economic studies. It also incorporates a weighted scoring process and which was used as one factor to assess included economic studies. This tool has not yet undergone extensive evaluation for broader use but provides a valuable starting point for critique.

In addition to assessment of criteria in the QHES, other factors are important in critical appraisal of studies from an epidemiologic perspective to assist in evaluation of generalizability and potential sources of study bias.

Such factors include:

- Are the interventions applied to similar populations (eg, with respect to age, gender, medical conditions, etc)? To what extent are the populations for each intervention comparable and are differences considered or accounted for? To what extent are population characteristics consistent with "real world" applications of the comparators?
- Are the sample sizes adequate so as to provide a reasonable representation of individuals to whom the technology would be applied?
- What types of studies form the basis for the data used in the analyses? Data (eg, complication rates) from randomized controlled trials or well-conducted, methodologically rigorous cohort studies for data collection are generally of highest quality compared with case series or studies with historical cohorts.
- Were the interventions applied in a comparable manner (eg, similar protocols, follow-up procedures, evaluation of outcomes, etc)?
- How were the data and/or patients selected or sampled (eg, a random selection of claims for the intervention from a given year/source or all claims)? What specific inclusion/exclusion criteria or processes were used?
- Were the outcomes and consequences of the interventions being compared comparable for each? (eg, were all of the relevant consequences/complications for each intervention considered or do they primarily reflect those for one intervention?)

Assessment of the overall strength of evidence for formal economic analyses does not appear to be documented in the literature. For the purposes of this HTA, overall strength was determined by:

- Quality of the individual studies: Where the majority of quality indicators described in the QHES met and were the methods related to patient/claim selection, patient population considerations and other factors listed above consistent with a high quality design?
- Number of formal analyses (3 or more)
- Consistency of findings and conclusions from analyses across studies.

QHES Instrument[131] **Study** _____

Questions	Points	Yes	No
1. Was the study objective presented in a clear, specific, and measurable manner?	**7**		
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	**4**		
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	**8**		
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	**1**		
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	**9**		
6. Was incremental analysis performed between alternatives for resources and costs?	**6**		
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	**5**		
8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	**7**		
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	**8**		
10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	**6**		
11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	**7**		
12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	**8**		
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	**7**		
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	**6**		
15. Were the conclusions/recommendations of the study justified and based on the study results?	**8**		
16. Was there a statement disclosing the source of funding for the study?	**3**		
TOTAL POINTS	**100**		

APPENDIX E. LEVEL OF EVIDENCE EVALUATION

KQ 2. Reliability studies evaluating instruments on FAI patients.

METHODOLOGICAL PRINCIPLE	Naal	Rothenfluh
Study Design		
Prospective cohort design	■	■
Retrospective cohort design		
Case-control design		
Broad spectrum of patients with expected condition	■	
Adequate description methods for replication	■	■
Blinded performance of tests, measurements, or interpretation	■	■
Second test/interpretation performed independently of first	■	■
Evidence Level	**I**	**II**

Blank box indicates criterion not met, could not be determined or information not reported by author

KQ 3. Critical appraisal

Methodological principle	Beaule (2007)	Beck (2004)	Brunner (2009)	Byrd (2009)	Clohisy (2010)	Eijer (2001)	Emara (2011)	Flecher 2011
Study design								
Randomized controlled trial								
Cohort study								
Case-series	■	■	■	■	■	■	■	■
Statement of concealed allocation*								
Intent-to-treat*								
Independent or blind assessment								
Complete follow-up of ≥85%		■	■		■	■	■	
Adequate sample size								
Controlling for possible confounding								
Evidence Class	**IV**	**IV**	**IV**	**IV**	**IV**	**IV**	**IV**	**IV**

*Applies to randomized controlled trials only.

KQ 3. Critical appraisal continued.

Methodological principle	Gedouin, Duperron (2010)	Gedouin, May (2010)	Graves (2009)	Hartmann (2009)	Horisberger (2010)	Horisberger (2010)	Ilizaliturri (2007)
Study design Randomized controlled trial Cohort study Case-series	■	■	■	■	■	■	■
Statement of concealed allocation*							
Intent-to-treat*							
Independent or blind assessment							
Complete follow-up of ≥85%	■	■		■	■	■	■
Adequate sample size							
Controlling for possible confounding							
Evidence Class	**IV**	**IV**	**IV**	IV	IV	IV	IV

*Applies to randomized controlled trials only.

KQ 3. Critical appraisal continued.

Methodological principle	Ilizaliturri (2008)	Javed (2011)	Larson (2008)	Laude (2009)	Lincoln (2009)	May (2007)	Murphy (2004)	Naal (2010)
Study design Randomized controlled trial Cohort study Case-series	■	■	■	■	■	■	■	■
Statement of concealed allocation*								
Intent-to-treat*								
Independent or blind assessment								
Complete follow-up of ≥85%		■				■		■
Adequate sample size								
Controlling for possible confounding								
Evidence Class	IV	IV	IV	IV	IV	IV		IV

*Applies to randomized controlled trials only.

KQ 3. Critical appraisal continued.

Methodological principle	Nassif (2010)	Peters (2010)	Philppon (2007)	Philppon (2008)	Philppon (2009)	Philppon (2010)
Study design Randomized controlled trial Cohort study Case-series	■	■	■	■	■	■
Statement of concealed allocation*						
Intent-to-treat*						
Independent or blind assessment						
Complete follow-up of ≥85%		■	■	■		■
Adequate sample size						
Controlling for possible confounding						

Evidence Class	IV	IV	IV	IV	IV	IV

*Applies to randomized controlled trials only.

Methodological principle	Pierannunzi (2007)	Ribas (2007)	Saw (2004)	Siebenrock (2003)	Singh (2010)	Stahelin (2008)	Yun (2009)
Study design							
Randomized controlled trial							
Cohort study							
Case-series	■	■	■	■	■	■	■
Statement of concealed allocation*							
Intent-to-treat*							
Independent or blind assessment							
Complete follow-up of ≥85%			■	■	■	■	■
Adequate sample size							
Controlling for possible confounding							
Evidence Class	IV	IV	IV	IV	IV	IV	IV

APPENDIX F. SUPPLEMENTAL DATA FOR KQ 2

Table F1. Descriptions of outcomes instruments used in studies of FAI.

Outcome measure	Clinician or patient reported	Instrument type	Components	Score range	Interpretation
Harris Hip Score (HHS)[46]	Clinician	Disease specific	4 subscales (16 items) • Pain (44 points) • Function (47 points) • Deformity (4 points) • Range of motion (5 points)	0–100	Excellent: 90–100 Good: 80–89 Fair: 70–79 Poor: <70
Modified HHS (MHHS)[18]	Clinician	Disease specific	2 subscales • Pain (44 points) • Function (47 points)	0–100 (total points multiplied by 1.1 to achieve final score)	Excellent: 90–100 Good: 80–89 Fair: 70–79 Poor: <70
Hip Outcome Score (HOS)[88]	Patient	Disease specific	2 subscales (28 items; 0–4 points each): • Activities of daily living (19 items*; 68 points) • Sports (9 items; 36 points)	0% – 100%; proportion of maximum potential score† for each subscale (total score not used)	Lower score = greater disability
Merle D'Aubigne (MA)[30]	Clinician	Disease specific	3 subscales (3 items) • Pain (6 points) • Mobility (6 points) • Walking ability (6 points)	0–12†	†Very good: 11–12 Good: 10 Medium: 9 Fair: 8 Poor: <7
Non-Arthritic Hip Score (NAHS)[21]	Patient	Disease specific	4 subscales (20 items; 0–4 points each) • Pain (20 points) • Mechanical symptoms (16 points) • Physical function (20 points) • Level of activity (24 points)	0–100 (total points multiplied by 1.25 to achieve final score)	Lower score = greater disability
WOMAC (Western Ontario and McMaster Universities OA index)[11]	Patient	Disease specific	3 subscales (24 items) • Pain (20 points) • Stiffness (8 points) • Physical function (68 points)	0–96	Higher score = greater disability

* HOS: the sitting and putting on sock items are not scored.
† based on the number of items answered (except those answered "nonapplicable"). For example, if all 9 items in the sports subscale are answered, the maximum potential score is 36, but if only 8 items are answered, the maximum score possible used to calculate the percentage is 32.
‡ MA final score: the pain and walking ability scores are summed and then adjusted down by 1–2 grades based on the mobility score for the final clinical grade.

Functional outcome measures used in FAI studies

Clinician-based outcome measures:

- The **Merle d'Aubigne Hip Score (MA)**[30] was most frequently used to evaluate function based on pain, mobility, and ability to walk. This outcome measures has not been validated in the FAI (or labral tear) population. However, outcomes of the MA have been validated in patients with acetabular fracture[104] and in those undergoing total hip replacement[133], and was shown to be reliable in patients with coxarthrosis that were candidates for total hip arthroplasty.[71]
- The **Harris Hip Score (HHS)**[46] consists of four subscales: pain, function, deformity, and range of motion. While the HHS has not been validated in FAI patients, studies have shown that this outcome measure is valid in the following populations: total hip arthroplasty[40,125,138], total hip replacement[133], and acetabular fracture.[104] The HHS has been shown to be reliable in patients with total hip arthroplasty[71,125] by two of three studies and responsive in total hip arthroplasty[56,138] and replacement patient populations.[133]

Patient-reported outcome measures:

- The **Modified Harris Hip Score (MHHS)**[18] reports only the pain and function subscales of the original HHS. The MHHS has high convergent construct validity compared with the clinician-reported Harris Hip Score in patients who had undergone one or two hip arthroplasties at least two years prior.[86]
- The **Hip Outcome Score (HOS)**[88] was developed to evaluate outcomes of patients with hip acetabular tears who may be functioning at a wide range of ability. It reports separate scores for activities of daily living and sports and is discussed in more detail in Key Question 2.
- The **Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index**[11] assesses function based on three subscales: pain, stiffness, and physical function. This outcome measure has been validation tested in the FAI population and is discussed in more detail as part of Key Question 2.
- The **Non-Arthritic Hip Score (NAHS)**[21] was derived in part from the WOMAC index and evaluates pain, physical function, mechanical symptoms, and level of activity. This outcome measure was designed specifically for younger, more active patients (aged 20–40 years) with hip pain and without a clear radiographic diagnosis. It has been validation tested in patients with labral tear, the details of which are presented in the results section for Key Question 2.

Table F2. Validity, reliability, and responsiveness of functional outcome measures

Outcome measure	Patient population tested in	Validity	Reliability	Responsiveness
MA	Patients with coxarthrosis and candidates for total hip arthroplasty (N = 35) (59 years; 49% male)[71]	not tested	+	not tested
	Patients with acetabular fracture (N = 450) (44 years; sex NR)[104]	+	not tested	not tested
	Patients undergoing total hip replacement (N = 61) (50 years; 33% male)[133]	+	not tested	–
HHS	Patients with total hip arthroplasty (N = 58) (71 years; 34% male)[125]	+	+	not tested
	Patients with total hip arthroplasty (N = 78) (62 years; 55% male)[138]	+	–	+
	Patients with hip osteoarthritis (N = 75) (72 years; 27% male)[56]	not tested	not tested	+
	Patients with coxarthrosis and candidates for total hip arthroplasty (N = 35) (59 years; 49% male)[71]	not tested	+	not tested
	Patients with total hip arthroplasty (N = 100) either cemented (n = 54) (71 years; 43% male) or uncemented (n = 46) (49 years; 50% male)[40]	+	not tested	not tested
	Patients undergoing total hip replacement (N = 61) (50 years; 33% male)[133]	+	not tested	+
	Patients with acetabular fracture (N = 450) (44 years; sex NR)[104]	+	not tested	not tested
MHHS*	Patients with 1–2 total hip arthroplasties (≥ 1 year postop) (N = 36) (69 years; 31% male)[86]	+	not tested	not tested
HOS	See Key Question 2 for validation in FAI population			
WOMAC	See Key Question 2 for validation in FAI population			
NAHS	See Key Question 2 for validation in FAI population			

* The version of the MHHS that was validated omitted the public transportation question (worth 1 points). Thus the maximum number of points was 90 (versus 91 in the more commonly used mHHS), which was then converted to a scale of 0–100.

Table F3. Descriptions of outcomes instruments used in studies of FAI.

Outcome measure	Clinician or patient reported	Instrument type	Components	Score range	Interpretation
Harris Hip Score (HHS)[46]	Clinician	Disease specific	4 subscales (16 items) • Pain (44 points) • Function (47 points) • Deformity (4 points) • Range of motion (5 points)	0–100	Excellent: 90–100 Good: 80–89 Fair: 70–79 Poor: <70
Modified HHS (MHHS)[18]	Clinician	Disease specific	2 subscales • Pain (44 points) • Function (47 points)	0–100 (total points multiplied by 1.1 to achieve final score)	Excellent: 90–100 Good: 80–89 Fair: 70–79 Poor: <70
Hip Outcome Score (HOS)[88]	Patient	Disease specific	2 subscales (28 items; 0–4 points each): • Activities of daily living (19 items*; 68 points) • Sports (9 items; 36 points)	0% – 100%; proportion of maximum potential score† for each subscale (total score not used)	Lower score = greater disability
Merle D'Aubigne (MA)[30]	Clinician	Disease specific	3 subscales (3 items) • Pain (6 points) • Mobility (6 points) • Walking ability (6 points)	0–12†	†Very good: 11–12 Good: 10 Medium: 9 Fair: 8 Poor: <7
Non-Arthritic Hip Score (NAHS)[21]	Patient	Disease specific	4 subscales (20 items; 0–4 points each) • Pain (20 points) • Mechanical symptoms (16 points) • Physical function (20 points) • Level of activity (24 points)	0–100 (total points multiplied by 1.25 to achieve final score)	Lower score = greater disability
WOMAC (Western Ontario and McMaster Universities OA index)[11]	Patient	Disease specific	3 subscales (24 items) • Pain (20 points) • Stiffness (8 points) • Physical function (68 points)	0–96	Higher score = greater disability

* HOS: the sitting and putting on sock items are not scored.

† based on the number of items answered (except those answered "nonapplicable"). For example, if all 9 items in the sports subscale are answered, the maximum potential score is 36, but if only 8 items are answered, the maximum score possible used to calculate the percentage is 32.

‡ MA final score: the pain and walking ability scores are summed and then adjusted down by 1–2 grades based on the mobility score for the final clinical grade.

Table F4. Demographics of studies validating outcome measures in FAI, labral tear, or hip arthroscopy patients.

Author (year)	Study design (LoE)	Outcome measure evaluated	Follow-up (% followed)	Demographics	Patient characteristics	Interventions	Inclusion/Exclusion
Martin (2006)	Prospective cohort study (cross-sectional design)	Hip Outcome Score (HOS)	Length f/u: n/a % f/u: NR*	N = 507* Male: 45.8% Mean age: 38 ± 13 (range, 13, 66) years	• Labral tear as primary diagnosis • Mean duration of symptoms: 3.4 ± 5 years (range, 11 days, 29 years) • Patient-reported current level of function: Normal: 3% Nearly normal: 26% Abnormal: 51% Severely abnormal: 20% • Arthroscopic surgery for labral tear: 52% (263/507) Mean length of time between surgery completion of questionnaires: 6.7 months (range, 2 days, 3.86 years) • Previous hip surgery (other): NR • Comorbidities: all patients reported that their hip condition was their primary limiting factor.	• Questionnaire filled out at office visit	Inclusion • Patients under the care of a single orthopedic surgeon who specializes in the treatment of musculoskeletal hip-related disorders and acetabular labral tears in particular (October 2003 – December 2004). Exclusion • Non-English speaking patients. • Patients without a labral tear. • Patients who had a high number of items that could not be scored (ie., marking "nonapplicable" or leaving question blank) (ADL subscale: ≥ 6/19 items; sports subscale: ≥ 3 items).
Martin (2007)	Prospective cohort study	HOS	Length f/u: n/a 34% (116/337) returned questionnaire†	N = 337 Male: 48%† Mean age: 42 ± 14 (range, 14, 79) years†	• Arthroscopic surgery: 100% Mean length of time between surgery completion of questionnaires: 3.1 ± 0.49 (range, 2, 4.6) years • Previous hip surgery (other): NR • Diagnosis: NR	• Questionnaire mailed to patients who had undergone hip arthroscopy more than 2 years prior. • Hip arthroscopy procedures: Labral debridement (89%) Psoas release (64%) Capsular modification (60%) Ligamentum teres debridement (55%) Microfracture (28%) Iliotibial band release (12%) Osteoplasty (9%) Labral repair (7%)	Inclusion • Patients who underwent hip arthroscopy by the senior author between August 2001 and August 2003. Exclusion • None

Author (year)	Study design (LoE)	Outcome measure evaluated	Follow-up (% followed)	Demographics	Patient characteristics	Interventions	Inclusion/Exclusion
Martin (2008)	Retrospective cohort study with prospectively collected data	HOS	Mean length of time between completion of questionnaires (done pre- and postoperatively) : 7 months ± 96 days (range, 55, 420 days) % f/u: NR‡	N = 126 Male: 47% Mean age: 41 ± 16 (range, 13, 80) years Change group§ n = 108 Stable group§ n = 18	• Arthroscopic surgery: 100% • Diagnosis: NR • Previous hip surgery (other): NR	• Questionnaire completed preoperatively and at 6 months' follow-up (or the last measurement taken closest to it was used if 6 months follow-up data were unavailable). • Hip arthroscopy procedures: Labral debridement and/or labral tear repair (91%) Osteoplasty for FAI (60%) Debridement-microfracture for chondral lesions (51%) Capsular tightening for capsular laxity (37%) (All patients underwent more than one of the above procedures.) All patients underwent standard postoperative rehabilitation.	Inclusion • Patients who underwent hip arthroscopy and evaluation by the senior author between March 2005 and April 2006. • Musculoskeletal hip pathology appropriate for arthroscopic intervention. • Patients were prospectively part of a larger ongoing study and had records available for review. Exclusion • Primary lumbopelvic pathology, advanced hip arthrosis, or other conditions for which arthroscopic hip surgery was contraindicated. • Patients who did not meet the criteria for the "change group" or the "stable group"§.
Naal (2011)	Prospective cohort study	Hip Outcome Score (HOS)	Length f/u: n/a HOS ADL subscale: 100% (85/85)** HOS sport subscale: 99% (84/85)**	N = 85 Male: 58% Mean age: 33 ± 12 years	• FAI patients undergoing surgery • Previous hip surgery: NR	• Questionnaire filled out at (and/or before) first office visit/consultation Surgery for FAI: surgical approach used depended on the type and extent of the pathology. *Open approach (surgical dislocation)* (n = 57 (41 males)), preferred in patients with: • Large head-neck deformities (i.e., high alpha-angle) • Very muscular male patients • Complex labral lesions in those with pincer-type FAI *Surgical hip arthroscopy* (n = 28 (8 males)), preferred in patients with: • Female sex with less extended pathologies • Cam-type FAI	Inclusion • Consecutive patients undergoing surgery for FAI in two centers in Germany (April – September 2009) • Diagnosis of FAI based on patient history, clinical exam (reduced flexion and internal rotation and positive impingment test), conventional radiographs (anteriorposterior pelvis and cross-table lateral view), and obligatory magnetic resonance imaging (intra-articular gadolinium contrast of the involved hip) Exclusion • None

Author (year)	Study design (LoE)	Outcome measure evaluated	Follow-up (% followed)	Demographics	Patient characteristics	Interventions	Inclusion/Exclusion
Rothenfluh (2008)	Prospective case-control study	Western Ontario and McMasters Universities Osteoarthritis Index (WOMAC)	Length f/u: n/a FAI/OA pts: 157/200 filled out form completely	N = 400 (FAI + OA + healthy controls) Male: 48%†† Mean age: 36.8 ± 7.7 years†† FAI n = 100†† Male: 45%†† Mean age: 31.7 ± 9.7 years†† OA n = 57†† Male: 49%†† Mean age: 60.3 ± 11.7 years†† Matched healthy controls: n = 200 Male: 49.5% Mean age: 32.6 ± 5.6 years	• FAI or OA patients at their first office visit • Previous hip surgery: 0% • Participation in sporting activities: 81%†† FAI: 83%†† OA: 35%†† Matched healthy control: 85%	• Questionnaire filled out at first office visit/consultation	Inclusion • FAI: positive anterior or posterior impingement test, correlated pathomorphology of the hip radiographically (decreased offset, femoral head asphericity and/or deep retroverted acetabulum) and by arthro-MRI and no radiographic signs of OA (joint space loss, hypersclerosis, cysts, osteophytes) according to Tönnis. • OA: diagnosed by clinical exam and radiography (Tönnis grade > 1). • No prior hip surgery; no comborbidities that may affect ambulation or cause pain. • 200 age- and gender- matched control patients with no hip pain were randomly selected out of a pool of 16,191 questionnaires. Exclusion • None
Christensen (2003)	Prospective cohort study	Nonarthritic Hip Score (NAHS)	Length f/u: n/a Internal consistency and validity: 90% (43/48) Reliability: 100% (17/17)	Internal consistency and validity: N = 48 Male: 40% Mean age: 33 (range, 16–45) years Reliability N = 17 (additional patients) Male: 35% Mean age: 32 (range, NR) years	• Patients with chronic hip pain (≥ 6 months) unresponsive to other therapies (see inclusion/exclusion critieria)	• Questionnaire filled out at (and/or before) first office visit/consultation	Inclusion • Consecutive patients referred for hip pain that had been present for 6 months and who had not improved with nonsteroidal anti-inflammatory drugs, therapy, or injections. • Normal results from plain radiographic (anterior view of the pelvis, anterior and lateral views of the affected hip). Exclusion • None

FAI: femoroacetabular impingement; OA: osteoarthritis

* 507 patients were included in the analysis, but it was not clear how many patients had been given the questionnaire. In addition, the authors excluded patients who had too many items on their questionnaire that could not be scored (see inclusion/exclusion criteria) but the number of patients who were excluded for this reason was not reported.

† Of the 116 patients who returned the questionnaire, 9 reported having surgery after August 2003 and were excluded. Thus a total of 107 patients were included for analysis. Demographic information was reported for these 107 patients.

‡ 126 patients were included in the analysis, but it was not clear how many patients had been given the questionnaire.
§ The "change group" consisted of patients whose condition changed as measured by the patient reporting being "much improved" or "somewhat improved" and having a "normal" or "nearly normal" level of functioning. The "stable group" consisted of patients whose condition remained stable as measured by the patient reporting being "unchanged" and having a "abnormal" or "severely abnormal" level of functioning.
** For the ADL subscale of the HOS questionnaire, 75/85 filled out the form completely, while 8/85 patients left one item blank and 2/85 left two or three items blank. For the sport subscale, 65/85 patients filled out the form completely, while 14/85 patients left one item blank, 3/85 left two items blank, and 1/85 left three items blank. The latter patient was excluded from analysis. The ADL and sport subscales consist of 17 and 9 scored items, respectively, and the final score of each is a percentage of the total possible score (which is adjusted appropriately if answers are left blank) for that subscale.
†† Demographic information provided for patients after loss to follow-up (i.e., the 43 FAI or OA patients who did not completely fill out the questionnaire).

Table F5. Outcome measures validated in FAI, labral tear, or hip arthroscopy patients.

Study	Outcome measure	Evaluation	Results
Martin (2006)	HOS ADL	Test for unidimensionality (PRELIS software)	n = 430 patients (85%) assessed as PRELIS requires the use of complete data α value set to .005 because there were 10 comparisons • gender: *P* = .94 (NS) • age: *P* = .009 (NS) • duration of symptoms: *P* = .7 (NS) • time between surgery and data collection: *P* = .012 (NS) • current rating of function: *P* = .18 (NS)
	HOS sport	Test for unidimensionality (PRELIS software)	n = 343 patients (68%) assessed as PRELIS requires the use of complete data α value set to .005 because there were 10 comparisons • gender: *P* < .0005 (ratio of females:males was lower in the group with no missing data compared with the group with one or two missing responses) • age: *P* = .58 (NS) • duration of symptoms: *P* = .37 (NS) • time between surgery and data collection: *P* = .39 (NS) • current rating of function: *P* = .56 (NS)
	HOS ADL	Test for unidimensionality (factor loading of individual items)	Factor analysis of the 19-item ADL subscale indicated that the items loaded on two factors (eigenvalues > 1): item 3 (putting on socks and shoes), and item 11 (sitting for 15 minutes). Subsequent factor analysis of the modified 17-item ADL subscale (i.e., without items 3 or 11, above)) loaded on one factor, which accounted for 68% of the variance (eigenvalue of 11.6). **Scores from the modified 17-item ADL subscale were used for validation testing.**
	HOS sport	Test for unidimensionality (factor loading of individual items)	Factor analysis of the 9-item sports subscale loaded on one factor, which accounted for 80.3% of the variance (eigenvalue of 7.1). **Scores from the 9-item sports subscale were used for validation testing.**
	HOS ADL	Item characteristic curves	Items that did <u>not</u> have well-fitting curves: • getting into a car • going up steps • going down steps • going up and down curves These items were considered for exclusion and test information function was recalculated separately with each item deleted. For each item, a decrease in information was found throughout the range of ability and the items were thus retained to maximize the ability of the HOS to measure across patients' ranges of abilities.
	HOS sport	Item characteristic curves	Items that did <u>not</u> have well-fitting curves: • none
	HOS ADL	Internal consistency	Cronbach α = 0.96 SEM = 2.8 90% CI: ± 4.8
	HOS sport	Internal consistency	Cronbach α = 0.95 SEM = 2.3 90% CI: ± 3.8
	HOS ADL	Construct validity	• SF-36 PCS: *r* = 0.76 (*P* < .0005) • SF-36 PCSS: *r* = 0.74 (*P* < .0005) • SF-36 MCS: *r* = 0.27 (*P* < .0005) • SF-36 MCSS: *r* = 0.18 (*P* < .0005)
	HOS sport	Construct validity	• SF-36 PCS: *r* = 0.72 (*P* < .0005) • SF-36 PCSS: *r* = 0.68 (*P* < .0005) • SF-36 MCS: *r* = 0.23 (*P* < .0005) • SF-36 MCSS: *r* = 0.10 (*P* < .0005)

Martin (2007)	HOS ADL	Construct validity	• SF-36 PCS: $r = 0.86$ ($P < .005$ compared with measures of mental functioning scores) • SF-36 PCSS: $r = 0.80$ ($P < .005$ compared with measures of mental functioning scores) • SF-36 MCS: $r = 0.41$ • SF-36 MCSS: $r = 0.17$
	HOS sport	Construct validity	• SF-36 PCS: $r = 0.84$ ($P < .005$ compared with measures of mental functioning scores) • SF-36 PCSS: $r = 0.81$ ($P < .005$ compared with measures of mental functioning scores) • SF-36 MCS: $r = 0.43$ • SF-36 MCSS: $r = 0.18$
	HOS patient-rated function	Descriptive results (mean ± SD)	• "Normal" (n = 26): 23% • "Nearly normal" (n = 45): 42% • "Abnormal" (n = 24): 24% • "Severely abnormal" hip function (n = 7): 6% • No answer (n = 5): 5%
	HOS patient-rated surgical outcome	Descriptive results (mean ± SD)	• "Excellent/good" (n = 86): 80% • "Fair/poor" (n = 20): 19% • No answer (n = 1): 1%
	HOS ADL	Descriptive results (mean ± SD)	Activity level • "Normal" (n = NR): 96 ± 5 (range, 76, 100) • "Nearly normal" (n = NR): 89 ± 8 (range, 68, 100) • "Abnormal" (n = NR): 64 ± 17 (range, 35, 90) • "Severely abnormal" (n = NR): 31 ± 10 (range, 19, 47) ($P < .05$ between each group) Surgical outcome • "Excellent/good" (n = NR): 85 ± 19 (range, 22, 100) • "Fair/poor" (n = NR): 63 ± 21 (range, 19, 88) ($P < .05$ between each group) Age • Below median age (44.2 years): 87 ± 13 (range, 53, 100) • Above median age (44.2 years): 74 ± 25 (range, 19, 100) ($P < .05$ between each group)
	HOS sport	Descriptive results (mean ± SD)	Activity level • "Normal" (n = NR): 94 ± 7 (range, 78, 100) • "Nearly normal" (n = NR): 87 ± 16 (range, 39, 100) • "Abnormal" (n = NR): 40 ± 16 (range, 6, 56) • "Severely abnormal" (n = NR): 6 ± 4 (range, 0, 11) ($P < .05$ between each group) Surgical outcome • "Excellent/good" (n = NR): 72 ± 29 (range, 0, 100) • "Fair/poor" (n = NR): 45 ± 21 (range, 6, 91) ($P < .05$ between each group) Age • Below median age (44.2 years): 78 ± 20 (range, 25, 100) • Above median age (44.2 years): 55 ± 33 (range, 0, 100) ($P < .05$ between each group)
Martin (2008)	HOS ADL	Descriptive results (mean ± SD)	Mean scores: • Change group* (n = 108), preoperative: 65.3 ± 19 (range, 13, 100) • Change group* (n = 108), follow-up: 87.7 ± 14 (range, 17, 100) • Stable group* (n = 18), preoperative: 65.3 ± 25 (range, 17, 100) • Stable group* (n = 18), follow-up: 68.7 ± 22 (range, 25, 100)

	HOS sport	Descriptive results (mean ± SD)	Mean scores: • Change group* (n = 108), preoperative: 65.3 ± 19 (range, 13, 100) • Change group* (n = 108), follow-up: 87.7 ± 14 (range, 17, 100) • Stable group* (n = 18), preoperative: 65.3 ± 25 (range, 17, 100) • Stable group* (n = 18), follow-up: 68.7 ± 22 (range, 25, 100)
	HOS ADL	Reliability (test-retest)	• Test (stable group*, preoperative (n = 18)): 65.3 ± 25 (range, 17, 100) • Retest (stable group*, follow-up (n = 18)): 65.3 ± 25 (range, 17, 100) • ICC: 0.98 • MDC†: ± 3 points
	HOS sport	Reliability (test-retest)	• Test (stable group*, preoperative (n = 18)): 65.3 ± 25 (range, 17, 100) • Retest (stable group*, follow-up (n = 18)): 68.7 ± 22 (range, 25, 100) • ICC: 0.92 • MDC†: ± 3 points
	HOS ADL	Responsiveness	Difference in mean scores (preoperative to follow-up): • Change group* (n = 108): 22.4 ± 18 (range, -57, 76) • Stable group* (n = 18): 3.7 ± 7.3 (range, -9, 23) Effect size‡ = 1.2 Group-by-time interaction: $F_{1,125} = 21.4$ ($P < .0005$)
	HOS sport	Responsiveness	Difference in mean scores (preoperative to follow-up): • Change group* (n = 108): 34.5 ± 26.2 (range, -24, 100) • Stable group* (n = 18): -3.7 ± 13.2 (range, -25, 30) Effect size‡ = 1.5 Group-by-time interaction: $F_{1,88} = 33.5$ ($P < .0005$)
	HOS ADL	MCID	• Area under ROC curve: 0.88 (95% CI, 0.80, 0.95) MCID§ of 9 points: • Sensitivity: 0.82 • Specificity: 0.89
	HOS sport	MCID	• Area under ROC curve: 0.90 (95% CI, 0.83, 0.97) MCID§ of 6 points: • Sensitivity: 0.85 • Specificity: 0.87
Naal (2011)	HOS-D ADL	Reliability (test-retest)	Subset of n = 33 patients assessed Median time between test and retest: 10 days Mean scores: • Test: 78 ± 18 points • Retest: 77 ± 18 • ICC: 0.94 (95% CI, 0.80, 0.97) • SEM: ± 4 (95% CI, 3,6) points • MDC**: 11 points
	HOS-D sport	Reliability (test-retest)	Subset of n = 33 patients assessed • Median time between test and retest: 10 days Mean scores: • Test: 53 ± 23 points • Retest: 56 ± 24 • ICC: 0.89 (95% CI, 0.80, 0.95) • SEM: ± 8 (95% CI, 6,11) points • MDC**: 22 points
	HOS-D ADL	Internal consistency	• Cronbach α: 0.95††
	HOS-D sport	Internal consistency	• Cronbach α: 0.91††
	HOS-D ADL	Construct validity	• UCLA activity scale: $r = 0.62$ ($P < .001$) • WOMAC pain: $r = –0.81$ ($P < .001$) • WOMAC function: $r = –0.90$ ($P < .001$) • WOMAC stiffness: $r = –0.63$ ($P < .001$) • WOMAC total: $r = –0.90$ ($P < .001$) • OHS: $r = –0.85$ ($P < .001$) • SF-12 PCS: $r = 0.79$ ($P < .001$) • SF-12 MCS: $r = –0.08$ ($P = .479$)

	HOS-D sport	Construct validity	• UCLA activity scale: $r = 0.58$ ($P < .001$) • WOMAC pain: $r = –0.62$ ($P < .001$) • WOMAC function: $r = –0.71$ ($P < .001$) WOMAC stiffness: $r = –0.48$ ($P < .001$) • WOMAC total: $r = –0.70$ ($P < .001$) • OHS: $r = –0.70$ ($P < .001$) • SF-12 PCS: $r = 0.72$ ($P < .001$) • SF-12 MCS: $r = 0.08$ ($P = .455$)
	HOS-D ADL	Floor and ceiling effects‡‡	• Floor effects: none • Ceiling effects: 21%
	HOS-D sport	Floor and ceiling effects‡‡	• Floor effects: 18% • Ceiling effects: 12%

	HOS-D ADL	Descriptive results (mean ± SD)	• All patients (n = 85): 70.7 ± 18.4 (range, 34.2, 100)
			• Hip arthroscopy patients (n = 28): 65.0 ± 16.0 (range, 40.8, 93.4)
			• Surgical dislocation patients (n = 57): 73.0 ± 19.0 (range, 34.2, 100)
			• Males (n = 49): 77.2 ± 18.6
			  • Females (n = 36): 61.0 ± 13.4 ($P < .001$ compared with males)
			• Patients < median age (31 years): 75.1 ± 17.9
			  • Patients > median age (31 years): 65.7 ± 18.0 ($P = .018$ compared with patients < median age)
			• "Normal" hip function§§ (n = 2): 96.6 ± 4.7***
			• "Nearly normal" hip function§§ (n = 20): 83.8 ± 12.2***
			• "Abnormal" hip function§§ (n = 43): 66.7 ± 17.3***
			• "Severely abnormal" hip function§§ (n = 18): 62.1 ± 17.5***

	HOS-D sport	Descriptive results (mean ± SD)	• All patients (n = 85): 47.1 ± 23.3 (range, 6.3, 100)
			• Hip arthroscopy patients (n = 28): 43.4 ± 22.9 (range, 9.4, 90.6)
			• Surgical dislocation patients (n = 57): 48.8 ± 23.5 (range, 6.3, 100)
			• Males (n = 49): 53.6 ± 25.3
			  • Females (n = 36): 37.4 ± 15.8 ($P < .001$ compared with males)
			• Patients < median age (31 years): 50.9 ± 24.4
			  • Patients > median age (31 years): 42.8 ± 21.5 ($P = .108$ compared with patients < median age)
			• "Normal" hip function§§ (n = 2): 95.3 ± 6.6)**
			• "Nearly normal" hip function§§ (n = 20): 64.0 ± 16.7**
			• "Abnormal" hip function§§ (n = 43): 41.1 ± 19.4**
			• "Severely abnormal" hip function§§ (n = 18): 34.3 ± 20.1**

	UCLA activity scale	Descriptive results (mean ± SD)	• All patients (n = 85): 6.3 ± 2.1 (range, 2, 10)
	WOMAC pain	Descriptive results (mean ± SD)	• All patients (n = 85): 31.4 ± 23.2 (range, 0, 88.0)
	WOMAC function	Descriptive results (mean ± SD)	• All patients (n = 85): 26.3 ± 22.4 (range, 0, 82.4)
	WOMAC stiffness	Descriptive results (mean ± SD)	• All patients (n = 85): 28.9 ± 25.8 (range, 0, 80.0)
	WOMAC total	Descriptive results (mean ± SD)	• All patients (n = 85): 27.6 ± 21.8 (range, 0, 81.3)
	OHS	Descriptive results (mean ± SD)	• All patients (n = 85): 26.1 ± 7.9 (range, 12.0, 44.0)
	SF-12 PCS	Descriptive results (mean ± SD)	• All patients (n = 85): 40.3 ± 9.8 (range, 23.4, 57.2)
	SF-12 MCS	Descriptive results (mean ± SD)	• All patients (n = 85): 50.8 ± 10.7 (range, 27.7, 65.7)
Rothenfluh (2008)	WOMAC-D†††	Rasch analysis	Item fit residual (mean ± SD) • FAI (n = 98): 0.123 ± 1.298 • OA (n = 56): 0.313 ± 1.154 • FAI & OA (n = 154): 0.124 ± 1.681 Person fit residual (mean ± SD) • FAI: –0.118 ± 1.403 • OA: –0.367 ± 2.001

		• FAI & OA: –0.309 ± 1.73 Chi square interaction (value (df); *P*) • FAI: 86.107 (*P* = .000607) • OA: 78.332 (*P* = .00371) • FAI & OA: 135.373 (*P* = .000001) Person separation index (PSI) • FAI: 0.94873 • OA: 0.94269 • FAI & OA: 0.94971 t-test for unidimensionality • FAI: 21.65% • OA: 18.87% • FAI & OA: 22.00% (Results suggest the construct is multidimensional) **Because patients with FAI and OA responded similarly, the data of FAI and OA patients were pooled for all subsequent analysis**
WOMAC-D††† pain subset, all items	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 138): 0.298 ± 0.983 Person fit residual (mean ± SD) • FAI & OA: –0.366 ± 1.129 Chi square interaction (value (df); *P*) • FAI & OA: 12.438 (*P* = .2568) Person separation index (PSI) • FAI & OA: 0.8841 t-test for unidimensionality • FAI & OA: 2.99% (Results suggest the construct is unidimensional)
WOMAC-D††† pain subset, item 3 (pain sitting/lying) removed‡‡‡	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 137): 0.179 ± 0.321 Person fit residual (mean ± SD) • FAI & OA: –0.425 ± 1.090 Chi square interaction (value (df); *P*) • FAI & OA: 8.640 (*P* = .3736) Person separation index (PSI) • FAI & OA: 0.86906 t-test for unidimensionality • FAI & OA: 3.01% (Results suggest the construct is unidimensional)
WOMAC-D††† function subset, all items	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 142): 0.020 ± 1.605 Person fit residual (mean ± SD) • FAI & OA: –0.362 ± 1.589 Chi square interaction (value (df); *P*) • FAI & OA: 57.089 (*P* = .00413) Person separation index (PSI) • FAI & OA: 0.9556 t-test for unidimensionality • FAI & OA: 18.44% (Results suggest the construct is multidimensional)
WOMAC-D††† function subset, items 7 (lying in bed), 24/25 (heavy/light chores), were removed§§§	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 139§§§): –0.057 ± 1.603 Person fit residual (mean ± SD) • FAI & OA: –0.286 ± 1.325 Chi square interaction (value (df); *P*) • FAI & OA: 49.818 (*P* = .01295) Person separation index (PSI) • FAI & OA: 0.9540 t-test for unidimensionality • FAI & OA: 16.79% (Results suggest the construct is multidimensional)

WOMAC-D††† function subset, items 7 (lying in bed), 24/25 (heavy/light chores), 20 (bending), 21/22 (putting on/off socks) were removed§§§,****	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 137§§§): –0.024 ± 1.521 Person fit residual (mean ± SD) • FAI & OA: –0.314 ± 1.262 Chi square interaction (value (df); *P*) • FAI & OA: 30.618 (*P* = .1042) Person separation index (PSI) • FAI & OA: 0.93770 t-test for unidimensionality • FAI & OA: 6.82% (Results suggest the construct is multidimensional)
WOMAC-D††† function subset, items 7 (lying in bed), 24/25 (heavy/light chores), 20 (bending), 21/22 (putting on/off socks), and 18 (getting on/off toilet) were removed §§§,****,††††	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 137§§§): 0.103 ± 1.229 Person fit residual (mean ± SD) • FAI & OA: –0.304 ± 1.264 Chi square interaction (value (df); *P*) • FAI & OA: 27.618 (*P* = .1187) Person separation index (PSI) • FAI & OA: 0.93040 t-test for unidimensionality • FAI & OA: 3.79% (Results suggest the construct is unidimensional; no DIF could be detected in the reduced subset for all person factors)
WOMAC-D††† combined pain + function subsets, items 3, 7, 24/15, 20, 21/22, 18, and stiffness items removed	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 147): 0.118 ± 1.361 • Person fit residual (mean ± SD) • FAI & OA: –0.267 ± 1.406 Chi square interaction (value (df); *P*) • FAI & OA: 50.724 (*P* = .0190) Person separation index (PSI) • FAI & OA: 0.94062 t-test for unidimensionality • FAI & OA: 13.89% (Results suggest the construct is multidimensional)
WOMAC-D††† combined pain + function subsets, items 1 (night pain), 3, 7, 24/25, 20, 21/22, 18, and stiffness items removed‡‡‡‡	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 140‡‡‡‡): 0.086 ± 1.226 • Person fit residual (mean ± SD) • FAI & OA: –0.255 ± 1.333 Chi square interaction (value (df); *P*) • FAI & OA: 43.973 (*P* = 0.4796) Person separation index (PSI) • FAI & OA: 0.94132 t-test for unidimensionality • FAI & OA: 11.43% (Results suggest the construct is multidimensional)
WOMAC-D††† combined pain + function subsets, items 1, 3, 7, 11 (getting out of bed) 24/25, 20, 21/22, 18, and stiffness items removed§§§§	Rasch analysis	Item fit residual (mean ± SD) • FAI & OA (n = 139‡‡‡‡): –0.111 ± 1.045 • Person fit residual (mean ± SD) • FAI & OA: –0.318 ± 1.286 Chi square interaction (value (df); *P*) • FAI & OA: 25.534 (*P* = 0.377) Person separation index (PSI) • FAI & OA: 0.93125 t-test for unidimensionality • FAI & OA: 6.62% (Results suggest the construct is unidimensional)
Individual items in 12-item WOMAC-D†††,*****	Descriptive results (mean ± SD)	• FAI (n = 100): 8.32 ± 7.32 • Normal population control group (n = 200): 0.39 ± 2.90 • *P* < .001; *t* = –8.5269 • Effect size: *r* = 0.71 (score > 0.5 indicates large difference between groups) • Post-hoc statistical power = 0.999983, which indicates the probability of

			a type II error is less than 0.0001 (which is below the accepted 0.2) • The high PSI of 0.93 and the high statistical power of downstream analysis of the raw score between groups suggests that the 12-item WOMAC can reliably discriminate between groups with a very low probability that a difference is not detected while present. • FAI (n = 100): 8.32 ± 7.32 • OA (n = 57): 16.23 ± 8.04 • *P* < .001; *t* = –7.7034 • Effect size: 0.45 • Post-hoc statistical power = 0.81
Christensen (2003)	NAHS	Reliability (test-retest)	n = 17 patients assessed (different subset of patients that those used to evaluate internal consistency and reliability) Mean time between test and retest: 5.5 (range, 1–16) days Mean scores: • Test: NR • Retest: NR Pearson correlation coefficient • Overall reliability: r = 0.96 (range for subscales, 0.87, 0.95) • Pain subscale: r = 0.92 (range for each question, 0.63, 0.9)††††† • Mechanical symptom subscale: r = 0.87 (range for each question, 0.72, 0.97) • Physical function subscale: r = 0.92 (range for each question, 0.84, 0.93) • Activity level subscale: r = 0.95 (range for each question, 0.81, 0.93)†††††
	NAHS	Internal consistency	n = 48 patients assessed Cronbach's coefficient alpha for each subscale • Pain subscale: α = 0.87 • Mechanical symptoms subscale: α = 0.69 • Physical function: α = 0.85 • Activity level subscale: α = 0.92
	NAHS	Validity	Mean scores ± SD; Pearson correlation coefficient (r) • NAHS: 56.0 ± 18.1 (range, 12.5, 92.5) • HHS (n = 46): 61.2 ± 16.6 (range, 24, 96) (r = 0.82 for NAHS/HHS) • NAHS pain subscale: mean score NR (r = 0.73 with HHS) • NAHS mechanical symptoms subscale: mean score NR (r = 0.61 with HHS) • NAHS physical function subscale: mean score NR (r = 0.73 with HHS) • NAHS activity level subscale: mean score NR (r = 0.76 with HHS) • SF-12 (n = 43): 81.9 ± 10.9 (range, 22, 56) (r = 0.59 for NAHS/SF-12) • SF-12 PCS: mean score NR (r = 0.37 for NAHS/SF-12 PCS) • SF-12 MCS: mean score NR (r = 0.51 for NAHS/SF-12 MCS)

DIF: differential item functioning; FAI: femoroacetabular impingement; HOS-D: Hip Outcome Score (German language version); HOS-ADL: HOS activities of daily living subscale; HOS-sport: HOS sport subscale; ICC: intraclass correlation coefficient; MCID: minimal clinically important difference; MCS: Mental Component Scale; MCSS: Mental Component Summary Score; MDC: minimal detectable change; NAHS: Nonarthritic Hip Score; OHS: Oxford Hip Score; PCS: Physical Component Scale; PCSS: Physical Component Summary Score; ROC: receiver operating characteristic; SEM: standard error of measurement; SF-12: Short Form-12;SF-36: Short Form-36; UCLA: University of California, Los Angeles; WOMAC: Western Ontario and McMasters Universities Osteoarthritis Index; WOMAC-D: WOMAC (German language version)

* The "change group" consisted of patients whose condition changed as measured by the patient reporting being "much improved" or "somewhat improved" and having a "normal" or "nearly normal" level of functioning. The "stable group" consisted of patients whose condition remained stable as measured by the patient reporting being "unchanged" and having a "abnormal" or "severely abnormal" level of functioning.
†MDC based on 95% CI
‡ Effect size = (mean change in score in the change group) / SD of preoperative scores.

§ MCID determined by whether the patient reported being "much improved" or "somewhat improved" and having a "normal" or "nearly normal" level of functioning.
** MDC = SEM x √2 x 1.9, which is the individual minimal change that could be thought of as "real" (i.e., not attributable to measurement error) with an acceptable level of probability.
†† removal of any item did not improve the Cronbach α.
‡‡ floor and ceiling effects calculated as the percentage of patients with the worst and best values; "worst" value = 0 points (actual worst end-anchor score) + MDC; "best" value = 100 points (actual best end-anchor score) – MDC.
§§ Patient- reported assessment of overall hip function.
*** all between-group differences (regarding overall hip function) were statistically significant except between patients who considered their hip function to be abnormal and those who considered it to be severely abnormal.
††† All items were graded by patients on a 7-point Likert scale rather than on the original 5-point scale in an attempt to increase the resolution of the scale. However, this resulted in "threshold disordering" and the thresholds had to be rescored back to the original 5-point scale. This resulted in ordered thresholds for all items except item 4, pain walking flat, which was subsequently rescored to a 4-point scale.
‡‡‡ DIF was detected for item 3 (pain sitting/lying) for the person factor disease, indicating patients with FAI responded differently to this item than did those with OA. Thus this item was subsequently removed.
§§§ Items 7 (lying in bed), 24 (heavy chores), and 25 (light chores) were indicated to have a significant chi square statistic after Bonferroni correction and were thus omitted from this analysis. Three patients also did not meet the criteria from individual person fit and were also removed.
**** Items 20-22 (bending, putting on/off socks) were removed as there were high correlations among items involving some kind of bending, which suggests that they may act as a second underlying concept and introduce multidimensionality.
†††† Item 18 (getting on/off toilet) still showed high fit residual and was thus removed.
‡‡‡‡ Item 1 (night pain) removed from the combined subsets as it showed a high fit residual and probability below the Bonferroni adjustement; six patients were also removed from analysis as they showed fit residuals outside the interval of ±2.5 for individual person fit.
§§§§ Item 11 (getting out of bed) removed as it had residual correlation with item 10 (arising from sitting), implying local response dependency.
***** 12- item WOMAC-D excludes items 1, 3, 7, 11, 18, 20–22, 24–25, and stiffness items as described in footnotes ‡‡ through §§§. This construct shows no DIF between person factors (FAI versus OA, sports, age groups, or sex) (data not reported).
††††† The reported mean Pearson value for pain subscale is outside the reported range of values; data reported here as in original report.

Table F6. Quality assessment of outcome measures evaluated in FAI /labral tear/ hip arthroscopy population (adapted from Lodhia et al. 2011/Terwee et al. 2007)

Instrument	Content validity	Internal consistency	Criterion validity	Construct validity	Reproducibility		Responsiveness	Floor/ceiling	Interpretability
					Agreement	Reliability			
HOS/ HOS-D	–	+	0	+	+	+	+	+	+
NAHS	+	?	+	+	0	?	0	+	0
12-item modified WOMAC	0	+	0	?	0	0	0	0	0

Plus sign indicates a positive rating, a question mark indicates an indeterminate rating, a minus sign indicates a negative rating, and 0 indicates no information available.

APPENDIX G. STUDY SUMMARIES FOR EFFECTIVENESS

Table G1. Failure, Conversion to THA, and OA Progression in Non- and Recreational Athletes.

Outcome	No. patients (no. hips)	Mean age years (range)	Male %	No. of hips with outcome (%)	FAI diagnosis (%)	Mean Follow-up (years)
ARTHROSCOPY						
Failure						
Ilizaturri 2008*	19 (19)	34 (27–43)	58	1 (5.3)	cam (100)	2.4 (2–3.0)
Philippon 2009†	112 (112)	40.6 (38–44)	45	10 (8.9)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Conversion to THA						
Byrd 2009	200 (207)	33 (NR)	69	1 (0.5)	cam (79.0); mixed (21.0)	1.3 (1–2)
Flecher 2011	23 (23)	34 (17–54)	61	0 (0)	NR	≥ 1
Gedouin, May 2010‡	110 (111)	31 (16–49)	71	5 (4.5)	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Horisberger 2010	88 (105)	40.9 (17–66)	68	9 (8.6)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Horisberger 2010§	20 (20)	47.3 (22–65)	80	10 (47.6)	cam (55.0); mixed (45.0)	3.0 (1.5–4.1)
Ilizaturri 2008	19 (19)	34 (27–43)	58	1 (5.3)	cam (100)	2.4 (2–3.0)
Javed 2011	40 (40)	65 (60–82)	65	7 (17.5)	cam (100)	2.5 (1–4.5)
Nassif 2010**	163 (178)	32.7 (15–56)	51	2 (1.2)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
Philippon 2009	112 (112)	40.6 (38–44)	45	10 (8.9)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Radiographic OA progression						
Gedouin, Duperron 2010	38 (38)	36 (24–64)	87	3 (8.3)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Gedouin, May 2010	110 (111)	31 (16–49)	71	0 (0)	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Ilizaturri 2008	19 (19)	34 (27–43)	58	1 (5.3)	cam (100)	2.4 (2–3.0)
OPEN DISLOCATION						
Failure						
Beaule 2007††	34 (37)	40.5 (19–54)	53	6 (16)	cam (100)	3.1 (2.1–5.0)
Beck 2004†	19 (19)	36 (21–52)	74	5 (26.3)	NR	4.7 (4.2–5.2)
Murphy 2004†	23 (23)	35	57	7 (30.4)	cam (44.0); pincer (4.0); mixed (52.0)	5.2 (2–12)
Peters 2010‡‡	94 (96)	28	59	6 (6.3)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Siebenrock 2003§§	22 (29)	23	66	1 (3.4)	pincer (100)	2.5 (2–4.1)
Conversion to THA						
Beaule 2007	34 (37)	40.5 (19–54)	53	0 (0)	cam (100)	3.1 (2.1–5.0)
Beck 2004	19 (19)	36 (21–52)	74	5 (26.3)	NR	4.7 (4.2–5.2)
Murphy 2004	23 (23)	35.4 (17–54)	57	7 (30.4)	cam (44.0); pincer (4.0); mixed (52.0)	5.2 (2–12)
Peters 2010	94 (96)	28 (14–51)	59	5 (5.2)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Siebenrock 2003	22 (29)	23 (14–41)	66	0 (0)	pincer (100)	2.5 (2–4.1)
Radiographic OA progression						
Beck 2004	19 (19)	36 (21–52)	74	2 (14.3)	NR	4.7 (4.2–5.2)
Peters 2010	94 (96)	28 (14–51)	59	25 (26)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
MINI-OPEN						
Failure						
Laude 2009†	97 (100)	33.4 (16–56	52	11 (11)	NR	4.9 (2.4–8.7)

Outcome	No. patients (no. hips)	Mean age years (range)	Male %	No. of hips with outcome (%)	FAI diagnosis (%)	Mean Follow-up (years)
Conversion to THA						
Clohisy 2010	41 (41)	34 (16–48)	80	0 (0)	cam (100)	2.2 (2–3)
Hartmann 2009	33 (34)	31 (15–47)	52	1 (3)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Laude 2009	97 (100)	33.4 (16–56	52	11 (11)	NR	4.9 (2.4–8.7)
Lincoln 2009	14 (16)	37 (17–51)	71	1 (5.3)	cam (63.0); mixed (37.0)	2.0 (1.3–3)
Ribas 2007	32 (35)	36.2 (23–48)	72	1 (2.9)	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
Radiographic OA progression						
Clohisy 2010	41 (41)	34 (16–48)	80	2 (5.7)	cam (100)	2.2 (2–3)
Laude 2009	97 (100)	33.4 (16–56	52	11 (11)	NR	4.9 (2.4–8.7)
Lincoln 2009	14 (16)	37 (17–51)	71	1 (5.3)	cam (63.0); mixed (37.0)	2.0 (1.3–3)

FAI: femoroacetabular impingement; NR: not reported. OA: osteoarthritis; THA: total hip arthroplasty.

*Failure defined as advanced OA and/or recommended THA.

†Failure defined as conversion to THA.

‡Three patients underwent hip resurfacing.

§One patient ended up with bilateral pathology and THA was recommended in both hips (Arthroscopy journal).

**One of the two patients underwent hip resurfacing.

††Failure defined as unsatisfactory outcome and no clinical improvement and/or worsening WOMAC score.

‡‡Failure defined as conversion to THA or worse HHS score.

§§Failure defined as fair results/residual pain.

Table G2. Patient- and Clinician-Reported Functional Outcomes in Non- or Recreational Athletes.

Outcome	No. patients (no. hips)	Mean age in years (range)	Male %	Mean pre-op score (± SD)	Mean Change Pre-Post (% mean change)	FAI diagnosis (%)	Mean follow-up (years)
ARTHROSCOPY							
PATIENT REPORTED							
Hip Outcome Score - ADLs							
Philippon 2008	16 (17)	15 (11–16)	12.5	58	36 (62.1)	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Philippon 2009	112 (112)	40.6 (38–44)	44.6	70	17.8 (25.4)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Hip Outcome Score - Sport							
Philippon 2008	16 (17)	15 (11–16)	12.5	33	56 (169.7)	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Philippon 2009	112 (112)	40.6 (38–44)	44.6	43	26 (60.5)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Nonarthritic Hip Score							
Brunner 2009	53 (53)	42 (17–66)	77.4	54.4	31.3 (57.5)	cam (58.5); mixed (41.5)	2.4 (2–3.2)
Flecher 2011	23 (23)	34 (17–54)	60.9	64	20 (31.3)	NR	≥ 1
Haviv, O'Donnell 2010*	82 (164)	29.4 (14–63)	81.7	68.5	21.5 (31.4)	cam (100)	2.2 (1.0–6.7)
Horisberger 2010†	88 (105)	40.9 (17–66)	68.0	56.7	27.9 (49.2)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Horisberger 2010‡	20 (20)	47.3 (22–65)	80.0	47.2	31.2 (66.1)	cam (55.0); mixed (45.0)	3.0 (1.5–4.1)
Javed 2011	40 (40)	65 (60–82)	65.0	62.1 ± 13.2	15.1 (24.3)	cam (100)	2.5 (1–4.5)
Philippon 2009	112 (112)	40.6 (38–44)	44.6		14	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Stahelin 2008	22 (22)	42 (18–67)	68.2	49	25 (51.0)	cam (100)	0.5 (NR)
WOMAC							
Flecher 2011	23 (23)	34 (17–54)	60.9	58	26 (44.8)	NR	≥ 1
Gedouin, Duperon 2010	38 (38)	36 (24–64)	86.8	55 ± 17	20 (36.4)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Ilizaliturri 2008	19 (19)	34 (27–43)	57.9	82 ± 9	7 (8.5)	cam (100)	2.4 (2–3.0)
Modified Harris Hip Score							
Haviv, O'Donnell 2010	82 (164)	29.4 (14–63)	81.7	72	21.5 (29.9)	cam (100)	2.2 (1.0–6.7)
Javed 2011	40 (40)	65 (60–82)	65.0	60.5 ± 16.3	19.2 (31.7)	cam (100)	2.5 (1–4.5)
Phillipon 2008	16 (17)	15 (11–16)	12.5	55	35 (63.6)	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Phillipon 2009	112 (112)	40.6 (38–44)	44.6	58.0	26.3 (45.3)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
CLINICIAN BASED							
Merle d'Aubigné hip score							
Flecher 2011	23 (23)	34 (17–54)	60.9	11	5 (45.5)	NR	≥ 1
Gedouin, Duperron 2010	38 (38)	36 (24–64)	86.8	14.6 ± 1.8	2.2 (15.1)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Harris Hip Score							
Byrd 2009§	200 (207)	33 (NR)	69.0	68	20 (29.4)	cam (79.0); mixed (21.0)	1.3 (1–2)
Flecher 2011	23 (23)	34 (17–54)	60.9	76	15.0 (19.7)	NR	≥ 1
Nassif 2010	163 (178)	32.7 (15–56)	42.9	64 ± 14	21 (32.8)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
OPEN DISLOCATION							
PATIENT REPORTED							
WOMAC							
Beaule 2007	34 (37)	40.5 (19–54)	52.9	61.2 ± 20	20.2 (33.0)	cam (100)	3.1 (2.1–5.0)

Outcome	No. patients (no. hips)	Mean age in years (range)	Male %	Mean pre-op score (± SD)	Mean Change Pre-Post (% mean change)	FAI diagnosis (%)	Mean follow-up (years)
UCLA Activity Score							
Beaule 2007	34 (37)	40.5 (19–54)	52.9	4.8 ± 1.9	2.7 (56.3)	cam (100)	3.1 (2.1–5.0)
CLINICIAN BASED							
Merle d'Aubigné hip score							
Beck 2004	19 (19)	36 (21–52)	73.7	14.1	2.4 (17)		4.7 (4.2–5.2)
Graves 2009	46 (48)	33 (18–51)	54.3	13 ± 1.7	3.8 (29.2)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
Murphy 2004	23 (23)	35.4 (17–54)	56.5	13.2 ± 1.5	3.7 (28.0)	cam (44.0); pincer (4.0); mixed (52.0)	5.2 (2–12)
Siebenrock 2003	22 (29)	23 (14–41)	86.4	14	2.9 (20.7)	pincer (100)	2.5 (2–4.1)
Harris Hip Score							
Peters 2010	94 (96)	28 (14–51)	58.5	67	24 (35.8)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Yun 2009	14 (15)	35.8 (22–54)	75.0	76	17 (22.4)	cam (60.0); mixed (40.0)	2.3 (1–10)
MINI-OPEN							
PATIENT REPORTED							
Nonarthritic Hip Score							
Clohisy 2010**	41 (41)	34 (16–48)	68.3	75.1 ± 14	15.1 (20.1)	cam (100)	2.2 (2–3)
Laude 2009	97 (100)	33.4 (16–56)	51.5		29.1	NR	4.9 (2.4–8.7)
Modified Harris Hip Score							
Clohisy 2010	41 (41)	34 (16–48)	68.3	63.8 ± 11.1	23.6 (37.0)	cam (100)	2.2 (2–3)
Hartmann 2009	33 (34)	31 (15–47)	51.5	63.9	21.2 (33.2)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
UCLA Activity Score							
Clohisy 2010††	41 (41)	34 (16–48)	68.3	6.1 ± 2.4	2.3 (37.7)	cam (100)	2.2 (2–3)
CLINICIAN BASED							
Merle d'Aubigné hip score							
Ribas 2007	32 (35)	36.2 (23–48)	71.9	13.8	3.1 (22.5)	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
Harris Hip Score							
Lincoln 2009	14 (16)	37 (17–51)	71.4	63.8 ± 5.1	12.3 (19.3)	cam (63.0); mixed (37.0)	2.0 (1.3–3)
CONSERVATIVE							
PATIENT REPORTED							
Nonarthritic Hip Score							
Emara 2011	37 (37)	33 (23–47)	73.0	72 ± 4	19 (26.4)	NR	2.2 (2.1–2.3)
CLINICIAN BASED							
Harris Hip Score							
Emara 2011	37 (37)	33 (23–47)	73.0	72 ± 6	19 (26.4)	NR	2.2 (2.1–2.3)

ADLs: Activities of Daily Living; FAI: femoroacetabular impingement; NR: not reported; UCLA: University of California Los Angeles; WOMAC: Western Ontario and McMaster Universities Arthritis Index.

*This study evaluated arthroscopy treatment for bilateral FAI and divided patients into two groups based upon the timing of their second FAI surgery. For the purpose of this report, this study was treated as a case-series and the preoperative and change scores reflect the average of the corresponding score following the second surgery.

†Clinical and Orthopedic Related Research (journal).

‡Arthroscopy (journal).

§Pre- and postoperative scores were estimated from a graph; mean improvement/change at last follow-up given in the text of the article.
**Preoperative NAHS score includes only 17/35 patients; postoperative score include 32/35 patients.
††Preoperative UCLA activity score includes only 21/35 patients; postoperative score includes 34/35 patients.

Table G3. Pain, Quality of Life, Patient Satisfaction, and Return to Normal Activities in Non- and Recreational Athletes.

	No. patients (no. hips)	Mean age years (range)	Male %	Mean pre-op score (± SD)	Mean Change Pre-Post (% mean change)	FAI diagnosis (%)	Mean follow-up (years)
ARTHROSCOPY							
Pain – VAS (0-10)							
Brunner 2009	53 (53)	42 (17–66)	77	5.7	4.2 (73.7)	cam (58.5); mixed (41.5)	2.4 (2–3.2)
Horisberger 2010*	88 (105)	40.9 (17–66)	68	5.5	4.0 (72.7)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Horisberger 2010†	20 (20)	47.3 (22–65)	80	6.0	4.2 (70.0)	cam (55.0); mixed (45.0)	3.0 (1.5–4.1)
Stahelin 2008	22 (22)	42 (18–67)	68	5.8 ± 2.1	4.4 (75.9)	cam (100)	0.5 (NR)
Patient Satisfaction (satisfied/very satisfied)							
Brunner 2009‡	53 (53)	42 (17–66)	77	50 (94.3)	-----	cam (58.5); mixed (41.5)	2.4 (2–3.2)
Gedouin, Duperron 2010§	38 (38)	36 (24–64)	87	30 (78.9)	-----	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Gedouin, May 2010**	110 (111)	31 (16–49)	71	85 (77.3)	----	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Javed 2011††	40 (40)	65 (60–82)	65	30 (90.9) ‡‡	----	cam (100)	2.5 (1–4.5)
Phillipon 2008§§	16 (17)	15 (11–16)	13	-----	mean 9 (range, 9–10)	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Phillipon 2009§§***	112 (112)	40.6 (38–44)	45	-----	median 9 (range NR)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Return to Normal Activities/Work							
Brunner 2009	53 (53)	42 (17–66)	77	31 (68.9)	-----	cam (58.5); mixed (41.5)	2.4 (2–3.2)
Javed 2011	40 (40)	65 (60–82)	65	30 (90.9) ‡‡	-----	cam (100)	2.5 (1–4.5)
Philippon 2008	16 (17)	15 (11–16)	13	16 (100)	-----	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Philippon 2009***	112 (112)	40.6 (38–44)	45	66 (73.3)	-----	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
OPEN DISLOCATION							
QUALITY OF LIFE							
SF-12 Physical							
Beaule 2007	34 (37)	40.5 (19–54)	53	37.3 ± 10.4	8.3 (22.3)	cam (100)	3.1 (2.1–5.0)
SF-12 Mental							
Beaule 2007	34 (37)	40.5 (19–54)	53	46.4 ± 11.4	4.8 (10.3)	cam (100)	3.1 (2.1–5.0)
Patient Satisfaction (satisfied/very satisfied)							
Beaule 2007‡	34 (37)	40.5 (19–54)	53	28 (82.4)	-----	cam (100)	3.1 (2.1–5.0)
MINI-OPEN							
Patient Satisfaction (satisfied/very satisfied)							
Hartmann†††	33 (34)	31 (15–47)	52	29 (90.6)	-----	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Return to Normal Activities/Work							
Ribas 2007	32 (35)	36.2 (23–48)	72	32 (100)	-----	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
CONSERVATIVE							
Pain – VAS (0-10)							
Emara 2011	37 (37)	33 (23–47)	73	6	4 (66.7)	NR	2.2 (2.1–2.3)

FAI: femoroacetabular impingement; SF-12: Short Form 12; VAS: Visual Analog Scale.
*Clinical and Orthopedic Related Research
†Arthroscopy

‡Patients satisfaction not defined/method used to assess not reported.
§Patients were questioned whether they were subjectively very satisfied, satisfied, or disappointed with their treatment.
**At the end of follow-up, patients were asked if they were disappointed, moderately satisfied, satisfied or very satisfied with the result of their operation.
††Patients were asked if they were satisfied or dissatisfied with the outcome of surgery and whether they would wish to have similar surgery on the contralateral hip with the same indication.
‡‡In the 33 patients who did not have THA.
§§Patients satisfaction was collected on a scale of 1-10, with 10 being very satisfied. The proportion of patients who were satisfied/very satisfied was not reported in these studies.
***Reported in the 90 patients who had 2 years follow-up.
†††Defined as the patient would undergo surgery again.

Table G4. Range of Motion in Non- and Recreational Athletes*.

Study	No. patients (no. hips)	Mean age years (range)	Male %	Mean Internal Rotation (°) (range or ±SD)		Mean Flexion (°) (range or ±SD)		FAI diagnosis (%)	Mean Follow-up (years)
				Pre	Post	Pre	Post		
ARTHROSCOPY									
Brunner 2009	53 (53)	42 (17–66)	77	6.0 (-20, 45)	19 (-5, 45)	107 (60-130)	122 (70–145)	cam (58.5), mixed (41.5)	2.4 (2–3.2)
Horisberger 2010	88 (105)	40.9 (17–66)	68	4.9 (-30, 30)	22.9 (-5, 50)	110 (60-150)	123.3 (70–150)	cam (54.3) mixed (45.7)	2.3 (1.3–4.1)
Horisberger 2010†	20 (20)	47.3 (22–65)	80	2.5 (-30, 20)	24.5 (5–40)	111 (90-135)	125 (115–130)	cam (55.0) mixed (45.0)	3.0 (1.5–4.1)
Stahelin 2008	22 (22)	42 (18–67)	68	4.5 (-20, 20)	22.3 (5–50)	108 (80-135)	124.1 (70–150)	cam (100)	0.5 (NR)
OPEN DISLOCATION									
Siebenrock 2003	22 (29)	23 (14–41)	66	11 (0–30)	21 (0–40)	99 (90-110)	106 (90–120)	pincer (100)	2.5 (2–4.1)
MINI-OPEN									
Lincoln	14 (16)	37 (17–51)	71	7.1 (±1.8)	12.3 (±2.0)	94.1 (±3.0)	110 (±11.9)	cam (63.0) mixed (37.0)	2.0 (1.3–3)
CONSERVATIVE									
Emara	37 (37)	33 (23–47)	73	9.4 (±0.3)	10 (±0.6)	95 (±0.4)	88 (±3.5)	NR	2.2 (2.1–2.3)

NR = not reported; SD = standard deviation.
*ROM was not reported in competitive athletes.
†ROM reported only in the 9 patients who had not yet undergone total hip replacement.

Table G5. Failure, Conversion to THA, and OA Progression in Competitive Athletes*.

Outcome	No. patients (no. hips)	Mean age years (range)	Male %	No. of hips with outcome (%)	FAI diagnosis (%)	Mean Follow-up (years)
OPEN DISLOCATION						
Radiographic OA progression						

Naal 2010	22 (30)	19.7 (16–25)	100	1 (3.3)	cam or mixed (100)	3.8 (1–6.6)

FAI: femoroacetabular impingement; NR: not reported; OA: osteoarthritis; THA: total hip arthroplasty.
*Failure and conversion to THA were not reported in any of the studies on competitive athletes.

Table G6. Patient- and Clinician-Reported Functional Outcomes in Competitive Athletes.

Outcome	No. patients (no. hips)	Mean age years (range)	Male %	Mean pre-op score (± SD)	Mean Change Pre-Post (% mean change)	FAI diagnosis (%)	Mean follow-up (years)
ARTHROSCOPY							
PATIENT REPORTED							
Nonarthritic Hip Score							
Singh 2010	24 (27)	22 (16–29)	100	81	15 (18.5)	cam (81.5); pincer (12.5); mixed (8.3)	1.8 (0.5–5)
Modified Harris Hip Score							
Philippon 2010	28 (28)	27 (18–37)	100	70	25 (35.7)	cam (32.1); mixed (67.9)	2.0 (1–3.5)
Singh 2010	24 (27)	22 (16–29)	100	86	10 (11.6)	cam (81.5); pincer (11.1); mixed (7.4)	1.8 (0.5–5)
OPEN DISLOCATION							
PATIENT REPORTED							
Hip Outcome Score - ADLs							
Naal 2010	22 (30)	19.7 (16–25)	100	NR	NR*	cam or mixed (100)	3.8 (1–6.6)
Hip Outcome Score - Sport							
Naal 2010	22 (30)	19.7 (16–25)	100	NR	NR†	cam or mixed (100)	3.8 (1–6.6)
UCLA Activity Score							
Naal 2010	22 (30)	19.7 (16–25)	100	NR	NR‡	cam or mixed (100)	3.8 (1–6.6)

ADLs: Activities of Daily Living; FAI: femoroacetabular impingement; NR: not reported; UCLA: University of California Los Angeles.
*No preoperative score reported. Only postoperative score reported at final follow-up: 94.5 ± 9.3.
†No preoperative score reported. Only postoperative score reported at final follow-up: 89.1 ± 16.0.
‡No preoperative score reported. Only postoperative score reported at final follow-up: 9.8 ± 0.8.

Table G7. QoL, Patient Satisfaction, Quality of Life and Return to Sports In Competitive Athletes.

	No. patients (no. hips)	Mean age years (range)	Male %	No. of hips with outcome (%)	FAI diagnosis (%)	Mean Follow-up (years)
ARTHROSCOPY						
Patient Satisfaction (satisfied/very satisfied)						
Phillipon 2010	28 (28)	27 (18–37)	100	median 10 (5–10)*	cam (32.1); mixed (67.9)	2.0 (1–3.5)
Singh 2010	24 (27)	22 (16–29)	100	24 (100)	cam (81.5); pincer (12.5); mixed (8.3)	1.8 (0.5–5)
Return to Sports						
Philippon 2007	45 (45)	31 (17–61)	93	35 (77.8)	cam (48.9); pincer (6.7); mixed (46.7)	1.6 (0.5–5.5)
Saw 2004	6 (6)	28 (24–32)	100	5 (83.3)	NR	NR
Singh 2010	24 (27)	22 (16–29)	100	23 (95.8)	cam (81.5); pincer (12.5); mixed (8.3)	1.8 (0.5–5)
OPEN DISLOCATION						
Quality of Life						
SF-12 Physical						
Naal 2010†	22 (30)	19.7 (16–25)	100	51.1 ± 8	cam or mixed (100)	3.8 (1–6.6)
SF-12 Mental						
Naal 2010†	22 (30)	19.7 (16–25)	100	54.3 ± 7.1	cam or mixed (100)	3.8 (1–6.6)
Patient Satisfaction (satisfied/very satisfied)						
Naal 2010	22 (30)	19.7 (16–25)	100	18 (81.8)	cam or mixed (100)	3.8 (1–6.6)
Return to Sports						
Naal 2010	22 (30)	19.7 (16–25)	100	21 (95.5)	cam or mixed (100)	3.8 (1–6.6)

FAI: femoroacetabular impingement; SF-12: Short Form 12.

*Proportion of patients/hips that were satisfied/very satisfied not reported. Satisfaction measured on a scale from 1 to 10, with 10 being very satisfied.

†No preoperative scores reported.

APPENDIX H. SUMMARIES OF STUDIES OF SAFETY

Table H1. Complications in case series reporting treatment for FAI in non- or recreational athletes

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
ARTHROSCOPY						
Reoperation other than THA						
Byrd 2009	200 (207)	33 (NR)	69.0	3 (1.4)	cam (79.0); mixed (21.0)	1.3 (1–2)
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Gedouin, Duperron 2010	38 (38)	36 (24–64)	86.8	0 (0)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Horisberger 2010 (arthro)	20 (20)	47.3 (22–65)	80.0	0 (0)	cam (55.0); mixed (45.0)	3.0 (1.5–4.1)
Haviv, O'Donnell 2010*	82 (164)	29.4 (14–63)	81.7	8 (9.8)	cam (100)	2.2 (1.0–6.7)
Javed 2011	40 (40)	65 (60–82)	65.0	0 (0)	cam (100)	2.5 (1–4.5)
May 2007	5 (5)	40 (27–48)	40.0	1 (20.0)	cam (100)	1.4 (1–2)
Nassif 2010	163 (178)	32.7 (15–56)	42.9	8 (4.9)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Trochanteric nonunion						
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Head-neck fracture						
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Gedouin, May 2010	110 (111)	31 (16–49)	71	1 (0.9)	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Horisberger 2010 (clin orthop)	88 (105)	40.9 (17–66)	68.0	0 (0)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Ilizaliturri 2007	13 (14)	31 (24–39)	46.2	0 (0)	cam (100)	2.5 (2–4)
Ilizaliturri 2008	19 (19)	34 (27–43)	57.9	0 (0)	cam (100)	2.4 (2–3.0)
Javed 2011	40 (40)	65 (60–82)	65.0	0 (0)	cam (100)	2.5 (1–4.5)
Larson 2008	96 (100)	34.7 (16–64)	56.3	0 (0)	cam (17%); pincer (28%); mixed (55%)	0.8 (0.25–3)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Nassif 2010	163 (178)	32.7 (15–56)	42.9	0 (0)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Avascular necrosis						
Ilizaliturri 2007	13 (14)	31 (24–39)	46.2	0 (0)	cam (100)	2.5 (2–4)
Ilizaliturri 2008	19 (19)	34 (27–43)	57.9	0 (0)	cam (100)	2.4 (2–3.0)
Javed 2011	40 (40)	65 (60–82)	65.0	0 (0)	cam (100)	2.5 (1–4.5)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Nassif 2010	163 (178)	32.7 (15–56)	42.9	0 (0)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
Philippon 2008	16 (17)	15 (11–16)	12.5	0 (0)	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Osteonecrosis						
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Horisberger 2010 (clin orthop)	88 (105)	40.9 (17–66)	68.0	0 (0)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Larson 2008	96 (100)	34.7 (16–64)	56.3	0 (0)	cam (17%); pincer (28%); mixed (55%)	0.8 (0.25–3)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Heterotopic Ossification						
Byrd 2009	200 (207)	33 (NR)	69.0	1 (0.5)	cam (79.0); mixed (21.0)	1.3 (1–2)
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Gedouin, May 2010	110 (111)	31 (16–49)	71	3 (2.7)	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Larson 2008	96 (100)	34.7 (16–64)	56.3	6 (6.0)	cam (17%); pincer (28%); mixed (55%)	0.8 (0.25–3)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Nassif 2010	163 (178)	32.7 (15–56)	42.9	1 (0.6)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Infection						
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Haviv, O'Donnell 2010*	82 (164)	29.4 (14–63)	81.7	0 (0)	cam (100)	2.2 (1.0–6.7)
Ilizaliturri 2008	19 (19)	34 (27–43)	57.9	0 (0)	cam (100)	2.4 (2–3.0)
Javed 2011	40 (40)	65 (60–82)	65.0	0 (0)	cam (100)	2.5 (1–4.5)
Larson 2008	96 (100)	34.7 (16–64)	56.3	0 (0)	cam (17%); pincer (28%); mixed (55%)	0.8 (0.25–3)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Nassif 2010	163 (178)	32.7 (15–56)	42.9	0 (0)	cam (67.0); pincer (11.0); mixed (11.0)	1.9 (1–5.4)
Philippon 2008	16 (17)	15 (11–16)	12.5	0 (0)	cam (11.8); pincer (31.3); mixed (56.3)	1.4 (1–2.0)
Philippon 2009	112 (112)	40.6 (38–44)	44.6	0 (0)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
DVT/PE						
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Haviv, O'Donnell 2010*	82 (164)	29.4 (14–63)	81.7	0 (0)	cam (100)	2.2 (1.0–6.7)
Javed 2011	40 (40)	65 (60–82)	65.0	0 (0)	cam (100)	2.5 (1–4.5)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Philippon 2009	112 (112)	40.6 (38–44)	44.6	0 (0)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Neurovascular						

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Any						
Flecher 2011	23 (23)	34 (17–54)	60.9	0 (0)	NR	≥ 1
Gedouin, Duperron 2010	38 (38)	36 (24–64)	86.8	0 (0)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Haviv, O'Donnell 2010*	82 (164)	29.4 (14–63)	81.7	0 (0)	cam (100)	2.2 (1.0–6.7)
Ilizaliturri 2007	13 (14)	31 (24–39)	46.2	0 (0)	cam (100)	2.5 (2–4)
May 2007	5 (5)	40 (27–48)	40.0	0 (0)	cam (100)	1.4 (1–2)
Stahelin 2008	22 (22)	42 (18–67)	68.2	0 (0)	cam (100)	0.5 (NR)
Nerve palsy						
Gedouin, Duperron 2010	38 (38)	36 (24–64)	86.8	0 (0)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Parathesias						
Gedouin, Duperron 2010	38 (38)	36 (24–64)	86.8	0 (0)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Horisberger 2010 (clin ortho)	88 (105)	40.9 (17–66)	68.0	9 (8.6)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Javed 2011	40 (40)	65 (60–82)	65.0	1 (2.5)	cam (100)	2.5 (1–4.5)
Philippon 2009	112 (112)	40.6 (38–44)	44.6	0 (0)	cam (20.5); pincer (2.7); mixed (76.8)	2.3 (2–2.9)
Neuropraxia						
Byrd 2009	200 (207)	33 (NR)	69.0	2 (1.0)	cam (79.0); mixed (21.0)	1.3 (1–2)
Gedouin, Duperron 2010	38 (38)	36 (24–64)	86.8	0 (0)	cam (44.7); pincer (10.5); mixed (44.7)	1.3 (0.5–3.0)
Gedouin, May 2010	110 (111)	31 (16–49)	71	1 (0.9)	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Horisberger 2010 (clin ortho)	88 (105)	40.9 (17–66)	68.0	2 (1.9)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Larson 2008	96 (100)	34.7 (16–64)	56.3	1 (1.0)	cam (17%); pincer (28%); mixed (55%)	0.8 (0.25–3)
Superficial tear of labia minora						
Gedouin, May 2010	110 (111)	31 (16–49)	71	1 (0.9)	cam (36.9); pincer (11.7); mixed (51.4)	1.0 (0.5–1.5)
Horisberger 2010 (clin ortho)	88 (105)	40.9 (17–66)	68.0	1 (1.0)	cam (54.3); mixed (45.7)	2.3 (1.3–4.1)
Temporal hypesthesia						
Horisberger 2010 (athro)	20 (20)	47.3 (22–65)	80.0	2 (10.0)	cam (55.0); mixed (45.0)	3.0 (1.5–4.1)
Hypoesthesia						
Stahelin	22 (22)	42 (18–67)	68.2	6 (27.3)	cam (100)	0.5 (NR)
Abdominal Fluid???						
OPEN DISLOCATION						
Reoperation other than THA						
Beaule 2007	34 (37)	40.5 (19–54)	52.9	0 (0)	cam (100)	3.1 (2.1–5.0)
Eijer 2001	8 (8)	33.3 (12–64)	25.0	1 (12.5)	cam (100)	1.4 (0.9–2.1)
Murphy 2004	23 (23)	35.4 (17–54)	56.5	1 (4.3)	cam (44.0); pincer (4.0); mixed (52.0)	5.2 (2–12)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Siebenrock 2003	22 (29)	23 (14–41)	86.4	3 (10.3)	pincer (100)	2.5 (2–4.1)
Yun 2009	14 (15)	35.8 (22–54)	75.0	3 (20.0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Loss of fixation						
Beaule 2007	34 (37)	40.5 (19–54)	52.9	1 (2.7)	cam (100)	3.1 (2.1–5.0)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Siebenrock 2003	22 (29)	23 (14–41)	86.4	1 (3.4)	pincer (100)	2.5 (2–4.1)
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Trochanteric nonunion						
Murphy 2004	23 (23)	35.4 (17–54)	56.5	0 (0)	cam (44.0); pincer (4.0); mixed (52.0)	5.2 (2–12)
Peters 2010	94 (96)	28 (14–51)	58.5	1 (1.0)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	3 (20)	cam (60.0); mixed (40.0)	2.3 (1–10)
Head-neck fracture						
Graves 2009	46 (48)	33 (18–51)	54.3	0 (0)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
Peters 2010	94 (96)	28 (14–51)	58.5	0 (0)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Avascular necrosis						
Beck 2004	19 (19)	36 (21–52)	73.7	0 (0)	NR	4.7 (4.2–5.2)
Graves 2009	46 (48)	33 (18–51)	54.3	0 (0)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
Murphy 2004	23 (23)	35.4 (17–54)	56.5	0 (0)	cam (44.0); pincer (4.0); mixed (52.0)	5.2 (2–12)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Osteonecrosis						
Beaule 2007	34 (37)	40.5 (19–54)	52.9	0 (0)	cam (100)	3.1 (2.1–5.0)
Eijer 2001	8 (8)	33.3 (12–64)	25.0	0 (0)	cam (100)	1.4 (0.9–2.1)
Peters 2010	94 (96)	28 (14–51)	58.5	0 (0)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Heterotopic Ossification						
Beaule 2007	34 (37)	40.5 (19–54)	52.9	1 (2.7)	cam (100)	3.1 (2.1–5.0)
Graves 2009	46 (48)	33 (18–51)	54.3	9 (18.8)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Infection						

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Eijer 2001	8 (8)	33.3 (12–64)	25.0	0 (0)	cam (100)	1.4 (0.9–2.1)
Graves 2009	46 (48)	33 (18–51)	54.3	0 (0)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
DVT/PE						
Graves 2009	46 (48)	33 (18–51)	54.3	0 (0)		
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Neurovascular						
Any						
Yun 2009	14 (15)	35.8 (22–54)	75.0	0 (0)	cam (60.0); mixed (40.0)	2.3 (1–10)
Nerve palsy						
Graves 2009	46 (48)	33 (18–51)	54.3	0 (0)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
Peters 2010	94 (96)	28 (14–51)	58.5	0 (0)	cam (34.4); pincer (6.3); mixed (59.4)	2.2 (1.5–8)
Pierannunzii 2007	8 (8)	30 (26–39)	75.0	0 (0)	cam (12.5); pincer (50.0); mixed (37.5)	NR
Symptomatic hardware						
Beaule 2007	34 (37)	40.5 (19–54)	52.9	9 (24.3)	cam (100)	3.1 (2.1–5.0)
Graves 2009	46 (48)	33 (18–51)	54.3	2 (4.3)	cam (75.0); mixed (25.0)	1.6 (0.5–2.8)
MINI-OPEN						
Reoperation other than THA						
Clohisy 2010	41 (41)	34 (16–48)	68.3	0 (0)	cam (100)	2.2 (2–3)
Laude 2009	97 (100)	33.4 (16–56)	51.5	15 (15.0)	NR	4.9 (2.4–8.7)
Lincoln 2009	14 (16)	37 (17–51)	71.4	1 (6.3)	cam (63.0); mixed (37.0)	2.0 (1.3–3)
Ribas 2007	32 (35)	36.2 (23–48)	71.9	1 (2.9)	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
Trochanteric nonunion						
Hartmann 2009	33 (34)	31 (15–47)	51.5	0 (0)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Laude 2009	97 (100)	33.4 (16–56)	51.5	0 (0)	NR	4.9 (2.4–8.7)
Head-neck fracture						
Clohisy 2010	41 (41)	34 (16–48)	68.3	0 (0)	cam (100)	2.2 (2–3)
Hartmann 2009	33 (34)	31 (15–47)	51.5	0 (0)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Laude 2009	97 (100)	33.4 (16–56)	51.5	1 (1.0)	NR	4.9 (2.4–8.7)
Avascular necrosis						
Laude 2009	97 (100)	33.4 (16–56)	51.5	0 (0)	NR	4.9 (2.4–8.7)
Osteonecrosis						

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Clohisy 2010	41 (41)	34 (16–48)	68.3	0 (0)	cam (100)	2.2 (2–3)
Hartmann 2009	33 (34)	31 (15–47)	51.5	0 (0)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Ribas 2007	32 (35)	36.2 (23–48)	71.9	0 (0)	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
Heterotopic Ossification						
Clohisy 2010	41 (41)	34 (16–48)	68.3	4 (11.4)	cam (100)	2.2 (2–3)
Hartmann 2009	33 (34)	31 (15–47)	51.5	6 (18.2)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Laude 2009	97 (100)	33.4 (16–56)	51.5	1 (1.0)	NR	4.9 (2.4–8.7)
Ribas 2007	32 (35)	36.2 (23–48)	71.9	0 (0)	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
Infection						
Clohisy 2010	41 (41)	34 (16–48)	68.3	1 (2.9)	cam (100)	2.2 (2–3)
Laude 2009	97 (100)	33.4 (16–56)	51.5	2 (2.0)	NR	4.9 (2.4–8.7)
DVT/PE						
Clohisy 2010	41 (41)	34 (16–48)	68.3	1 (2.9)	cam (100)	2.2 (2–3)
Neurovascular						
Any						
Clohisy 2010	41 (41)	34 (16–48)	68.3	0 (0)	cam (100)	2.2 (2–3)
Ribas 2007	32 (35)	36.2 (23–48)	71.9	6 (17)	cam (51.4); pincer (22.9); mixed (52.7)	2.4 (1.5–3.3)
Nerve palsy						
Hartmann 2009	33 (34)	31 (15–47)	51.5	2 (6.1)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Parathesias						
Hartmann 2009	33 (34)	31 (15–47)	51.5	17 (51.5)	cam (20.6); mixed (79.4)	1.3 (0.5–2.3)
Neuropraxia						
Lincoln 2009	14 (16)	37 (17–51)	71.4	7 (43.8)	cam (63.0); mixed (37.0)	2.0 (1.3–3)
CONSERVATIVE						
Reoperation other than THA						
Emara 2011	37 (37)	33 (23–47)	73.0	4 (10.8)	NR	2.2 (2.1–2.3)

DVT/PE: deep vein thrombosis/pulmonary embolism; FAI: femoracetabular impingement; THA: total hip arthroplasty

*This study evaluated arthroscopy treatment for bilateral FAI and divided patients into two groups based upon the timing of their second FAI surgery. For the purpose of this report, this study was treated as a case-series and the preoperative and change scores reflect the average of the corresponding score following the second surgery.

Table H2. Complications in studies reporting on treatment for FAI in competitive athletes.

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
ARTHROSCOPY						
Reoperation other than THA						
Philippon 2007	45 (45)	31 (17–61)	93.3	5 (11.1)	cam (48.9); pincer (6.7); mixed (46.7)	1.6 (0.5–5.5)
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Trochanteric nonunion						
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Head-neck fracture						
Philippon 2007	45 (45)	31 (17–61)	93.3	0 (0)	cam (48.9); pincer (6.7); mixed (46.7)	1.6 (0.5–5.5)
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Avascular necrosis						
Philippon 2007	45 (45)	31 (17–61)	93.3	0 (0)	cam (48.9); pincer (6.7); mixed (46.7)	1.6 (0.5–5.5)
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Osteonecrosis						
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Heterotopic Ossification						
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Infection						
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
DVT/PE						
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
Neurovascular						
Any						
Philippon 2007	45 (45)	31 (17–61)	93.3	0 (0)	cam (48.9); pincer (6.7); mixed (46.7)	1.6 (0.5–5.5)
Saw 2004	6 (6)	28.2 (24–32)	100	0 (0)	NR	NR
OPEN DISLOCATION						
Reoperation other than THA						
Naal 2010	22 (30)	19.7 (16–25)	100	1 (3.3)	cam or mixed (100)	3.8 (1–6.6)
Screw removal						
Naal 2010	22 (30)	19.7 (16–25)	100	6 (20.0)	cam or mixed (100)	3.8 (1–6.6)

DVT/PE: deep vein thrombosis/pulmonary embolism; FAI: femoracetabular impingement; NR: not reported; THA: total hip arthroplasty.

Table H3. Complications in cohorts reporting on efficacy of FAI

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Labral Debridment vs. Labral Refixation						
Reoperation other than THA						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Larson 2009*	Group 1: 34 (36) Group 2: 37 (39)	Group 1: 31 (16–57) Group 2: 27 (16–56)	Group 1: 74 Group 2: 62	Group 1: 4 (11.1)† Group 2: 1 (2.6)	roup 1: pincer (16.7); mixed (83.3) roup 2: pincer (15.4); mixed (84.6)	Group 1: 1.8 (1.0–3.0) Group 2: 1.4 (1.0–2.0)
Loss of fixation						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Trochanteric nonunion						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Head-neck fracture						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Avascular necrosis						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Osteonecrosis						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Heterotopic Ossification						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Larson 2009*	Group 1: 34 (36) Group 2: 37 (39)	Group 1: 31 (16–57) Group 2: 27 (16–56)	Group 1: 74 Group 2: 62	Group 1: 3 (8.3) Group 2: 0 (0)	roup 1: pincer (16.7); mixed (83.3) roup 2: pincer (15.4); mixed (84.6)	Group 1: 1.8 (1.0–3.0) Group 2: 1.4 (1.0–2.0)
Infection						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
DVT/PE						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
Neurovascular						
Any						
Espinosa 2006*	Group 1: 20 (25) Group 2: 32 (35)	Total: 30 (20–40)	Total: 63	Group 1: 0 (0) Group 2: 0 (0)	NR	2.0 (NR)
No Osteoplasty vs. Osteoplasty						
Reoperation other than THA						

Outcome	No. Patients (no. hips)	Mean age in years (range)	% Male	No. Hips (% hips) with complication	FAI diagnosis (%)	Mean Follow-up in years (range)
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Trochanteric nonunion						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Head-neck fracture						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Avascular necrosis						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Osteonecrosis						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Heterotopic Ossification						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Infection						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
DVT/PE						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Neurovascular						
Any						
Bardakos 2008‡	Group 1: 47 (47) Group 2: 24 (24)	Group 1: 35 (27–46) Group 2: 33 (27–41)	Group 1: 49 Group 2: 58	Group 1: 0 (0) Group 2: 0 (0)	Cam (100)	Overall: >1
Non-Operative Treatment vs. Arthroscopy or Open Dislocation vs. THA						
Reoperation other than THA						
Jager 2004§	Group 1: 9 (10) Group 2: 6 (8) Group 3: 2 (4)	Group 1: 34.5 (NR) Group 2: 27.3 (NR) Group 3: 49.5 (NR)	Total: 76.5	Group 1: 0 (0) Group 2: 1 (16.7) Group 3: 0 (0)	Cam (100)	Group 1: 1.4 (NR) Group 2: 1.8 (NR) Group 3: 2.2 (NR)
Hematoma						
Jager 2004§	Group 1: 9 (10) Group 2: 6 (8) Group 3: 2 (4)	Group 1: 34.5 (NR) Group 2: 27.3 (NR) Group 3: 49.5 (NR)	Total: 76.5	Group 1: 0 (0) Group 2: 1 (16.7) Group 3: 0 (0)	Cam (100)	Group 1: 1.4 (NR) Group 2: 1.8 (NR) Group 3: 2.2 (NR)
Arthroscopic Partial Labral Resection and Chondroplasty vs. Same Plus Mini-Open						
Reoperation other than THA						
Nepple 2009**	Group 1: 23 (23) Group 2: 25 (25)	Group 1: 37 (NR) Group 2: 33 (NR)	Group 1: 52 Group 2: 68	Group 1: 3 (13.0) Group 2: 0 (0)	Cam (100)	Group 1: 2.3 (1.0–4.0) Group 2: 1.7 (1.0–4.0)

DVT/PE: deep vein thrombosis/pulmonary embolism; FAI: femoracetabular impingement; NR: not reported; THR: total hip replacement.

*Group 1 = Labral debridement (historical controls); Group 2 = Labral refixation.
†Includes 2 of the patients with heterotopic ossification; both revisions.
‡Group 1 = No osteoplasty (historical controls); Group 2 = Osteoplasty.
§ Group 1 = Nonoperative treatment (physiotherapy and anti-inflammatory cyclooxygenase-2 (COX-2) inhibitor drugs); Group 2 = open dislocation or arthroscopy; Group 3 = total hip replacement.
**Group 1 = Arthroscopic partial labral resection and chondroplasty (historical controls); Group 2 = same procedure as Group 1 but modified to address the anterolateral femoral cam impingement lesion with a limited open osteochondroplasty.

Table H4. Summary of overall complication rate in cohorts reporting on efficacy of various treatments for FAI.

	Labral Debridement vs. Labral Refixation				No Osteoplasty vs. Osteoplasty				Conservative treatment vs. Arthroscopy or Open Dislocation vs. THA				Arthroscopic Partial Labral Resection and Chondroplasty vs. Same Plus Mini-Open		
Complication	**No. Studies**	**No. Hips**	**%**		**No. Studies**	**No. Hips**	**%**		**No. Studies**	**No. Hips**	**%**		**No. Studies**	**No. Hips**	**%**
Reoperation*	2	Group 1: 61 Group 2: 74	Group 1: 6.6 Group 2: 1.4		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		1	Group 1: 10 Group 2: 8 Group 3: 4	Group 1: 0 Group 2: 16.7 Group 3: 0		1	Group 1: 23 Group 2: 25	Group 1: 13.0 Group 2: 0
Trochanteric nonunion	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
Head-neck fracture	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
AVN	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
ON	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
HO	2	Group 1: 61 Group 2: 74	Group 1: 4.9 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
Infection	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
DVT/PE	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
Neurological†	1	Group 1: 25 Group 2: 35	Group 1: 0 Group 2: 0		1	Group 1: 47 Group 2: 24	Group 1: 0 Group 2: 0		- - -	- - -	- - -		- - -	- - -	- - -
Other‡	- - -	- - -	- - -		- - -	- - -	- - -		1	Group 1: 10 Group 2: 8 Group 3: 4	Group 1: 0 Group 2: 16.7 Group 3: 0		- - -	- - -	- - -

AVN: avascular necrosis; DVT/PE: deep vein thrombosis/pulmonary embolism; HO: heterotopic ossification; ON: osteonecrosis.
*Excluding THA.

†Including nerve plasy, parathesias, and neuropraxia, and other.
‡Including superficial tear of the labia minora, temporal hypesthesia, and symptomatic hardware.

APPENDIX I. CLINICAL PEER REVIEWERS

Reviewer	Areas of expertise
Paul A. Manner, MD, FRCSC Assistant Professor University of Washington School of Medicine Department of Orthopaedics	• Orthopedic surgeon • Assistant Professor, University of Washington School of Medicine, Department of Orthopaedics and Sports • Adult reconstruction and arthroplasty
John R. Green, III, M.D. Associate Professor, Chief, Sports Medicine University of Washington Sports Medicine Clinic Department of Orthopaedics & Sports Medicine	• Orthopedic surgeon • Associate Professor, University of Washington School of Medicine, Department of Orthopaedics • Chief, Sports Medicine Clinic • Arthroscopic surgery of the shoulder, elbow, hip, knee and ankle