# Osteoarthritis
## Getting the picture

- Willem Paul Gielis -

# Osteoarthritis
## Getting the picture

**Willem Paul Gielis**

**Osteoarthritis** Getting the picture

PhD thesis, Utrecht University, The Netherlands



| | |
|---|---|
| Cover | Willem Paul Gielis & Roy Sanders. Based on '*Le Petit Prins*' by Antoine de Saint-Exupéry |
| Lay-out | Roy Sanders |
| Printed by | Gildeprint Drukkerijen, Enschede |
| ISBN | 978-94-6419-428-9 |

# Osteoarthritis
## Getting the picture

## Artrose, het plaatje snappen

(met een samenvatting in het Nederlands)

**Proefschrift**

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 27 januari 2022 des middags te 4.15 uur

door

**Willem Paul Gielis**
geboren op 26 november 1990
te Breda

## Promotoren:

Prof. dr. ir. H. Weinans
Prof. dr. P.A. de Jong

## Copromotoren:

Dr. W. Foppen
Dr. R.J.H. Custers

## Beoordelingscommissie:

Prof. dr. R.M. Castelein
Prof. dr. ir. M.A. Viergever
Prof. dr. F.P.J.G. Lafeber
Prof. dr. M. Kloppenburg
Prof. dr. F.W. Roemer

This work was supported by Reuma Nederland (LLP-22) and the APPROACH project. APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking, composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' kind contribution.

*Ik draag dit werk op aan Julia,*
*mijn kleine prinses*

## Table of contents

# CHAPTER 1

## General introduction and thesis outline

## A giant burden

The musculoskeletal system is the organ system that gives humans the ability to move. Bones are the frame and muscles are the motors. To allow for motion between bones, joints are essential. They cover boney ends and allow for low friction movement under heavy loads, whilst warranting stability. Worldwide, musculoskeletal diseases are the second cause of years lived in disability after mental and substance use disorders.[1] These numbers are rapidly increasing due to aging and an uprise of obesity.

In musculoskeletal diseases, osteoarthritis (OA) is the second biggest culprit, after neck and lower back pain. OA causes pain, stiffness and loss of function in the joints of patients and impairs their quality of life. A disease modifying drug for OA does not exist. Non-surgical therapy strategies are limited to patient education, exercise therapy, pain medication and bracing. Additionally, a range of intra-articular injections is available. Unfortunately, with limited or only temporal effects. For a selected group of patients, joint preserving surgery is an option. However, most patients with end-stage OA are condemned to surgical joint replacement. In 2018, 77,521 joints were surgically replaced in the Netherlands, mostly due to OA.[2] As younger and younger patients are implanted with artificial joints, the need for revision surgery, with a much higher complication and lower success rate, will be a future problem.[3]

## The definition of OA

### The OARSI, the main research foundation in OA, defines the disease as follows:

"*OA is a disorder involving movable joints characterized by cell stress and extracellular matrix degradation initiated by micro- and macro-injury that activates maladaptive repair responses including pro-inflammatory pathways of innate immunity. The disease manifests first as a molecular derangement (abnormal joint tissue metabolism) followed by anatomic, and/or physiologic derangements (characterized by cartilage degradation, bone remodeling, osteophyte formation, joint inflammation and loss of normal joint function), that can culminate in illness.*"

This definition shows that OA is no longer viewed as a simple wear-and-tear disease of the cartilage. Furthermore, it shows that it is difficult to grasp all facets of the disease in an easy to comprehend definition. The joint is formed by more tissues than cartilage alone and a homeostasis between these tissues is essential to

prevent deterioration. Joint homeostasis is acquired and preserved by mechanical and chemical interaction between the tissues of the joint and the body.[4] We will elaborate on the main tissues in the joint.

### Tissues of the joint

***Cartilage*** is the smooth and resilient tissue, which provides low friction movement between the bones. It comprises a network of type II collagen arches and proteoglycans that attract water. The attraction of water is the main driver of compressive strength of the cartilage. Within this avascular and aneural network, the cartilage houses a small number of cells "trapped" in lacuna. Due to these properties, the regenerative capacity of cartilage is limited.

The ***synovium*** forms the joint capsule and controls synovial fluid volume and composition. Additionally, it nourishes the chondrocytes and expresses both pro and anti- inflammatory cytokines via the synovial fluid.

***Ligaments*** made of collagenous fibers bridging the joint and connecting the bones. The connections are mobile, but only allow for movement in a defined range of motion. Together with the muscles and tendons, ligaments stabilize the joint and protect the cartilage from exorbitant loads. Additionally, some joints contain a labrum or menisci, made from fibrous cartilaginous tissue, that provide further structural integrity to the joint.

An often-neglected tissue is the ***subchondral bone***, a rigid tissue that supports the cartilage both mechanically and chemically. It is highly vascularized and houses nerves. However, the osteochondral junction, the transitional zone between soft cartilage and rigid subchondral bone, seals off the cartilage from nerves and vessels. This makes joint loading a painless process. The vertical part of the arches of collagen type II pass through it, anchoring the cartilage. The superficial cartilage is mainly nourished by the synovium and synovial fluid. The deeper parts however, are nourished by the subchondral bone, as cytokines transfer through the junction.

### Pathologic mechanisms in OA

***Initial damage*** to any tissue of the joint, due to excessive or erroneous loading, hypoxia or inflammation (albeit infectious) starts an active response in the joint. A joint in healthy hemostasis may have sufficient repair capacity to repair the damage. When hemostasis is frail, repair reactions may overshoot, damaging the joint further than the initial damage. When this process is repeated and inflammation, necrosis, and structural deterioration accumulate (as seen in **Figure 1**), OA

becomes a fact. OA is a heterogeneous cascade of pathophysiological mechanisms and in every patient some of these mechanisms are more important than others. Recently, theories on phenotypes in OA emerged in literature. These phenotypes try to describe clusters of pathophysiologic mechanisms in OA patients. However, these phenotypes are not yet clear-cut and there will be gray areas between them. Below, the main pathophysiologic mechanisms in OA are outlined.

***Healthy Loading*** of the joint is essential for keeping the joint healthy. By loading and unloading, synovial fluid is being circulated, nutrients are being pumped into the cartilage and degradation products out.[5,6] The joint tissues can be trained to withstand high loading. A striking example was published by Schütz et al..[7] Ultra-long-distance runners competing in the TransEurope FootRace were followed by a mobile MRI in their 64-day race over 4,486-kilometer. In the first 1,100 kilometers, an increase in T2-mapping values was found, corresponding with cartilage degradation and/or softening. After a stable phase between 1,100 and 2,800 kilometers, a decrease of T2 values was found at the 3,700 km mark. Seven months after the endurance race, T2 values decreased even further returning to their original ranges. This study clearly shows the adaptive capabilities of the cartilage in highly trained individuals and that cartilage does not simply wear out.

***Trauma*** is a short moment of overload, most often during sports, that may cause damage to the joint tissues and initiate osteoarthritic cascades. Meniscal tears, cruciate ligament tears and focal cartilage defects are frequent and well-known examples. These injuries will give an eight times higher risk for OA on average within 19 years.[8]

***Chronic overload*** poses another threat to the joint. For example, certain occupations may be strenuous on specific joints. Mining involves a fair amount of squatting and lifting weight while squatted. These movements increase the risk for knee OA to such a degree that the term Miners knee was proposed in literature.[9] Of course, many other causes for chronic overloading exists.

***Morphology*** is a result of genetics and loading. Already at a young age, excess loading can alter the joints morphology. In a cohort study on young pre-professional soccer players, a third of the participants between 12 and 13 years old developed an irregular osseous prominence at the antero-superior head-neck junction of the hip joint, named Cam-lesion.[10] In later life, these Cam-lesions may lead to excessive forces on the labrum, acetabular rim and cartilage This makes hips with a Cam-lesion highly susceptible to OA.[11,12] Other morphological variations that predispose for OA have been explored using Statistical Shape Modelling.

A technique that can accurately quantify shape variations in 2D radiographs or 3D scans. For the hip, a narrower and longer neck and a shallower (dysplastic) acetabulum are other examples of morphological variations that predispose for OA.[13] OA knees are characterized by widening and flattening of the femur and tibia and narrowing of the notch.[14] These characteristics seem to be more of a result of the disease as opposed to risk factors. Shape variations predisposing for knee OA have been described less extensively.

***Joint malalignment*** causes an unequal distribution of loads within the joint and is another risk factor for OA. The relationship between varus/valgus alignment of the lower extremity and knee OA is the best studied example.[15–18] There is a natural variance of the leg alignment in the general population. Most healthy knees have an alignment between 3 degrees of varus and 1 degree of valgus, with an average of two degrees of varus.[19,20] In that average situation, 65 to 80 percent of the load during gait is transferred through the medial compartment.[19] Malalignment of the lower limb causes the cartilage, meniscus and subchondral bone in the medial (varus) or lateral (valgus) compartment of the knee to endure higher stresses.[21–23] Consequently, malalignment is strongly related to symptomatic unilateral knee OA and radiographic progression [24]. As cartilage and meniscal tissue in the stressed compartment degrade, the angle of malalignment further increases, amplifying pathological loads. This vicious circle may be ended by an orthopedic surgeon in three ways which are discussed later in the general introduction.

***Obesity*** is an important risk factor for OA. Cartilage surface and volume are strongly correlated to the height, but not the weight of a person.[25,26] Furthermore, cartilage thickness is not correlated to any anthropomorphic metrics.[26] The weight of a child does not affect the amount of cartilage formed in childhood, however activity does.[27] On the other hand, each kilo of body weight strains the cartilage with a four unit increased load during walking.[28] Obesity is the main modifiable risk factor for knee OA, however the relation with hip OA is less clear. Furthermore, obesity *is* a risk factor for OA in non-weight-bearing joints, such as the hands. This is caused by the dyslipidemia and general inflammatory state of adipose tissue in obesity which will be discussed later.[29]

***Cytokines*** are immunomodulating proteins and are key in regulating inflammatory processes in the joint. Broadly speaking they can be divided into catabolic or proinflammatory (like Interleukine-1 and TNF-α) and anabolic or anti-inflammatory cytokines (like Insulin-Like Growth Factor 1). The balance between catabolic and anabolic cytokines will change as a reaction to damage of the intra-

articular tissues. Elaborating each cytokine is outside the scope of the present thesis. However, it is important to mention the role that two important risk factors for OA, namely aging and obesity, have on the inflammatory balance. Aging itself does not cause OA, however it makes the joint more susceptible for homeostatic imbalances. Of course, the cumulative chance for damage grows with time, but this does not explain exponential growth in prevalence starting from the age of 50. The shear modulus of cartilage decreases with age, making it more susceptible to harmful loading.[30] Oxidative stress, compromised autophagy of chondrocytes, a buildup of senescent chondrocytes, a decrease of Sirtuins, mitochondrial DNA damage and accumulating AGEs are all factors that potentially disrupt the balance between catabolic and anabolic cytokines. The homeostatic buffer is smaller in older individuals, the risk for a catabolic synovial environment and irreversible structural damage is higher. Obese individuals generally have more subcutaneous and visceral fat tissue. This tissue contains adipocytes and activated macrophages, which release inflammatory cytokines and adipokines systemically, causing a low-grade systemic inflammation, contributing to an increased inflammatory response in joints.[31] Additionally, high systemic levels of triglycerides and ox-LDL, and perturbations in the HDL metabolic pathways found in obesity may disrupt the joint homeostasis.[29] Fat tissue can also be found inside joints. The infrapatellar (Hoffa's) fat pad is the most prominent example. Hoffa's fat pad has been researched extensively as a cause for anterior knee pain and degenerative joint disease. However, the role of articular fad pads in normal joint homeostasis remains to be elucidated. Fat pads are present in multiple joints, including in the metacarpal and hip joints.[32,33] Fat pads may excrete multiple growth-factors, cytokines and adipokines into the joint. The levels of adipokines in synovial fluid are correlated to body mass index (BMI) as well as knee OA severity.[34] Interestingly, the structure of Hoffa's fat pad does not fluctuate with caloric intake, even in extreme cases like obesity and cachexia.[35] However, in end-stage OA the functionality of the adipocytes is decreased, and fibrosis and macrophage infiltration in Hoffa are increased.[36] As Hoffa's fat pad, and other articular fat pads, contain abundant sensory nerves and can produce inflammatory proteins, they may be a key communicator between the biomechanical and inflammatory homeostasis.

Integrity of the ***subchondral bone*** Due to animal studies and advanced MRI, we now have understanding of the role of periarticular bone in early OA. Excessive forces on the subchondral bone cause microfractures. In a joint with signs of early OA these microfractures trigger an inflammatory response and neovascularization. The balance between osteoblasts and osteoclasts is disturbed and hypo-mineralized,

porous, vascularized and nerve rich fibrous tissue is formed.[37–39] As nerves and vessels breach the osteochondral junction, pain stimuli and inflammatory cytokines can be transferred from cartilage to bone.[40]

During ***menopause*** hormone levels of women drastically change. Especially estrogen levels decrease strongly. Sex differences in the prevalence of OA are site specific.[41] The risk for knee OA sky-rockets after menopause. A decrease in estrogen may cause pro-inflammatory cytokine expression. This effect may be abated by hormone supplementation.[42] Furthermore, a decrease in estrogen impairs the bone mineralization units and shifts the bone remodeling balance towards resorption.[43] This process may have harmful effects the subchondral bone integrity.

***Angiogenesis*** is an important factor in OA. Harrison and colleagues described already in 1953 that OA may be primarily caused by the reaction of juxta-chondral blood vessel on cartilage damage.[44] The growth of osteophytes is preceded by vascular ingrowth in the joint. Hypervascularity of the joint weakens the structure of the bone and lowers the load capacity.[45] This enlarges the chance on further damage. Due to increased vascularization, catabolic cytokines reach the joint more easily, propelling the detrimental attempt of repair. The condition of the supplying vessels is also important for the joint. Atherosclerosis has a strong connection to OA.[46] The supplying vessel needs a minimal throughput to nourish the joint and prevent hypoxia and malnourishment. Additionally, cytokines released by the vascular wall of supplying vessels may impact joint. In atherosclerosis, VEGF and InterLeukin-6 are examples of upregulated cytokines.[47,48] Both cytokines are associated with OA. If they are upregulated in the supplying vessels they will pass through the joint and possibly disrupt the pro-, anti-inflammatory balance.[49,50] It is difficult to study the relationship between vascular disease and OA, as risk factors strongly overlap. Furthermore, immobility and inflammation caused by OA may have an aggravating effect on vascular damage. Therefore, it will be hard to define causal pathways.

***Heredity*** has a strong influence on the risk for OA, with an estimated heritability of different types of OA ranging between 45 and 75%.[51] Until now, 30 genomes have been identified.[52,53] These genomes give insights into possible molecular pathways in OA, and may also be used in stratifying OA patients.

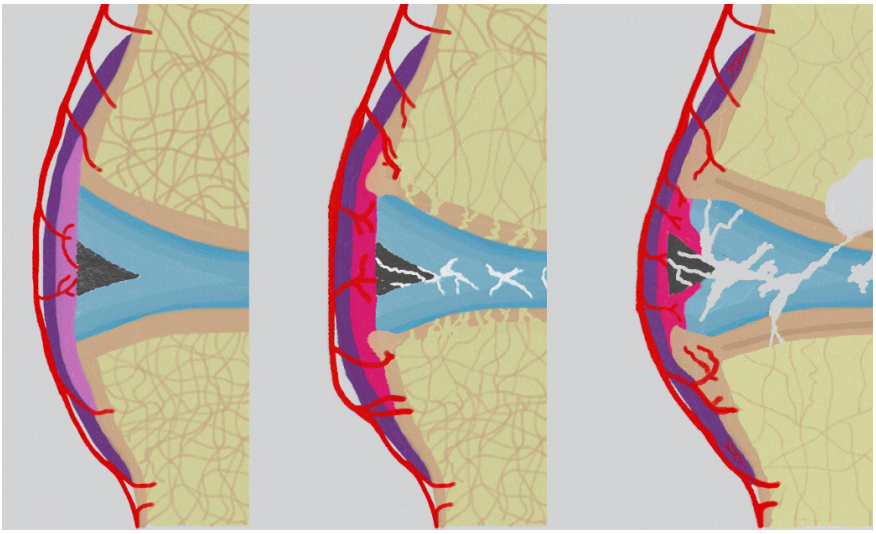


**Figure 1.** A schematic illustration of structural changes in the different tissues involved in OA. From left to right, a healthy, early OA and end-stage OA joint are depicted. In early OA, initial cartilage (blue) and meniscal damage (gray) induces angiogenesis (red) and synovitis (pink). Inflammatory cytokines disrupt the joint homeostasis. Microcracks in the subchondral bone (yellow) cause the bone to remodel, resulting in a thinner and more porous subchondral plate. As a reaction to angiogenesis and erroneous loading, osteophytes grow. In end-stage OA, the cartilage is damaged further by a combination of erroneous loading and catabolic cytokines. Regions of bare subchondral bone result and intra-osseous cysts (white) develop. The meniscus is extruded from the joint, osteophytes further develop along with tendinitis (purple). The joint widens and range of motion is limited. As physical activity declines, the bones become osteopenic.

## Phenotyping in OA

Phenotypes may be used to stratify patients based on both pathological pathways and distribution of OA throughout the body. The idea of phenotypes in OA is becoming more and more accepted, but definitions of phenotypes are still in their infancy. An example of an OA phenotype is the active subchondral bone phenotype. A proposed candidate biomarker for active structural degradation of the subchondral bone is the bone marrow lesion (BML). BMLs can be visualized on MRI scans as hyperintense lesions in the subchondral bone on T2-weighted sequences. Only little is known on what BMLs are exactly. They are associated to the presence of pain in OA and are highly predictive of future cartilage loss and the need for surgical joint replacement.[54,55] BMLs show angiogenesis and high bone turn-over in histology.[56] Therefore, BMLs may be indicative for a phenotype of OA in which structural degradation of the subchondral bone is a driving pillar.

Laslett et al. used this phenotype in a RCT testing the effect of bisphosphonates on knee OA.[57] A single infusion of 5mg of zoledronic acid reduced pain and BML size compared to placebo at 6-month follow-up.

Phenotypes as suggested by Karsdal et al in 2014 divide OA patients based on the main pathologic driver: cartilage; bone and synovitis.[58] However, they pose the problem that phenotypes may rather stratify disease stages as opposed to difference between pathological pathways. Herrero et al. suggest in their 2016 review, that four clinical phenotypes of OA exist: biomechanical; osteoporotic (estrogen dependent); metabolic and inflammatory.[59] Dell'Isola et al. wrote a systematic review on evidence for phenotypes in 2016 and suggests six phenotypes, namely chronic pain; inflammatory; metabolic syndrome; bone and cartilage metabolism; mechanical overload and minimal joint disease.[60] In a subsequent study, their group managed to phenotype 84% of 599 knee OA patients of the OsteoArthritisInitiative Database. 20% of the cases fitted into multiple phenotypes and were labelled "complex" OA. This shows the feasibility of phenotyping in knee OA in a longitudinal database.[61] Mobasheri et al. suggest that phenotyping alone might not be enough and that aim should be to find endotypes for OA.[62] Endotypes differ to phenotypes, as an endotype is also a subtype of a disease or condition, but is defined by a distinct functional or pathobiological mechanism. A phenotype that starts as biomechanical driven, may transfer into a more inflammatory phenotype in later stages of the disease, particularly in obese individuals. Such complete disease pathways may define OA endotypes. Patients within certain endotypes may present themselves at different disease stages in clinical care, and phenotypic characteristics may overlap at that point. Therefore, it will be a major challenge to define endotypes in OA. The first step will be studying the natural course of different phenotypes in OA and develop selection and follow-up criteria for OA phenotypes. It will be important to position the pathophysiological mechanisms within a specific joint in relation to age, sex, and systemic conditions or comorbidities, but also to OA in other joints.

## Diagnosis

In clinical care, knee and hip OA are mainly diagnosed through anamnesis, physical examination and standard radiography. An age of 45+ years, activity related pain and morning stiffness are the most important hints.[63] A decline in range of motion, crepitation, tenderness of the joint line and broadening of the joint can be assessed during physical examination. Imaging plays a key role in the diagnosis, research and treatment of OA. On standard radiographic exams the classic signs of OA are

joint space narrowing, osteophytes, subchondral sclerosis and in later stage joint deformation. Standard radiography is the most widely used modality for OA. It is a quick and cheap technology, easy to acquire and easy to interpret. Additional imaging in the form of MRI scans, CT scans or ultrasound are usually not routinely used in clinical care to evaluate OA, but may help in diagnosing ligamentous or meniscal lesions.

In research, the radiographic criteria described by Kellgren and Lawrence are the most popular with a cut-off point of ≥2 resembling OA.[64] The drawback of this and similar criteria is that they are very well at defining end-stage disease, where different pathological pathways eventually come together. However, in early disease, where patients may still be phenotypable, these systems lack in sensitivity and specificity.[65,66] These radiographic scores have a limited reproducibility, are insensitive to changes, and have a moderate correlation to arthroscopically assessed cartilage damaged.[67] Furthermore, the correlation between clinical and radiographic criteria is not always consistent.[68] Criteria that combine clinical criteria with radiographic features, such as the ACR criteria, suffer from similar problems as their radiography-only counterparts.[69] MRI based scores such as the semi-quantitative MOAKS, quantitative cartilage volume and T2-mapping are promising techniques that can adequately measure degradation of the joint, but cut-offs for disease vs healthy/normal do not exist. OA is often a polyarticular disease and OA criteria focus only on one joint or group of joints. Altman et al. proposed a grading system for hand, knee and hip OA, using similar criteria based on standard radiography.[70] However, many joints are not included in their criteria. Information on the burden of OA in different joints within a patient, using a standardized and uniform method, can be very valuable in the process of phenotyping OA patients. Especially, when systemic biochemical markers are used, as the levels of these will be influenced by the OA process in all joints.

## Treatment

The treatment of OA should have a holistic approach.[63,71] OA has a different impact on each individual. The disease may affect a range of different joints, follow different pathological pathways, and vary between stable and fast progressive phases. On the other hand, each patient has different demands of their joints. For example. a 28-year-old, elite athlete with post-traumatic OA symptoms will have different requirements of his knee compared to an 88-year-old pensioned epicurean.

Evidence-based non-surgical therapies in OA are self-management and education,

weight loss, supervised exercise therapy, injury prevention, walking aids, pain medication, oral, topical and intra-articular anti-inflammatory drugs.[72] All these therapies have proven to lower symptoms and increase physical functioning. However, these are mainly based on a one-size-fits-all principle, and do not tackle specific pathogenic mechanism in the OA cascade. Subsequently, structural degradation of the joint is not prevented. Bracing may slow structural progression by unloading the diseased part of a joint, for instance the medial compartment in varus knee OA. [73] However, functional braces are often uncomfortable and irritate the skin, causing non-compliance in patients.[74] No true disease modifying drug for OA exists today.

Surgical treatment may alter the structural progression of OA. In the previous century, many arthroscopic surgeries were performed in OA knees. Degenerative meniscal tears were excided in addition to debridement and lavage. An RCT in 2002 was a big game changer, showing no beneficiary effect of these interventions as opposed to sham surgery.[75] Later studies confirmed that a partial meniscectomy in OA patients shows no benefit over physical therapy.[76,77] Prevention of knee OA in patients with cartilage defects may be reached by surgical therapies as micro fracture or autologous chondrocyte implantation.[78] In patients with instability of the knee due to an ACL tear, ACL reconstruction may reduce the risk for OA, but will not fully eliminate the increased risk. [79] In patients with femoral acetabular impingement (FAI) due to a Cam or pincer lesion, it is thought arthroscopic resection of the bone may postpone the onset of OA.[80] However, trails to back this theory are still lacking. A specific subpopulation with OA symptoms may benefit from correctional osteotomies. These are performed for malaligned hips, ankles, feet and knees and are one of the few options to intervene with the structural degradation in a partly OA joint. The osteotomy aims to partially unload the diseased part of the joint, and transfer more load through the healthy part of the joint. High tibial valgus osteotomies for unicompartmental medial knee OA was the most popular osteotomy for OA in recent years.[81] This procedure is mainly useful for active young non-obese patients with early OA of the medial knee, a healthy lateral compartment and varus alignment of the lower extremity. The alignment is measured on a Whole Leg Radiograph (WLR) and the surgical procedure is planned to correct to knee to a certain amount of valgus alignment. The amount of correction should be planned according to the type of patient. Early OA patients with a high demand sports activity may benefit from a smaller correction as opposed to more progressed OA patients with sports activities being limited to walking and cycling. In nine out of ten patients a surgical replacement

of the joint can be prevented for over ten years.[82] However, the post-operative recovery is challenging with a minimal period of 6 weeks of partial and an average return to work at 17 weeks.[83]

Most surgical procedures in OA joints involve replacing (a part of) the joint. While joint replacements have a very good effect on patient reported outcomes, they can also be seen as an internal amputation.[2] While low in prevalence, complications can have devastating consequences including 2-stage revisions, in which the patient will live with a temporal joint made from cement for 6 weeks. Furthermore, the joint prosthesis will wear over time and the only solution to fix wear is revision surgery. Revision surgery is accompanied with higher complication rates, higher costs and longer recovery periods compared to initial joint replacement surgery.

## Challenges

Due to a great socioeconomic impact and a lack of disease modifying drugs for OA, there is a major unmet need. Pharmaceutical companies gave up on OA due to many costly development lines that resulted in failed trials. Four underlying problems hamper efficient treatment trials in OA.

*First*, OA progression rates are very heterogeneous. Symptoms and structural radiographic parameters slowly progress over years. Some patients will experience sudden rapid progression, but it is very hard to predict which patients progress and during what period. Biomarkers to select candidates that will show diseases progression within a manageable trail duration of several years are lacking.[84] When only a small portion of trail candidates will progress without treatment, it is hard to show a treatment effect. Furthermore, accurate progression of OA progression may help to indicate "early OA" patients, without evident structural damage. In general, treatments are more effective in early disease.

*Second*, clinical trials rely on insensitive outcome measures. The journey from the first OA symptoms to a destructed joint might take decades. For example, the FDA does not accept surrogate biomarkers for structural progression of OA outside of radiographic joint space narrowing. Joint space narrowing is insensitive for change over a period of one or two years. Work has to be done to validate surrogate biomarkers for structural OA that allow trials of manageable duration.[85]

*Third*, our understanding of the pathology is limited. While important risk factors and disease processes are elucidated, many remain to be explored. New techniques in genetic sequencing, transcriptomic, proteomic, metabolomic analyses, 7 Tesla

MRI, gait analysis etcetera give us the opportunity to further enlighten pathological pathways. However, innovative use of existing techniques to explore the disease should not be abandoned.

*Fourth*, development plans frequently used a one-size-fits-all principle. While the clinical and structural presentation of end-stage OA is fairly uniform, heterogeneous pathological mechanisms are active in earlier disease stages. Most pre-clinical studies focus on treating only a single or a small number these mechanisms. If only a specific mechanism is countered and the treatment is tested in a clinical study using a "generic OA" population, it will have no effect in participants with different pathophysiologic mechanisms. This will dilute the true treatment effect. A solution would be to cluster pathophysiologic mechanisms into phenotypes (i.e. trauma induced, deteriorating subchondral bone, inflammatory induced etc) and test treatments only in patients with the corresponding phenotype. However, to date we do not have clear-cut phenotypes, let alone biomarkers to identify these phenotypes.

The APPROACH project aims to tackle these four problems for knee OA. This project is a cooperation between pharmaceutical companies, the European Union and university research teams. First, the data of big OA cohort studies was bundled for data mining. The goal of the data mining process was to develop an algorithm to select patients with a high probability of structural and/or pain progression. This algorithm was then used to include patients into a prospective cohort that includes almost all established biomarkers and heaps of novel biomarkers. These biomarkers will be collected together with patient reported outcome measures, physical exams and function tests during a two-year period. First at baseline, later at six-month follow-up and finally at two years. The period between baseline and six months will be used to find rapidly changing surrogate biomarkers for structural and pain progression at two-year follow-up. The prediction algorithm for structural and pain progression will be validated and refined. Furthermore, OA phenotypes based on the biomarkers will be described and biomarkers to select these phenotypes in future trials will be validated. Finally, it will be an enormous library with state-of-the-art data. An incredible chance to answer the abundance of research questions that remains.

## Outline of this thesis

There remains much to be studied and learned in OA. In a single thesis, we will not cure the disease, but hopefully make steps forward. This thesis has two parts. First, we focus on new imaging techniques to add to our toolbox for studying OA in patients. Second, we look into less studied pathological mechanisms in OA and try to find overlap with other diseases.

### Part I: Developing new imaging techniques to analyze OA in patients

Alignment of the lower extremity plays a pivotal role in knee OA. The alignment is historically measured as the mechanical Hip-Knee-Ankle angle (HKAA) on Whole Leg Radiographs (WLR). An easy to use, validated protocol to assure reproducible and accurate WLRs is lacking. In **Chapter 2,** we study the importance of patient positioning and X-ray beam height on the measured HKAA. Additionally, we propose a protocol for reproducible WLRs based on our findings. In **Chapter 3**, we study the reproducibility of the proposed protocol. While HKAA measured from WLRs is the gold standard for measuring lower extremity alignment, many OA cohorts lack this information. This is because WLRs require specialized equipment, trained personnel and cost more money compared to standard radiographs. In **Chapter 4**, we demonstrate how the Femoral-Tibial angle can be used to assess lower extremity alignment when WLRs are missing.

Uniform criteria to quantify OA throughout the body are lacking. However, information on total body OA status in research participants can be valuable. For instance, when studying the relationship between systemic biochemical markers and the progression of knee OA. Systemic biochemical markers are not joint specific. Once they are in the bloodstream, degradation markers origination from the hip are inseparable from degradation markers originating from the knee. Comparably, patient reported outcomes (e.g. EQ5D) and performance test (e.g. Timed Up and Go test) do not discriminate a bottle neck joint. In **Chapter 5** we describe how we developed and tested the OACT-score (OsteoArthritisComputedTomography-score) in order to quantify structural OA throughout all large joints and the spine.

### Part II: The multifactorial pathways to OA

Plain radiography is a cheap, fast and widely available modality to image joints. Standard radiography mainly excels in depicting boney changes in OA, but surrogate measures like joint space narrowing may give additional information on other tissues. Therefore, results of OA are cleary visualized on standard radiographs. However, it would be very worthwhile to be able to predict structural OA before

classic radiographic changes are visible. A tool used to predict the incidence of knee OA, in such cases, is bone texture analysis.[86–88] Additionally, bone texture can be used to predict progression of knee OA in defined OA cases.[87,89–92] Texture is a property that is hard to comprise in words, but seen intuitively in images as being coarse, grainy, rough, smooth or homogenous (**Figure 2**). Bone texture analysis can be used to quantify the subchondral bone and the resulting numbers are correlated to the density, arrangement, porosity and integrity of the trabecular bone.[93–97] Only a small (n=14) study was performed using bone texture analysis in hip OA. Papaloucas et al. showed that bone texture parameters change in 18 months in hip OA patients.[98] In contrast to knee OA, the value of bone textures parameters in predicting the incidence or progression of hip OA was not studied. In **Chapter 6**, we test the added value of bone texture parameters next to classic OA predictors to predict hip OA at ten-year follow-up.

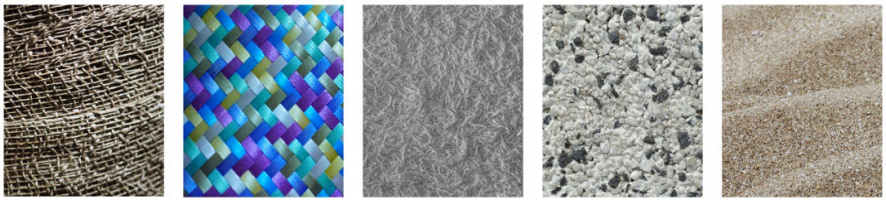


**Figure 2.** Differences in textures can be seen intuitively.

Baseline morphological variations of the hip are associated with the incidence of hip OA. However, the progression of hip OA changes the shape of the hip joint. We know that some morphological variations, such as the cam-lesion and hip dysplasia, originate in children.[99,100] Therefore, they can be seen as predictor or initiator of pathophysiologic mechanisms in hip OA. However, for other morphological variations, such as a varus femur or pincer impingement, it is unclear whether the morphological variation initiates the OA cascade or is a result of it. A classic chicken and egg dilemma. In **Chapter 7**, we try to unravel which morphological variations originate before structural changes of OA and which morphological variations can be deemed as a result of OA.

The spine, pelvis, hip and knees are connected as a chain and transfer loads from the head and torso to the lower extremities. The spinopelvic alignment and morphology will impact the loading and shear stress on the intervertebral discs

and the hip joints. Additionally, if the pelvis is tilted, the acetabular joints coverage of the hip joint may change, together with the subsequent risk for impingement. Pelvic tilt may also affect the location of loads in the hip joint. In **Chapter 8**, we study the role of the sagittal spinopelvic alignment and morphology in hip and knee OA and lumbar spine degeneration. In **Chapter 9**, we explore the role of sagittal spinopelvic alignment and morphology on clinical and radiographic parameters of hip joint impingement.

Multiple studies showed associations between morphological variations of the hip and the incidence of hip OA.[101] However, the added value of morphological variations on top of classic predictors of hip OA remains to be determined. In **Chapter 10**, we develop the Shape-Score, a score based on the morphology of the hip on standard radiograph, to predict hip OA. We test the added value of the Shape-Score besides baseline characteristics, clinical parameters and standard radiographic scores to predict hip OA at eight-year follow-up.

Tissue vascularization plays an important role in many diseases. As stated earlier, pathological mechanisms in atherosclerosis and OA overlap. Arterial calcifications in the lower extremity are clearly visualized on standard radiography. They are related to cardiovascular morbidity and mortality on a patient level and the severity of local atherosclerosis.[102,103] Understanding the relationship between vascular health and OA may guide towards new treatment targets for OA. In **Chapter 11**, we investigate the association between incidence of arterial calcifications and incidence of radiographic knee and hip osteoarthritis. This study was performed on both a patient-level and local-level per joint.

Pseudoxanthoma Elasticum (PXE) is a rare genetic disease caused by mutations in the ABCC6-gene. It causes skin lesions, retinal cracks, and vascular calcifications. In our hospital we noticed multiple patients were complaining about pain in their joints. As OA is the most common cause for joint symptoms, we aimed to study whether PXE patients have a higher prevalence of OA compared to hospital controls. Validating this clinical finding may open up a new research possibility towards pathways in OA. In **Chapter 12**, we investigate whether patients with PXE are more at risk for developing osseous signs of OA.

**Table 1.** Research aims

| | |
|---|---|
| | **Part I**<br>*Developing new imaging techniques to analyze OA in patients* |
| Chapter 2 | (I) To find the effects of leg rotation, knee flexion and altering X-ray beam heights on the measured Hip-Knee-Ankle angle.<br>(II) To implement these findings into a new Whole Leg Radiography protocol. |
| Chapter 3 | To test the reproducibility of the newly implemented Whole Leg Radiography protocol, using a test-retest principle. |
| Chapter 4 | (I) To develop an automated image analysis pipeline to measure the Femoral-Tibial angle from a standard knee radiograph.<br>(II) To analyze the performance of various Femoral-Tibial angle definitions in predicting the Hip-Knee-Ankle angle as measured on a full-limb radiograph. |
| Chapter 5 | (I) To develop a reliable scoring system for assessing structural osteoarthritis burden in large joints and the spine.<br>(II) To demonstrate the inter-observer reliability of said scoring system. |
| | **Part II**<br>*The multifactorial pathways to OA* |
| Chapter 6 | To assess the ability of radiography-based bone texture variables in proximal femur and acetabulum to predict incident radiographic hip OA over a ten-year period. |
| Chapter 7 | To investigate morphological variations of the hip joint and to explore whether they are the cause or a result of hip OA. |
| Chapter 8 | To explore the relationship between sagittal pelvic morphology and the development of the most common degenerative lumbar, hip, and knee pathologies, and the corresponding clinical outcome scores. |
| Chapter 9 | To evaluate the relation between sagittal pelvic morphology in relation to hip-joint anatomy and range-of-motion, and the onset of radiographic signs of FAI and hip OA |
| Chapter 10 | To develop an automated workflow based on hip shape to improve personalized risk prediction for hip OA. |
| Chapter 11 | To investigate the association between incidence of arterial calcifications and incidence of radiographic knee and hip OA. |
| Chapter 12 | To investigate whether patients with Pseudoxanthoma Elasticum are more at risk for developing osseous signs of OA. |

## References


1. Sebbag E, Felten R, Sagez F, Sibilia J, Devilliers H, Arnaud L. The world-wide burden of musculoskeletal diseases: A systematic analysis of the World Health Organization Burden of Diseases Database. *Ann Rheum Dis*. 2019;78(6):844-848. doi:10.1136/annrheumdis-2019-215142

2. *Annual Report Dutch Arthroplasty Registration (LROI)*.; 2018. http://www.lroi-rapportage.nl/.

3. Bayliss LE, Culliford D, Monk AP, et al. The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *Lancet*. 2017;389(10077):1424-1430. doi:10.1016/S0140-6736(17)30059-4

4. Saris DBF. Joint Homeostasis in Tissue Engineering for Cartilage Repair. 2002.

5. Zhang L, Szeri A. Transport of neutral solute in articular cartilage: effects of loading and particle size. *Proc R Soc A Math Phys Eng Sci*. 2005;461(2059):2021-2042. doi:10.1098/rspa.2005.1461

6. Evans RC, Quinn TM. Dynamic compression augments interstitial transport of a glucose-like solute in articular cartilage. *Biophys J*. 2006;91(4):1541-1547. doi:10.1529/biophysj.105.080366

7. Schütz U, Ehrhardt M, Göd S, Billich C, Beer M, Trattnig S. A mobile MRI field study of the biochemical cartilage reaction of the knee joint during a 4,486 km transcontinental multistage ultra-marathon using T2* mapping. *Sci Rep*. 2020;10(1):1-15. doi:10.1038/s41598-020-64994-2

8. Snoeker B, Turkiewicz A, Magnusson K, et al. Risk of knee osteoarthritis after different types of knee injuries in young adults: A population-based cohort study. *Br J Sports Med*. 2020;54(12):725-730. doi:10.1136/bjsports-2019-100959

9. McMillan G, Nichols L. Osteoarthritis and meniscus disorders of the knee as occupational diseases of miners. *Occup Environ Med*. 2005;62(8):567-575. doi:10.1136/oem.2004.017137

10. Agricola R, Heijboer MP, Ginai AZ, et al. A cam deformity is gradually acquired during skeletal maturation in adolescent and young male soccer players: A prospective study with minimum 2-year follow-up. *Am J Sports Med*. 2014;42(4):798-806. doi:10.1177/0363546514524364

11. Eijer H, Hogervorst T. Femoroacetabular impingement causes osteoarthritis of the hip by migration and micro-instability of the femoral head. *Med Hypotheses*. 2017;104:93-96. doi:10.1016/j.mehy.2017.05.035

12. Agricola R, Heijboer MP, Bierma-Zeinstra SM a., Verhaar J a. N, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). *Ann Rheum Dis*. 2013;72(6):918-923. doi:10.1136/annrheumdis-2012-201643

13. Agricola R, Leyland KM, Bierma-Zeinstra SMA, et al. Validation of statistical shape modelling to predict hip osteoarthritis in females: Data from two prospective cohort studies (Cohort Hip and Cohort Knee and Chingford). *Rheumatol (United Kingdom)*. 2015;54(11):2033-2041. doi:10.1093/rheumatology/kev232

14. Neogi T, Bowes MA, Niu J, et al. Magnetic resonance imaging-based three-dimensional bone shape of the knee predicts onset of knee osteoarthritis: data from the osteoarthritis initiative. *Arthritis Rheum*. 2013;65(8):2048-2058. doi:10.1002/art.37987

15. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The Role of Knee Alignment in Disease Progression and Functional Decline in Knee Osteoarthritis. *JAMA*. 2001;286(2):188. doi:10.1001/jama.286.2.188

16. Cerejo R, Dunlop DD, Cahue S, Channin D, Song J, Sharma L. The influence of alignment on risk of knee osteoarthritis progression according to baseline stage of disease. *Arthritis Rheum*. 2002;46(10):2632-2636. doi:10.1002/art.10530

17. Brouwer GM, Tol AW Van, Bergink AP, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. *Arthritis Rheum*. 2007;56(4):1204-1211. doi:10.1002/art.22515

18. Cicuttini F, Wluka A, Hankin J, Wang Y. Longitudinal study of the relationship between knee angle and tibiofemoral cartilage volume in subjects with knee osteoarthritis. *Rheumatology*. 2004;43(3):321-324. doi:10.1093/rheumatology/keh017

19. Kumar D, Manal KT, Rudolph KS. Knee joint loading during gait in healthy controls and individuals with knee osteoarthritis. *Osteoarthr Cartil*. 2013;21(2):298-305. doi:10.1016/j.joca.2012.11.008

20. Hsu RW, Himeno S, Coventry MB, Chao EY. Normal axial alignment of the lower extremity and load-bearing distribution at the knee. *Clin Orthop Relat Res*. 1990;(255):215-227. http://www.ncbi.nlm.nih.gov/pubmed/2347155.

21. Beckwée D, Vaes P, Shahabpour M, Muyldermans R, Rommers N, Bautmans I. The influence of joint loading on bone marrow lesions in the knee. *Am J Sports Med*. 2015;43(12):3093-3107. doi:10.1177/0363546514565092

22. Englund M, Felson DT, Guermazi A, et al. Risk factors for medial meniscal pathology on knee MRI in older US adults: a multicentre prospective cohort study. *Ann Rheum Dis*. 2011;70(10):1733-1739. doi:10.1136/ard.2011.150052

23. Mazzuca SA, Brandt KD, Lane KA, Chakr R. Malalignment and subchondral bone turnover in contralateral knees of overweight/obese women with unilateral osteoarthritis: Implications for bilateral disease. *Arthritis Care Res (Hoboken)*. 2011;63(11):1528-1534. doi:10.1002/acr.20574

24. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The Role of Knee Alignment in Disease Progression and Functional Decline in Knee Osteoarthritis. *JAMA*. 2001;286(2):188. doi:10.1001/jama.286.2.188

25. Eckstein F, Winzheimer M, Hohe J, Englmeier KH, Reiser M. Interindividual variability and correlation among morphological parameters of knee joint cartilage plates: Analysis with three-dimensional MR imaging. *Osteoarthr Cartil*. 2001;9(2):101-111. doi:10.1053/joca.2000.0365

26. Hudelmaier M, Glaser C, Englmeier K-H, Reiser M, Putz R, Eckstein F. Correlation of knee-joint cartilage morphology with muscle cross-sectional areas vs. anthropometric variables. *Anat Rec*. 2003;270A(2):175-184. doi:10.1002/ar.a.10001

27. Jones G, Ding C, Glisson M, Hynes K, Ma D, Cicuttini F. Knee articular cartilage development in children: A longitudinal study of the effect of sex, growth, body composition, and physical activity. *Pediatr Res*. 2003;54(2):230-236. doi:10.1203/01.PDR.0000072781.93856.E6

28. Messier SP, Gutekunst DJ, Davis C, DeVita P. Weight loss reduces knee-joint loads in overweight and obese older adults with knee osteoarthritis. *Arthritis Rheum*. 2005;52(7):2026-2032. doi:10.1002/art.21139

29. Thijssen E, Van Caam A, Van Der Kraan PM. Obesity and osteoarthritis, more than just wear and tear: Pivotal roles for inflamed adipose tissue and dyslipidaemia in obesity-induced osteoarthritis. *Rheumatol (United Kingdom)*. 2014;54(4):588-600. doi:10.1093/rheumatology/keu464

30. Peters AE, Akhtar R, Comerford EJ, Bates KT. The effect of ageing and osteoarthritis on the mechanical properties of cartilage and bone in the human knee joint. *Sci Rep*. 2018;8(1):1-13. doi:10.1038/s41598-018-24258-6

31. Urban H, Little CB. The role of fat and inflammation in the pathogenesis and management of osteoarthritis. 2018;(February):10-21. doi:10.1093/rheumatology/kex399

32. Clavert P, Dosch C, Wolfram R, Kahn L. New findings on intermetacarpal fat pads : anatomy and imaging. 2006:351-354. doi:10.1007/s00276-006-0106-z

33. Eymard F, Pigenet A, Citadelle D, et al. Knee and hip intra-articular adipose tissues ( IAATs ) compared with autologous subcutaneous adipose tissue : a speci fi c phenotype for a central player in osteoarthritis. 2017:1142-1148. doi:10.1136/annrheumdis-2016-210478

34. Gandhi R, Takahashi M, Syed K, Davey JR, Mahomed NN. Relationship between Body Habitus and Joint Leptin Levels in a Knee Osteoarthritis Population. 2010;(March):329-333. doi:10.1002/jor.21000

35. Labusca L. The Unexplored Role of Intra-articular Adipose Tissue in the Homeostasis and Pathology of Articular Joints. 2018;5(March). doi:10.3389/fvets.2018.00035

36. Harasymowicz NS, Clement ND, Azfer A, Burnett R, Salter DM, Simpson AHWR. Regional Differences Between Perisynovial and Infrapatellar Adipose Tissue Depots and Their Response to Class II and Class III Obesity in Patients With Osteoarthritis. 2017;69(7):1396-1406. doi:10.1002/art.40102

37. Hunter DJ, Gerstenfeld L, Bishop G, et al. Bone marrow lesions from osteoarthritis knees are characterized by sclerotic bone that is less well mineralized. *Arthritis Res Ther*. 2009;11(1):R11. doi:10.1186/ar2601

38. Sniekers YH, Intema F, Lafeber FPJG, et al. A role for subchondral bone changes in the process of osteoarthritis; a micro-CT study of two canine models. *BMC Musculoskelet Disord*. 2008;9(1):20. doi:10.1186/1471-2474-9-20

39. Mastbergen SC, Lafeber FPJG. Changes in Subchondral Bone Early in the Development of Osteoarthritis. 2011;63(9):2561-2563. doi:10.1002/art.30306

40. Wen C, Lu WW, Chiu KY. Importance of subchondral bone in the pathogenesis and management of osteoarthritis from bench to bed. *J Orthop Transl*. 2014;2(1):16-25. doi:10.1016/j.jot.2013.11.004

41. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthr Cartil*. 2005;13(9):769-781. doi:10.1016/j.joca.2005.04.014

42. Martín-Millán M, Castañeda S. Estrogens, osteoarthritis and inflammation. *Jt Bone Spine*. 2013;80(4):368-373. doi:10.1016/j.jbspin.2012.11.008

43. Khosla S, Oursler MJ, Monroe DG. Estrogen and the skeleton. *Trends Endocrinol Metab*. 2012;23(11):576-581. doi:10.1016/j.tem.2012.03.008

44. Harrison MH, Schajowicz F, Trueta J. Osteoarthritis of the hip: a study of the nature and evolution of the disease. *J Bone Joint Surg Br*. 1953;35-B(4):598-626. http://www.ncbi.nlm.nih.gov/pubmed/13108925. Accessed June 23, 2016.

45. Mapp PI, Walsh DA. Mechanisms and targets of angiogenesis and nerve growth in osteoarthritis. *Nat Rev Rheumatol*. 2012;8(7):390-398. doi:10.1038/nrrheum.2012.80

46. Bierma-Zeinstra SMA, Waarsing JH. The role of atherosclerosis in osteoarthritis. *Best Pract Res Clin Rheumatol*. September 2018. doi:10.1016/J.BERH.2018.08.006

47. Camaré C, Pucelle M, Nègre-Salvayre A, Salvayre R. Angiogenesis in the atherosclerotic plaque. *Redox Biol*. 2017;12:18-34. doi:10.1016/j.redox.2017.01.007

48. Reiss AB, Siegart NM, De Leon J. Interleukin-6 in atherosclerosis: atherogenic or atheroprotective? *Clin Lipidol*. 2017;12(1):14-23. doi:10.1080/17584299.2017.1319787

49. Hamilton JL, Nagao M, Levine BR, Chen D, Olsen BR, Im HJ. Targeting VEGF and Its Receptors for the Treatment of Osteoarthritis and Associated Pain. *J Bone Miner Res*. 2016;31(5):911-924. doi:10.1002/jbmr.2828

50. Livshits G, Zhai G, Hart DJ, et al. Interleukin-6 is a significant predictor of radiographic knee osteoarthritis: The Chingford Study. *Arthritis Rheum*. 2009;60(7):2037-2045. doi:10.1002/art.24598

51. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: Genetics. *Osteoarthr Cartil*. 2004;12(SUPLL.):39-44. doi:10.1016/j.joca.2003.09.005

52. Zengini E, Hatzikotoulas K, Tachmazidou I, et al. Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. *Nat Genet*. 2018;50(4):549-558. doi:10.1038/s41588-018-0079-y

53. Uhalte EC, Wilkinson JM, Southam L, Zeggini E. Pathways to understanding the genomic aetiology of osteoarthritis. *Hum Mol Genet*. 2017;26(R2):R193-R201. doi:10.1093/hmg/ddx302

54. Hosny S, Gademsetty C, Smith T, Hing C. A Systematic Review and Meta-analysis Investigating the Association Between Bone Marrow Lesions in People with Osteoarthritis. *Curr Rheumatol Rev*. 2014;9(3):146-157. doi:10.2174/1573397109031401301142l5

55. Yusuf E, Kortekaas MC, Watt I, Huizinga TWJ, Kloppenburg M. Do knee abnormalities visualised on MRI explain knee pain in knee osteoarthritis? A systematic review. *Ann Rheum Dis*. 2011;70(1):60-67. doi:10.1136/ard.2010.131904

56. Shabestari M, Kise NJ, Landin MA, et al. Enhanced angiogenesis and increased bone turnover characterize bone marrow lesions in osteoarthritis at the base of the thumb. *Bone Jt Res*. 2018;7(6):406-413. doi:10.1302/2046-3758.76.BJR-2017-0083.R3

57. Laslett LL, Doré DA, Quinn SJ, et al. Zoledronic acid reduces knee pain and bone marrow lesions over 1 year: a randomised controlled trial. *Ann Rheum Dis*. 2012;71(8):1322-1328. doi:10.1136/annrheumdis-2011-200970

58. Karsdal MA, Christiansen C, Ladel C, Henriksen K, Kraus VB, Bay-Jensen AC. Osteoarthritis - a case for personalized health care? *Osteoarthr Cartil*. 2014;22(1):7-16. doi:10.1016/j.joca.2013.10.018

59. Herrero-Beaumont G, Roman-Blas JA, Bruyère O, et al. Clinical settings in knee osteoarthritis: Pathophysiology guides treatment. *Maturitas*. 2017;96:54-57. doi:10.1016/j.maturitas.2016.11.013

60. Dell'Isola A, Allan R, Smith SL, Marreiros SSP, Steultjens M. Identification of clinical phenotypes in knee osteoarthritis: a systematic review of the literature. *BMC Musculoskelet Disord*. 2016;17(1):425. doi:10.1186/s12891-016-1286-2

61. Dell'Isola A, Steultjens M. Classification of patients with knee osteoarthritis in clinical phenotypes: Data from the osteoarthritis initiative. *PLoS One*. 2018;13(1). doi:10.1371/journal.pone.0191045

62. Mobasheri A, Saarakkala S, Finnilä M, Karsdal MA, Bay-Jensen AC, van Spil WE. Recent advances in understanding the phenotypes of osteoarthritis. *F1000Research*. 2019;8. doi:10.12688/f1000research.20575.1

63. NOV. Richtlijn diagnostiek heup- of knieklachten. https://richtlijnendatabase.nl/richtlijn/artrose_in_heup_of_knie/diagnostiek_heup-_of_knieartrose.html. Published 2019.

64. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. http://www.scopus.com/inward/record.url?eid=2-s2.0-62949246527&partnerID=tZOtx3y1.

65. Muñoz-García N, Cordero-Ampuero J, Madero-Jarabo R. Diagnostic Accuracy of Magnetic Resonance Images and Weight-Bearing Radiographs in Patients With Arthroscopic-Proven Medial Osteoarthritis of the Knee. *Clin Med Insights Arthritis Musculoskelet Disord*. 2020;13. doi:10.1177/1179544120938369

66. Peat G, Thomas E, Duncan R, Wood L, Hay E, Croft P. Clinical classification criteria for knee osteoarthritis: Performance in the general population and primary care. *Ann Rheum Dis*. 2006;65(10):1363-1367. doi:10.1136/ard.2006.051482

67. Wright RW, Ross JR, Haas AK, et al. Osteoarthritis classification scales: Interobserver reliability and arthroscopic correlation. *J Bone Jt Surg - Am Vol*. 2014;96(14):1145-1151. doi:10.2106/JBJS.M.00929

68. Kinds MB, Welsing PMJ, Vignon EP, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. *Osteoarthritis Cartilage*. 2011;19(7):768-778. doi:10.1016/j.joca.2011.01.015

69. 69. Altman RD. Criteria for the classification of osteoarthritis of the knee and hip. *Scand J Rheumatol Suppl*. 1987;65:31-39. http://www.ncbi.nlm.nih.gov/pubmed/3317807. Accessed May 3, 2016.

70. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage*. 2007;15 Suppl A:A1-56. doi:10.1016/j.joca.2006.11.009

71. National Institute for Health and Care Excellence. Osteoarthritis: care and management. 2019. https://www.nice.org.uk/guidance/conditions-and-diseases/musculoskeletal-conditions/arthritis.

72. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee. *Arthritis Care Res*. 2020;72(2):149-162. doi:10.1002/acr.24131

73. Reeves ND, Bowling FL. Conservative biomechanical strategies for knee osteoarthritis. *Nat Rev Rheumatol*. 2011;7(2):113-122. doi:10.1038/nrrheum.2010.212

74. Squyer E, Stamper DL, Hamilton DT, Sabin JA, Leopold SS. Unloader knee braces for osteoarthritis: Do patients actually wear them? *Clin Orthop Relat Res*. 2013;471(6):1982-1991. doi:10.1007/s11999-013-2814-0

75. Bruce Moseley J, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347(2):81-88. doi:10.1056/NEJMoa013259

76. Katz JN, Brophy RH, Chaisson CE, et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *N Engl J Med*. 2013;368(18):1675-1684. doi:10.1056/NEJMoa1301408

77. Van De Graaf VA, Noorduyn JCA, Willigenburg NW, et al. Effect of Early Surgery vs Physical Therapy on Knee Function among Patients with Nonobstructive Meniscal Tears: The ESCAPE Randomized Clinical Trial. *JAMA - J Am Med Assoc*. 2018;320(13):1328-1337. doi:10.1001/jama.2018.13308

78. de Windt TS, Vonk LA, Brittberg M, Saris DBF. Treatment and Prevention of (Early) Osteoarthritis Using Articular Cartilage Repair-Fact or Fiction? A Systematic Review. *Cartilage*. 2013;4(3 SUPPL.). doi:10.1177/1947603513486560

79. Paschos NK. Anterior cruciate ligament reconstruction and knee osteoarthritis. *World J Orthop*. 2017;8(3):212-217. doi:10.5312/wjo.v8.i3.212

80. Mella C, Villalón IE, Núñez Á, Paccot D, Díaz-Ledezma C. Hip arthroscopy and osteoarthritis: Where are the limits and indications? *Sicot-J*. 2015;1:27. doi:10.1051/sicotj/2015027

81. Gao L, Madry H, Chugaev D V., et al. Advances in modern osteotomies around the knee: Report on the Association of Sports Traumatology, Arthroscopy, Orthopaedic surgery, Rehabilitation (ASTAOR) Moscow International Osteotomy Congress 2017. *J Exp Orthop*. 2019;6(1):1-14. doi:10.1186/s40634-019-0177-5

82. Kim JH, Kim HJ, Lee DH. Survival of opening versus closing wedge high tibial osteotomy: A meta-Analysis. *Sci Rep*. 2017;7(1):1-7. doi:10.1038/s41598-017-07856-8

83. Faschingbauer M, Nelitz M, Urlaub S, Reichel H, Dornacher D. Return to work and sporting activities after high tibial osteotomy. *Int Orthop*. 2015;39(8):1527-1534. doi:10.1007/s00264-015-2701-2

84. Felson D, Niu J, Sack B, Aliabadi P, McCullough C, Nevitt MC. Progression of osteoarthritis as a state of inertia. *Ann Rheum Dis*. 2013;72(6):924-929. doi:10.1136/annrheumdis-2012-201575

85. Katz JN, Neogi T, Callahan LF, et al. Disease modification in osteoarthritis; pathways to drug

approval. *Osteoarthr Cartil Open*. 2020;2(2):100059. doi:10.1016/j.ocarto.2020.100059

86. Kraus VB, Collins JE, Charles HC, et al. Predictive Validity of Radiographic Trabecular Bone Texture in Knee Osteoarthritis: The Osteoarthritis Research Society International/Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. *Arthritis Rheumatol*. 2018;70(1):80-87. doi:10.1002/art.40348

87. Janvier T, Jennane R, Valery A, et al. Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: data from the Osteoarthritis Initiative: Tibial bone texture & knee OA progression. *Osteoarthr Cartil*. 2017;25(2):259-266. doi:10.1016/j.joca.2016.10.005

88. Podsiadlo P, Nevitt MC, Wolski M, et al. Baseline trabecular bone and its relation to incident radiographic knee osteoarthritis and increase in joint space narrowing score: directional fractal signature analysis in the MOST study. *Osteoarthr Cartil*. 2016;24(10):1736-1744. doi:10.1016/j.joca.2016.05.003

89. Kraus VB, Feng S, Wang SC, et al. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. *Arthritis Rheum*. 2009;60(12):3711-3722. doi:10.1002/art.25012

90. Kraus VB, Feng S, Wang S, et al. Subchondral bone trabecular integrity predicts and changes concurrently with radiographic and magnetic resonance imaging-determined knee osteoarthritis progression. *Arthritis Rheum*. 2013;65(7):1812-1821. doi:10.1002/art.37970

91. Podsiadlo P, Cicuttini FM, Wolski M, Stachowiak GW, Wluka AE. Trabecular bone texture detected by plain radiography is associated with an increased risk of knee replacement in patients with osteoarthritis: A 6 year prospective follow up study. *Osteoarthr Cartil*. 2014. doi:10.1016/j.joca.2013.10.017

92. Woloszynski T, Podsiadlo P, Stachowiak GW, Kurzynski M, Lohmander LS, Englund M. Prediction of progression of radiographic knee osteoarthritis using tibial trabecular bone texture. *Arthritis Rheum*. 2012;64(3):688-695. doi:10.1002/art.33410

93. Hirvasniemi J, Thevenot J, Kokkonen HT, et al. Correlation of Subchondral Bone Density and Structure from Plain Radiographs with Micro Computed Tomography Ex Vivo. *Ann Biomed Eng*. 2016;44(5):1698-1709. doi:10.1007/s10439-015-1452-y

94. Ollivier M, Le Corroller T, Blanc G, et al. Radiographic bone texture analysis is correlated with 3D microarchitecture in the femoral head, and improves the estimation of the femoral neck fracture risk when combined with bone mineral density. *Eur J Radiol*. 2013;82(9):1494-1498. doi:10.1016/j.ejrad.2013.04.042

95. Ranjanomennahary P, Ghalila SS, Malouche D, et al. Comparison of radiograph-based texture analysis and bone mineral density with three-dimensional microarchitecture of trabecular bone. *Med Phys*. 2011;38(1):420-428. doi:10.1118/1.3528125

96. Steines D, Liew SW, Arnaud C, et al. Radiographic trabecular 2D and 3D parameters of proximal femoral bone cores correlate with each other and with yield stress. *Osteoporos Int*. 2009;20(11):1929-1938. doi:10.1007/s00198-009-0908-z

97. Le Corroller T, Pithioux M, Chaari F, et al. Bone texture analysis is correlated with three-dimensional microarchitecture and mechanical properties of trabecular bone in osteoporotic femurs. *J Bone Miner Metab*. 2013;31(1):82-88. doi:10.1007/s00774-012-0375-z

98. Papaloucas CD, Ward RJ, Tonkin CJ, Buckland-Wright C. Cancellous bone changes in hip osteoarthritis: A short-term longitudinal study using fractal signature analysis. *Osteoarthr Cartil*. 2005;13(11):998-1003. doi:10.1016/j.joca.2005.06.009

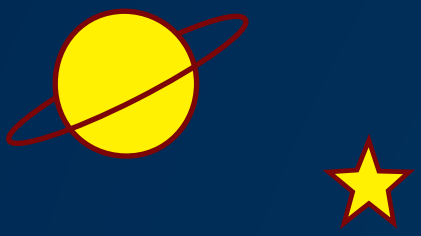
99. Agricola R, Bessems JHJM, Ginai AZ, et al. The Development of Cam-Type Deformity in Adolescent and Young Male Soccer Players. *Am J Sports Med*. 2012;40(5):1099-1106. doi:10.1177/0363546512438381

100. Kotlarsky P, Haber R, Bialik V, Eidelman M. Developmental dysplasia of the hip: What has changed in the last 20 years? *World J Orthop*. 2015;6(11):886-901. doi:10.5312/wjo.v6.i11.886

101. Buuren MMA Van, Arden NK, Bierma-zeinstra SMA, et al. Statistical shape modeling of the hip and the association with hip osteoarthritis : a systematic review. 2020;(December):1-12. doi:10.1016/j.joca.2020.12.003

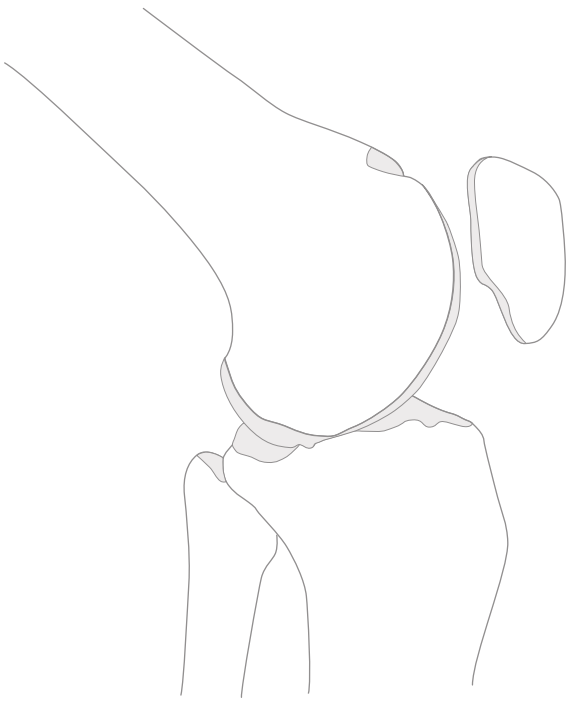
102. Chowdhury MM, Makris GC, Tarkin JM, et al. Lower limb arterial calcification (LLAC) scores in patients with symptomatic peripheral arterial disease are associated with increased cardiac mortality and morbidity. *PLoS One*. 2017;12(9). doi:10.1371/journal.pone.0182952

103. Rocha-Singh KJ, Zeller T, Jaff MR. Peripheral arterial calcification: Prevalence, mechanism, detection, and clinical implications. *Catheter Cardiovasc Interv*. 2014;83(6):E212. doi:10.1002/ccd.25387



# Part I

## Developing new imaging techniques to analyze OA in patients



# CHAPTER 2

# The need for a standardized whole leg radiograph protocol: the effects of knee flexion, leg rotation, and x-ray beam height

*Journal of Cartilage & Joint Preservation, September 2021, doi.org/10.1016/j.jcjp.2021.100022*

Huu C. Nguyen, Willem Paul Gielis, Nienke van Egmond, Harrie Weinans, Kees H. Slump, Ralph J.B. Sakkers, Roel J.H. Custers

## Abstract

### Background and aim

Lower limb malalignment is a major risk factor for knee osteoarthritis (OA) and is mainly diagnosed using the Hip Knee Ankle Angle (HKA). Therefore, accurate HKA measurements are indispensable. We aimed to study the effects of knee flexion, leg rotation, and X-ray beam height on the accuracy of the HKA measurement. We aimed to convert our findings into a guideline for obtaining whole leg radiographs (WLR) in favour of accuracy and reproducibility.

### Methods

An in vitro experiment was designed using sawbones (in 5° varus) of the whole lower limb, fixated in different leg rotation angles, knee flexion angles, HKAs and three different X-ray beam heights.

### Results

The HKA measurement error was 1° per 20° of leg rotation without flexion ($p<0.01$). When 5° of flexion was added, the HKA measurement error was 0.8° per 20° rotation ($p<0.01$). When the leg was in 15° flexion, the HKA measurement error was 4° per 20° rotation ($p<0.01$). X-ray beam height did not cause any significant measurement errors ($p=0.348$).

### Conclusion

This study showed that leg rotation only can lead to clinically relevant measurement errors when exceeding 9°. When there is 15° of knee flexion and leg rotation the error becomes approximately 3°. Varying X-ray beam heights within a range of 10 cm does not affect the accuracy. Based on these findings, we propose guidelines for system setup and patient positioning during a WLR that is easy to apply and aims at minimizing errors when measuring the HKA.

## Introduction

Osteoarthritis (OA) is a common and disabling condition, with worldwide estimates suggesting that 500 million people are currently affected.[1] OA is a multifactorial joint disorder, associated with changes to all tissues in the knee joint, causing pain, stiffness, deformity, and disability.[2] Malalignment of the lower limb (varus or valgus) is an important aetiology of OA. The gold standard for diagnosing malalignment is weight-bearing whole leg radiography (WLR), on which the hip knee ankle angle (HKA) can be measured.[3,4] Accurate leg alignment diagnosis is of most importance when OA patients are indicated for total knee arthroplasty or realignment surgery.[5,6] WLRs form the basis for a reliable pre-operative plan for osteotomies.[5]

Patient positioning during WLRs is an inducer of HKA measurement errors. Known affecters are: knee flexion and extension, foot rotation, hip rotation, weight-bearing, and foot positioning.[7-13] Multiple WLR protocols have been suggested, the one by Paley and Herzenberg being the most popular.[4,5,14,15] In this protocol, the patellae is used as key landmark to determine the true antero-posterior (AP) plane. Heavily relying on the skillset of each X-ray technician, pointing a slightly convex patella bone straightforward.[5,16] Other important aspects such as upper body positioning, imaging system setup (X-ray source height, distance, and calibration tools), and patient instructions (foot positioning and weight-bearing) are not described by the protocol.[5] This method is even more difficult after total knee arthroplasties, with reported HKA measurement errors up to 3.5°.[14,17] Thereby, patellar malalignment is quite common in cases of arthritic knees with a varus malalignment, biasing the HKA measurement when aligned on the patella.[18]

Reproducibility of HKA measurements performed on WLRs is very high.[19] However, there is little evidence on the test-retest reproducibility of obtaining WLRs. Odenbring et al. performed a test-retest study of the measured HKA on a very small scale including 8 lower limbs, with a mean difference of 1.3°.[20] Even this group made use of a meticulous method to determine the AP plane in every individual, by obtaining a lateral radiograph where the posterior aspects of the femoral condyles should be superimposed. [20]

Next to patient positioning, the height of the X-ray beam may also impact on HKA accuracy.[21] Katsui et al. reported measurement errors of the tibial medial malleolus angle and the tibial bimalleolus angle (ankle joint) on weight-bearing radiographs, due to different positioning of the X-ray beam angles.[22] Errors could be up to 2.4° when the projection angle differed 10°.[22]

Inaccurate HKA measurements caused by the lack of standardization can result in misdiagnosis, inaccurate preoperative planning, and erroneous assessment of achieved surgical corrections (rubbish in is rubbish out). The radiographic technique should be consistent and accurate, with a desired accuracy of 0.45°, reported by Jones et al., as the needed wedge accuracy to achieve target corrections in high tibial osteotomy.[23]

Therefore, the purpose of the current study is to find the effects of leg rotation, knee flexion and altering X-ray beam heights on the measured HKA. These results can be used to improve current guidelines for obtaining WLRs with improved accuracy and reproducibility of the measured HKA.

## Methods

### Materials

Solid foam sawbones (Sawbones Europe AB, Malmoe, Sweden) representing a left leg, including one femur, tibia, fibula, talus, calcaneus, and forefoot were used. The model had mean femoral anteversion of 15° and tibiofibular torsion of 30°.[24,25] Metal spheres of 4 mm were placed in the femoral head, tibial spines, and talus, representing the landmarks for HKA measurement to ensure reproducible image analysis. Two Kirschner wires were implanted in the tibia and femur representing the Akagi line and the transepicondylar line (**Figure 1**), used as landmarks for the AP plane.[26]

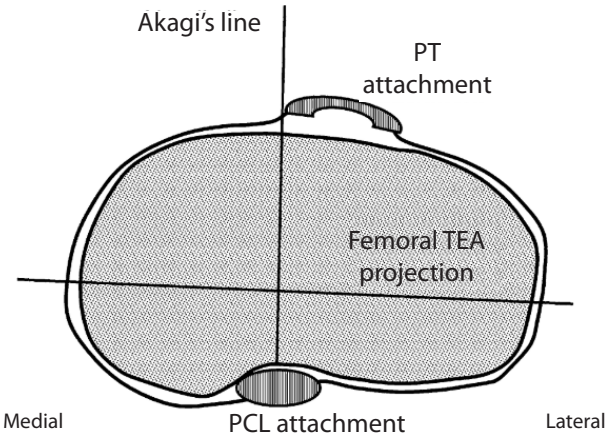



**Figure 1:** Akagi line runs from the centre of the posterior cruciate ligament (PCL) insertion to the medial border of the tuberosity. The femoral transepicondylar (TEA) is projected on the tibial surface.[26]

An Ilizarov frameset fixated the sawbones in a predetermined position, in such a way that the HKA was 5° varus or valgus and the knee flexion could be adjusted between 0 to 15°.[27] The setup is illustrated in **Figure 2**.

**Figure 2:** Measurement setup in the projection radiography room.

## Radiography system

This study used a Philips DigitalDiagnost v4.0, with a fixed X-ray beam height during acquisition where the source pivoted and aimed towards the upper, middle, and lower parts of the limb. The fixed distance between the detector plate and X-ray beam source was set to 265 cm. The X-ray settings were equal to the protocol for scanning patients with kV set at 81 and varying mAs. The radiographic system contained a laser pointer (at the X-ray beam source) directed to the lead measurement tape indicating the X-ray beam height.

First a reference radiograph, as shown in **Figure 3**, was made with the sawbone model set to 5° varus, 0° leg rotation, and 0° knee flexion, with X-ray beam height centred on the joint space. From this reference position, different combinations of leg-rotation (from -10° to 10°), knee flexion (from 0° to 15°) and X-ray beam height (from 0 cm to 10 cm) were applied. Rotation of the leg was described as positive (external) and negative values (internal). The different parameter variations are listed in **Table 1**.

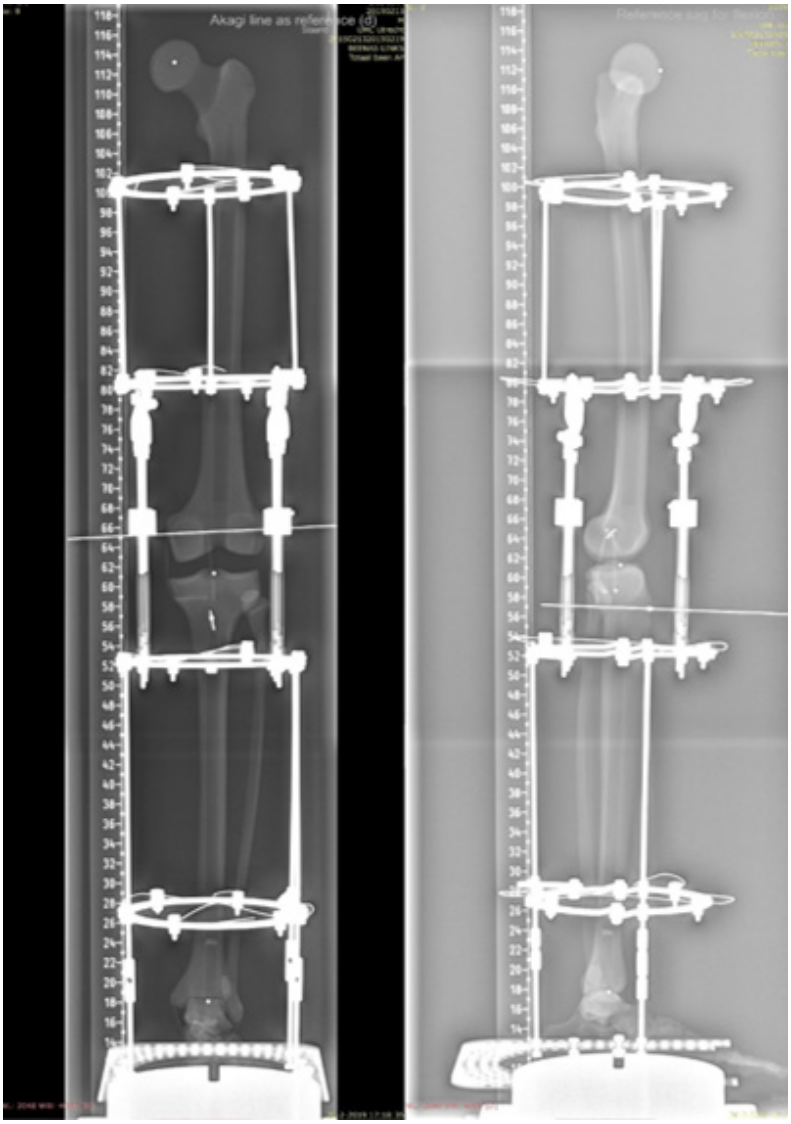


**Figure 3:** Whole leg radiographs exported from PACS image viewer.

**Table 1.** Parameter variations in the experiment.

| Varus/Valgus | Leg rotation | Knee flexion | X-ray beam height |
|---|---|---|---|
| 5° varus | 0°, 5°, 10°, -5°, -10° | 0°, 5°, 15° | Knee-joint, 5cm above, 10cm above |
| 5° valgus | 0°, 5°, 10°, -5°, -10° | 0° | Knee-joint, 5cm above, 10cm above |

### Hip-Knee-Angle measurement

HKAs were measured with one decimal place, using PACS IDS7 19.3 (Sectra AB, Linköping, Sweden) by annotating the metal spheres in the femoral head, tibial spine, and talus. Mechanical lateral distal femoral angle (mLDFA) is the lateral angle formed between the mechanical axis line of the femur and the knee joint line of the femur in the frontal plane.[5] This study did not include mechanical medial proximal tibial (mMPTA) angles for analyses. The setup did not allow the tibial sawbone to flex.

One observer (CN) rated the images twice randomly on independent moments with one week in between, to obtain intra-observer reliability. A second observer (WG) measured the HKAs of the 5° varus radiographs to obtain inter-observer reliability.

### Statistical analysis

The relationship between the measured HKA and knee flexion and/or leg rotation was determined using multivariable and univariable linear regression (SPSS version 25.0, Chicago, IL, USA). The intra-observer reliability was tested for agreement using a two-way mixed Intraclass Correlation (ICC) for absolute agreement. A two-way random ICC for absolute agreement was used to test the inter-rater agreement.

## Results

### Intra- & inter-rater reliability

The HKA ICC for intra-observer reliability was perfect 1.000 (95% CI 0.999 – 1.000) and excellent for the mLDFA 0.993 (95% CI 0.986 – 0.996). The HKA inter-rater reliability was nearly perfect with an ICC of 0.999 (95% CI 0.998 – 0.999).

### Effects of beam height, knee flexion and leg rotation

Multivariable linear regression analyses of leg rotation and knee-flexion on the measured HKA and mLDFA resulted in excellent significant correlations. Both knee flexion (within the range 0° to 15°) and leg rotation (within range -15° to 10°) showed a linear relationship with the measured HKA and mLDFA (**Figures 4** and **5**). Knee flexion and leg rotation had a significant interaction ($p<0.001$), meaning that the effect of leg rotation is affected by the amount of knee flexion and vice versa. External leg rotation in combination with flexion caused the HKA and mLDFA to be overestimated with greater errors under higher flexion (**Figure 4** and **5**). Leg rotation

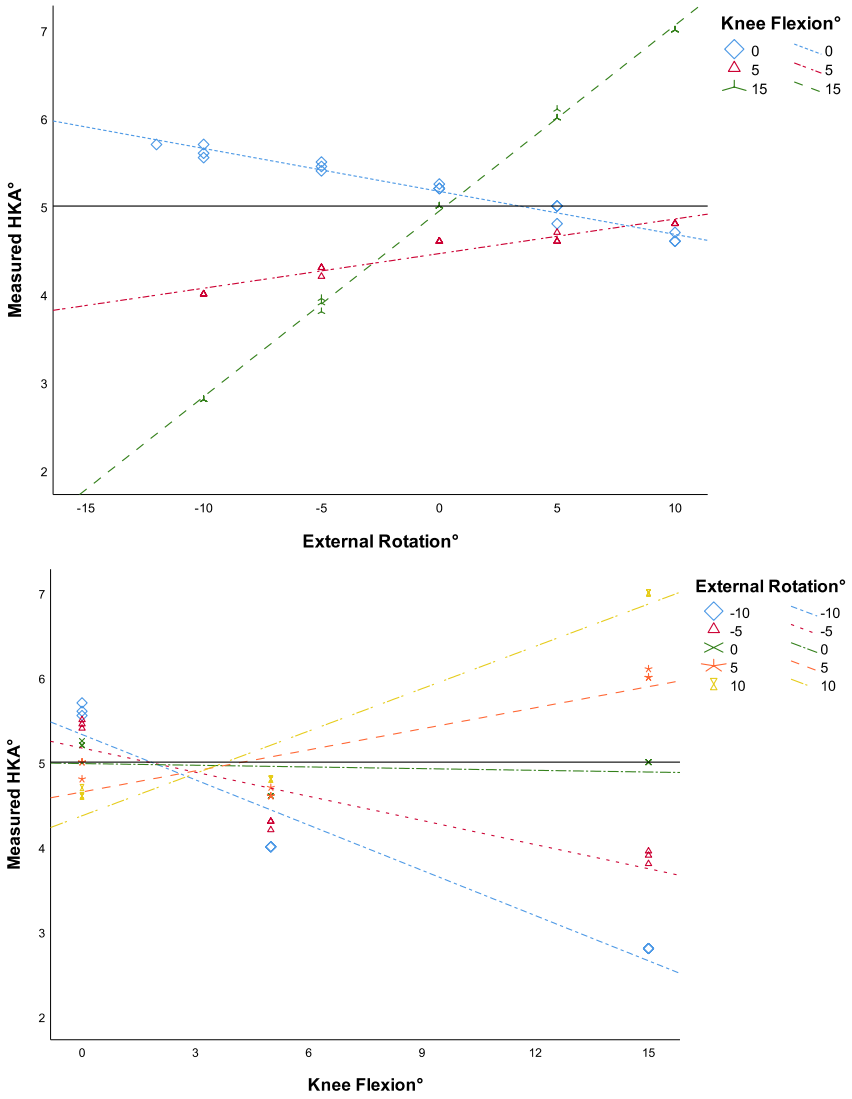



**Figure 4:** Scatterplot of the measured HKA (on 5° varus model) for knee flexion angles (upper figure) and leg rotation angles (lower figure). Each measurement was performed using three different X-ray beam heights, which are the different markers on the same leg rotation and knee flexion combinations. Multivariable linear regression analyses of leg rotation and knee-flexion on the measured HKA showed excellent significant correlations ($p<0.001$). Leg rotation alone affected the measured HKA significantly ($p<0.001$).

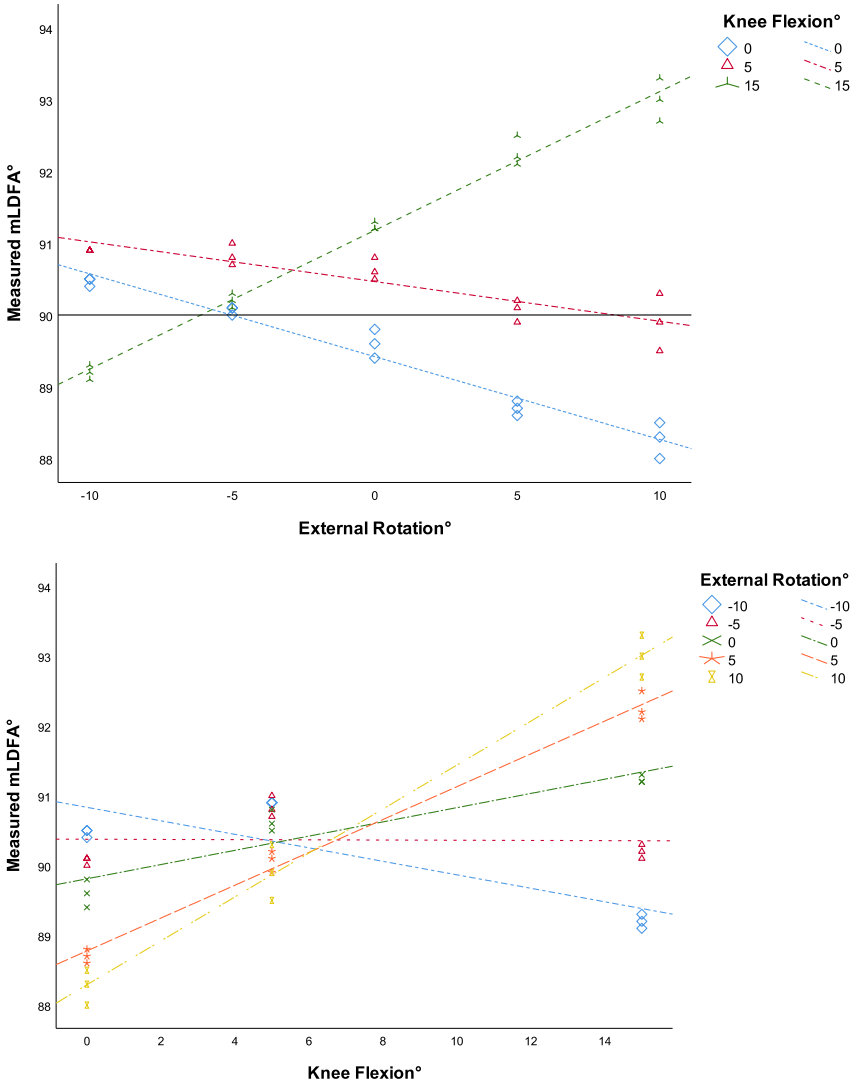



**Figure 5:** Scatterplot of the measured mLDFA (on 90° mLDFA model) for knee flexion angles and leg rotation angles. Each measurement was performed using three different X-ray beam heights, which are the different markers on the same leg rotation and knee flexion combinations. Multivariable linear regression analyses of leg rotation and knee-flexion on the measured mLDFA showed excellent significant correlations ($p<0.001$). Leg rotation alone affected the measured mLDFA significantly ($p<0.001$).

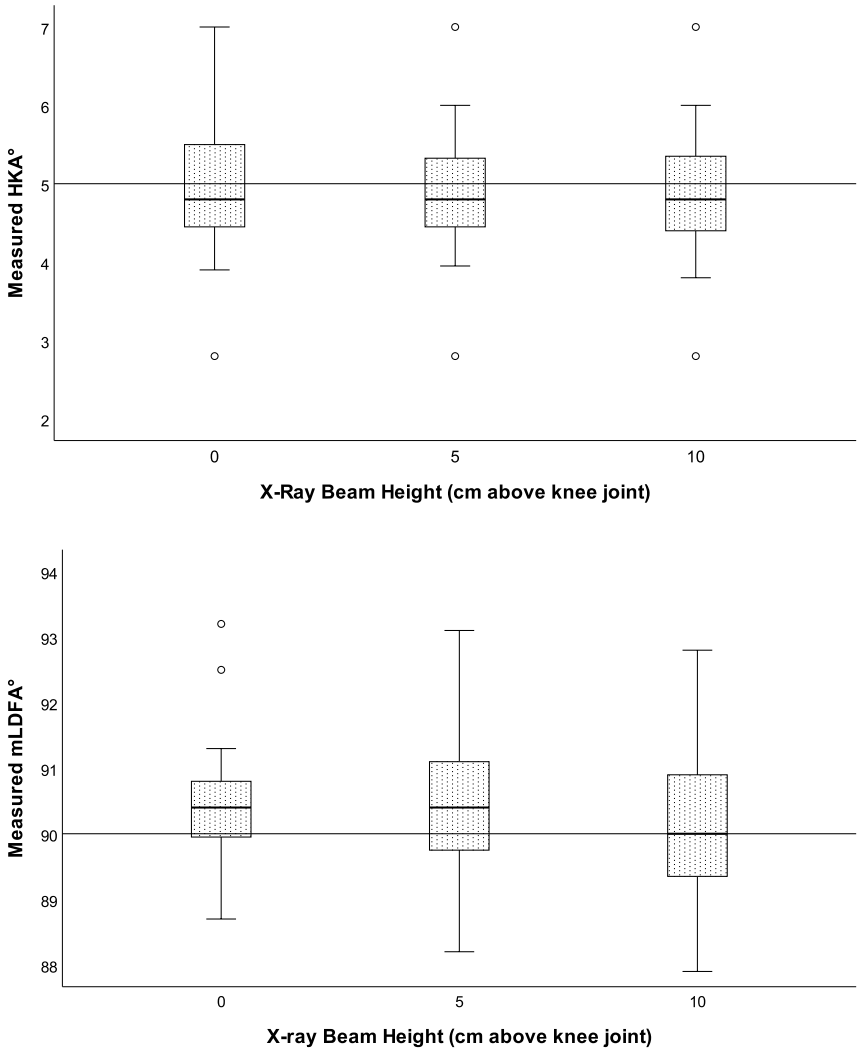



**Figure 6:** Boxplot of the effect of different X-ray Beam Heights on the HKA and mLDFA measurements. All measurements were included in the boxplot (with different knee flexion and leg rotation angles). The horizontal black line displays the set 5° varus and 89° mLDFA angle of the model. X-ray beam height of 5 cm ($p = 0.695$) and 10 cm ($p = 0.424$) above knee joint did not influence the measured HKA on a significant scale. Also, X-ray beam height of 5 cm ($p = 0.549$) and 10 cm ($p = 0.093$) above knee joint did not influence the measured mLDFA on a significant scale.

alone affected both the measured HKA and mLDFA significantly ($p<0.001$). X-ray beam height of 5 cm ($p = 0.695$) and 10 cm ($p = 0.424$) above knee joint did not affect the measured HKA and mLDFA on a significant scale (**Figure 6**). Also, X-ray beam height of 5 cm ($p = 0.549$) and 10 cm ($p = 0.093$) above knee joint did not affect the measured mLDFA on a significant scale.

## Discussion

This study showed that leg rotation alone can lead to clinically relevant measurement errors when exceeding 9°. When 15° knee flexion was combined with 15° leg rotation the HKA measurement error became 3°, thereby grossly exceeding the desired osteotomy accuracy of 0.45° of wedge opening.[23] Different X-ray beam heights did not affect the measured HKA, regardless of flexion and rotation.

Significant effects of leg rotation on the measured HKA were expected and our results correspond with previously published literature.[8,10] Radtke et al. conducted a study using sawbones and found a similar effect of leg rotation, with 0.0558° measurement error as a result of 1° leg rotation, which is about the same as the 1° measurement error per 20° of leg rotation reported in our study.[10] Brouwer et al. conducted a comparable study investigating the relationship between leg rotation, knee flexion and their effect on the measured HKA. Knee flexion was manipulated to a cadaver leg with a transepicondylar rod to control leg rotation.[8] They concluded that the measured HKA is only significantly affected when leg rotation and knee-flexion are combined.[8] But our research showed that there are already clinically relevant effects of leg rotation even with full knee extension. Indeed, when knee flexion was combined with leg rotation the HKA measurement error turned out to be huge. The difference could be explained by the fact that Brouwer et al. only described the HKA in whole degrees.[8] But we aimed to be precise on a tenth of a degree, which means that the desired osteotomy accuracy of 0.45° is achievable.[23]

Both Radtke et al. and Brouwer et al. did not report the possible effect of different X-ray beam heights on the HKA measurement error.[8,10] We investigated the possible effects of using different X-ray beam heights on the measured HKA, which showed no significant effects. Meaning in clinical care it is not necessary to standardize X-ray Beam Heights to obtain reliable and reproducible WLRs.

Pre-operative planning of lower limb osteotomy surgery requires insight in the mechanical medial proximal tibial angle (mMPTA) and the mechanical lateral distal femoral angle (mLDFA).[5] This study analysed the behaviour of mLDFA, and the results were comparable to the HKA measurement errors. Greatest mLDFA

measurement errors occur when the knee is flexed in 15° and combined with 10° leg rotation, with approximately the same 3° discrepancy. Unfortunately, due to the hinge setup used in this study, only the femur could move in the sagittal plane and therefore mimic knee flexion. The tibia was fixed in the sagittal plane, leading to the exclusion of mMPTA measurements in knee flexion and leg rotation conditions. This study also included measurements performed on a 5° valgus stance. The same effect was observed in terms of the effect of knee flexion, leg rotation, and X-ray beam height on the measured HKA when the model was in 5° varus.

Cooke and Sheehy proposed a WLR protocol with the purpose of eliminating leg rotation, at the same time accounting for torsional deformity of the tibia.[14] They proposed that practitioners align each leg using a rotating platform for each foot. Each platform would be fixed to a certain amount of rotation, determined by flexing the knee and observing the frontal plane while making sure that the flexion plane is in line with the X-ray beam.[14] Making it very time-consuming and impractical, while this method heavily relies on the skillset of each practitioner.[17] The Paley and Herzenberg protocol is prone to non-reproducible radiographs, as it relies on the skillset of different X-ray technicians to rotate the knee (using the patellae) the same way as in clinical care.[28] Also, patellar malalignment is quite common in arthritic knees, especially in cases with varus deformities.[18] The Paley and Herzenberg protocol is even more difficult in cases with total knee arthroplasties, with HKA measurement errors up to 3.5°.[17] Probably postoperative swelling and misleading surgical incisions, caused technicians difficulties to exactly centre the patella. Another important finding of this study is the overall average internal rotation of the lower limbs on WLRs, as the patella is located slightly lateral. Centring the patella as instructed in the Paley and Herzenberg protocol requires an internally rotated lower limb.[17]

WLR acquisition guidelines should focus on eliminating leg rotation to minimize the possible effect of knee flexion and account for the mean tibial rotation. It should deliver reproducible radiographs, while being quick and easy to perform by X-ray technicians with fixed positioning for the feet and leg rotation. Firstly, a consensus is needed on which anatomical about which anatomical landmark is easy to define and useable for knee joint rotation assessments. A viable landmark to define proximal tibial rotation on CT scans is the Akagi line, which can represent the antero-posterior alignment of the knee joint.[29] Unfortunately, this line cannot be determined during a physical examination. But the angle between the Akagi line and longitudinal axes of the feet in neutral stance is around 10°.[25,30-32] We believe that aligning each knee straightforward using anatomical landmarks as the

patella, malleoli, or condyles based on experiences of the x-ray technician is very prone to inconsistencies and will result in measurement errors.[14,17,18]

For reproducible WLR acquisition, we propose that patients are positioned in maximum knee extension, which is more straightforward to apply by X-ray technicians compared to a certain knee flexion angle. The feet are pointed outwards with 10° of rotation and with 10 centimetres between the centre of their heels. The angle of 10° is situated between the longitudinal axes of the foot.[25,30–34] Practitioners thereby control the hip rotation, by placing the upper body in a straightforward position.[7,8,10–13] No handlebars or supports are allowed to ensure full weight-bearing. [9,34,35] The practitioners additionally instructs the patient to distribute the weight equally over both leg.[9,35] Each WLR is made bilaterally and remarks about the acquisition should be annotated in the radiograph. We started a test-retest study, aiming to describe the scan rescan reproducibility of the proposed WLR positioning protocol.

Aligning the feet at 10° to define the antero-posterior alignment does not account for individual variances in tibiofibular torsion and femoral anteversion. However, the standard deviation of tibiofibular torsion and femoral anteversion within the population of both angles is below 9°, which means that approximately 68% of the patients show rotational variances below 9°.[24,25] Our results show that leg rotation up to ±9° causes a ±0.45° difference in measured HKA, which is an acceptable error in a medial open wedge high tibial osteotomy.[8,10,23,36] This rule of thumb is also substantiated by results of Kawahara et al. and Jud et al.[33,36] But with the proposed guidelines we aim on a more reproducible basis for obtaining WLRs. This benefits the post-operative assessment of realized lower limb osteotomies and future studies relying on reproducible WLRs.

When there is a suspicion of a large lower limb torsional deformity (tibia and/or femur), we suggest using 3D imaging techniques in the work-up. Indications for rotational errors on an AP knee radiograph are no (or too much) tibia-fibular overlap and femoral condyle asymmetry.[37] When using the proposed protocol with fixed feet, the left and right knee should be presented in the same manner. In cases with differences between left and right, further analyses for possible rotational deformities should be conducted. Also, when there is a suspicion of a hyperextension or a fixed flexion deformity during physical examination, the WLR becomes unreliable.

Our study has many limitations. First, our model could not represent the knee joint kinematics in terms of soft tissue compression and tension. During knee

flexion, knee joint kinematics are complex due to muscle contractions and ligament tensions. Our sawbone model did not incorporate these parameters and the influence of weight-bearing. However, we tried to mimic the knee-joint articulation by positioning the hinge points of the Ilizarov frame slightly above the knee-joint (on Blumensaat line).[38] We checked every position with a frontal antero-posterior for varus alignment and a sagittal radiograph for flexion. Future research should also consider the role of hyperextension in lower limb alignment measurement errors. Second, leg rotation was simulated by rotating the whole sawbone model. In reality, the lower limb has multiple bones and joints which can cause or compensate for leg rotation. Both the ankle and femoral joint can rotate internally/externally. Third, we did not include the mechanical tibia angle as described by Paley, due to model limitations.[5] Future studies should investigate the reproducibility of positioning protocols, including the standard measurement protocol proposed by Paley. Fourth, our model did not include the patella, which is important in the WLR protocol of Paley.

## Conclusion

This study showed that leg rotation only can lead to clinically relevant measurement errors when exceeding 9°. When there is 15° of knee flexion and leg rotation the error becomes approximately 3°. Varying X-ray beam heights within a range of 10 cm does not affect the accuracy. Based on these findings, we propose guidelines for system setup and patient positioning during a WLR that is easy to apply and aims at minimizing errors when measuring the HKA.

## Acknowledgements

Our group thanks the Orthopaedic Technicians and the X-ray technicians for the received support.

## References


1. Hunter DJ. Osteoarthritis in 2020 and beyond: a Lancet Commission. 2020;396(10264):2.
2. Glyn-Jones S, Palmer AJR, Agricola R, et al. Osteoarthritis. The Lancet. 2015;386(9991):376-387. doi:10.1016/S0140-6736(14)60802-3
3. Lee YS, Kim MG, Byun HW, Kim SB, Kim JG. Reliability of the imaging software in the preoperative planning of the open-wedge high tibial osteotomy. Published online 2015:846-851. doi:10.1007/s00167-013-2700-z
4. Sheehy L, Cooke TD V. Radiographic assessment of leg alignment and grading of knee osteoarthritis: A critical review. World J Rheumatol. 2015;5(2):69-81. doi:10.5499/wjr.v5.i2.69
5. Paley D. Normal Lower Limb Alignment and Joint Orientation. In: Principles of Deformity Correction. Springer Berlin Heidelberg; 2002:1-18.
6. Zahn RK, Renner L, Perka C, Hommel H. Weight-bearing radiography depends on limb loading. Knee Surg Sports Traumatol Arthrosc. Published online 2018:1-7. doi:10.1007/s00167-018-5056-6
7. Sanfridsson J, Ryd L, Svahn G, Fridén T, Jonsson K. Radiographic measurement of femorotibial rotation in weight-bearing: The influence of flexion and extension in the knee on the extensor mechanism and angles of the lower extremity in a healthy population. Acta Radiol. 2001;42(2):207-217. doi:10.1080/028418501127346512
8. Brouwer RW, Jakma TSC, Brouwer KH, Verhaar JAN. Pitfalls in determining knee alignment: a radographic cadaver study. J Knee Surg. 2007;20(3):210-215.
9. Specogna A V., Birmingham TB, Hunt MA, et al. Radiographic measures of knee alignment in patients with varus gonarthrosis: Effect of weightbearing status and associations with dynamic joint load. Am J Sports Med. 2007;35(1):65-70. doi:10.1177/0363546506293024
10. Radtke K, Becher C, Noll Y, Ostermeier S. Effect of limb rotation on radiographic alignment in total knee arthroplasties. Arch Orthop Trauma Surg. 2010;130(4):451-457. doi:10.1007/s00402-009-0999-1
11. Lee YS, Lee BK, Lee SH, Park HG, Jun DS, Moon DH. Effect of foot rotation on the mechanical axis and correlation between knee and whole leg radiographs. Knee Surg Sports Traumatol Arthrosc. 2013;21(11):2542-2547. doi:10.1007/s00167-013-2419-x
12. Cho YM, Ko YJ, Lee W. Relationships among foot position , lower limb alignment , and knee adduction moment in patients with degenerative knee osteoarthritis. J Phys Ther Sci. 2015;27(1):1-4.
13. Hunt MA, Fowler PJ, Birmingham TB, Jenkyn TR, Giffin JR. Foot rotational effects on radiographic. 2006;49(6):401-406.
14. Cooke TD V, Sled EA. Optimizing limb position for measuring knee anatomical axis alignment from standing knee radiographs. J Rheumatol. 2009;36(3):472-477. doi:10.3899/jrheum.080732
15. Sanfridsson J, Ryd L, Eklund K, Kouvaras Y, Jonsson K. Angular configuration of the knee: Comparison of conventional measurements and the QUESTOR precision radiography system. Acta Radiol. 1996;37(5):633-638. doi:10.3109/02841859609177689

16. Fox A, Wanivenhaus F, Rodeo S. The Basic Science of the Patella: Structure, Composition, and Function. J Knee Surg. 2012;25(02):127-142. doi:10.1055/s-0032-1313741

17. Maderbacher G, Baier C, Benditz A, et al. Presence of rotational errors in long leg radiographs after total knee arthroplasty and impact on measured lower limb and component alignment. Int Orthop. 2017;41(8):1553-1560. doi:10.1007/s00264-017-3408-3

18. Harrison MM, Cooke TD, Fisher SB, Griffin MP. Patterns of knee arthrosis and patellar subluxation. Clin Orthop. 1994;(309):56-63.

19. Bowman A, Shunmugam M, Watts AR, Bramwell DC, Wilson C, Krishnan J. Inter-observer and intra-observer reliability of mechanical axis alignment before and after total knee arthroplasty using long leg radiographs. Knee. 2016;23(2):203-208. doi:10.1016/j.knee.2015.11.013

20. Odenbring S, Berggren AM, Peil L. Roentgenographic Assessment of the Hip-Knee-Ankle Axis in Medial Gonarthrosis. Clin Orthop. 1993;(289):195-196.

21. Brage ME, Holmes JR, Sangeorzan BJ. The influence of x-ray orientation on the first metatarsocuneiform joint angle. Foot Ankle Int. 1994;15(9):495-497. doi:10.1177/107110079401500907

22. Katsui R, Fujii T. Change of the X-ray Beam Angle may Influence Ankle Image of Weight-bearing Anteroposterior View: Trial to Evaluate Ankle Joint on Standing Whole-leg Radiograph. Clin Res Foot Ankle. 2015;04(01):1-5. doi:10.4172/2329-910X.1000184

23. Jones LD, Brown CP, Jackson W, Monk AP, Price AJ. Assessing accuracy requirements in high tibial osteotomy: a theoretical, computer-based model using AP radiographs. Knee Surg Sports Traumatol Arthrosc. 2017;25(9):2952-2956. doi:10.1007/s00167-016-4092-3

24. Bråten M, Terjesen T, Rossvoll I. Femoral anteversion in normal adults: Ultrasound measurements in 50 men and 50 women. Acta Orthop. 1992;63(1):29-32. doi:10.3109/17453679209154844

25. Turgut A, Göktürk E, Seber S, Köse N, Hazer B, Günal I. Rotational profile of the lower extremity and foot progression angle: computerized tomographic examination of 50 male adults. Arch Orthop Trauma Surg. 2002;120(5-6):255-258. doi:10.1007/s004020050459

26. Aglietti P, Sensi L, Cuomo P, Ciardullo A. Rotational position of femoral and tibial components in TKA using the femoral transepicondylar axis. Clin Orthop. 2008;466(11):2751-2755. doi:10.1007/s11999-008-0452-8

27. Spiegelberg B, Parratt T, Dheerendra SK, Khan WS, Jennings R. Ilizarov principles of deformity correction History of Ilizarov. Published online 2010:101-105. doi:10.1308/003588410X12518836439326

28. Ahrend M-D, Finger F, Grünwald L, Keller G, Baumgartner H. Improving the accuracy of patient positioning for long-leg radiographs using a Taylor Spatial Frame mounted rotation rod. Arch Orthop Trauma Surg. 2021;141(1):55-61. doi:10.1007/s00402-020-03460-0

29. Saffarini M, Nover L, Tandogan R, et al. The original Akagi line is the most reliable: a systematic review of landmarks for rotational alignment of the tibial component in TKA. Knee Surg Sports Traumatol Arthrosc. 2018;27(4):1018-1027. doi:10.1007/s00167-018-5131-z

30. Chotanaphuti T, Srisawasdi R, Rattanaprichavej P, Laoruengthana A. The rotational axis of the tibia and relationship to the tibial torsion in varus osteoarthritic knee. J Med Assoc Thail Chotmaihet Thangphaet. 2012;95 Suppl 1:6-11.

31. Isman R, Inman V, Poor P. Anthropometric studies of the human foot and ankle. Bull Prosthet Res. 1969;11:97-129. doi:10.1007/s11695-014-1382-z

32. Lamm BM, Stasko PA, Gesheff MG, Bhave A. Normal Foot and Ankle Radiographic Angles, Measurements, and Reference Points. J Foot Ankle Surg. 2016;55(5):991-998. doi:10.1053/j.jfas.2016.05.005

33. Jud L, Trache T, Tondelli T, Fürnstahl P, Fucentese SF, Vlachopoulos L. Rotation or flexion alters mechanical leg axis measurements comparably in patients with different coronal alignment. Knee Surg Sports Traumatol Arthrosc. 2019;(0123456789). doi:10.1007/s00167-019-05779-7

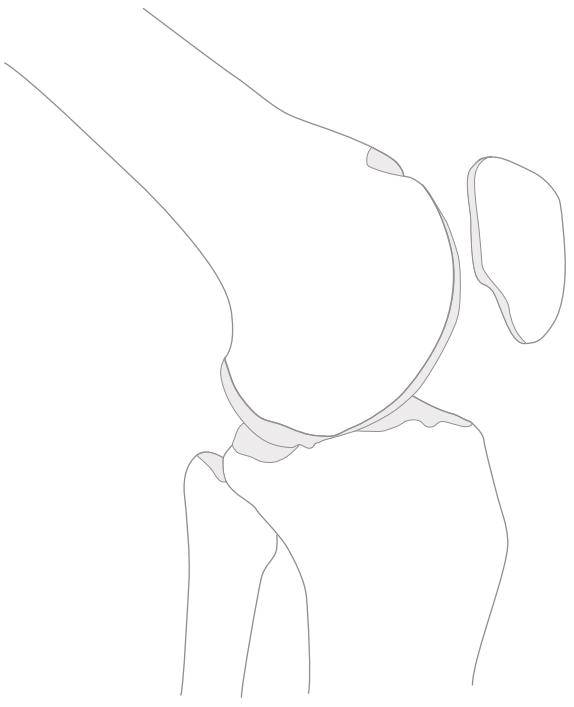
34. Jud L, Roth T, Fürnstahl P, Vlachopoulos L, Sutter R, Fucentese SF. The impact of limb loading and the measurement modality (2D versus 3D) on the measurement of the limb loading dependent lower extremity parameters. BMC Musculoskelet Disord. 2020;21(1):418. doi:10.1186/s12891-020-03449-1

35. Schoenmakers DAL, Feczko PZ, Boonen B, Schotanus MGM, Kort NP, Emans PJ. Measurement of lower limb alignment: there are within-person differences between weight-bearing and non-weight-bearing measurement modalities. Knee Surg Sports Traumatol Arthrosc. 2017;25(11):3569-3575. doi:10.1007/s00167-017-4636-1

36. Kawahara S, Mawatari T, Matsui G, Mizu H, Akasaki Y. Malrotation of whole - leg radiograph less than 10 degrees does not influence preoperative planning in open - wedge high tibial osteotomy. 2020;(February):1-7. doi:10.1002/jor.24845

37. Maderbacher G, Schaumburger J, Baier C, et al. Predicting knee rotation by the projection overlap of the proximal fibula and tibia in long-leg radiographs. Knee Surg Sports Traumatol Arthrosc. 2014;22(12):2982-2988. doi:10.1007/s00167-014-3327-4

38. Çakmak M, Şen C, Eralp L, Balci HI, Civan M, eds. Basic Techniques for Extremity Reconstruction. Springer International Publishing; 2018. doi:10.1007/978-3-319-45675-1



# CHAPTER 3

# A new protocol for obtaining whole leg radiographs with excellent reproducibility

*Journal of Cartilage & Joint Preservation*
*Submitted for publication*

Huu C. Nguyen, Nienke van Egmond, Mario Hevesi,
Harrie Weinans, Willem Paul Gielis, Roel J.H. Custers,

## Abstract

### Background and aim

Whole Leg Radiographs (WLR) are the gold standard for diagnosing malalignment and for pre-operative osteotomy planning. A widely accepted standardized protocol for patient positioning during WLRs does not exist. Positioning can affect the reproducibility of the measured hip knee angle (HKA), resulting in insufficient diagnostics and preoperative plans. We developed an easy-to-use WLR protocol by standardizing patient positioning and focusing on reproducibility. This study aims on testing this reproducibility.

### Patients and methods

This study enrolled 30 patients for a test-retest analysis. Each patient underwent two bilateral WLRs on the same day using the investigated positioning protocol. Three observers measured the HKA, mechanical medial proximal tibial angle (mMPTA), mechanical lateral distal femoral angle (mLDFA), and joint line convergence angle (JLCA) on the two radiographs. Twice each, with one week between.

### Results

The intra-observer and inter-observer reliabilities were excellent, with intraclass correlation coefficients (ICCs) between 0.990 and 0.996. The ICCs between the measured HKA (0.985), mMPTA (0.922), and mLDFA (0.903) on the two separate radiographs were excellent. The ICC between the JLCA measured on the first and second WLR was moderate with 0.632. The mean absolute error between the HKA, mMPTA, mLDFA, and JLCA measurements on the first and second WLR were respectively: 0.442°, 0.783°, 0.828°, and 0.794°.

### Conclusion

The investigated novel WLR positioning protocol produced excellent and reproducible HKA measurements, with clinically acceptable degrees of error. We recommend applying this easy-to-use protocol when obtaining WLRs for osteotomy planning.

## Introduction

Osteoarthritis (OA) is one of the most common chronic pathologies and a leading cause of disability and pain.[1] Among adults above 60 years of age, the prevalence of symptomatic knee OA is around 10% in men and 13% in women.[2] Worldwide estimates suggest that 250 million people are currently affected, with substantial associated socioeconomic costs.[3]

An important aetiology of knee OA is a varus or valgus malalignment, causing progressive unicompartmental knee pathology.[4,5] Given this, young and active patients with malalignment benefit particularly from corrective osteotomies as it concerns.[6] A corrective osteotomy is a joint preserving technique that can postpone joint replacement surgery for up to 10 years in more than 85% of patients.[3,7] By proceeding with corrective osteotomy, the age of the patient at primary total knee arthroplasty potentially significantly increases, resulting in decreased likelihood of revision surgery later in life.[8]

Diagnosing and planning corrective osteotomies requires whole leg radiographs (WLR).[5,9] It is important that the radiographic technique is consistent and accurate. Jones et al. reported a desired accuracy of 0.45° as the needed wedge accuracy to achieve target corrections in high tibial osteotomy (HTO).[10]

But to the best of our knowledge, no widely accepted standardized protocol for patient positioning during WLRs has been reported.[11,12] This is of clinical importance given that the reproducibility of the measured hip knee angle (HKA) on a WLR is influenced by positional differences of patients, like foot positioning, knee flexion, leg rotation, and weight-bearing.[13–18] This results in insufficient preoperative planning and assessment of achieved surgical corrections (rubbish in is rubbish out). Causing under- and severe overcorrections, which leads to failing HTO treatments with reported surgical accuracies of ±3°.[19–21]

Odenbring et al. performed a test-retest study including only 8 patients using a WLR protocol (which was clinically not viable), resulting in a mean error of 1.3° and exceeding the desired surgical accuracy.[10,22]

A new easy to use WLR positioning protocol was developed and introduced in clinical care, with a focus on standardizing patient positioning and reproducibility. This study aims on testing the reproducibility of the implemented WLR protocol, using a test-retest principle. We hypothesize that with this new protocol, WLRs are reproducible and within the desired published osteotomy accuracy of 0.45°.

## Methods

This prospective explorative study was approved by the ethics board of the UMC Utrecht hospital (METC number 19-474) on 27 November 2019. Patients were recruited at the Mobility Clinic of the UMC Utrecht, which is a tertiary orthopaedic referral centre for knee joint preserving treatments, including cartilage repair techniques, osteotomies, and knee joint distraction[23–25]. Patients were eligible when they had an appointment at the outpatient's clinic with a scheduled WLR. Exclusion criteria were; age under 18 years, inability to read, communicate, and/or speak the Dutch language, pregnancy, and patients incapable of providing informed consent. When willing to participate, patients signed an informed consent. Included patients underwent a second WLR after their appointment at the outpatient clinic, minimally half an hour after the initial examination. Both acquisitions followed the same standardized, novel positioning protocol. Resulting in two separate episodes of positioning and acquisition within the same day.

Patients were positioned with their knees in full extension. Feet were positioned with a distance of 10 centimetres between the heels, and aligned in 10 degrees of external rotation. This is achieved by placing the feet on the positioning template (**Figure 1**). X-ray technicians subsequently adjusted the hip rotation, by aligning the upper body and pelvis in a straightforward (AP) position. No handlebars or supports were employed and patients were asked to place their hands alongside their body to ensure full weight-bearing. The X-ray technicians additionally instructed the patient to distribute their weight equally to each leg. All WLRs were performed with the bilateral lower extremities captured on the radiographs, with a radiopaque measurement tape positioned behind the patients for subsequent image calibration. **Figure 2** illustrates a patient undergoing WLR with our standardized positioning protocol.

Each WLR is obtained using the Philips DigitalDiagnost v4.0 (Philips, Amsterdam, the Netherlands). The fixed distance between the detector plate and X-ray beam source was set to 265 cm. The X-ray settings were kept the same for each patient, with kV set at 81 and varying mAs.

Radiographic measurements of the HKA were done for each leg separately by three observers in PACS IDS7 19.3 (Sectra AB, Linköping, Sweden), with the module OrthoStation. Two orthopaedic surgeons (RC&NE) and a researcher (CN) performed the measurements twice, with one week in between. The two radiographs of each patient were analysed separately and in a random order.

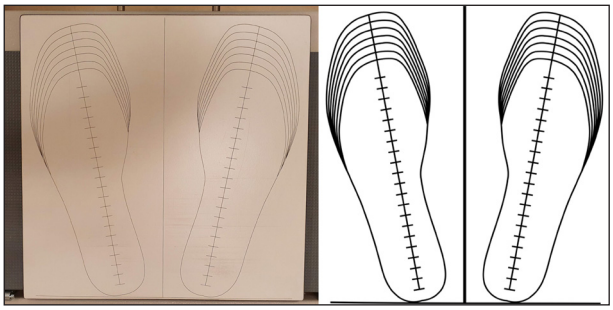


**Figure 1.** Template for feet positioning during a WLR, as part of the WLR positioning protocol. Feet are pointed outwards in 10 degrees between the midlines and placed 10 from each other from the centre of the heels. This template is engraved onto a durable Trespa® board (right picture).

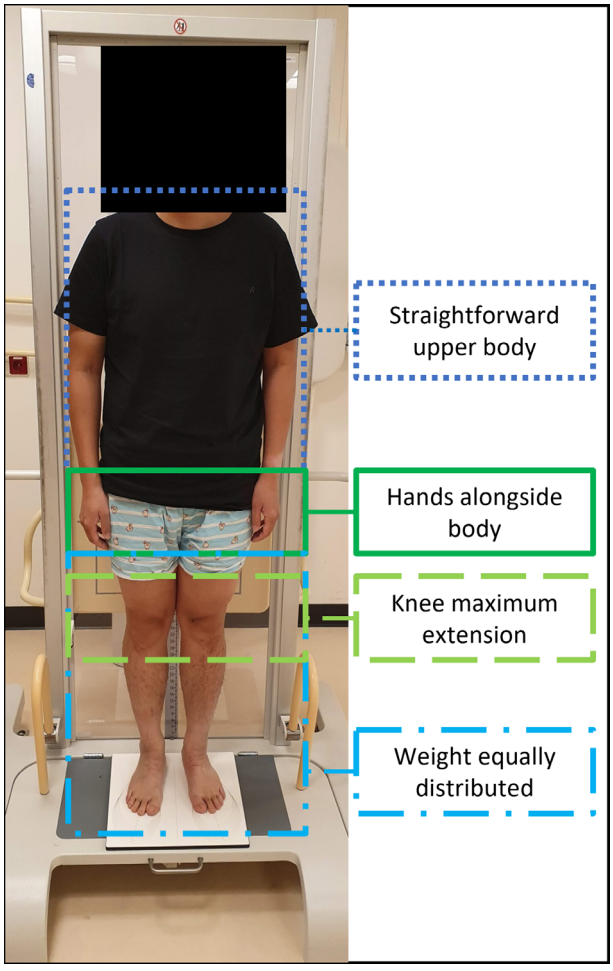



**Figure 2.** Patient positioning following WLR protocol.

HKA was determined as the angle between the line from the middle of the femoral head to the middle of the trochlea, and the line from the centre of the tibial spines to the centre of the talus (**Figure 3**). The HKA was measured with one decimal place accuracy.

OrthoStation provides a semi-automatic method to determine the joint line convergence angle (JLCA), mechanical medial proximal tibial angle (mMPTA), and mechanical lateral distal femoral angle (mLDFA). The tool requires input from the observers, selecting the centre of the femoral head and talus, followed by marking the joint lines of the tibia and femur (**Figure 3**). The mMPTA, mLDFA, and JLCA were provided in whole numbers.

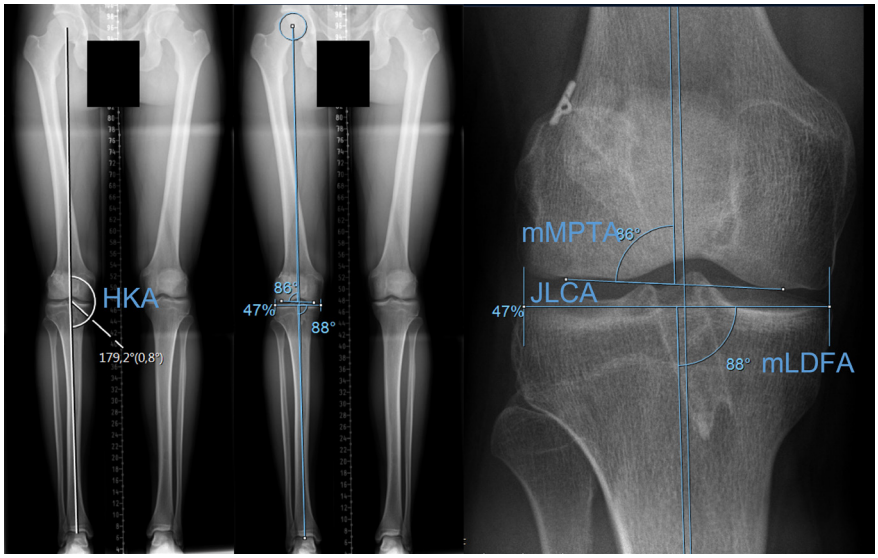



**Figure 3.** Leg geometry measurements on a WLR in PACS IDS7 19.3 (Sectra AB, Linköping, Sweden). Left image illustrates the Hip Knee Angle (HKA) measurement, the right image illustrates the mechanical medial proximal tibial angle (mMPTA), mechanical lateral distal femoral angle (mLDFA), and the joint line convergence angle (JLCA).

## Statistical analysis

Intra-observer reliability was tested using a two-way mixed Intraclass Correlation (ICC) for absolute agreement. The inter-observer reliabilities were tested using a two-way random ICC for absolute agreement. Test-retest agreement was calculated using a two-way random ICC and Bland-Altman analyses. The errors between the

measured parameters on the two WLRs and different observers were reported as mean (95%-CI interval). Due to a non-normal right skewed distribution of the absolute errors, these values were reported as mean (bootstrapped resampled 95%-CI intervals). Statistical significance was set at alpha = 0.05. All statistical calculations were performed in SPSS Statistics (IBM, version 25.0.0.2.).

Power calculations for reproducibility studies are not straightforward and ill-studied26. Following the guideline as proposed by Bujanga and Baharum, we need only 14 legs to find a significant difference between an ICC of 0.7 and 0.9. We however deemed this number as very small and aimed to reach a narrower confidence interval27. Following the advice of Cicchetti we used three highly skilled readers to assess 60 legs in 30 patients, allowing us to include a range of alignments, male and females 28.

## Results

For this study 31 patients signed an informed consent. We had to exclude 1 patient due to a no show for the retest radiograph. Therefore, 30 patients with 30 bilateral WLRs taken on two separate time-points (60 measured WLRs, 120 legs in total) were included in this study. The study included 15 males and 15 women, with a median age of 34.5 (18-61) years, and mean BMI of 25.8 (SD 3.2). Mean HKA was 179.25°, mean mMPTA was 87.06°, mean mLDFA was 87.31°, and mean JLCA was 1.32° (**Figure 4**).

Test results of the intra-observer and inter-observer reliabilities are listed in **Table 1**. Intra-observer and inter-observer reliabilities of the HKA, mMPTA, and mLDFA were excellent. Intra-observer and inter-observer reliabilities of the JLCA were fair to good.

The mean absolute differences between the three observers of the measured HKA, mMPTA, mLDFA, and JLCA were respectively: 0.491° (CI 0.430° – 0.552°), 0.889° (CI 0.761° – 1.013°), 0.922° (CI 0.806° – 1.056), and 0.931° (CI 0.805° – 1.069°).

All results of the test-retest analyses are listed in **Table 2**. The test-retest ICCs between the measured HKA, mMPTA, and mLDFA on the first and second radiograph were excellent. The ICC between the measured JLCA on the first and second WLR was moderate. The mean absolute test-retest errors were for the measured HKA 0.442° (CI 0.387°-0.498°), mMPTA 0.783° (CI 0.683°-0.878°), mLDFA 0.828° (CI 0.722°-0.944°), and JLCA 0.794° (CI 0.683°-0.911°).

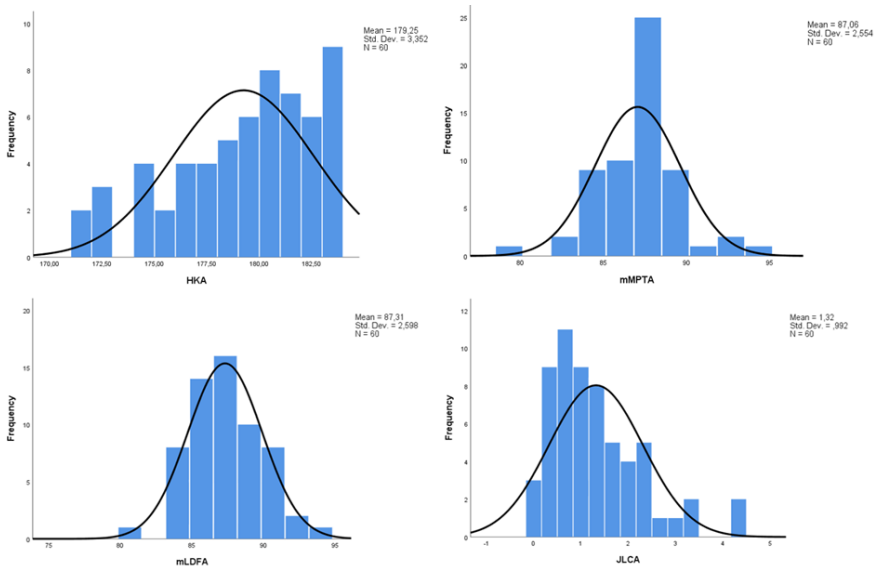



**Figure 4.** Histograms of the population mean measured HKA, mMPTA, mLDFA, mLDFA, and JLCA. Reported normal distributions according to Bellemans et al. are: HKA = 178.67 ± 2.34, mMPTA = 87.04 ± 2.07, mLDFA = 87.90 ± 1.74, and JLCA = 0.51 ± 1.0538.

**Table 1:** Intra-observer and inter-bserver reliabilities for the different measurements.

| | ICC | 95% confidence interval |
|---|---|---|
| **HKA** | | |
| Observer 1 | 0.990* | 0.986-0.993 |
| Observer 2 | 0.977* | 0.968-0.984 |
| Observer 3 | 0.996* | 0.994-0.998 |
| Inter-observer reliability (1, 2, & 3) | 0.982* | 0.973-0.988 |
| **mMPTA** | | |
| Observer 1 | 0.974* | 0.964-0.982 |
| Observer 2 | 0.903* | 0.864-0.932 |
| Observer 3 | 0.945* | 0.922-0.961 |
| Inter-observer reliability (1, 2, & 3) | 0.906* | 0.861-0.936 |
| **mLDFA** | | |
| Observer 1 | 0.912* | 0.874-0.939 |
| Observer 2 | 0.850* | 0.791-0.893 |
| Observer 3 | 0.925* | 0.895-0.947 |
| Inter-observer reliability (1, 2, & 3) | 0.871* | 0.824-0.907 |
| **JLCA** | | |
| Observer 1 | 0.629* | 0.507-0.726 |
| Observer 2 | 0.463* | 0.300-0.599 |
| Observer 3 | 0.676* | 0.565-0.763 |
| Inter-observer reliability (1, 2, & 3) | 0.507* | 0.395-0.611 |

Inter-observer and intra-observer reliabilities with significant (*) p-values when below 0.05.

**Table 2** Calculated Test-Retest errors in mean absolute degrees and 95% Confidence Interval (CI) for the HKA, mMPTA, mLDFA, and JLCA.

| | Mean absolute error | 95%- CI | ICC | 95%-CI |
|---|---|---|---|---|
| HKA | 0.442° | 0.387°-0.498° | 0.985* | 0.980-0.989 |
| mMPTA | 0.783° | 0.683°-0.878° | 0.922* | 0.896-0.941 |
| mLDFA | 0.828° | 0.722°-0.944° | 0.903* | 0.871-0.927 |
| JLCA | 0.794° | 0.683°-0.911° | 0.632* | 0.534-0.712 |

Intraclass Correlations between the measurements performed on the test and retest radiographs, with significant (*) p-values below 0.05

**Figure 5** illustrates the Bland-Altman test-retest analyses of the HKA, mMPTA, mLDFA, and JLCA, measured on the first and second WLR. With no significant systemic biases. A mean error between the HKA measurements on two separate WLRs of 0.01°, and the 95% limits of agreement between 1.15° and -1.13°.

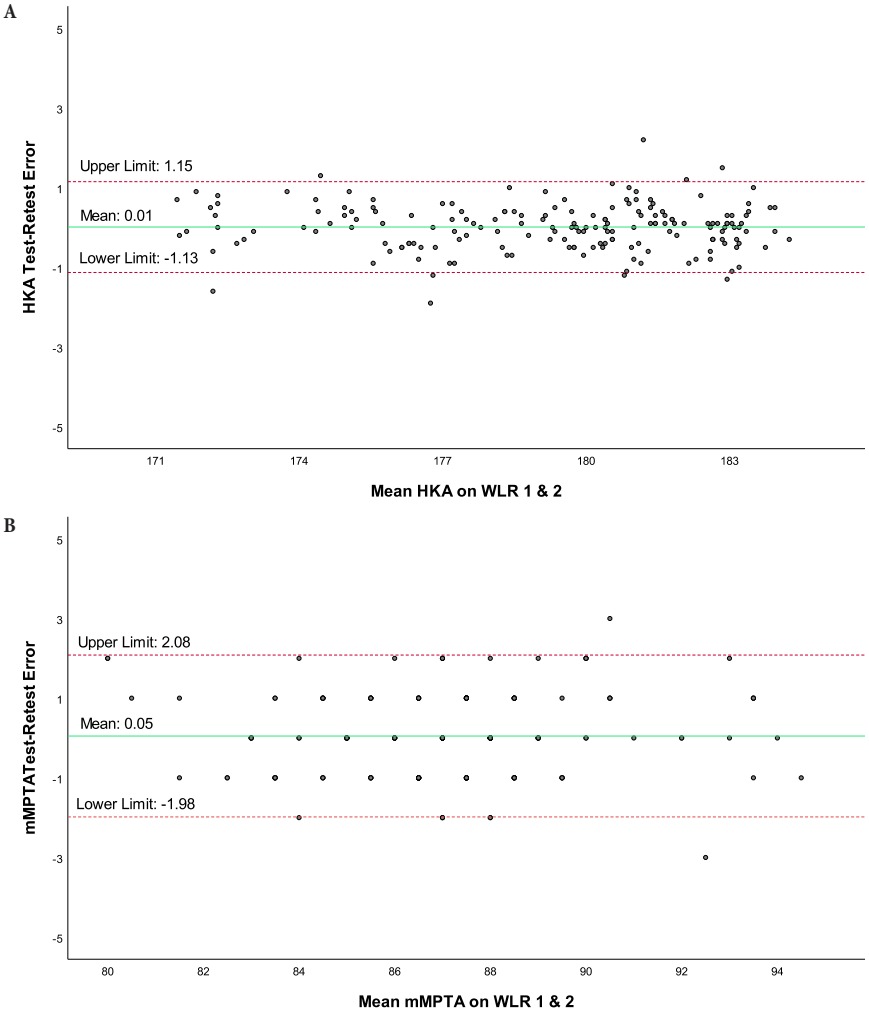



**Figure 5.** Bland Altman analyses of the Test-Retest results of the measured HKA (**A**), mMPTA (**B**), mLDFA (**C**) and JLCA (**D**) on two WLRs made at two different time points.

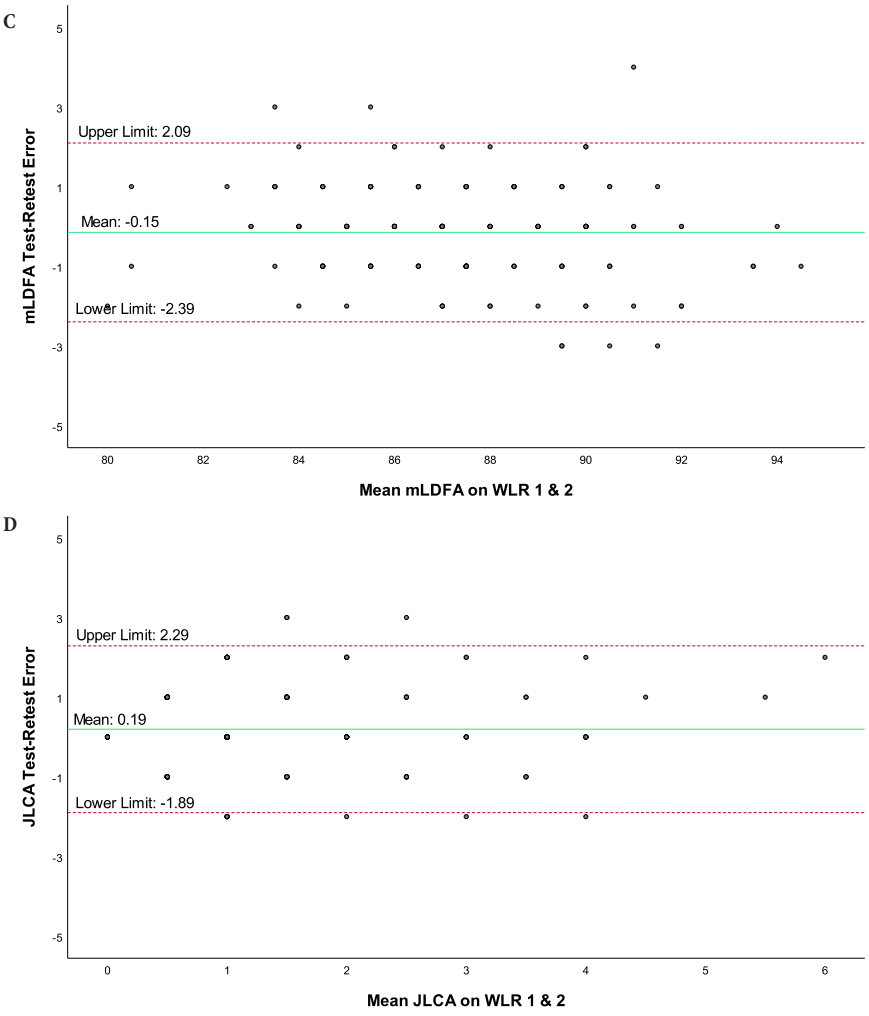



**Figure 5.** Continued

## Discussion

This study demonstrated successful introduction of an easy-to-use protocol for obtaining WLRs in a real-world clinical setting. The objective of the present study was to examine the test-retest reliability of our WLR positioning protocol. We observed excellent ICCs for HKA, mMPTA, and mLDFA measured on the first and second WLRs in this test-retest study. A moderate ICC was observed for JLCA. The absolute errors of HKA, mMPTA, mLDFA, and JLCA between the first and second WLR were below the inter-observer variabilities.

To the best of our knowledge, only one small study with eight participants has been published with the aim to optimize test-retest reproducibility of WLRs. Odenbring et al. used a protocol, positioning the patient on one leg and 10 degrees of knee flexion. Making sure that the knee was straightforward by superimposing the dorsal aspects of the femoral condyles, using a lateral fluoroscopic control.[22] This resulted in a test-retest mean absolute error of 1.3°, which is clinically inadequate, particularly given that many osteotomies are performed for deformities of 5° - 10°.[10,22] In contrast, our protocol, which is clinically readily employable, achieves a mean reliability of 0.442°, which is well within the clinically desired accuracy for a HTO.[10]

Jones et al. performed an in-silico study with the objective of describing the ideal accuracy for HTO. This was of clinical relevance given that it reported the desired measurement accuracies for lower limb geometry used in this study. Jones et al. found an ideal accuracy of 0.45° in order to achieve sufficient target corrections at the time of tibial osteotomy. Their findings support that the reliability of our WLR positioning protocol is a clinically sufficient and an employable solution for the current clinical need for accurate and reproducible WLRs. Of note, our WLR positioning protocol was developed with the aim on user-friendliness, while achieving clinically sufficient reproducibility. The protocol is inexpensive to implement and easy to perform by the X-ray technicians.

A surgical accuracy of 0.45° can be achieved when using patient specific instrumentation (PSI) intraoperatively, however this requires an additional CT-scan and associated radiation.[29,30] Further analyses of our results demonstrated that in 59% of cases the test-retest error was ≤0.45° and 95% of cases had test-retest error ≤1.15°. Of note, current standard HTO treatment for varus deformity and unicompartimental OA is performed without 3D analyses or intraoperative use of PSI, with a resulting accuracy of ±3°, as reported by Van den Bempt et al.[20]. This supports that the measurement errors of the proposed WLR protocol

are substantially smaller than the subsequent technical surgical accuracy of conventional HTO, highlighting the clinical utility of the protocol.

In order to maximise user-friendliness, some positional parameters of the patients were not controlled by the investigated protocol. Weight-bearing was based on patient perception of equal weight distribution on both legs. This can change between pre- and post-operative WLRs, with possible post-operative pain up to a year present after osteotomy in a unilateral fashion[31]. A force plate under our feet template could be a viable solution to this phenomenon, assisting in assessing whether the ground reaction force is equal and at the centre of the body.[32]



However, little is known about the amount of changing weight distribution following knee surgery and associated recovery. Also, for first implementation of our WLR protocol, this means added costs and implementation of complicated systems at the time of otherwise standard and established care in the radiology department. A force plate would heavily impact the user-friendliness and time efficiency of the protocol. Our suggestion is to obtain post-operative WLRs 4 months after surgery, when most patients are able to endure full weight-bearing on the operated limb during the radiograph moment.

Knee flexion was also only indirectly controlled in our protocol by asking patients to stand in full knee extension. Previous studies have demonstrated that a substantial degree of leg rotation of 10° or more is needed to alter HKA measurements in a clinically relevant fashion.[10,16,33] Where isolated knee flexion did not induce clinically relevant measurement errors.[14,33] Our protocol therefore focuses on obtaining WLRs with standardized and controlled leg rotation.[33]

However, up to one-third of knee OA patients present knee flexion contracture of 5° or higher and are unable to fully extend the knee[34]. While knee flexion of only 5° combined with 10° leg rotation can induce substantial measurement errors of 1° or more, which pleas for controlling leg rotation during WLR acquisition.[33,35] Of note, controlling leg rotation using a fixed floor template is not suitable for every patient. Patients with a high degree of rotational deformity in the femur and/or tibia are not eligible for the WLR positioning protocol presented.

Indications for rotational errors on a AP knee radiograph are no (or too much) tibia-fibular overlap and femoral condyle asymmetry.[36] When patients do show rotational deformities above 10°, the surgeon should consider making a new WLR with feet in a different position, compensating for the rotation deformity in the lower limb. Additionally, we recommend to consider computed tomography (CT) to assess lower limb rotation in such cases.

Our results show a moderate test-retest ICC for JLCA. At the same time, the mean absolute error in degrees (0.794°) is comparable or lower than the test-retest error of mMPTA (0.783°) and mLDFA (0.828°). The moderate ICC for JLCA can be explained by the normal distribution of the JLCA, which is between 0° and 2°.[37] This is narrower than the distribution of mMPTA and mLDFA suggesting that the absolute error of 0.794° is relatively higher than that of mMPTA and mLDFA given its narrowed underlying distribution. Furthermore, the measured angles of JLCA in the radiologic program used are provided in round numbers, creating an even narrower distribution for subsequent mean absolute error calculations.

This study has some limitations. First, we used whole number measurements of mMPTA, mLDFA, and JLCA as OrthoStation provides these values in whole numbers. Therefore, it was not possible to analyse these parameters within one decimal place of accuracy. Nevertheless, given our excellent test-retest values, we believe we have obtained a clinically very reliable method for making WLRs. Second, we did not perform rotational analyses of the lower limbs. This could be valuable in order to study the effects of rotational deformities in the lower limb on the measured geometry. An implied limitation of employing single view, standardized AP radiographs is the lack of 3-dimensional information regarding the tibial and femoral rotation and knee flexion contracture.[33,35] Future studies should consider measuring the test-retest mMPTA, mLDFA, and JLCA at and beyond one decimal place accuracy, understanding that further accuracy may be mathematically significant but of clinically decreasing relevance. Additionally, future studies should consider adding 3D analyses for the measurement of tibial and femoral rotation.[33,35]

## Conclusion

The novel WLR positioning protocol investigated produced excellent and reproducible HKA measurements, with clinically acceptable degrees of error. We recommend applying this easy-to-use protocol when obtaining WLRs for osteotomy planning.

## Funding

This study was supported by a grant from the Dutch Arthritis Society (project number: 19-2-102).

## References


1. Allen KD, Golightlya YM. Epidemiology of osteoarthritis: state of the evidence. *Curr Opin Rheumatol*. 2015;27(3):276-283. doi:10.1097/BOR.0000000000000161.Epidemiology

2. Zhang Y, Jordan J. Epidemyology of Osteoarthritis. *Clin Geriatr Med*. 2010;26(3):355-369. doi:10.1016/j.cger.2010.03.001.Epidemiology

3. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. *Lancet*. 2019;393(10182):1745-1759. doi:10.1016/S0140-6736(19)30417-9

4. Palmer JS, Jones LD, Monk AP, et al. Varus alignment of the proximal tibia is associated with structural progression in early to moderate varus osteoarthritis of the knee. *Knee Surgery, Sport Traumatol Arthrosc*. 2020. doi:10.1007/s00167-019-05840-5

5. Issa SN, Dunlop D, Chang A, et al. Full-limb and knee radiography assessments of varus-valgus alignment and their relationship to osteoarthritis disease features by magnetic resonance imaging. *Arthritis Care Res*. 2007;57(3):398-406. doi:10.1002/art.22618

6. Stukenborg-Colsman C, Wirth CJ, Lazovic D, Wefer A. High tibial osteotomy versus unicompartmental joint replacement in unicompartmental knee joint osteoarthritis: 7-10-Year follow-up prospective randomised study. *Knee*. 2001;8(3):187-194. doi:10.1016/S0968-0160(01)00097-7

7. Salmon JH, Rat AC, Sellam J, et al. Economic impact of lower-limb osteoarthritis worldwide : a systematic review of cost-of-illness studies. *Osteoarthr Cartil*. 2020;24(9):1500-1508. doi:10.1016/j.joca.2016.03.012

8. Bayliss LE, Culliford D, Monk AP, et al. The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *Lancet*. 2017;389(10077):1424-1430. doi:10.1016/S0140-6736(17)30059-4

9. Lee YS, Kim MG, Byun HW, Kim SB, Kim JG. Reliability of the imaging software in the preoperative planning of the open-wedge high tibial osteotomy. 2015:846-851. doi:10.1007/s00167-013-2700-z

10. Jones LD, Brown CP, Jackson W, Monk AP, Price AJ. Assessing accuracy requirements in high tibial osteotomy: a theoretical, computer-based model using AP radiographs. *Knee Surgery, Sport Traumatol Arthrosc*. 2017;25(9):2952-2956. doi:10.1007/s00167-016-4092-3

11. Sheehy L, Cooke TD V. Radiographic assessment of leg alignment and grading of knee osteoarthritis: A critical review. *World J Rheumatol*. 2015;5(2):69-81. doi:10.5499/wjr.v5.i2.69

12. Cooke TD V, Sled EA. Optimizing limb position for measuring knee anatomical axis alignment from standing knee radiographs. *J Rheumatol*. 2009;36(3):472-477. doi:10.3899/jrheum.080732

13. Sanfridsson J, Ryd L, Svahn G, Fridén T, Jonsson K. Radiographic Measurement of Femorotibial Rotation in Weight-Bearing. *Acta radiol*. 2001;42(2):207-217. doi:10.1080/028418501127346512

14. Brouwer RW, Jakma TSC, Brouwer KH, Verhaar JAN. Pitfalls in determining knee alignment: a radographic cadaver study. *J Knee Surg*. 2007;20(3):210–215.

15. Specogna A V., Birmingham TB, Hunt MA, et al. Radiographic measures of knee alignment in patients with varus gonarthrosis: Effect of weightbearing status and associations with dynamic joint load. *Am J Sports Med*. 2007;35(1):65-70. doi:10.1177/0363546506293024

16. Radtke K, Becher C, Noll Y, Ostermeier S. Effect of limb rotation on radiographic alignment in total knee arthroplasties. *Arch Orthop Trauma Surg*. 2010;130(4):451-457. doi:10.1007/s00402-009-0999-1

17. Lee YS, Lee BK, Lee SH, Park HG, Jun DS, Moon DH. Effect of foot rotation on the mechanical axis and correlation between knee and whole leg radiographs. *Knee Surgery, Sport Traumatol Arthrosc*. 2013;21(11):2542-2547. doi:10.1007/s00167-013-2419-x

18. Cho YM, Ko YJ, Lee W. Relationships among foot position , lower limb alignment , and knee adduction moment in patients with degenerative knee osteoarthritis. *J Phys Ther Sci*. 2015;27(1):1-4.

19. Ollivier B, Berger P, Depuydt C, Vandenneucker H. Good long - term survival and patient - reported outcomes after high tibial osteotomy for medial compartment osteoarthritis. *Knee Surgery, Sport Traumatol Arthrosc*. 2020. doi:10.1007/s00167-020-06262-4

20. Van den Bempt M, Van Genechten W, Claes T, Claes S. How accurately does high tibial osteotomy correct the mechanical axis of an arthritic varus knee? A systematic review. *Knee*. 2016;23(6):925-935. doi:10.1016/j.knee.2016.10.001

21. Gao L, Madry H, Chugaev D V., et al. Advances in modern osteotomies around the knee: Report on the Association of Sports Traumatology, Arthroscopy, Orthopaedic surgery, Rehabilitation (ASTAOR) Moscow International Osteotomy Congress 2017. *J Exp Orthop*. 2019;6(1):1-14. doi:10.1186/s40634-019-0177-5

22. Odenbring S, Berggren AM, Peil L. Roentgenographic assessment of the hip-knee-ankle axis in medial gonarthrosis. A study of reproducibility. *Clin Orthop Relat Res*. 1993;289:195-196.

23. Korpershoek J V., Vonk LA, Kester EC, et al. Efficacy of one-stage cartilage repair using allogeneic mesenchymal stromal cells and autologous chondron transplantation (IMPACT) compared to nonsurgical treatment for focal articular cartilage lesions of the knee: Study protocol for a crossover randomiz. *Trials*. 2020;21(1):1-11. doi:10.1186/s13063-020-04771-8

24. Jansen MP, Mastbergen SC, van Heerwaarden RJ, et al. Knee joint distraction in regular care for treatment of knee osteoarthritis: A comparison with clinical trial data. *PLoS One*. 2020;15(1):1-15. doi:10.1371/journal.pone.0227975

25. van Egmond N, van Grinsven S, van Loon CJM, Gaasbeek RD, van Kampen A. Better clinical results after closed- compared to open-wedge high tibial osteotomy in patients with medial knee osteoarthritis and varus leg alignment. *Knee Surgery, Sport Traumatol Arthrosc*. 2016;24(1):34-41. doi:10.1007/s00167-014-3303-z

26. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64(1):96-106. doi:10.1016/j.jclinepi.2010.03.002

27. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Arch Orofascial Sci*. 2017;12(1):1-11.

28. Cocchetti D V. Sample size requirements for increasing the precision of reliability estimates: Problems and proposed solutions. *J Clin Exp Neuropsychol*. 1999;21(4):567-570. doi:10.1076/jcen.21.4.567.886

29. Jacquet C, Chan-Yu-Kin J, Sharma A, Argenson JN, Parratte S, Ollivier M. "More accurate correction using "patient-specific" cutting guides in opening wedge distal femur varization osteotomies. *Int Orthop*. 2019;43(10):2285-2291. doi:10.1007/s00264-018-4207-1

30. Victor J, Premanathan A. Virtual 3D planning and patient specific surgical guides for osteotomies around the knee: a feasibility and proof-of-concept study. *Bone Joint J*. 2013;95-B(11):153-158. doi:10.1302/0301-620X.95B11.32950

31. Jansen MP, Besselink NJ, van Heerwaarden RJ, et al. Knee Joint Distraction Compared with High Tibial Osteotomy and Total Knee Arthroplasty: Two-Year Clinical, Radiographic, and Biochemical Marker Outcomes of Two Randomized Controlled Trials. *Cartilage*. 2019;100. doi:10.1177/1947603519828432

32. Struzik A, Zawadzki J. Leg stiffness during phases of countermovement and take-off in vertical jump. *Acta Bioeng Biomech*. 2013;15(2):113-118. doi:10.5277/abb130213

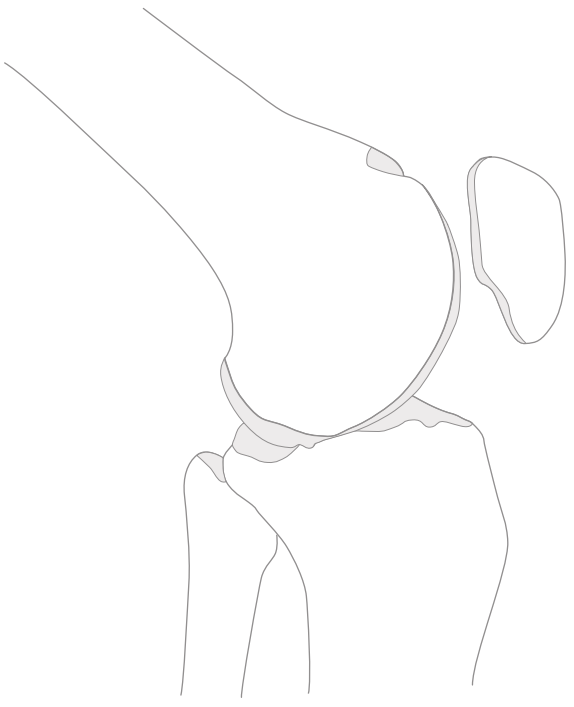
33. Nguyen HC, Gielis WP, van Egmond N, et al. The Need for a Standardized Whole Leg Radiograph Protocol: the Effects of Knee Flexion, Leg Rotation, and X-ray Beam Height. *J Orthop Res*. 2020;(under review).

34. Ritter MA, Lutgring JD, Davis KE, Berend ME, Pierson JL, Meneghini RM. The Role of Flexion Contracture on Outcomes in Primary Total Knee Arthroplasty. *J Arthroplasty*. 2007;22(8):1092-1096. doi:10.1016/j.arth.2006.11.009

35. Jud L, Trache T, Tondelli T, Fürnstahl P, Fucentese SF, Vlachopoulos L. Rotation or flexion alters mechanical leg axis measurements comparably in patients with different coronal alignment. *Knee Surgery, Sport Traumatol Arthrosc*. 2019;(0123456789). doi:10.1007/s00167-019-05779-7

36. Maderbacher G, Schaumburger J, Baier C, et al. Predicting knee rotation by the projection overlap of the proximal fibula and tibia in long-leg radiographs. *Knee Surgery, Sport Traumatol Arthrosc*. 2014;22(12):2982-2988. doi:10.1007/s00167-014-3327-4

37. Paley D. Normal Lower Limb Alignment and Joint Orientation. In: *Principles of Deformity Correction*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002:1-18.



# CHAPTER 4

# Predicting the mechanical Hip-Knee-Ankle angle accurately from standard knee radiographs, a cross-validation experiment in 100 patients

*Acta orthopaedica,*
*June 2020, doi.org/10.1080/17453674.2020.1779516*

Willem Paul Gielis, Hassan Rayegan, Vahid Arbabi,
Seyed Y. Ahmadi Brooghani, Claudia Lindner , Tim F. Cootes,
Pim A. de Jong, Harrie Weinans, Roel J.H. Custers

## Abstract

### Background and aim

Being able to predict the hip-knee-ankle angle (HKAA) from standard knee radiographs allows studies on malalignment in cohorts lacking full-limb radiography. We aimed to develop an automated image analysis pipeline to measure the femoro-tibial angle (FTA) from standard knee radiographs and test various FTA definitions to predict the HKAA.

### Patients and methods

We included 110 pairs of standard knee and full-limb radiographs. Automatic search algorithms found anatomic landmarks on standard knee radiographs. Based on these landmarks, the FTA was automatically calculated according to 9 different definitions (6 described in the literature and 3 newly developed). Pearson and intra-class correlation coefficient (ICC)) were determined between the FTA and HKAA as measured on full-limb radiographs. Subsequently, the top 4 FTA definitions were used to predict the HKAA in a 5-fold cross-validation setting.

### Results

Across all pairs of images, the Pearson correlations between FTA and HKAA ranged between 0.83 and 0.90. The ICC values from 0.83 to 0.90. In the cross-validation experiments to predict the HKAA, these values only decreased minimally. The mean absolute error for the best method to predict the HKAA from standard knee radiographs was 1.8° (SD 1.3°).

### Conclusion

We showed that the HKAA can be automatically predicted from standard knee radiographs with a fair accuracy and high correlation compared to the true HKAA. Therefore, this method enables research of the relationship between malalignment and knee pathology in large (epidemiological) studies lacking full-limb radiography.

## Introduction

The mechanical axis of the lower limb, which determines the knee (mal)alignment, is historically measured using the hip-knee-ankle-angle (HKAA), an angle between the mechanical axes of the femur and tibia. The femoral axis runs through the centers of the femoral head and knee joint. The tibial axis runs through the centers of the knee and the ankle joints (**Figure 1**).

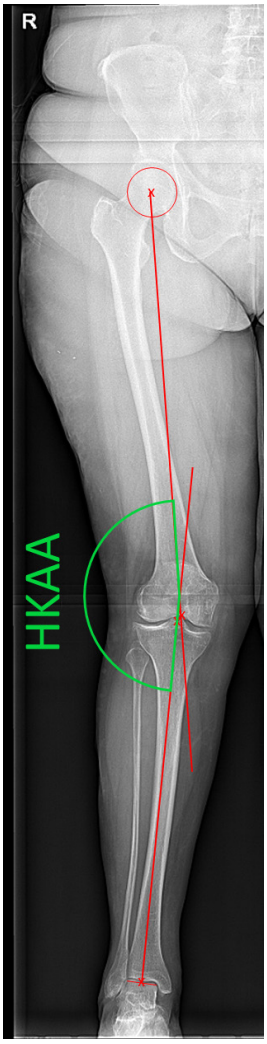



**Figure 1.** Measurement of the Hip-Knee-Ankle angle on full limb radiograph

The Hip-Knee-Ankle angle (HKAA, in green) is measured between two axes (in red). One axis runs from the middle of the femoral head to the middle of the femoral notch, and a second axis from the middle of the tibial notch, to the middle of the talar head.

A standard knee radiograph is one of the primary tools in the diagnostic process of knee complaints and it is made for a majority of patients. Correspondingly, many epidemiological studies focusing on the knee include standard knee radiographs. However, to verify and measure involvement of malalignment in the pathophysiology, the HKAA should be measured. This requires a full-limb radiograph (**Figure 1**). Compared to a standard knee radiograph, full-limb radiography involves higher costs, the need of specialized equipment and a larger effective radiation dose for the patient. These are important reasons for knee OA cohort studies to not include full-limb radiographs. As standard knee radiographs are available for the majority of patients with knee complaints and participants of epidemiological knee focused studies, it is desirable to have a method for defining knee (mal)alignment from a standard knee radiograph.

The femoral-tibial angle (FTA), an angle between the anatomic axes of the femur and tibia (**Figure 2**), can be used to predict the mechanical axis from a standard knee radiograph. The FTA, is an important measurement that can predict the development of knee OA.[1,2] Multiple definitions for the FTA have been proposed (**Table 1**).[2–12] However, a direct comparison between all FTA definitions on the same data is lacking. Additionally, no studies used cross- or external validation to confirm results. As such, there is no consensus on which FTA definition should be used to predict the HKAA.

In addition to morphological measurements, statistical shape modelling is used in OA research to quantify variation in joint shape. Multiple studies showed the shape of a joint is a major factor in the incidence and progression of OA.[13–15] A key step in the statistical shape modelling process is to outline the structures of interest in the medical images (e.g. radiographs) using anatomical landmarks. Manually placed landmarks on a set of radiographs can be used to train automated search models to automatically place the respective points on new unseen images, paving the road to analyzing large datasets. [16,17] Furthermore, the landmark positions obtained by the search models can easily be used to calculate morphological measurement such as joint space width or the FTA.

The ability to predict the HKAA using automated FTA measurements from standard knee radiographs, would make studies on malalignment feasible in large cohorts that lack full-limb radiography. This study aimed (i) to develop an automated image analysis pipeline to measure the FTA from a standard knee radiograph; and (ii) to analyze the performance of various FTA definitions in predicting the HKAA as measured on a full-limb radiograph.

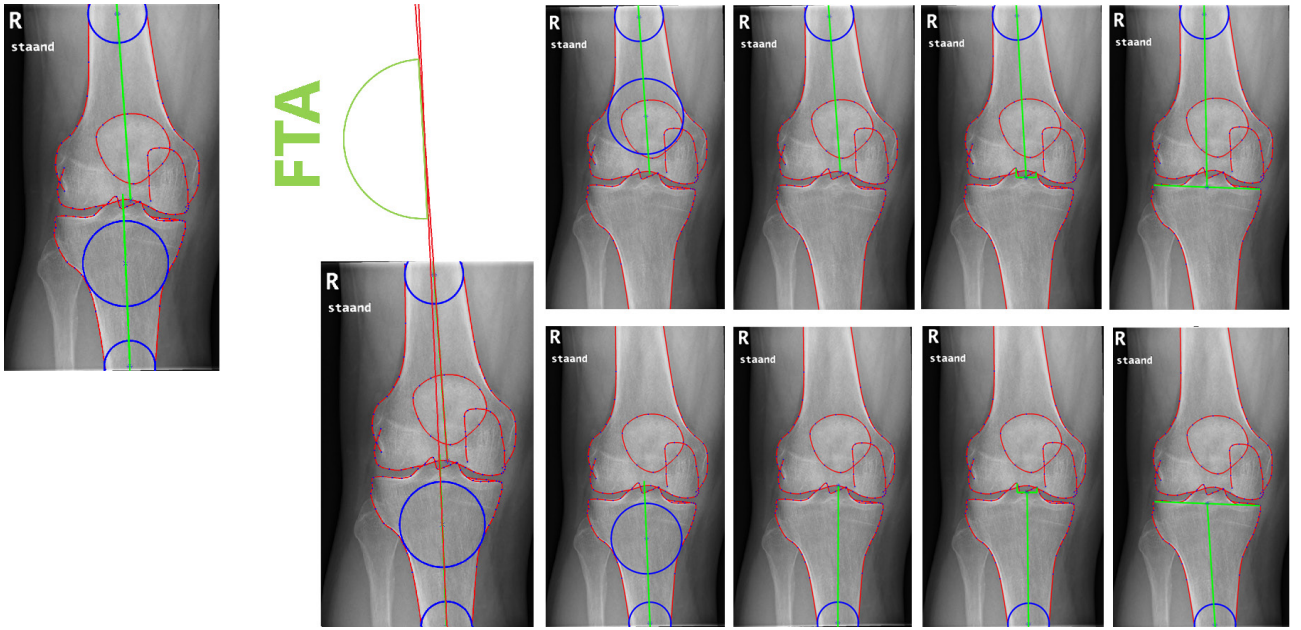



**Figure 2.** Measurement of the Femoral-Tibial-Angle on standard knee radiographs
Definitions of the femoral axis from left to right (**top**): **Fem1** Mid-shaft at approximately 10 cm proximal of the femoral notch + mid-shaft in the area where the meta- and epiphysis meet [7,11,18] and the amount of correction required are essential when planning limb realignment surgery. The hip-knee-ankle angle (HKA; **Fem2** Mid-shaft at approximately 10 cm proximal of the femoral notch + center of the femoral notch [5,9,11,12] cost effectiveness and minimal radiation exposure. Our goal was to investigate the performance metrics of currently used methods of FTA measurement to determine whether a specific protocol could be recommended based on these results. METHODS Inter- and intra-rater reliability of FTA measurements were determined by intraclass correlation coefficient (ICC; **Fem3** Mid-shaft at approximately 10 cm proximal of the femoral notch + base of the tibial spines [3–7,10,18]and the amount of correction required are essential when planning limb realignment surgery. The hip-knee-ankle angle (HKA; **Fem4** Mid-shaft at approximately 10 cm proximal of the femoral notch + middle of tibial plateau 5. Definitions of the tibial axis from left to right (**bottom**): **Tib1** Mid-shaft at approximately 10 cm distal of the base of the tibial spines + mid-shaft in the area where the meta- and epiphysis meet [7,11,18] and the amount of correction required are essential when planning limb realignment surgery. The hip-knee-ankle angle (HKA; **Tib2** Mid-shaft at approximately 10 cm distal of the base of the tibial spines + center of the femoral notch 5; **Tib3** Mid-shaft at approximately 10 cm distal of the base of the tibial spines + base of the tibial spines [3–8,10–12,18] and the amount of correction required are essential when planning limb realignment surgery. The hip-knee-ankle angle (HKA; **Tib4** Mid-shaft at approximately 10 cm distal of the base of the tibial spines + middle of tibial plateau 5. The 2 pictures on the left show the measurement of the FTA using method 2 for the femoral axis and method 1 for the tibial axis on a standard AP knee radiograph from the present data set.

**Table 1.** Correlations between femoral-tibial angle and hip-knee-ankle angle as reported in the literature

| Landmarks | Radiograph | Pearson Correlation | Authors | Year |
|---|---|---|---|---|
| Fem3 + Tib3 | AP extended | 0.26 | Zampogna et al. | 2015 |
| Fem1 + Tib1 | AP extended | 0.71 | | |
| Fem3 + Tib3 | AP extended | 0.81 | Colebatch et al. | 2009 |
| Fem2 + Tib2 | PA semiflexed * | 0.50 | Mc Daniel et al. | 2010 |
| Fem3 + Tib3 | PA semiflexed * | 0.65 | | |
| Fem4 + Tib4 | PA semiflexed * | 0.55 | | |
| Fem3 + Tib3[x] | PA semiflexed * | 0.64 | | |
| Fem2 + Tib3 | PA semiflexed * | 0.59 | | |
| Fem3 + Tib3 | PA semiflexed * | 0.86 | Issa et al. | 2007 |
| Fem2 + Tib3 | PA semiflexed * | 0.66 | Felson et al. | 2009 |
| Fem1 + Tib1[#] | PA semiflexed * | 0.76 | Iranpour-Boroujeni et al. | 2014 |
| Fem3 + Tib3 | PA semiflexed * | 0.68 | | |
| Fem3 + Tib3 | PA semiflexed * | 0.75 | Kraus et al. | 2005 |
| Fem3 + Tib3 | Full limb radiograph | 0.65 | | |
| Fem3 + Tib3 | Full limb radiograph | 0.88 | Hinman | 2006 |
| Fem2 + Tib3 | Full limb radiograph | 0.34 | van Raaij et al. | 2009 |
| Fem1 + Tib1 | Full limb radiograph | 0.65 | | |
| Fem2 + Tib3 | Full limb radiograph | 0.88 | Sheehy et al. | 2012 |
| Fem3 + Tib3 | Full limb radiograph | 0.93 | Navali et al. | 2012 |

*Positioning aided with Synaflexer frame
[x]Slight variation where the tips of the tibial spines are used instead of the base.
[#]Slight variation where the bottom point at the femur is determined using the middle femoral condyls instead of the shaft

## Methods

### Patients

We included 100 full-limb (50 males) radiographs, acquired for clinical care at the department of Orthopaedic Surgery of the UMC Utrecht, the Netherlands, in a consecutive series between March and November 2017. All patients were 40 years or older. For inclusion at least 1 standard knee radiograph made on the same day had to be available. We excluded patients with femoral and tibial deformities due to fractures, surgeries (including joint replacement of knee, hip or ankle and osteotomies) and developmental disorders. When radiographs of both legs were available for 1 subject, both were included in the study.

### Radiographic acquisition

Weight-bearing extended anteroposterior full-limb radiographs, with the patella facing straight towards the X-ray tube were made. On the same day, weight-bearing extended anteroposterior knee radiographs with the patella facing forward were made. Standard knee radiographs were assessed by a medical researcher (WPG) for Kellgren-Lawrence (KL) grades. Before the assessment, the rater completed the tutorial for KL grading by Hayes et al..[19] The tutorial includes 19 cases and an answer sheet to test the effect of the tutorial. The square weighted kappa for inter-rater reliability between WPG and the answer sheet was 0.969.

### Alignment measurements

The mechanical HKAA was used as gold standard and was measured on the full-limb radiographs [20]. An axis was drawn from the middle of the femoral head to the center of the femoral notch. A second axis was drawn from the base of the tibial spines to the center of the ankle joint. The HKAA was defined as the angle between these axes (**Figure 1**).[21]

The FTA was measured on standard knee radiographs as the angle between the axis of the femur and tibia. We used a bespoke search model in BoneFinder® (www.bone-finder.com, Centre for Imaging Sciences, The University of Manchester, UK) to automatically outline the distal femur, patella and proximal tibia using 111 landmarks.[17] All automatically obtained landmarks were checked and manually corrected if needed. The identified landmarks were used to automatically calculate the FTA. For measuring the femoral and tibial axes, four definitions each were considered based on previous literature (**Figure 2**).[2–11] Nine combinations between the femoral and tibial axes measurements were used to calculate the FTA (Fem1Tib1, Fem1Tib3, Fem1Tib4, Fem2Tib1, Fem2Tib2, Fem2Tib3, Fem2Tib4, Fem3Tib3, Fem4Tib4). A varus angle is displayed as a negative number, a valgus angle as a positive number. As the standard knee radiographs were not calibrated for absolute distance, we used the width of the femoral condyles and tibial plateau to place circles needed for FTA measurements at approximately 10 cm from the joint line (**Figure 2**). Based on data from previous work we used a ratio of 1.52 for the femur and 1.42 for the tibia. [22] We used the center of a circle touching the medial and lateral cortex to determine the mid-shaft points.

### Statistical analysis

We used Pearson correlation coefficients to study which FTA method has the strongest correlation with the HKAA. Across all image pairs, we predicted the

HKAA from the FTA using linear regression models. A simple model using only one FTA predictor, a model including a quadratic term and a model including sex were considered. Using these predictions, we calculated two-way mixed single measures intra-class correlation coefficients for absolute agreement (ICC) between predicted HKAA and observed HKAA.

For the four FTA definitions with the strongest correlation to the HKAA, we performed 5-fold cross-validation experiments on the same data set. We randomly distributed the dataset in five parts, each accounting for 20% of the cases. In each fold, we calculated a linear regression formula to predict the HKAA based on the FTA using 80% of the data and used it to predict the HKAA in the remaining 20% of the cases. We repeated this process 5 times, so each case will have a predicted HKAA. For this dataset, we present the Pearson correlation and ICC between the predicted HKAA and gold standard. In addition, we present a Bland-Altman plot displaying absolute measurement errors of predicted HKAA vs the gold standard in our cross-validation experiments. As no similar experiment was published previously, a valid sample size calculation was not possible and we applied the minimum of 100 cases as suggested by Vergouwe et al.. [23] We chose to include 50 males and 50 females, to account for sex-specific differences.

### Ethics, funding, and potential conflicts of interest

All radiographs were anonymized, and a waiver of consent was obtained from the local medical ethical committee (no. 17-760/C). This work was supported by Reuma Nederland (LLP-22) and the APPROACH project. APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement n°115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme(FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu. C. Lindner and T.F. Cootes were funded by the Engineering and Physical Sciences Research Council, UK (EP/M012611/1) and by the Medical Research Council, UK (MR/S00405X/1). Drs. Cootes and Lindner have a patent US 9928443, EP 2893491 issued.

## Results

Of 100 full-limb radiographs, 11 had radiographs of both knees available and 89 had only 1 knee radiograph available. 1 knee radiograph was of insufficient quality to perform FTA measurements and was excluded, resulting in 110 full-limb / standard knee radiograph pairs. The mean age was 54 (SD 7.4) and 53 knees were

male. Of all knees, 9 were KL 0, 23 were KL 1, 30 were KL 2, 36 were KL3 and 12 were KL 4.

### Correlation between FTA and HKAA measurements across all pairs of images.

Across all pairs of images, the Pearson correlations between FTA and HKAA ranged between 0.83 and 0.90 (**Table 2**). The ICC values from 0.83 to 0.90. The best correlations between HKAA and FTA measurements were found using the FTA defined as a femoral axis between the mid-shaft of the femur (approximately 10 cm above the joint line) and the femoral notch (Fem2), and a tibial axis running through 2 points in the mid-shaft of the tibia (approximately 4 cm and 10 cm beneath the tibial plateau (Tib1). Linear regression to predict the HKAA using the

**Table 2.** Pearson correlation coefficients and intra-class correlations (ICC) between FTA and HKAA measurements (across all pairs of images)

| Method | Pearson correlation | ICC | 95%-CI |
|---|---|---|---|
| Fem1Tib1 | 0.88 | 0.87 | (0.82 – 0.91) |
| Fem1Tib3 | 0.86 | 0.86 | (0.80 – 0.90) |
| Fem1Tib4 | 0.86 | 0.86 | (0.80 – 0.90) |
| Fem2Tib1 | 0.90 | 0.90 | (0.85 – 0.93) |
| Fem2Tib2 | 0.87 | 0.860 | (0.80 – 0.90) |
| Fem2Tib3 | 0.89 | 0.89 | (0.84 – 0.92) |
| Fem2Tib4 | 0.89 | 0.89 | (0.84 – 0.92) |
| Fem3Tib3 | 0.84 | 0.83 | (0.76 – 0.88) |
| Fem4Tib4 | 0.83 | 0.82 | (0.74 – 0.87) |

**Table 3.** Pearson correlation coefficients and intra-class correlations (ICC) between FTA and HKAA measurements (cross-validation experiments)

| Method | Pearson correlation | ICC | 95%-CI |
|---|---|---|---|
| Fem1Tib1 | 0.88 | 0.87 | (0.82 – 0.91) |
| Fem2Tib1 | 0.90 | 0.90 | (0.85 – 0.93) |
| Fem2Tib3 | 0.89 | 0.89 | (0.84 – 0.92) |
| Fem2Tib4 | 0.89 | 0.87 | (0.80 – 0.91) |

optimal FTA method (Fem2Tib1) produced the formula: HKAA = -2.182 + FTA *0.995. The mean absolute error between the predicted HKAA and the observed HKAA was 1.7° (SD 1.2°, range 0.1 – 5.4).

**Correlation between FTA and HKAA predictions in cross-validation experiments.**

The correlation statistics found in the cross-validation setting were comparable with those found across all pairs of images, albeit minimally weakened (**Table 3**). Again, the combination of femoral axis 2 and tibial axis 1 showed the best correlation (Pearson correlation 0.90, ICC 0.90). The mean absolute error between the predicted HKAA and the observed HKAA was 1.8° (SD 1.3°, range 0.1 – 5.3) in the cross-validation setting. The Bland-Altman plot depicts the error between the observed HKAA (gold standard) and the predicted HKAA in the cross-validation setting (**Figure 3**). No systematic errors or outliers were found in this plot. Although females were more probable to have a valgus alignment compared to males, the error between observed HKAA and predicted HKAA was similar between sexes

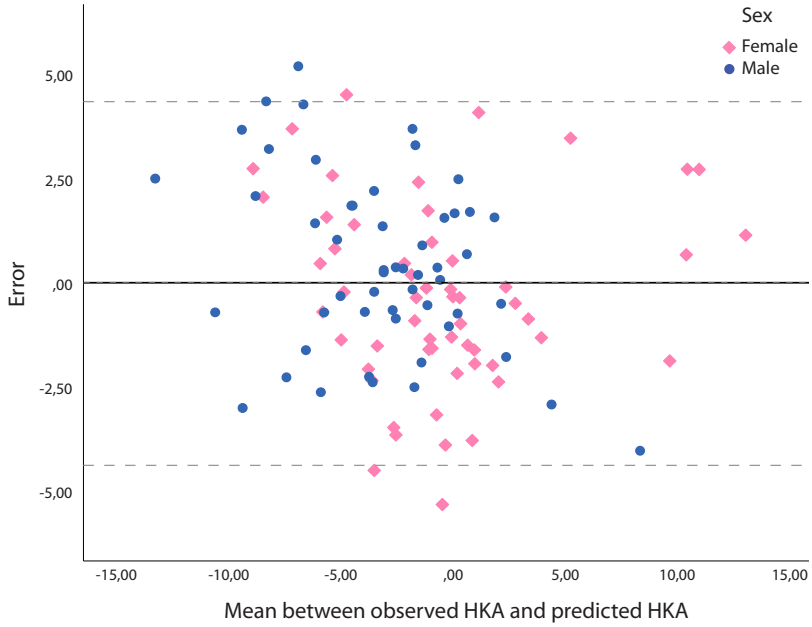



**Figure 3.** Bland-Altman plot depicting the error between the observed HKAA (gold standard) and the predicted HKAA in the cross-validation setting. Negative numbers represent the degree of varus alignment and positive numbers represent the degree of valgus alignment.

(p = 0.9). Linear regression models containing an interaction between FTA and sex, or a quadratic term, performed slightly better across all pairs of images, but performed slightly worse in the cross-validation experiments (data not shown).

## Discussion

This study showed that the mechanical HKAA can be predicted from a standard knee radiograph using our automated pipeline. We used several FTA definitions and compared their performance in predicting the HKAA. The best performing FTA definition used a femoral axis between the mid-shaft of the femur (approximately 10 cm above the joint line) and the femoral notch, and a tibial axis running through 2 points in the mid-shaft of the tibia (approximately 4cm and 10 cm beneath the tibial plateau). This combination to measure FTA had not been reported in the literature.

Compared to previous work, the Pearson correlation coefficient between FTA and HKAA measurements (0.83 to 0.90) was high across all pairs of images (**Tables 1**).[2–12] The results of the cross-validation can be used to estimate the performance of the HKAA predictions in new cases (**Table 2**). In the cross-validation the Pearson correlation between the automatically calculated FTA and the predicted HKAA was 0.90. The mean absolute error was 1.8° (SD 1.3°). To the best of our knowledge the performance of the predicted HKAA based on FTA has not been reported in any cross- or external validation studies.

The automatically calculated FTA provides an easy tool to study the influence of varus/valgus malalignment in OA cohorts or trials for which standard knee radiographs are available. We only tested the FTA produced by our automatic analysis pipeline to predict mechanical HKAA. However, the pipeline may be used to collect other measurements automatically and enables the rapid analysis of a collection of measurements for a large number of radiographs. The search algorithms we used were trained on only a small database (293 knees). Small corrections to the landmarks had to be made, costing approximately 1 minute per radiograph. In the future we expect the search model to have sufficient accuracy to run fully automatically without the need for manual correction. A database containing around 1000 knees should be sufficient to achieve this.[24]

While the standard AP radiograph is most often used in the clinics, numerous OA studies use a radiograph or use semi flexed PA radiographs. This technique aims to compensate for the tibial slope and give a more accurate read on the joint space width. The generalizability of our methods to PA radiographs should be

checked, and the formula to calculated the predicted HKAA may need adaption. Additionally, some of the FTA definition may be applied in knees with a prosthesis. Notably the definition using only landmarks in the femoral and tibial shafts. However, due to prosthesis placement, translation between the joint center and the femur and/or tibia or changes of the joint angle might occur. This has not been validated as we only studied native knees.

Clinically, a validated method to measure leg malalignment from standard knee radiographs would be very useful, as this would make a large proportion of full-limb radiographs unnecessary. A full-limb radiograph has several disadvantages, such as higher costs, more radiation, and the need for specialized equipment. However, it is important to question whether the mean observed error of 1.8° is of sufficient accuracy for clinical applications.[25] suggested that a 3-degree accuracy in measuring the mechanical HKAA is sufficient, as this resembles the precision of a correction osteotomy. To our knowledge the scan-rescan error for determining the HKAA using full-limb is only described in one study including 8 cases. The authors reported a mean error of 1.3°, but their measurements are rounded to the full degree.[26] found a correlation of 0.91 when comparing standard HKAA radiography to the novel QUESTOR method (using a specific positioning platform and software to perform and analyze the full limb radiography). A statistically significant mean difference of 0.7° in HKAA between double and single leg weight-bearing full-limb radiographs was reported by Yazdanpanah et al. but they did not report the mean absolute error.[27] More research is needed to investigate the scan-rescan reliability of the HKAA from full-limb radiographs.

Our study has a number of strengths. We used standardized clinical radiographs, with a protocol feasible in clinical care. We included an equal number of males and females. We tested a large number of FTA definitions using the same set of radiographs to directly compare their performance. Finally, we used a cross-fold validation experiment to test our predictions in unseen radiographs. The main limitation of our study is that the reliability of the gold standard (the HKAA) is poorly studied. However, the HKAA is the most commonly used measurement to determine the mechanical angle of the lower extremity in both research studies and clinical care.

## Conclusion

We have developed an automated image analysis pipeline to calculate the FTA from standard knee radiographs. We directly compared multiple FTA definitions and tested their performance in predicting the HKAA, as measured from full-limb radiographs. The best performing FTA definition correlated strongly with the HKAA and predicted it with high accuracy. The proposed image analysis pipeline can be used for epidemiological research on lower-limb alignment in cohorts with standard knee radiographs.

## References


1. Moyer R, Wirth W, Duryea J, Eckstein F. Anatomical alignment, but not goniometry, predicts femorotibial cartilage loss as well as mechanical alignment: data from the Osteoarthritis Initiative. Osteoarthr Cartil. 2016;24(2):254-261. doi:10.1016/j.joca.2015.08.016

2. Brouwer GM, Tol AW Van, Bergink AP, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. Arthritis Rheum. 2007;56(4):1204-1211. doi:10.1002/art.22515

3. Issa SN, Dunlop D, Chang A, et al. Full-limb and knee radiography assessments of varus-valgus alignment and their relationship to osteoarthritis disease features by magnetic resonance imaging. Arthritis Care Res. 2007;57(3):398-406. doi:10.1002/art.22618

4. Navali AM, Bahari LAS, Nazari B. A comparative assessment of alternatives to the full-leg radiograph for determining knee joint alignment. Sports Med Arthrosc Rehabil Ther Technol. 2012;4(1):40. doi:10.1186/1758-2555-4-40

5. McDaniel G, Mitchell KL, Charles C, Kraus VB. A comparison of five approaches to measurement of anatomic knee alignment from radiographs. Osteoarthr Cartil. 2010;18(2):273-277. doi:10.1016/j.joca.2009.10.005

6. Kraus VB, Vail TP, Worrell T, McDaniel G. A comparative assessment of alignment angle of the knee by radiographic and physical examination methods. Arthritis Rheum. 2005;52(6):1730-1735. doi:10.1002/art.21100

7. Zampogna B, Vasta S, Amendola A, et al. Assessing Lower Limb Alignment: Comparison of Standard Knee Xray vs Long Leg View. Iowa Orthop J. 2015;35:49-54. http://www.ncbi.nlm.nih.gov/pubmed/26361444. Accessed December 21, 2017.

8. Colebatch AN, Hart DJ, Zhai G, Williams FM, Spector TD, Arden NK. Effective measurement of knee alignment using AP knee radiographs. Knee. 2009;16(1):42-45. doi:10.1016/J.KNEE.2008.07.007

9. Felson DT, Cooke TD V, Niu J, et al. Can anatomic alignment measured from a knee radiograph substitute for mechanical alignment from full limb films? Osteoarthr Cartil. 2009;17(11):1448-1452. doi:10.1016/j.joca.2009.05.012

10. Hinman RS, May RL, Crossley KM. Is there an alternative to the full-leg radiograph for determining knee joint alignment in osteoarthritis? Arthritis Rheum. 2006;55(2):306-313. doi:10.1002/art.21836

11. van Raaij TM, Brouwer RW, Reijman M, Bierma-Zeinstra SMA, Verhaar JAN. Conventional knee films hamper accurate knee alignment determination in patients with varus osteoarthritis of the knee. Knee. 2009;16(2):109-111. doi:10.1016/J.KNEE.2008.10.003

12. Sheehy L, Felson D, Zhang Y, et al. Does measurement of the anatomic axis consistently predict hip-knee-ankle angle (HKA) for knee alignment studies in osteoarthritis? Analysis of long limb radiographs from the multicenter osteoarthritis (MOST) study. Osteoarthr Cartil. 2011;19(1):58-64. doi:10.1016/j.joca.2010.09.011

13. Agricola R, Leyland KM, Bierma-Zeinstra SMA, et al. Validation of statistical shape modelling to predict hip osteoarthritis in females: data from two prospective cohort studies (Cohort Hip and Cohort Knee and Chingford). Rheumatology. 2015;54(11):2033-2041. doi:10.1093/rheumatology/kev232

14. Haverkamp DJ, Schiphof D, Bierma-Zeinstra SM, Weinans H, Waarsing JH. Variation in joint shape of osteoarthritic knees. Arthritis Rheum. 2011;63(11):3401-3407. doi:10.1002/art.30575

15. Waarsing JH, Kloppenburg M, Slagboom PE, et al. Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. Arthritis Rheum. 2011;63(5):1349-1354. doi:10.1002/art.30288

16. Lindner C, Thiagarajah S, Wilkinson J, The arcOGEN Consortium, Wallis G, Cootes T. Fully Automatic Segmentation of the Proximal Femur Using Random Forest Regression Voting. IEEE Trans Med Imaging. 2013;32(8):1462-1472. doi:10.1109/TMI.2013.2258030

17. Lindner C, Thiagarajah S, Wilkinson JM, The acrOGEN Consortium, Wallis G a., Cootes TF. Development of a fully automatic shape model matching (FASMM) system to derive statistical shape models from radiographs: Application to the accurate capture and global representation of proximal femur shape. Osteoarthr Cartil. 2013;21(10):1537-1544. doi:10.1016/j.joca.2013.08.008

18. Iranpour-Boroujeni T, Li J, Lynch JA, Nevitt M, Duryea J. A new method to measure anatomic knee alignment for large studies of OA: data from the Osteoarthritis Initiative. Osteoarthr Cartil. 2014;22(10):1668-1674. doi:10.1016/J.JOCA.2014.06.011

19. Hayes B, Kittelson A, Loyd B, Wellsandt E, Flug J, Stevens-lapsley J. Assessing Radiographic Knee Osteoarthritis : An Online Training Tutorial for the Kellgren-Lawrence Grading Scale. Assoc Am Med Coll. 2016:18-21. doi:10.15766/mep_2374-8265.10503

20. Moreland JR, Bassett LW, Hanker GJ. Radiographic analysis of the axial alignment of the lower extremity. J Bone Jt Surg - Ser A. 1987;69(5):745-749. doi:10.2106/00004623-198769050-00016

21. Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The Role of Knee Alignment in Disease Progression and Functional Decline in Knee Osteoarthritis. JAMA. 2001;286(2):188. doi:10.1001/jama.286.2.188

22. Wesseling J, Boers M, Viergever MA, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. Int J Epidemiol. 2016;45(1):36-44. doi:10.1093/ije/dyu177

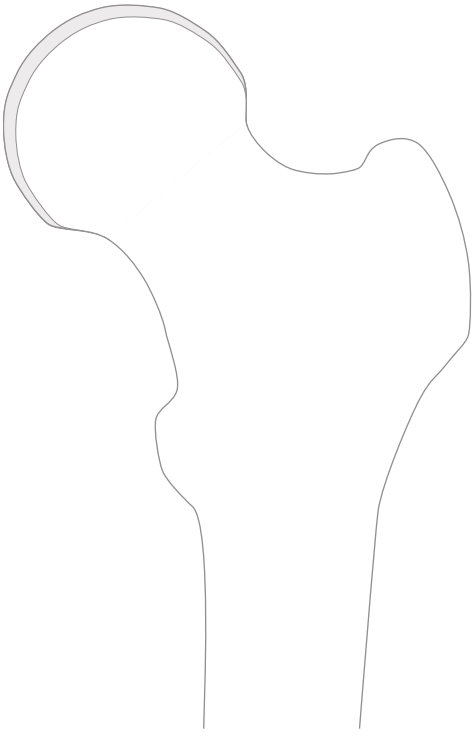
23. Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol. 2005;58(5):475-483. doi:10.1016/J.JCLINEPI.2004.06.017

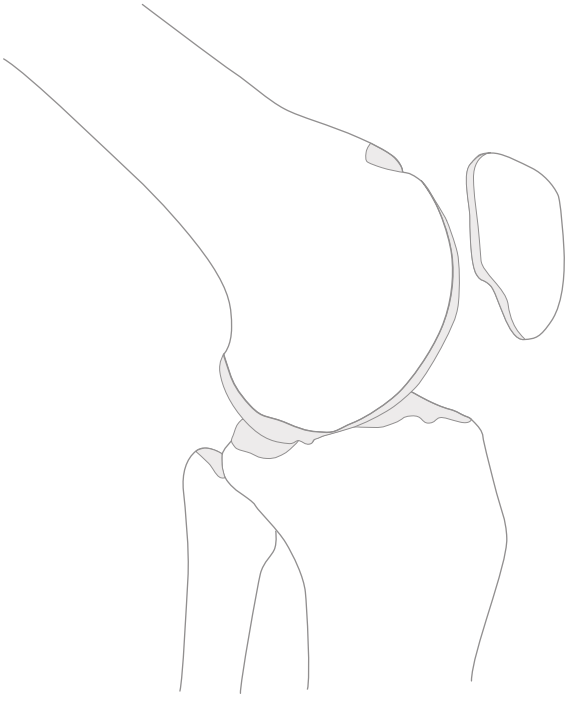
24. Gielis WP, Weinans H, Welsing PMJ, et al. An automated workflow based on hip shape improves personalized risk prediction for hip osteoarthritis in the CHECK study. Osteoarthr Cartil. 2019. doi:10.1016/j.joca.2019.09.005

25. Odenbring S, Berggren AM, Peil L. Roentgenographic assessment of the hip-knee-ankle axis in medial gonarthrosis. A study of reproducibility. Clin Orthop Relat Res. 1993;289:195-196.

26. Sanfridsson J, Ryd L, Eklund K, Kouvaras Y, Jonsson K. Angular configuration of the knee: Comparison of conventional measurements and the QUESTOR precision radiography system. Acta radiol. 1996;37(5):633-638. doi:10.3109/02841859609177689

27. Yazdanpanah O, Mobarakeh MK, Nakhaei M, Baneshi MR. Comparison of Double and Single Leg Weight-Bearing Radiography in Determining Knee Alignment. Arch bone Jt Surg. 2017;5(3):174-180. http://www.ncbi.nlm.nih.gov/pubmed/28656165. Accessed June 27, 2018.





# CHAPTER 5

# Scoring osteoarthritis reliably in large joints and the spine using whole-body CT: osteoarthritis computed tomography-score (OACT-Score)

*Journal of Personalized Medicine, Special issue*
*Submitted for publication*

Willem Paul Gielis, Harrie Weinans,
Frank J. Nap, Frank W. Roemer, Wouter Foppen

## Abstract

### Background and aim

A standardized method to assess structural osteoarthritis (OA) burden thorough the body lacks from literature. Such a method can be valuable in developing personalized treatments for OA. We developed a reliable scoring system to evaluate OA in large joints and the spine —the OsteoArthritis Computed Tomography (OACT) score.

### Patients and Methods

We used a convenience sample of 197 whole-body low-dose non-contrast CTs. We created an atlas, containing example images as reference points for training and scoring. Each joint was graded between 0–3. The total OA burden was calculated by summing scores of individual joints. Intra- and inter-observer reliability was tested 25 randomly selected scans (N = 600 joints). Intra-observer reliability and inter-observer reliability between three observers was assessed using intraclass correlation coefficient (ICC) and square-weighted kappa statistics.

### Results

The square-weighted kappa for intra-observer reliability for OACT-score at joint-level ranged from 0.79 to 0.95; the ICC for the total OA grade was 0.97 (95%-CI, 0.94 to 0.99). Square-weighted kappa for interobserver reliability ranged from 0.48 to 0.95; the ICC for the total OA grade was 0.95 (95%-CI, 0.90 to 0.98).

### Conclusion

The OACT score, a new reproducible CT-based grading system reflecting OA burden in large joints and the spine, has a satisfactory reproducibility. The atlas can be used for research purposes, training, educational purposes and systemic grading of OA on CT-scans.

## Introduction

Osteoarthritis (OA) is a leading cause of disability worldwide, with the estimated socioeconomic burden being 1%–2.5% of the gross national product in Western countries.[1] Until now, the search for a disease modifying drug for OA has failed. A key factor for this failure is the use of a one-size-fits-all principle in the development and testing of potential treatments. End-stage osteoarthritis is a fairly uniform disease, but etiological pathways in early disease vary strongly. There is a desire to group OA patients into phenotypes, with the ultimate aim of finding the right treatment for the right patient.[2] The APPROACH study aims to describe these different phenotypes for knee OA and validate models to predict disease progression within these phenotypes.[3] This allows for more patient specific treatments and more efficient clinical trials. The APPROACH study includes knee specific parameters, including patient reported outcome measures (e.g., knee specific questionnaires), physical examination (e.g., knee range of motion), and imaging features (e.g., knee MRI). Additionally, more generic parameters are measured, such as general quality of life, physical performance (e.g., 40 m fast paced walk test) and biochemical marker levels in serum and urine. OA is often a polyarticular disease and the relationship between the latter parameters and knee OA will be heavily influenced by the overall OA burden in the body. However, there is no efficient and standardized method to assess this burden.[4,5]

Radiography is widely used for visualizing and grading structural OA. However, it has limited sensitivity for detecting structural damage because of its projectile nature; repeatability is also an issue as positioning errors are common (e.g., wide variations in joint space measurements due to inconsistent flexion of the knee).[6] Magnetic resonance imaging (MRI) is excellent for visualizing the different tissues within a joint, but it is expensive and time consuming; for example, to obtain good-quality MRI images of multiple joints, the patient would need to lie still for hours. However, CT has several advantages. It uses ionizing radiation to produce a three-dimensional (3D) tomographic images, without the projection limitations of radiography, and is known for its excellent visualization of bone. Advances, such as iterative reconstruction have substantially reduced exposure to ionizing radiation and scanning time.[7,8] Low-dose CT scans provide valuable information on the bony aspects of the joints, with a relatively high signal-to-noise ratio. Whole-body Low-dose CT (WBLDCT) scans, with a scan time of less than one minute and an effective radiation dose <3 mSv for a 70 kg adult male, are increasingly used for evaluation various conditions.

In this study, we aim to develop and describe a WBLDCT-based scoring system to quantify OA burden throughout the body. We believe that the score—the OsteoArthritis Computed Tomography (OACT) score—will be especially useful for research towards personalized OA treatments. We assess the inter- and intra-reader agreement of the new score and present an atlas, with extensive image examples, that can be used for training and educational purposes, for uniform grading of OA on CT-scans.

## Methods

### Patients and image acquisition

The scoring system was developed using a convenience sample of 197 WBLDCTs acquired for diagnosis or for attenuation correction in PET/CTs in the UMC Utrecht, Utrecht, The Netherlands, between June 2011 and November 2015; the scanning was performed as part of workup for suspected cancer and vascular or infectious disease. Scans were acquired in the supine position without any contrast enhancement, with 64 × 0.625-mm collimation, 120 kV, and dose modulation with a reference of 40 mAs; the estimated effective dose was <3.0 mSv for a 70-kg adult male. Reconstructions in the axial plane were made with 1-mm slices and 0.7-mm increments. Joints with metallic implants were excluded. This study was approved by the local institutional review board (protocol number 15/446-C), with waiver of the need for informed consent.

### Image assessment

The Picture Archiving and Communication System (PACS IDS7 19.3.12; SECTRA) was used to produce multiplanar view reconstructions. Using the 197 scans we created a feasible and reproducible system for grading the severity of OA in each of the major joints. Then, a reference atlas was composed that could be used to teach new readers the scoring definitions. Finally, we tested intra- and inter-observer reproducibility on a subset of 25 randomly selected scans (which included a total of 600 joints).

We aimed to grade all large synovial diarthrodial joints, intervertebral discs (IVD), and facet joints. The elbow was frequently positioned outside the field of view and was therefore excluded. Degenerative disc disease (DDD) of the IVD differs from OA, as IVDs are fibrocartilaginous and not synovial joints. However, the biochemical and radiological features of DDD closely resemble those of OA.[4] Many previous OA studies have assessed the lumbar spine but, as other researchers have suggested, DDD in the cervical and thoracic spine also needs to be considered.[9,10]

We first performed a thorough literature search to locate CT-based scoring systems for OA of different joints. If no viable CT-based scoring system was found, we modified the standard radiography–based scores for use on CT images. If no viable scoring system was available for a joint, we developed a new system using the classic radiographic OA characteristics (joint space narrowing, osteophytosis, sclerosis, and subchondral cysts). Each joint was graded on a scale of 0 to 3; thus, four grades were possible. The goal was to develop a scoring system that could be used to score all joints in a single patient within 15 min. The process of development of the scoring system for each joint is described below. The scoring of each joint was discussed in multiple sessions between a group consisting of a MD researcher with 5 years of experiences in medical imaging of OA (WPG), a radiologist in training with a subspecialization in musculoskeletal radiology (WF), and a fellowship-trained musculoskeletal radiologist with 6 years of experience (FJN) and an associate professor, section chief of Musculo-Skeletal Research and attending Radiologist with extensive experience in developing radiologic scores (FWR) The supplementary atlas (supplementary materials), which contains extensive examples, can be used for training and also serves as a reference for scoring. **Figure 1** presents an overview of the tibiofemoral joint, and **Figure 2** shows different grades of tibiofemoral OA.

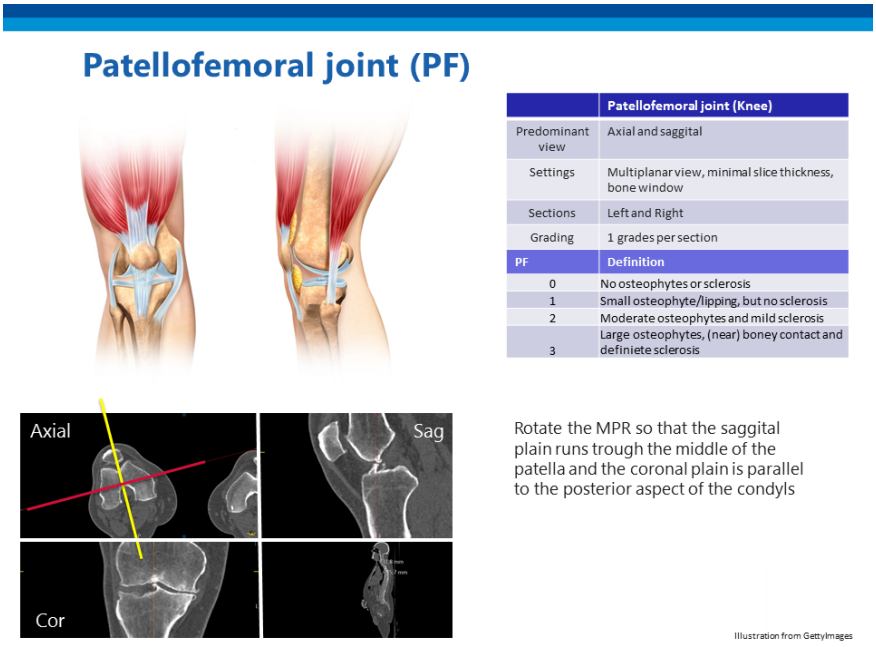



**Figure 1.** An example from the atlas showing the overview for scoring tibiofemoral osteoarthritis.

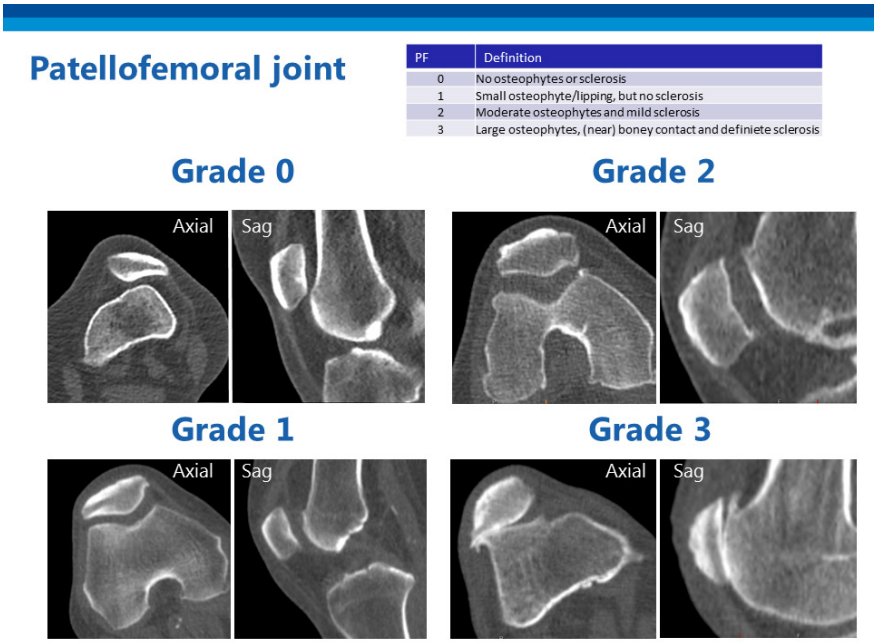



**Figure 2.** Example images from the atlas showing different grades of tibiofemoral osteoarthritis.

## Upper Extremity

### Acromioclavicular Joint

Our literature search located a single grading system for acromioclavicular joint degeneration.[11] Using 108 cadaveric joints, Stenlund et al. created a radiographic score that demonstrated satisfactory correlation with macroscopic morphological grade. However, this system was not tested for reproducibility. We used the radiographic characteristics identified by Sterlund et al. to create four grades (**Table 1**).

### Glenohumeral Joint

We did not find a validated CT-based grading system for glenohumeral OA. Therefore, we based our score on the widely used and reliable system proposed by Samilson and Prieto that scores OA according to the size of inferior humeral osteophytes on radiographs (**Table 1**).[12,13] As CT images offer 3D visualization of the joint, we considered osteophytes everywhere in the glenohumeral joint, i.e., inferior, anterior, and posterior humeral and glenoidal.

**Table 1.** Definition of OACT scores for individual joints.

| | |
|---|---|
| **Acromioclavicular joint** | |
| 0 | No osteophytes or joint space narrowing (JSN) |
| 1 | Lipping and/or possible JSN |
| 2 | Definite osteophytes and/or JSN |
| 3 | Definite osteophytes and/or JSN and sclerosis and/or cysts and/or bony deformities |
| **Glenohumeral joint** | |
| 0 | No osteophytes or definite JSN |
| 1 | Osteophyte measured less than 3 mm |
| 2 | Osteophyte measured between 3 and 7 mm, slight joint irregularity |
| 3 | Osteophyte measured more than 7 mm, definite JSN and/or irregularity. |
| **Degenerative disc disease** | |
| 0 | Score 0–2 (Based on disc space narrowing, osteophytes, end plate regularity and sclerosis) |
| 1 | Score 3–5 |
| 2 | Score 6–8 |
| 3 | Score 9–10 |
| **Facet joint** | |
| 0 | Normal facet joint space width (JSW) (2–4 mm) |
| 1 | Narrowing of facet JSW (<2 mm) and small osteophytes and/or mild hypertrophy of the articular process |
| 2 | Narrowing of facet JSW (<2 mm) and moderate osteophytes and/or moderate hypertrophy of the articular process and/or mild subarticulare bone erosions |
| 3 | Narrowing of facet JSW (<2 mm) and large osteophytes and/or severe hypertrophy of the articular process and/or severe subarticulare bone erosions and/or subchondral cysts |
| **Hip joint** | |
| 0 | Score 0–1 (Based on joint space narrowing, osteophytes, and cysts) |
| 1 | Score 2–3 |
| 2 | Score 4–5 |
| 3 | Score 6–7 |
| **Tibiofemoral joint** | |
| 0 | Score 0–1 (Based on joint space narrowing, osteophytes, and cysts) |
| 1 | Score 2–3 |
| 2 | Score 4–5 |
| 3 | Score 6–7 |
| **Patellofemoral joint** | |
| 0 | No osteophytes, joint space narrowing (JSN)/sclerosis |
| 1 | Small osteophyte/lipping and mild JSN, but no defined sclerosis |
| 2 | Moderate osteophytes, moderate JSN and possible sclerosis |
| 3 | Large osteophytes, (near) boney contact and defined sclerosis |
| **Ankle joint** | |
| 0 | No clinical evidence of OA; joint space integrity fully intact |
| 1 | Mild; osteophyte formation/lipping, possible joint space narrowing |
| 2 | Moderate; joint space narrowing evident, obvious osteophyte formation and some sclerosis/cystic changes |
| 3 | Severe; near absence of joint space, severe osteophyte/cyst formation, deformity of bone |

## Spine

### Degenerative Disc Disease

The system proposed by Lane et al. for grading degenerative disease of the thoracic and lumbar spine is convenient and reliable.[14,15] We modified it for use on CT images of the cervical, thoracic, and lumbar spine (**Table 1**). In addition to sclerosis, we considered endplate irregularity, which can be evaluated on CT, as a sign of disease involvement of cartilaginous and bony endplates. Extensive grading 21 spinal levels would be too time consuming, thus, a concise screening of the spine is performed to identify the two most affected levels within the cervical, thoracic, and lumbar regions. For these levels the extensive grading is performed. If these scores are low, this means that degenerative changes in the whole spinal region and therefore we expect limited impact on on systemic biomarker levels and quality of life measurements.

### Facet Joint OA

We incorporated the grading system created by Weishaupt et al. for the lumbar facet joint OA (an adaption of the original scoring system proposed by Pathria et al.) in our score, extending its application to the cervical and thoracic spine also.[16,17] We recommend the sagittal view for an easier, faster and more reproducible evaluation. Only the two most affected levels within each region are extensively graded (**Table 1**).

## Lower Extremity

### Hip

Turmezei et al. published a CT grading system for hip OA.[18] This system is highly detailed and time-consuming. In our experience, it takes about 5–10 min for an experienced reader to score 2 hips. The learning curve was long for new readers. We did not find any other grading systems for hip OA on CT and modified the score of Turmezei et al. it to obtain a more straightforward four-grade score based on their principles (**Table 1**).

### Knee—Tibiofemoral

We found no validated CT-based grading system for knee OA. A combination of characteristics of radiographic OA as described by Kellgren and Lawrence and, more recently, by Altman et al. (joint space narrowing, osteophytosis, and subchondral cysts) was used to create the four-grade score (**Table 1**).[19,20]

### Knee—Patellofemoral

Scoring of patellofemoral joint OA was based on the grades described by Jones et al..[21] CT is acquired with extended knees, causing the patella to be located proximal to the femoral notch; in this position, it is difficult to accurately measure joint space narrowing. Therefore, we opted for a combined score that considered osteophytosis, sclerosis, and diminishment of the joint space (**Table 1**).

### Ankle

The CT scoring system and atlas as published by Cohen et al. was used for grading ankle OA (**Table 1**).[22]

### Total OA Grade

To test the eliability of a total score for OA in the large joints and the spine, a total OA score was calculated by summing the scores of the individual joints. Therefore, with each joint scored on a scale of 0–3, the total score could range from 0 to 72. (**Table 1**).

## Testing Reproducibility

To test intra-observer reproducibility, a medical doctor and researcher with 4 years of experience (WPG) scored the same subset of 25 randomly selected WBLDCTs twice, with an interval of at least 1 week in between. To test inter-observer reproducibility, a radiologist in training, with a subspecialization in musculoskeletal radiology (WF) and a fellowship-trained musculoskeletal radiologist with 6 years of experience (FJN), scored the same random sample of 25 scans independently. The atlas was used as reference for the grading system. In accordance with the Guidelines for Reporting Reliability and Agreement Studies, reliability was tested using Cohen's kappa for binominal grade, squared weighted kappa for ordinal grade, and two-way intraclass correlation coefficient (ICC) for consistency for the total OA score.[23,24] Kappa values were interpreted according to Landis and Koch: i.e., 0–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; 0.81–1 almost perfect agreement.[25] Agreement was tested using absolute agreement percentages for binominal and ordinal grades and Bland–Altman and Jones plots for continuous values.[26,27] All analyses were carried out in R version 3.4.4 (https://cran.r-project.org/) using the irr package, version 0.84.

## Results

The 197 scans used for the development of the atlas were acquired from a sample comprising 43% males (85/197). The mean age (SD) of the patients was 54 (±15) years. Indications for scanning included vasculitis (n = 106), suspected infection (n = 57), and suspected malignancy (n = 34). The 25 scans included in the reliability analyses were from a patient subset that comprised 44% males (11/25). The mean age (SD) of the patients was 54 (±17) years. Indications for scanning were vasculitis (n = 15), suspected infection (n = 8), and suspected malignancy (n = 2). Within the test set, OA grades 0 to 3 were found in all joints, except for the hip and ankle, where only grades 0 to 2 were found (**Table 2**). Most joints were graded as having no OA or only mild OA, which is to be expected in a random sample of hospital. One ankle could not be scored due to beam-hardening artifacts caused by screws.

**Table 2.** Frequency of grades per joint (n= 25 patients)

| Joint | 0 (No) | 1 (Mild) | 2 (Moderate) | 3 (Severe) |
|---|---|---|---|---|
| Acromioclavicular, N(%) | 24 (48) | 10 (20) | 5 (10) | 11 (22) |
| Glenohumeral, N(%) | 37 (74) | 7 (14) | 3 (6) | 3 (6) |
| Intervertebral Disc, N(%) | 48 (32) | 47 (31) | 33 (22) | 22 (15) |
| Facet, N(%) | 91 (61) | 37 (25) | 7 (5) | 15 (10) |
| Hip, N(%) | 33 (66) | 13 (26) | 4 (8) | 0 (0) |
| Knee, N(%) | 25 (50) | 13 (26) | 8 (16) | 4 (8) |
| Patellofemoral, N(%) | 25 (50) | 15 (30) | 5 (10) | 5 (10) |
| Ankle[1], N(%) | 26 (54) | 19 (38) | 4 (8) | 0 (0) |

All subscores are presented in the atlas.

### Intra- and Interobserver Reliability for Total OA Grade

Intra-observer reliability for total OA grade was excellent, with an ICC of 0.97 (95% CI, 0.93 to 0.99). The Bland–Altman plot showed an even spread of errors between the first and second observation, with a mean error of −3.5 (SD, 3.4). Inter-observer reliability for total OA grade was also excellent, with an ICC of 0.94 (95% CI, 0.86 to 0.98). ICCs for inter-observer reliability were comparable between observer pairs of different proficiency levels, 0.95 between WPG and WF, 0.93 between WPG and FJN, and 0.97 between WF and FJN. The Jones plot showed an even spread of errors between all observers, with WF giving grades around the mean, FJN giving lower grades on average, and WPG giving higher grades on average (**Figure 3**).

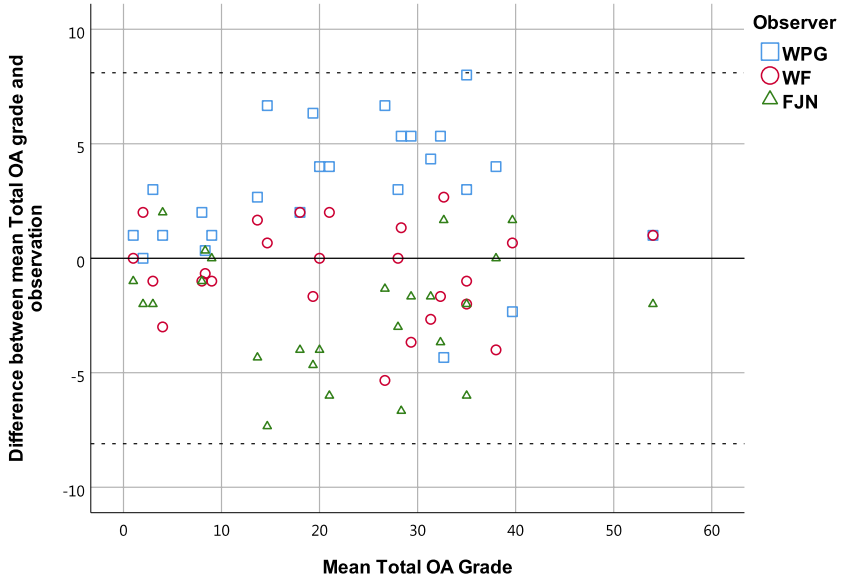



**Figure 3.** Jones plot depicting the difference between each observation of the different readers and the mean observed score for the total OA grade. The interrupted lines show the 95% limits of agreement.

## Intra- and Interobserver Reliability for OACT Scores for Individual Joints

Intra-observer reliability of the OA grades for individual joints was substantial to almost perfect, with the kappa values ranging from 0.79 to 0.95 and absolute percentage agreement, ranging from 67% to 92% (**Table 3**). Inter-observer reliability of the OA grades for individual joints was moderate to almost perfect, with the kappa values ranging from 0.48 to 0.95 and absolute percentage agreement ranging from 36% to 90% (**Table 3**). Table A1 shows the intra- and inter-observer reliability for grading of individual OA characteristics (joint space narrowing, osteophytosis, and so on).

**Table 3.** Intra- and interobserver reliability as weighted kappa (percentage of absolute agreement) for OACT scores for individual joints.

| Joints | Reader 1 (intra) | Reader 1 vs. Reader 2 | Reader 1 vs. Reader 3 | Reader 2 vs. Reader 3 |
|---|---|---|---|---|
| Acromioclavicular | 0.84 (80) | 0.87 (74) | 0.75 (62) | 0.82 (68) |
| Glenohumeral | 0.95 (92) | 0.69 (72) | 0.58 (38) | 0.50 (48) |
| Intervertebral Disc | 0.85 (67) | 0.80 (61) | 0.80 (68) | 0.77 (53) |
| Facet | 0.90 (85) | 0.68 (64) | 0.66 (57) | 0.66 (57) |
| Hip | 0.85 (88) | 0.53 (68) | 0.65 (64) | 0.48 (64) |
| Knee | 0.84 (72) | 0.85 (68) | 0.73 (50) | 0.64 (36) |
| Patellofemoral | 0.94 (88) | 0.95 (90) | 0.79 (60) | 0.78 (64) |
| Ankle | 0.79 (84) | 0.74 (80) | 0.56 (65) | 0.49 (63) |

Reader 1: Medical doctor and researcher; Reader 2: Radiologist in training with a subspecialization in musculoskeletal radiology; Reader 3: Fellowship-trained musculoskeletal radiologist with five years of experience.

## Discussion

The OACT score described here—a new reproducible WBLDCT-based grading system for OA in large joints and the spine—was developed for research purposes. In this first step, we introduce the scoring methods and present a reference atlas with multiple example images. The atlas can be used as a reference for training new readers, educational purposes and systemic grading of OA on CT-scans. We demonstrated a satisfactory intra-observer reliability and decent inter-observer reliability. The use of WBLDCT for this goal is associated with short scanning time with comparatively low-level exposure to ionizing radiation (effective radiation dose <3 mSv for a 70-kg adult male). Furthermore, with this newly developed grading system, it is possible to reliably assess overall structural burden of OA in a patient within 15 min.

There is still no disease modifying drug for OA, mainly because drug development focused on finding a one-size-fits-all drug. Drug development and evaluation will have a higher chance of success if it is focused on specific structural phenotypes of OA. The selection criteria for these OA phenotypes has to be determined. The APPROACH study uses a combination of established and novel biomarkers to develop stratification models that can help select the appropriate therapy for each knee OA patient.[3] Many parameters, such as quality of life, physical performance and biochemical markers levels in serum or urine are affected by the disease burden of other joints.[4,28-31] These parameters potentially impact the efficacy of drug development and evaluation in OA. In the APPROACH study, the OACT

score helps to phenotype OA patients and correct for confounding at the patient level when assessing the relation between systemic biomarkers, and e.g., knee OA. Besides structural progression, disease burden is an important marker for treatment success. Eventually the OACT-score will help improve patient selection for OA observational studies and clinical trials that include clinical outcome parameters. The clinical relevance needs to be established before clinical application may be considered. This has been the case for many other scoring-based assessment instruments in the field of OA that were primarily developed in the context of MRI evaluation.[32,33] Future studies should test the validity of the OACT-score against clinical outcome parameters and other biomarkers.

In our sample the total OACT score showed excellent intra- and inter-observer reliability (ICC, 0.97, and 0.94, respectively). To our knowledge, this is the first study test to reliability for an OA grade at patient level. However, we would like to stress that summing separate ordinal grades has limitations; for example, this would result in multiple low-grade joints being equivalent to a single high-grade joint. For future studies, the weighting factors for composing a total score, reflecting OA throughout the body, should be altered to the goal of the specific study. Systemic cartilage degradation markers or global quality of life measurements could be used to assess the influences of the different joints on the total OA burden in future studies. Adding the OA scores of the joints of the hands and feet would undoubtedly improve the value of the scoring system; however, we did not do so because of the variable positioning of the hands and feet in the CT images in our study. Validated radiographic scores for OA of the hands and feet could be used in combination with the OACT score for a more complete assessment of total OA burden in the body.[34]

The reliability results are in the expected range for a semi-quantitative radiological score for OA. For the acromioclavicular joint, we found substantial to almost perfect reliability. No other CT-based study is available for comparison. For the glenohumeral joint, inter-observer reliability was moderate to substantial, while the intra-observer reliability was almost perfect. We expect the moderate intra-observer reliability to be caused by the high prevalence of no and mild glenohumeral OA, as this emphasizes the decision between the presence of no, or a small (<3 mm) osteophyte. Again, no CT-based studies are available for comparison. We found almost perfect intra-observer reliability and substantial to almost perfect inter-observer reliability for DDD. No CT-based studies are available for comparison. While, OA and DDD are different entities, the response to mechanical loading, symptoms and matrix degradation pattern are highly correlated.[35] Therefore, we

chose to include DDD in our score. Based on the aim of their study, researcher may decide to in- or exclude DDD.

Pathria et al. tested the inter-observer reliability of their CT-based scoring system for facet joint OA and reported a kappa value of 0.46, while Weishaupt et al. reported a weighted kappa of 0.60; the overall percentage agreement was 63%, and 51%, respectively.[16,17] These results were comparable to our results, where the weighted kappa values ranged from 0.66 to 0.68 and absolute percentage agreement ranged from 57% to 64%.

Turmezei et al. tested the reliability of their CT grading system for hip OA and reported a weighted kappa of 0.74 and 0.75 for intra- and inter-observer reliability, respectively. [18] We simplified their scoring system to enhance grading speed and reliability for new readers and found a weighted kappa of 0.85 for intra-observer reliability and between 0.48 and 0.65 for inter-observer reliability. The lower inter-observer reliability in our study may be due to the very low prevalence of hip OA in our study population (8% with moderate OA or higher) compared to the study population of Turmezei et al., which was selected to include the full spectrum of hip OA.

For both patella and knee OA, we found almost perfect intra-observer reliability and substantial to almost perfect inter-observer reliability. For the ankle joint, we found moderate to substantial inter-observer agreement. Cohen et al. introduced an atlas for grading ankle osteoarthritis on CT and reported an ICC of 0.851 and unweighted kappa of 0.582 in a population of specifically selected scans. As such, a valid comparison with our results is not possible.

Our scoring system has several limitations. First, it does not consider OA in the elbows, hands, and feet. The elbow was not included in our score as it was positioned outside the field of view in a large number of scans. However, it should be noted that elbow OA is rare, with a prevalence of only ~2%.[36] Second, we used semi-quantitative grades. However, it must be noted that semi-quantitative grading enabled scoring a full WBLDCT in 15 min. Third, WBLDCT is obtained with the patient lying supine; assessment of joint space is influenced by the lack of weight bearing. The development of weight-bearing CT-scan will hopefully counter this problem in the near future. Fourth, WBLDCT can clearly visualize bony changes, but soft tissue degeneration (e.g., meniscal and capsule tears) will be missed. Fifth, concurrent pathology such as diffuse idiopathic skeletal hyperostosis may aggravate OA scores. Grading systems for such concurrent diseases could be used along with the OA scores to further characterize individuals.[37-39] Sixth, CT involves exposure to

possibly harmful ionizing radiation. Due to technical advances, including iterative reconstruction, the effective radiation dose of the WBLDCT was around ≤3 mSv, which approximates one year of background radiation.[40] The exact risk for excess death by cancer to a given effective radiation dose is difficult to determine. Using the rule of 5% excess mortality per 1 Sv, each WBLDCT may be accompanied by a 0.00015% excess risk for cancer mortality.[41] Determining the sample size for a reproducibility study using weighted kappa statistics is not straightforward.[24] We deemed a sample of 25 as appropriate since this results in a minimum of 50 joints per analysis and a total time invested for training and scoring of ~10 h per reader. For the analysis of the total OA grade, only 25 cases were available, which partly explains the high standard deviations in the Bland–Altman and Jones plots.

**Conclusions**

To summarize, we introduce the OACT score, a WBLDCT-based reproducible grading system for large-joint OA burden in the body. The OACT score can be used as an outcome measure in OA research or to correct for the influence of total OA burden on patient reported outcomes and biochemical marker levels.

**Supplementary Materials**

The reference atlas is available by scanning the QR code below:


https://www.mdpi.com/2075-4426/11/1/5

**Funding**

This work was supported by Reuma Nederland (LLP-22) and the APPROACH project. APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement n°115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu.

**Acknowledgments**

We would like to thank Chris van Kesteren for his work on the graphical abstract.

## References


1. Hunter, D.J.; Schofield, D.; Callander, E. The individual and socioeconomic impact of osteoarthritis. Nat. Rev. Rheumatol. 2014, 10, 437–441, doi:10.1038/nrrheum.2014.44.
2. Van Spil, W.E.; Kubassova, O.; Boesen, M.; Bay-Jensen, A.C.; Mobasheri, A. Osteoarthritis phenotypes and novel therapeutic targets. Biochem. Pharmacol. 2019, 165, 41–48, doi:10.1016/j.bcp.2019.02.037.
3. APPROACH Project. Available online: https://www.approachproject.eu (accessed on 8 March 2019).
4. Kraus, V.B.; Kepler, T.B.; Stabler, T.; Renner, J.; Jordan, J. First qualification study of serum biomarkers as indicators of total body burden of osteoarthritis. PLoS ONE 2010, 5, e9739, doi:10.1371/journal.pone.0009739.
5. Meulenbelt, I.; Kloppenburg, M.; Kroon, H.M.; Houwing-Duistermaat, J.J.; Garnero, P.; Hellio-Le Graverand, M.-P.; DeGroot, J.; Slagboom, P.E. Clusters of biochemical markers are associated with radiographic subtypes of osteoarthritis (OA) in subject with familial OA at multiple sites. The GARP study. Osteoarthr. Cartil. 2007, 15, 379–85, doi:10.1016/j.joca.2006.09.007.
6. Kinds, M.B.; Vincken, K.L.; Hoppinga, T.N.; Bleys, R.L.A.W.; Viergever, M.A.; Marijnissen, A.C.A.; Welsing, P.M.J.; Lafeber, F.P.J.G. Influence of variation in semiflexed knee positioning during image acquisition on separate quantitative radiographic parameters of osteoarthritis, measured by Knee Images Digital Analysis. Osteoarthr. Cartil. 2012, doi:10.1016/j.joca.2012.04.016.
7. Willemink, M.J.; Leiner, T.; de Jong, P.A.; de Heer, L.M.; Nievelstein, R.A.J.; Schilham, A.M.R.; Budde, R.P.J. Iterative reconstruction techniques for computed tomography part 2: Initial results in dose reduction and image quality. Eur. Radiol. 2013, 23, 1632–1642, doi:10.1007/s00330-012-2764-z.
8. Willemink, M.J.; de Jong, P.A.; Leiner, T.; de Heer, L.M.; Nievelstein, R.A.J.; Budde, R.P.J.; Schilham, A.M.R. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur. Radiol. 2013, 23, 1623–1631, doi:10.1007/s00330-012-2765-y.
9. Weiler, C.; Schietzsch, M.; Kirchner, T.; Nerlich, A.G.; Boos, N.; Wuertz, K. Age-related changes in human cervical, thoracal and lumbar intervertebral disc exhibit a strong intra-individual correlation. Eur. Spine J. 2012, 21 (Suppl. 6), S810–S818, doi:10.1007/s00586-011-1922-3.
10. O'Neill, T.W.; McCloskey, E.V.; Kanis, J.a.; Bhalla, A.K.; Reeve, J.; Reid, D.M.; Todd, C.; Woolf, A.D.; Silman, A.J. The distribution, determinants, and clinical correlates of vertebral osteophytosis: A population based survey. J. Rheumatol. 1999, 26, 842–848.
11. Stenlund, B.; Marions, O.; Engström, K.F.; Goldie, I. Correlation of macroscopic osteoarthrotic changes and radiographic findings in the acromioclavicular joint. Acta Radiol. 1988, 29, 571–576, doi:10.1177/028418518802900516.
12. Elsharkawi, M.; Cakir, B.; Reichel, H.; Kappe, T. Reliability of radiologic glenohumeral osteoarthritis classifications. J. Shoulder Elbow Surg. 2013, 22, 1063–7, doi:10.1016/j.jse.2012.11.007.

13. Samilson, R.L.; Prieto, V.; Dislocation arthropathy of the shoulder. J. Bone Joint Surg. Am. 1983, 65, 456–60.

14. Lane, N.E.; Nevitt, M.C.; Genant, H.K.; Hochberg, M.C. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. J. Rheumatol. 1993, 20, 1911–1918.

15. Kettler, A.; Wilke, H.-J. Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. Eur. Spine J. 2006, 15, 705–718, doi:10.1007/s00586-005-0954-y.

16. Weishaupt, D.; Zanetti, M.; Boos, N.; Hodler, J. MR imaging and CT in osteoarthritis of the lumbar facet joints. Skelet. Radiol. 1999, 28, 215–219, doi:10.1007/s002560050503.

17. Pathria, M.; Sartoris, D.J.; Resnick, D. Osteoarthritis of the facet joints: Accuracy of oblique radiographic assessment. Radiology 1987, 164, 227–230, doi:10.1148/radiology.164.1.3588910.

18. Turmezei, T.D.; Fotiadou, A.; Lomas, D.J.; Hopper, M.a.; Poole, K.E.S. A new CT grading system for hip osteoarthritis. Osteoarthr. Cartil. 2014, 22, 1360–1366, doi:10.1016/j.joca.2014.03.008.

19. Altman, R.D.; Gold, G.E. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthr. Cartil. 2007, 15 (Suppl. A), A1–A56, doi:10.1016/j.joca.2006.11.009.

20. Kellgren, J.H.; Lawrence, J.S. Radiological assessment of osteo-arthrosis. Ann. Rheum. Dis. 1957, 16, 494–502.

21. Jones, A.C.; Ledingham, J.; Mcalindon, T.; Regan, M.; Hart, D.; Macmillan, P.J.; Doherty, M. Radiographic assessment of patellofemoral osteoarthritis. Ann. Rheum. Dis. 1993, 52, 655–658.

22. Cohen, M.M.; Vela, N.D.; Levine, J.E.; Barnoy, E.a. Validating a New Computed Tomography Atlas for Grading Ankle Osteoarthritis. J. Foot Ankle Surg. 2014, 54, 207–213, doi:10.1053/j.jfas.2014.06.017.

23. Fleiss, J.L.; Cohen, J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ. Psychol. Meas. 1973, 33, 613–619, doi:10.1177/001316447303300309.

24. Kottner, J.; Audigé, L.; Brorson, S.; Donner, A.; Gajewski, B.J.; Hróbjartsson, A.; Roberts, C.; Shoukri, M.; Streiner, D.L. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J. Clin. Epidemiol. 2011, 64, 96–106, doi:10.1016/J.JCLINEPI.2010.03.002.

25. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977, 33, 159, doi:10.2307/2529310.

26. Martin Bland, J.; Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986, 327, 307–310, doi:10.1016/S0140-6736(86)90837-8.

27. Jones, M.; Dobson, A.; O'Brian, S. A graphical method for assessing agreement with the mean between multiple observers using continuous measures. Int. J. Epidemiol. 2011, 40, 1308–1313, doi:10.1093/ije/dyr109.

28. Mahir, L.; Belhaj, K.; Zahi, S.; Azanmasso, H.; Lmidmani, F.; El Fatimi, A. Impact of knee osteoarthritis on the quality of life. Ann. Phys. Rehabil. Med. 2016, 59, e159, doi:10.1016/j.rehab.2016.07.355.

29. Glazebrook, M.; Daniels, T.; Younger, A.; Foote, C.J.; Penner, M.; Wing, K.; Lau, J.; Leighton, R.; Dunbar, M. Comparison of Health-Related Quality of Life Between Patients with End-Stage Ankle and Hip Arthrosis. J. Bone Jt. Surg.-Am. Vol. 2008, 90, 499–505, doi:10.2106/JBJS.F.01299.

30. Gautschi, O.P.; Smoll, N.R.; Corniola, M.V.; Joswig, H.; Chau, I.; Hildebrandt, G.; Schaller, K.; Stienen, M.N. Validity and reliability of a measurement of objective functional impairment in lumbar degenerative disc disease: The Timed Up and Go (TUG) test. Neurosurgery 2016, 79, 270–278, doi:10.1227/NEU.0000000000001195.

31. Dobson, F.; Hinman, R.S.; Hall, M.; Terwee, C.B.; Roos, E.M.; Bennell, K.L. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: A systematic review. Osteoarthr. Cartil. 2012, 20, 1548–1562, doi:10.1016/j.joca.2012.08.015.

32. Peterfy, C.G.; Guermazi, A.; Zaim, S.; Tirman, P.F.J.; Miaux, Y.; White, D.; Kothari, M.; Lu, Y.; Fye, K.; Zhao, S.; et al. Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis. Osteoarthr. Cartil. 2004, 12, 177–190, doi:10.1016/j.joca.2003.11.003.

33. Hunter, D.J.; Guermazi, A.; Lo, G.H.; Grainger, A.J.; Conaghan, P.G.; Boudreau, R.M.; Roemer, F.W. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthr. Cartil. 2011, 19, 990–1002, doi:10.1016/j.joca.2011.05.004.

34. Visser, A.W.; Bøyesen, P.; Haugen, I.K.; Schoones, J.W.; van der Heijde, D.M.; Rosendaal, F.R.; Kloppenburg, M. Radiographic scoring methods in hand osteoarthritis—A systematic literature search and descriptive review. Osteoarthr. Cartil. 2014, 22, 1710–1723, doi:10.1016/j.joca.2014.05.026.

35. Rustenburg, C.M.E.; Emanuel, K.S.; Peeters, M.; Lems, W.F.; Vergroesen, P.-P.A.; Smit, T.H. Osteoarthritis and intervertebral disc degeneration: Quite different, quite similar. JOR Spine 2018, 1, e1033, doi:10.1002/jsp2.1033.

36. Gramstad, G.D.; Galatz, L.M. Management of Elbow Osteoarthritis. J. Bone Jt. Surg. 2006, 88, 421–430, doi:10.2106/jbjs.e.00568.

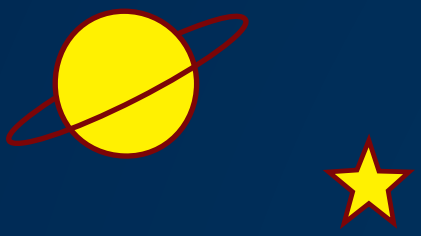
37. Kuperus, J.S.; Oudkerk, S.F.; Foppen, W.; Mohamed Hoesein, F.A.; Gielis, W.P.; Waalwijk, J.; Regan, E.A.; Lynch, D.A.; Oner, F.C.; de Jong, P.A.; et al. Criteria for Early-Phase Diffuse Idiopathic Skeletal Hyperostosis: Development and Validation. Radiology 2019, 291, 420–426, doi:10.1148/radiol.2019181695.

38. Meinberg, E.; Agel, J.; Roberts, C.; Karam, M.; Kellam, J. Fracture and Dislocation Classification Compendium—2018. J. Orthop. Trauma 2018, 32, S1–S10, doi:10.1097/BOT.0000000000001063.

39. Chanchairujira, K.; Chung, C.B.; Kim, J.Y.; Papakonstantinou, O.; Lee, M.H.; Clopton, P.; Resnick, D. Intervertebral Disk Calcification of the Spine in an Elderly Population: Radiographic Prevalence, Location, and Distribution and Correlation with Spinal Degeneration. Radiology 2004, 230, 499–503, doi:10.1148/radiol.2302011842.

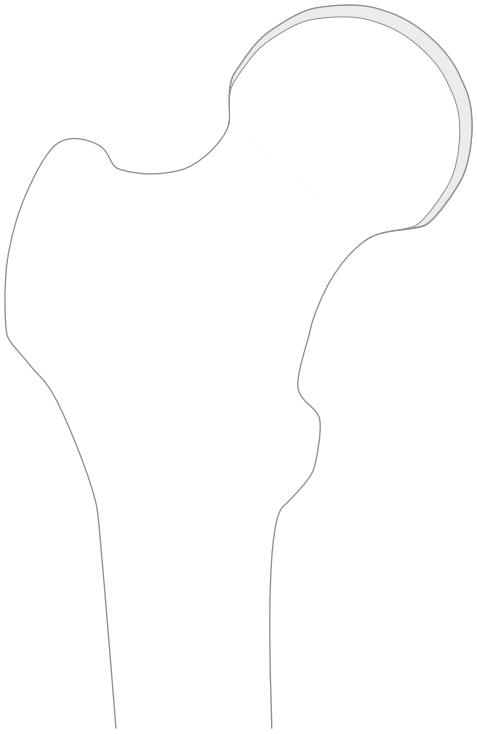
40. Mettler, F.a; Huda, W.; Yoshizumi, T.T.; Mahesh, M. Effective doses in radiology and diagnostic nuclear medicine: A catalog. Radiology 2008, 248, 254–263, doi:10.1148/radiol.2481071451.

41. Lin, E.C. Radiation risk from medical imaging. Mayo Clin. Proc. 2010, 85, 1142–1146.



# PART II

## The multifactorial pathways to OA



# CHAPTER 6

# Bone texture analysis for prediction of incident radiographic hip osteoarthritis using machine learning: data from the cohort hip and cohort knee (CHECK) study

*Osteoarthritis and Cartilage*
*June 2019, doi.org/10.1016/j.joca.2019.02.796*

Jukka Hirvasniemi, Willem Paul Gielis, Saeed Arbabi, Rintje Agricola, Willem Evert van Spil, Vahid Arbabi, Harrie Weinans

## Abstract

### Background and aim

To assess the ability of radiography-based bone texture variables in proximal femur and acetabulum to predict incident radiographic hip osteoarthritis (rHOA) over a 10 years period.

### Patients and methods

Pelvic radiographs from CHECK (Cohort Hip and Cohort Knee) at baseline (987 hips) were analyzed for bone texture using fractal signature analysis in proximal femur and acetabulum. Elastic net (machine learning) was used to predict the incidence of rHOA (including Kellgren-Lawrence grade (KL) ≥ 2 or total hip replacement (THR)), joint space narrowing score (JSN, range 0-3), and osteophyte score (OST, range 0-3) after 10 years. Performance of prediction models was assessed using the area under the receiver operating characteristic curve (ROC AUC).

### Results

Of the 987 hips without rHOA at baseline, 435 (44%) had rHOA at 10-year follow-up. Of the 667 hips with JSN grade 0 at baseline, 471 (71%) had JSN grade ≥ 1 at 10-year follow-up. Of the 613 hips with OST grade 0 at baseline, 526 (86%) had OST grade ≥ 1 at 10-year follow-up. AUCs for the models including age, gender, and body mass index to predict incident rHOA, JSN, and OST were 0.59, 0.54, and 0.51, respectively. The inclusion of bone texture variables in the models improved the prediction of incident rHOA (ROC AUC 0.68 and 0.71 when baseline KL was also included in the model) and JSN (ROC AUC 0.62), but not incident OST (ROC AUC 0.52).

### Conclusion

Bone texture analysis provides additional information for predicting incident rHOA or THR over 10 years.

## Introduction

Plain radiography is a cheap, fast and widely available imaging method for osteoarthritis (OA). Bony changes can be clearly seen on plain radiographs and provide useful information about bone deformities, density, and structure. A plain radiograph is a projection (summation) through a three-dimensional structure and this is one main limitation of this imaging method. However, it has been shown that radiography-based bone texture is significantly related with the three-dimensional structure of bone.[1-5]

Medical image analysis often involves interpretation of tissue appearance, *e.g.*, smooth, grainy, rough, or homogenous. These image properties are related to the spatial arrangement of pixel intensities in images, *i.e.,* image texture, and can be quantified using texture analysis[6]. Radiography-based texture analysis of the proximal femur has been applied for example in osteoporosis and in the assessment of femoral neck fracture risk.[7, 8] However, in OA research, the majority of studies analyzing bone texture are concentrated on the knee, using mostly fractal-based texture analysis methods.[1, 9-13] There is evidence that tibial trabecular bone texture can be used to predict both development and progression of OA as well as total knee replacement.[14-21] Only one study applied fractal signature analysis (FSA) on hip radiographs to quantify trabecular bone changes in subjects with prevalent hip OA and reported changes in fractal dimension of femoral head between baseline and 18 months follow-up.[22] However, the sample size of that study was relatively small (14 subjects) and the follow-up rather short.

Given the previous results showing that FSA can be applied on hip radiographs[22] and that bone density related variables from dual energy X-ray absorptiometry (DXA) contribute to the risk and progression of hip OA,[23-26] we hypothesize that radiography-based bone texture gives additional information in predicting the development of radiographic hip OA (rHOA). Consequently, our aim was to create a method for automated assessment of bone texture in proximal femur and acetabulum from plain hip radiographs and to assess the ability of these bone texture variables to predict incident rHOA.

## Methods

### Study cohort

Data from CHECK (Cohort Hip and Cohort Knee) cohort was used in this study.[27] CHECK is a prospective cohort study of 1002 subjects initiated to study

the course of early knee and hip OA. Data was collected in ten medical centers in The Netherlands. Subjects were recruited by general practitioners and via advertisements. At baseline, subjects were aged 45-65 years, had first onset of pain or stiffness in hip(s) and/or knee(s), and had never or not longer than 6 months ago consulted a physician for these complaints. Subjects with a pathological condition other than early OA that could explain symptoms were excluded. The study was approved by medical ethics committees of all ten participating centers and written informed consent was obtained from all participants.

Plain pelvic anterior-posterior radiographs and clinical data at baseline and 10-year follow-up were used in the current study. Subjects with missing data (radiographs, demographics, clinical examination), KL grade ≥ 2 at baseline, and/or with insufficient radiograph quality (artefacts or underexposed) were excluded (**Figure 1**). As such, the final subset for assessing incident rHOA (KL grade ≥ 2 or total hip replacement (THR)) within the period from baseline to 10 years included 987 hips from 601 subjects (**Table 1**).

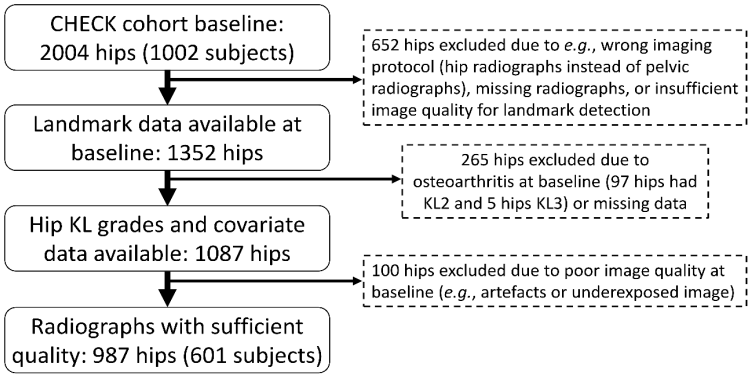



**Figure 1.** Flowchart of the selection of CHECK subjects and hips for the current study.

## Acquisition and grading of the radiographs

Weight-bearing, anterior-posterior pelvic radiographs were acquired according to a standardized protocol. A wedge was used to assure 15-degree internal rotation in feet. The source – detector distance was 120 cm, and the X-ray beam was centered on the superior part of the pubic symphysis.

**Table 1a.** Characteristics of the subjects included in the study.

| | Baseline | 10-year |
|---|---|---|
| **Anthropometric variables** | | |
| Age [years] | 55.7 (5.2) | |
| BMI [kg/m$^2$] | 25.9 (3.8) | |
| Male Sex | 160 (16.2%) | |
| **KL grade distribution** | | |
| 0 | 720 (72.9%) | 131 (13.3%) |
| 1 | 267 (27.1%) | 421 (42.7%) |
| 2 | | 389 (39.4%) |
| 3 | | 15 (1.5%) |
| 4 | | - |
| THR | | 31 (3.1%) |
| **JSN grade distribution** | | |
| 0 | 667 (100%) | 196 (29.4%) |
| 1 | | 438 (65.7%) |
| 2 | | 21 (3.1%) |
| 3 | | 2 (0.3%) |
| THR | | 10 (1.5%) |
| **OST grade distribution** | | |
| 0 | 613 (100%) | 87 (14.2%) |
| 1 | | 348 (56.8%) |
| 2 | | 167 (27.2%) |
| 3 | | 5 (0.8%) |
| THR | | 6 (1.0%) |

**Table 1b.** Characteristics of the subjects per outcome group.

| Variable | Incident OA (KL ≥2) | | Incident JSN (grade ≥1) | | Incident OST (grade ≥1) | |
|---|---|---|---|---|---|---|
| | Controls (n = 552) | OA (n = 435) | Controls (n = 196) | JSN (n = 471) | Controls (n = 87) | OST (n = 526) |
| Age [years] | 55.0 (5.2) | 56.5 (5.1) | 55.1 (5.0) | 56.0 (5.1) | 55.3 (5.6) | 55.2 (5.1) |
| BMI [kg/m$^2$] | 25.9 (3.9) | 25.9 (3.6) | 25.5 (3.2) | 25.9 (3.8) | 25.5 (3.5) | 26.0 (3.8) |
| Male Sex | 71 (12.9%) | 89 (20.5%) | 23 (11.7%) | 67 (14.2%) | 9 (10.3%) | 74 (14.1%) |

BMI = Body mass index, KL = Kellgren-Lawrence, JSN = joint space narrowing, OST = osteophyte. All values are given as mean (standard deviation) or n (%), THR = Total hip replacement.

Hips were classified according to the KL grading scale[28] at baseline and 10-year follow-up. Superior and medial joint space narrowing (JSN) and superior and inferior osteophytes (OST) in acetabulum/femur were classified according to the Osteoarthritis Research Society International (OARSI) grading scale.[29] The scale for classifying the changes in JSN and OST was from 0 (normal) to 3 (severe

change).[29] Highest JSN and OST grades of the analyzed regions were used in the analyses. **Table 1** summarizes the distribution of the KL, JSN, and OST grades of included hips at baseline and 10-year follow-up.

### Selection of regions of interests

Prior to extraction of the regions of interests (ROIs), all images were resampled to have the same femoral head diameter in pixels as the smallest femoral head diameter on the data. Bicubic interpolation was used to ensure comparability of the structural variables, without producing as much artefacts as bilinear or nearest neighbor interpolation algorithms. The resampling was also needed because part of the baseline radiographs were digitized and saved in TIFF images (501 hip images were in DICOM and 486 in TIFF format) and the actual pixel size on the detector was not available. After resampling, to assess bone texture from the radiographs, 41 circular ROIs with 70 pixels diameter were extracted from femoral head and acetabulum (25 on femoral head and 16 on acetabulum) (**Figure 2**). Although previous studies have typically used rectangular ROIs, circular ROIs were used in the current study to better cover femoral head and acetabular area and to enable bone texture assessment in many different directions inside the ROI without losing pixels when rotating the ROI. ROI selection procedure was based on 14 out of 75 landmarks, which were manually placed on the proximal femur and pelvis in a previous study (**Figure 2**).[30] Two circles were fitted in femoral head and acetabulum for ROI placement using Least Squares Optimization algorithm, which calculates the center and radius of best fitting circle in an iterative process. Locations of the ROIs were determined after a robustness assurance step which guarantees that the same ROI number selects the same corresponding pixels on images despite the size and rotation differences among the images. Locations of the ROI1, ROI17, ROI25, and ROI26 were defined based on the center of the femoral head and the second-most lateral landmark on the acetabular rim (**Figure 2**). Other ROIs were automatically placed based on the locations of those ROIs.

### Bone texture analysis

Before texture analysis, images were median filtered with a 3x3 pixels filter to remove high-frequency noise and grayscale values were expanded to full dynamic range (0 – 255). Bone texture was assessed using the FSA method.[1, 10, 11] FSA produces fractal dimension values that are related to the roughness and complexity of the image. To calculate the fractal dimensions, the image was dilated and eroded with a rod-shaped, horizontally oriented, one-pixel wide structuring element. After that,

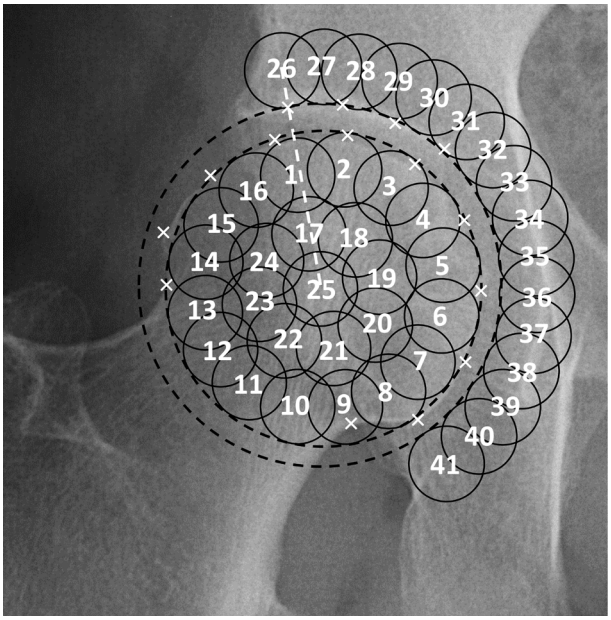



**Figure 2.** Location of regions of interest (ROIs). Landmarks that were used when fitting circles to the femoral head and acetabulum are indicated with white "x". The white dashed line shows the centers of the femoral head, ROI1, ROI25 and ROI26. Other ROIs were automatically placed based on the locations of those ROIs.

the volume, *V*, between dilated and eroded images was calculated. Calculations were repeated by varying the element length *r* from 2 to 5 pixels. The surface area, *A(r)*, was obtained from the Equation 1:

$$A(r) = (V(r) - V(r-1))/2 \qquad (1)$$

Subsequently, a log-log plot was constructed by plotting log of *A(r)* against log of *r*. Finally, the fractal dimension was estimated by the slope of regression line that fitted the three points in the log-log plot. High fractal dimension values are associated with high complexity of the image, whereas low complexity results in low fractal dimension values.

Because the orientation of the bone structures in femoral head and acetabulum varies, we assessed fractal dimensions in 18 different angles, *i.e.* from 0 to 170 degrees with 10 degrees increments, by rotating the ROI. To reduce the number of fractal dimension values (18 values for 41 ROIs = 738 features per hip), minimum

($FD_{min}$) and maximum ($FD_{max}$) fractal dimension values and their respective angles ($Angles_{min}$, $Angles_{max}$) per ROI were selected. Consequently, 164 texture features were used in the analyses.

### Statistical analyses

Logistic regression with generalized estimating equation (GEE) was used to assess the association of each baseline covariate (age, gender, body mass index (BMI), and baseline KL grade (KL 0 or KL1)) as well as bone texture variable with incident rHOA (KL grade ≥ 2 or THR), incident JSN (JSN grade ≥ 1 or THR), and incident OST (OST grade ≥ 1 or THR). GEE was used to take into account the potential correlation between bilateral measurements. Odds ratios (OR) and 95% confidence interval of the univariable models were reported. It should be noted that the OR is not directly interpretable as a relative risk, because the OR most likely overestimates the relative risk. The analyses were performed using IBM SPSS Statistics (version 25, IBM Corp., Armonk, NY, USA).

To prevent overfitting, machine learning was used for dimensionality reduction and to assess the predictive ability of the bone texture variables and baseline covariates. For the dimensionality reduction and prediction, a regularized logistic regression method called elastic net was used.[31, 32] The elastic net linearly combines the $L_1$ and $L_2$ penalties of lasso and ridge regression methods.[31, 32] The samples were randomly divided into a training and validation set (790 hips, 80% of the data) and a hold-out test set (197 hips, 20% of the data) by stratifying the proportion of the controls and subjects with incident rHOA at follow-up in each set. To optimize the ratio of the $L_1$ and $L_2$ penalties (α) and the strength of the penalty parameter (λ) of the elastic net, 10-fold cross-validation was performed. When α is close to zero, the elastic net approaches ridge regression, while when α is 1, lasso regression is performed. The combination of the α- and λ-parameters that had the highest area under the receiver operating characteristics curve (ROC AUC) in the cross-validation, was selected. After optimizing elastic net parameters, the predictive ability of 1) covariate model, 2) texture feature model, 3) covariate + baseline KL (KL0 or KL1) model, and 4) model with covariates, baseline KL, and texture features combined in the test set were assessed using ROC AUC value. The agreement between observed outcomes and predictions in the test set were assessed using calibration plots.[33] Analyses were repeated to predict incident rHOA among subjects with only KL0 or KL1 at baseline separately. Furthermore, incident JSN (JSN ≥ 1 or THR) among subjects with JSN grade 0 at baseline and incident OST (OST grade ≥ 1 or THR) among subjects with OST grade 0 at baseline were predicted with the

elastic net. Distribution of the proportion of incident rHOA cases was different between imaging centers. To remove the potential effect of imaging center to fractal dimension values, the variables were standardized with mean and standard deviation values of the center where the imaging was performed ($z = (x - \mu)/SD$, where x is the value of each measurement, μ and SD are the average and standard deviation of the variable at the center where the imaging was performed). Statistical analyses were performed using R (version 3.4.1) software with Caret[34] (version 6.0-47), pROC[35] (version 1.13.0), glmnet[31] (version 2.0-16), and CalibrationCurves[36] (version 0.1.2) packages.

## Results

Of the 987 hips without rHOA at baseline, 435 (44%) had developed incident rHOA (KL ≥ 2 or THR) at 10-year follow-up (**Table 1**). Of the 667 hips with JSN grade 0, 471 (71%) had JSN grade ≥ 1 or THR at 10-year follow-up. Of the 613 hips with OST grade 0, 526 (86%) had OST grade ≥ 1 or THR at 10-year follow-up.

The GEE models for covariates showed that higher age (OR: 1.06) and baseline KL grade (OR: 3.80) were associated with significantly ($p$<0.05) higher odds (**Table 2**). Female gender (OR: 0.54) was associated with lower odds of incident rHOA at the 10-year follow-up (**Table 2**). When looking at the univariable texture models, ORs for minimum fractal dimension ($FD_{min}$) variable models were statistically significant in 11/41 ROIs (OR range: 0.71 – 1.15), the maximum fractal dimension ($FD_{max}$) in 8/41 ROIs (OR range: 0.83 – 1.11), the angle associated with $FD_{min}$ ($Angles_{min}$) in 5/41 ROIs (OR range: 0.96 – 1.02), and the angle associated with $FD_{max}$ ($Angles_{max}$) in 12/41 ROIs (OR range: 0.93 – 1.05) (**Table 3**). For incident JSN grade ≥ 1, ORs for 9 texture variables (OR range: 0.64 – 1.43) were significant (**Supplementary Tables 1 and 2**). For incident OST grade ≥ 1, ORs for 5 texture variables (OR range: 0.78 – 1.08) were significant (**Supplementary Tables 3 and 4**).

The selected elastic net parameters and ROC AUC values for the covariate model, texture model, covariate + baseline KL model, and for the combined covariate, baseline KL, and texture feature model from 10-fold cross-validation are shown in **Table 4**. The model that included covariates, baseline KL, and texture features had the highest ROC AUC (0.73 [95% CI: 0.70 – 0.76]) in cross-validation. The variables that were selected based on the 10-fold cross-validation of the elastic net are listed in **Supplementary Table 5** and visualized in **Figure 3**. BMI was not selected in any of the models by the algorithm.

**Table 2.** Odds ratios (95% confidence interval) of the univariable covariate models to assess incident rHOA (KL ≥ 2 or THR).

| Predictor | Oddsratio (95%-Confidence interval) |
|---|---|
| Age (years) | 1.06 (1.03 – 1.09) * |
| Female sex | 0.54 (0.37 – 0.80) * |
| Baseline KL grade | 3.80 (2.79 – 5.16) * |
| BMI (kg/m2) | 1.00 (0.96 – 1.04) |

*p < 0.05, BMI = Body mass index, KL = Kellgren & Lawrence

**Table 3.** Odds ratios (95% confidence interval) of univariable texture models to assess incidence of rHOA (KL ≥ 2 or THR).

| ROI | FDmin | FDmax | Anglesmin | Anglesmax |
|---|---|---|---|---|
| 1 | 0.84 (0.74 – 0.96)* | 0.88 (0.78 – 1.01) | 1.02 (1.00 – 1.05)* | 0.93 (0.90 – 0.97)* |
| 2 | 0.84 (0.73 – 0.96)* | 0.88 (0.76 – 1.01) | 1.00 (0.97 – 1.04) | 0.98 (0.95 – 1.01) |
| 3 | 0.96 (0.85 – 1.09) | 0.88 (0.78 – 0.99)* | 0.98 (0.94 – 1.02) | 1.02 (0.99 – 1.05) |
| 4 | 1.01 (0.89 – 1.15) | 0.97 (0.86 – 1.10) | 0.97 (0.94 – 1.00)* | 1.00 (0.97 – 1.03) |
| 5 | 0.99 (0.87 – 1.13) | 1.00 (0.89 – 1.13) | 0.97 (0.94 – 0.99)* | 0.99 (0.95 – 1.03) |
| 6 | 0.83 (0.73 – 0.95)* | 0.87 (0.77 – 0.99)* | 0.99 (0.97 – 1.02) | 0.99 (0.96 – 1.02) |
| 7 | 0.95 (0.84 – 1.08) | 1.00 (0.88 – 1.13) | 0.99 (0.97 – 1.01) | 1.00 (0.96 – 1.03) |
| 8 | 1.02 (0.89 – 1.16) | 1.04 (0.92 – 1.18) | 1.00 (0.98 – 1.02) | 0.98 (0.95 – 1.01) |
| 9 | 1.01 (0.89 – 1.14) | 1.05 (0.93 – 1.19) | 0.99 (0.97 – 1.02) | 0.98 (0.96 – 1.00) |
| 10 | 0.81 (0.72 – 0.93)* | 0.88 (0.78 – 1.01) | 1.01 (0.98 – 1.05) | 0.96 (0.94 – 0.98)* |
| 11 | 0.89 (0.79 – 1.01) | 0.95 (0.84 – 1.07) | 1.01 (0.98 – 1.04) | 0.95 (0.93 – 0.98)* |
| 12 | 0.97 (0.85 – 1.09) | 1.04 (0.92 – 1.18) | 1.02 (0.99 – 1.04) | 0.95 (0.92 – 0.98)* |
| 13 | 0.94 (0.82 – 1.07) | 1.01 (0.89 – 1.15) | 1.01 (0.99 – 1.03) | 0.98 (0.94 – 1.01) |
| 14 | 1.03 (0.91 – 1.17) | 1.11 (0.98 – 1.26) | 1.02 (1.00 – 1.03) | 0.96 (0.93 – 0.99)* |
| 15 | 0.97 (0.85 – 1.10) | 0.96 (0.85 – 1.09) | 0.99 (0.97 – 1.01) | 1.02 (0.98 – 1.05) |
| 16 | 1.06 (0.93 – 1.20) | 0.92 (0.81 – 1.03) | 0.96 (0.93 – 0.98)* | 0.97 (0.93 – 1.02) |
| 17 | 1.03 (0.90 – 1.17) | 1.03 (0.91 – 1.17) | 1.00 (0.97 – 1.02) | 0.93 (0.90 – 0.97)* |
| 18 | 0.88 (0.77 – 1.00) | 0.92 (0.81 – 1.04) | 0.97 (0.94 – 1.01) | 0.98 (0.96 – 1.01) |
| 19 | 0.81 (0.71 – 0.93)* | 0.87 (0.76 – 0.99)* | 1.02 (0.99 – 1.06) | 0.99 (0.96 – 1.02) |
| 20 | 0.92 (0.81 – 1.04) | 1.00 (0.88 – 1.12) | 1.01 (0.98 – 1.04) | 0.98 (0.96 – 1.00) |
| 21 | 1.03 (0.91 – 1.17) | 0.97 (0.85 – 1.10) | 1.00 (0.96 – 1.04) | 0.98 (0.96 – 1.01) |
| 22 | 0.93 (0.83 – 1.05) | 0.88 (0.77 – 0.99)* | 0.97 (0.93 – 1.01) | 0.97 (0.94 – 0.99)* |
| 23 | 0.88 (0.78 – 1.00) | 0.92 (0.82 – 1.04) | 1.03 (0.99 – 1.06) | 0.95 (0.92 – 0.98)* |
| 24 | 1.05 (0.92 – 1.19) | 1.09 (0.96 – 1.24) | 1.01 (0.99 – 1.04) | 0.93 (0.90 – 0.97)* |
| 25 | 0.98 (0.87 – 1.12) | 1.01 (0.90 – 1.14) | 1.00 (0.94 – 1.05) | 0.95 (0.92 – 0.99)* |
| 26 | 0.96 (0.84 – 1.09) | 1.01 (0.90 – 1.15) | 1.03 (1.00 – 1.06) | 1.03 (0.99 – 1.07) |
| 27 | 0.79 (0.70 – 0.90)* | 0.85 (0.75 – 0.96)* | 1.00 (0.97 – 1.03) | 0.99 (0.97 – 1.00) |
| 28 | 0.78 (0.69 – 0.90)* | 0.89 (0.78 – 1.02) | 1.03 (1.00 – 1.06) | 0.99 (0.98 – 1.01) |
| 29 | 0.78 (0.68 – 0.89)* | 0.95 (0.84 – 1.08) | 0.98 (0.95 – 1.01) | 1.01 (0.99 – 1.04) |
| 30 | 0.84 (0.75 – 0.95)* | 0.98 (0.87 – 1.11) | 1.00 (0.97 – 1.03) | 0.99 (0.96 – 1.02) |
| 31 | 1.00 (0.88 – 1.13) | 0.97 (0.85 – 1.10) | 0.99 (0.96 – 1.01) | 0.99 (0.96 – 1.02) |
| 32 | 1.13 (0.99 – 1.30) | 0.92 (0.81 – 1.04) | 0.99 (0.97 – 1.01) | 1.05 (1.00 – 1.09)* |
| 33 | 1.14 (1.00 – 1.30) | 1.07 (0.94 – 1.22) | 0.98 (0.96 – 1.01) | 0.96 (0.90 – 1.03) |
| 34 | 1.09 (0.96 – 1.24) | 1.12 (0.98 – 1.28) | 1.00 (0.98 – 1.02) | 1.00 (0.89 – 1.14) |
| 35 | 1.00 (0.88 – 1.13) | 1.02 (0.90 – 1.16) | 1.02 (1.00 – 1.04) | 1.01 (0.90 – 1.13) |
| 36 | 1.01 (0.89 – 1.15) | 0.95 (0.84 – 1.08) | 1.02 (1.00 – 1.03) | 0.93 (0.81 – 1.07) |
| 37 | 1.02 (0.90 – 1.15) | 1.03 (0.91 – 1.17) | 1.01 (0.99 – 1.03) | 0.92 (0.80 – 1.05) |
| 38 | 1.00 (0.87 – 1.14) | 1.09 (0.96 – 1.24) | 0.99 (0.97 – 1.01) | 0.99 (0.93 – 1.06) |
| 39 | 1.15 (1.01 – 1.32)* | 1.07 (0.94 – 1.22) | 0.99 (0.97 – 1.01) | 0.99 (0.95 – 1.03) |
| 40 | 0.92 (0.81 – 1.04) | 0.86 (0.76 – 0.98)* | 0.98 (0.96 – 0.99)* | 1.04 (1.01 – 1.07)* |
| 41 | 0.71 (0.63 – 0.81)* | 0.83 (0.73 – 0.94)* | 0.98 (0.96 – 1.00) | 1.01 (0.98 – 1.03) |

*p < 0.05

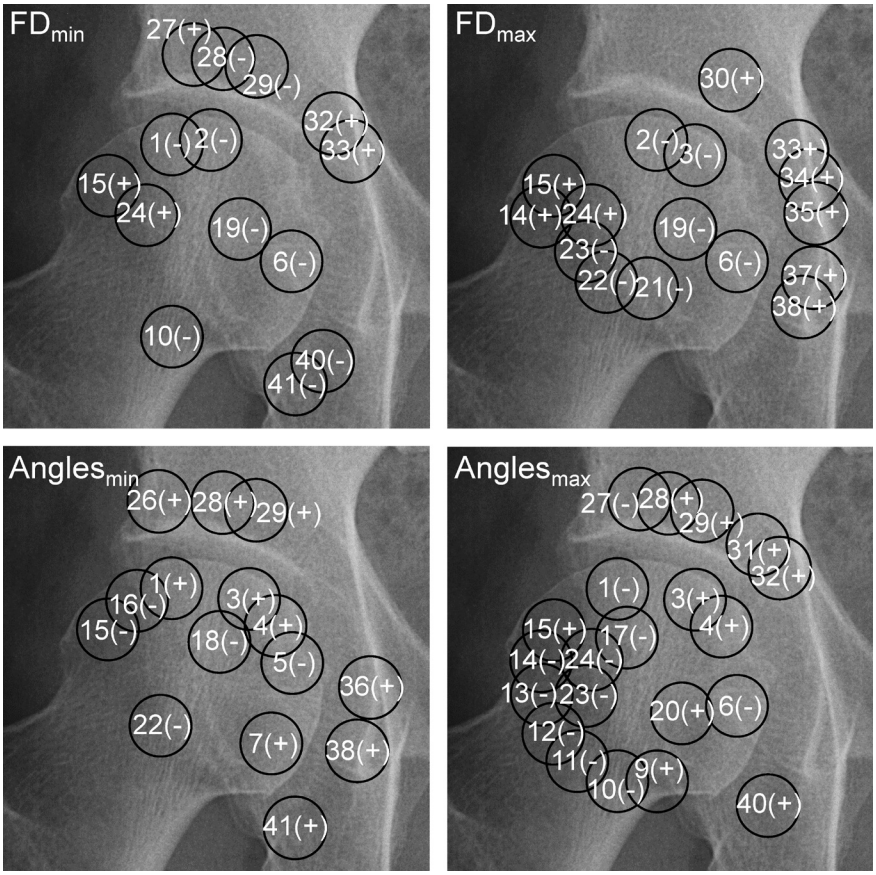



**Figure 3.** Location of regions of interest (ROIs) selected to the final elastic net model.

**Table 4.** Selected λ and α parameters for the elastic net from 10-fold cross-validation and areas under the receiver operating characteristics curve (ROC AUC) (95% confidence interval) for predicting incidence of rHOA (KL ≥ 2 or THR).

| Model | Selected λ | Selected α | ROC AUC in validation |
|---|---|---|---|
| Covariates (age, gender, BMI) | 0.096 | 0.15 | 0.61 (0.58 – 0.64) |
| Texture features | 0.059 | 0.10 | 0.68 (0.65 – 0.71) |
| Covariates + KL | 0.001 | 0.30 | 0.69 (0.64 – 0.74) |
| Texture features + Covariates + KL | 0.153 | 0.10 | 0.73 (0.70 – 0.76) |

*p < 0.05, BMI = Body mass index

When assessing the performance of the optimized elastic net models in test set, the combined covariate, baseline KL, and bone texture feature model had the highest AUC (0.71 [95% CI: 0.63 – 0.78]). ROC curves for the 1) covariate, 2) texture, 3) combined covariates and baseline KL, and 4) combined covariate, baseline KL, and texture feature models in the test set are shown in **Figure 4**. Calibration-in-the-large coefficients of the models varied from -0.42 to -0.31 and calibration slopes from 0.81 to 1.56 (**Supplementary figure 1**).

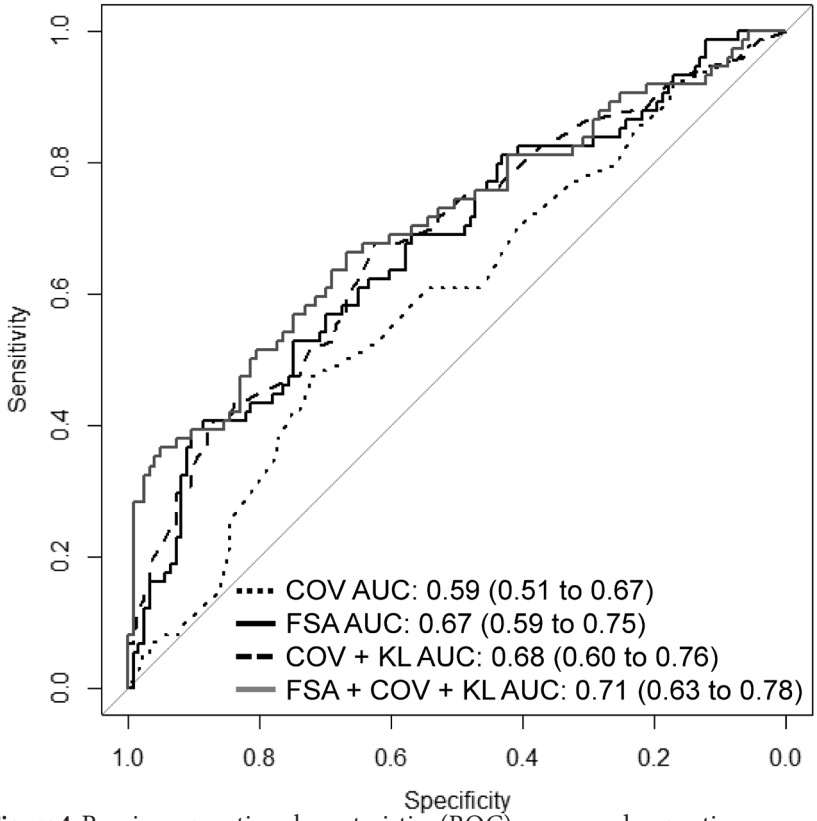



**Figure 4.** Receiver operating characteristics (ROC) curves and respective area under the curve (AUC) values for predicting incident radiographic hip osteoarthritis (rHOA) (KL ≥ 2 or total hip replacement (THR)) using 1) covariates (age, gender and body mass index), 2) texture variables from fractal signature analysis (FSA), 3) covariates and baseline KL grade, and 4) texture variables combined with covariates and KL grade.

When assessing subjects with KL0 or KL1 at baseline separately, adding the texture variables improved model performance in the test set. ROC AUC value increased from 0.51 [95% CI: 0.41 – 0.61] to 0.62 [95% CI: 0.53 – 0.72] among KL0 subjects (**Supplementary figure 2**) and from 0.54 [95% CI: 0.37 – 0.70] to 0.65 [95% CI: 0.48 – 0.81] among KL1 subjects (**Supplementary Figure 3**). Calibration plots of texture and covariate + texture models among KL0 subjects showed an agreement between actual and predicted risks for low-risk subjects, but failed for high-risk subjects (data not shown). For KL1 subjects, calibration plots showed a reasonable agreement between actual and predicted risks for high-risk subjects, but failed for low-risk subjects (data not shown). For analyses in subjects with KL0 at baseline, ROC AUC values in cross-validation were 0.60 (95% CI: 0.55 – 0.66), 0.60 (95% CI: 0.57 – 0.62), and 0.62 (95% CI: 0.59 – 0.64) for covariate, texture, and covariate + texture models, respectively. For analyses in subjects with KL1 at baseline, ROC AUC values in cross-validation were 0.62 (95% CI: 0.56 – 0.69), 0.71 (95% CI: 0.65 – 0.76), and 0.71 (95% CI: 0.63 – 0.79) for covariate, texture, and covariate + texture models, respectively.

When predicting incident JSN in the test set, the combined texture and covariate model had the highest ROC AUC value of 0.62 (95% CI: 0.52 – 0.72) (**Supplementary Figure 4**). Calibration plots of texture and covariate + texture models showed a reasonable agreement between actual and predicted risks (data not shown). ROC AUC values in cross-validation were 0.54 (95% CI: 0.50 – 0.58), 0.67 (95% CI: 0.63 – 0.71), and 0.67 (95% CI: 0.63 – 0.72) for covariate, texture, and covariate + texture models, respectively.

Selected models performed poorly when predicting incident OST in the test set (**Supplementary Figure 5**) and showed a poor calibration (data not shown). The highest ROC AUC value (0.52 [95% CI: 0.38 – 0.65]) was obtained with the combined texture and covariate model. ROC AUC values in cross-validation were 0.55 (95% CI: 0.45 – 0.65), 0.60 (95% CI: 0.52 – 0.68), and 0.60 (95% CI: 0.53 – 0.67) for covariate, texture, and covariate + texture models, respectively.

## Discussion

In this study, we created a method for the assessment of bone texture in proximal femur and acetabulum from plain pelvic radiographs and assessed the ability of bone texture to predict incident rHOA (KL ≥ 2 or THR). Fractal dimension was measured from 41 ROIs that were placed on femoral head and acetabulum. Inclusion of bone texture variables in the prediction model increased the ROC

AUC value of the model during cross-validation (from 0.69 to 0.73) and in the hold-out test set (from 0.68 to 0.71) as compared to the model with baseline patient characteristics and baseline KL grade.

As there were no previous data for the optimal location of ROIs for bone texture analysis in the hip, we decided to cover the whole femoral ROI and also incorporate the acetabulum in our analyses. As shown in a previous study for the knee, areas distal from the subchondral bone might also include relevant texture information.[15] Interestingly, majority of the relevant ROIs for the prediction of rHOA in our analyses were either at or next to the principal compressive trabeculae or close to the joint space (**Figure 3**). Depending on the ROI and variable, either higher or lower values were predictive for rHOA. For example, higher values in FDmin in ROIs 15, 24, 27, 32 and 33 and lower values in ROIs 1, 2, 6, 10, 19, 28, 29, 40, and 41 were predictive for rHOA. Lower fractal dimension values are likely associated with trabecular thickening or a reduction in trabecular number22. Changes in the angles associated with the fractal dimensions indicate the changes in the orientation of the trabeculae within the ROI. For example, higher Anglesmax values indicate that the maximum fractal dimension value (FDmax) was detected from higher angles. One possible explanation for the differences between the directions of the predictive values in the ROIs may be the adaptation of bone according to the daily loading of the joint. Some areas may experience higher loads whereas the loads may be reduced in other areas. Furthermore, subchondral bone sclerosis affects the values in areas that are near the joint space.

In a previous study assessing changes in FSA in OA hips, 14 subjects were followed 18 months. Fractal dimension of small and medium sized structures in the image were decreased during follow-up probably due to trabecular thickening or a reduction in trabecular number.[22] However, the relationship between FSA and development and/or progression of OA was not studied. Another study assessed texture on the hip joint space area and reported 88.9% classification accuracy between controls and OA subjects. However, that study was cross-sectional, did not assess femoral head or acetabular bone texture, and had a limited sample size (n = 64).[37]

In contrast to the scarce assessment of associations between bone texture and hip OA, the association between tibial bone texture and knee OA has been described in many papers.[14-18, 20, 21] ROC AUC values between 0.65 and 0.79 have been reported for models predicting progression or development of knee OA using bone texture and clinical covariates (e.g., age, BMI and gender).[14-16, 18, 21] Associations of covariates and bone texture with incident rHOA in our study are in line with these results. The ROC AUC to predict rHOA increased from 0.68 to 0.71 after

including bone texture to the model that included covariates and KL grading. Relatively low increase in ROC AUC may be because baseline KL grade alone is already a quite strong predictor of rHOA. Bone texture does not directly provide information about JSN, whereas JSN affects directly to KL grade, as this is included in its definition. However, trabecular bone structure is not evaluated in KL grading whereas bone texture analysis provides information about that.

When the ability to predict incident rHOA was compared between subjects with KL0 and KL1 at baseline, ROC AUC values were higher for KL1 subjects (ROC AUC values for the full model: 0.62 vs 0.65). These results suggest that in KL0 subjects there might not yet be changes that can be captured with bone texture analysis and predict incident rHOA.

In this study, the ability to predict incident JSN or OST in the hip was worse than that reported for the knee14. Our results showed that ROC AUC to predict incident JSN (0.62) was higher than incident OST (0.52). Distribution of controls and cases was more unbalanced for incident OST analyses, which may explain some of the difference. Another reason might be that OSTs were assessed from four different locations and combined into one variable, whereas JSN was assessed from two different locations and combined into one variable. Furthermore, JSN and OST are different characteristics of OA.

We decided to resample all images to same size (in pixels) based on the femoral head diameter, pre-process images with median filtering, and standardize FSA values within each center, because ten different medical centers with different proportions of incident rHOA cases ($p<0.05$) participated in the data collection and different X-ray machines were used for the imaging. We believe that the standardization lowers the possibility that the prediction models recognize a center based on texture values and use that information to improve the prediction. It should be noted that the performance of the models was very similar without pre-processing and/or standardization (0.01 – 0.03 difference in ROC AUCs, data not shown). The resampling of the images and the lack of actual pixel size of some images complicates the assessment of the actual scale that the FSA quantified here. However, based on the images that had the pixel size was available, the scale was around 600 μm.

Assessment of calibration (agreement between observed outcomes and predictions) in the test set indicated that predicted risks were on average slightly overestimated when predicting rHOA (calibration-in-the-large coefficients < 0). The model using only FSA variables showed some overfitting (calibration slope <

1) and other models some underfitting (calibration slope > 1). For the elastic net models, the combination of the α- and λ-parameters that had the highest ROC AUC in the cross-validation, were selected. We also tested selecting the largest λ-value such that error in cross-validation was within 1 standard error of the λ that had the highest ROC AUC, but calibration plots indicated that these models were in general underfitted (data not shown). Our sample size was too low for reliable assessment of the calibration intercept and slope when predicting other outcomes than incident rHOA (123 controls, 74 OA in the test set), because at least 100 events and 100 non-events have been recommended[36] and were not therefore calculated.

This study contains some limitations that need to be discussed. First, the relevance of bone changes to the OA disease process might differ between HOA phenotypes. In the current study, different phenotypes were mixed as there is no consensus on how to define OA phenotypes yet. Second, radiographic scoring is subjective, semi-quantitative, and a plain radiograph is a projection of 3-dimensional structure. Therefore, some OA changes may have been missed. Third, rHOA and THR were combined as an outcome, while they might be different. We decided to combine these outcomes due to low number of THR subjects. Fourth, ten different medical centers with different proportions of incident rHOA cases ($p<0.05$) participated in the data collection and different X-ray machines were used for the imaging, which may have affected texture analysis. However, FSA has been shown to be robust to the changes in imaging settings (e.g., exposure and pixel size)11 and FSA values were standardized within each center. We think that including data from multiple X-ray machines increases the generalizability of our results. Fifth, training, validation, and test sets were derived from CHECK and the model was not tested in another cohort. However, to reduce overfitting, the hold-out test set was not used in cross-validation and the optimal elastic net parameters were searched using 10-fold cross-validation. Sixth, some ROIs (e.g., ROIs 32-38 in acetabulum) may contain both cortical and trabecular bone and should be considered when interpreting the results. Seventh, the potential correlation between bilateral measurements was not taken into account in the elastic net analyses, which may have introduced some bias into our analyses. However, correlation between hips was taken into the account in the GEE analyses.

## Conclusion

In conclusion, bone texture analysis in proximal femur and acetabulum provides additional information when trying to predict incident rHOA or THR. Our results

suggest that bone texture variables could be valuable when building prediction tools for OA. Given the current results and the previous results on knee, a similar analysis approach could be tested in other joints (e.g., hands) as well in the future.

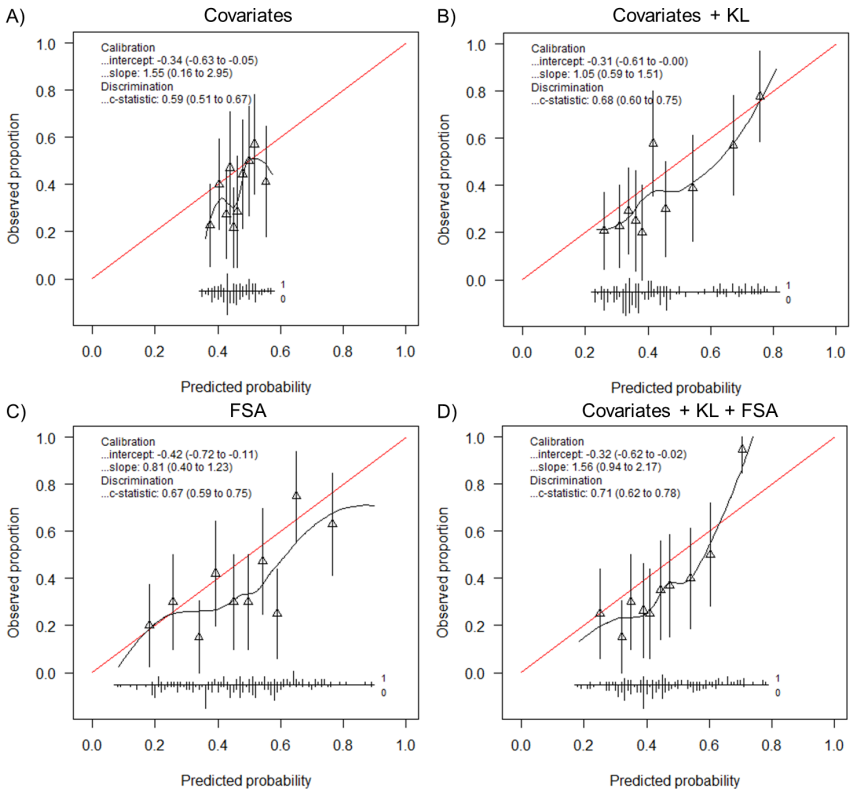
**Funding**

The research leading to these results has received funding from the Academy of Finland (project 308165). The funding sources had no role in the study design, data collection or analysis, interpretation of data, writing of the manuscript, or in the decision to submit the manuscript for publication.

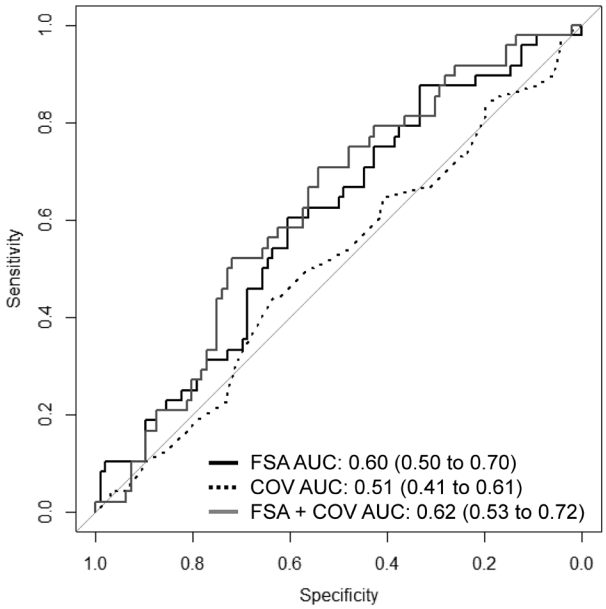
**Acknowledgements**

The CHECK-cohort study is funded by the Dutch Arthritis Foundation. Involved are: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen /Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede /Ziekenhuisgroep Twente Almelo; Reade Center for Rehabilitation and Rheumatology; St.Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

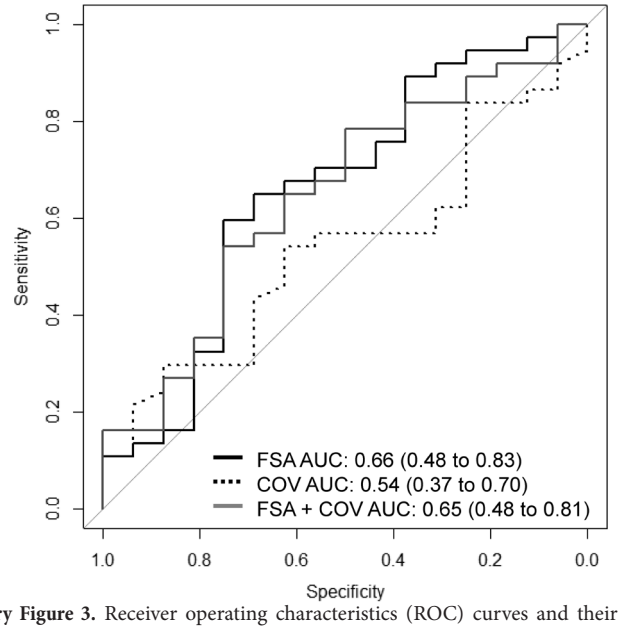
# Supplementary material



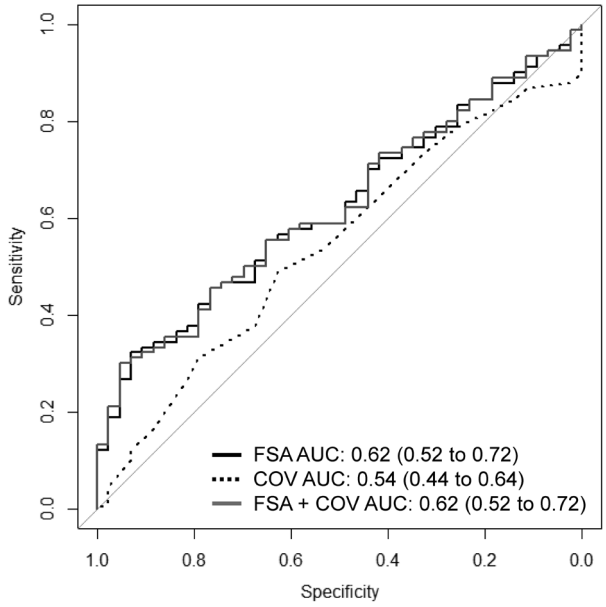

**Supplementary figure 1.** Calibration plots for models predicting incident rHOA (KL ≥ 2 or THR) using (**A**) covariates (age, gender and body mass index), (**B**) covariates and baseline KL grade, (**C**) texture variables from fractal signature analysis (FSA), and (**D**) texture variables combined with covariates and KL grade.



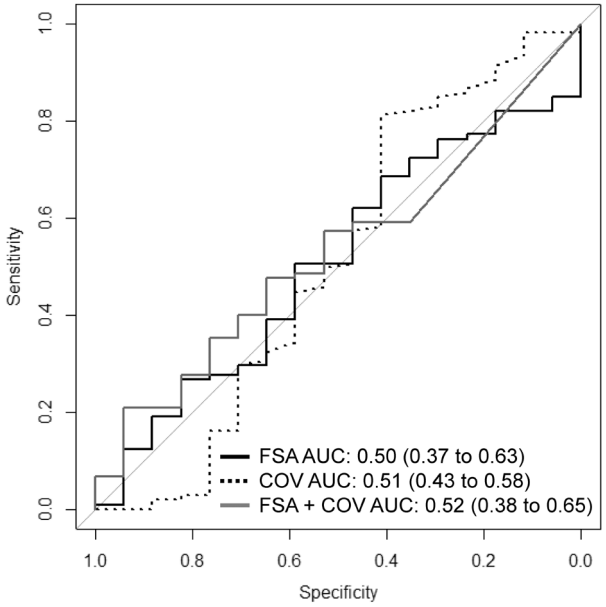

**Supplementary figure 2**. Receiver operating characteristics (ROC) curves and their respective area under the curve (AUC) values for predicting incident rHOA (KL ≥ 2 or THR) among subjects with KL0 grade at baseline using 1) texture variables from fractal signature analysis (FSA), 2) covariates (age, gender, and body mass index), and 3) texture variables combined with covariates.




**Supplementary Figure 3.** Receiver operating characteristics (ROC) curves and their respective area under the curve (AUC) values for predicting incident rHOA (KL ≥ 2 or THR) among subjects with KL1 grade at baseline using 1) texture variables from fractal signature analysis (FSA), 2) covariates (age, gender, body mass index), and 3) texture variables combined with covariates.




**Supplementary Figure 4.** Receiver operating characteristics (ROC) curves and their respective area under the curve (AUC) values for predicting incident joint space narrowing (JSN) among subjects with JSN grade 0 at baseline using 1) texture variables from fractal signature analysis (FSA), 2) covariates (age, gender, body mass index), and 3) texture variables combined with covariates.




**Supplementary Figure 5.** Receiver operating characteristics (ROC) curves and their respective area under the curve (AUC) values for predicting incident osteophytes (OST) among subjects with OST grade 0 at baseline using 1) texture variables from fractal signature analysis (FSA), 2) covariates (age, gender, body mass index), and 3) texture variables combined with covariates.

**Supplementary Table 1**. Odds ratios (95% confidence interval) of univariable covariate models to predict incident joint space narrowing (JSN grade ≥ 1 or THR).

| Predictor | Odds ratio |
|---|---|
| Age (years) | 1.04 (0.99 – 1.08) |
| Female gender | 0.60 (0.18 – 2.04) |
| Body mass index (kg/m2) | 1.03 (0.98 – 1.08) |

**Supplementary Table 2**. Odds ratios (95% confidence interval) of univariable texture models to predict incident joint space narrowing (JSN grade ≥ 1 or THR).

| ROI | FDmin | FDmax | Anglesmin | Anglesmax |
|---|---|---|---|---|
| 1 | 0.99 (0.71 – 1.38) | 0.95 (0.72 – 1.27) | 1.03 (0.98 – 1.09) | 1.04 (0.93 – 1.17) |
| 2 | 0.83 (0.60 – 1.16) | 0.88 (0.68 – 1.14) | 0.96 (0.90 – 1.02) | 1.00 (0.92 – 1.09) |
| 3 | 0.83 (0.60 – 1.16) | 0.64 (0.52 – 0.79)* | 0.93 (0.86 – 1.02) | 1.02 (0.95 – 1.10) |
| 4 | 1.05 (0.74 – 1.50) | 0.87 (0.62 – 1.21) | 1.02 (0.92 – 1.12) | 0.96 (0.88 – 1.04) |
| 5 | 1.00 (0.74 – 1.35) | 0.81 (0.63 – 1.04) | 0.96 (0.87 – 1.04) | 1.01 (0.91 – 1.12) |
| 6 | 0.78 (0.59 – 1.02) | 0.74 (0.57 – 0.98)* | 1.01 (0.95 – 1.06) | 1.01 (0.91 – 1.13) |
| 7 | 0.93 (0.65 – 1.33) | 0.83 (0.62 – 1.10) | 1.02 (0.97 – 1.06) | 0.98 (0.89 – 1.07) |
| 8 | 1.27 (0.91 – 1.78) | 0.97 (0.66 – 1.41) | 0.98 (0.94 – 1.02) | 0.95 (0.87 – 1.04) |
| 9 | 0.95 (0.70 – 1.30) | 1.20 (0.88 – 1.65) | 0.99 (0.93 – 1.06) | 1.03 (0.97 – 1.08) |
| 10 | 0.92 (0.65 – 1.29) | 0.84 (0.59 – 1.20) | 1.05 (0.96 – 1.15) | 0.98 (0.93 – 1.02) |
| 11 | 1.03 (0.69 – 1.54) | 1.01 (0.75 – 1.36) | 0.97 (0.90 – 1.05) | 0.96 (0.89 – 1.05) |
| 12 | 0.97 (0.72 – 1.30) | 1.04 (0.81 – 1.33) | 0.97 (0.91 – 1.03) | 0.98 (0.90 – 1.08) |
| 13 | 0.87 (0.68 – 1.1) | 0.95 (0.73 – 1.25) | 1.00 (0.95 – 1.04) | 1.16 (1.02 – 1.32)* |
| 14 | 1.19 (0.81 – 1.73) | 1.29 (0.92 – 1.82) | 1.04 (0.98 – 1.10) | 1.01 (0.91 – 1.11) |
| 15 | 1.43 (1.03 – 2.00)* | 1.05 (0.78 – 1.41) | 0.94 (0.90 – 0.99)* | 0.94 (0.84 – 1.05) |
| 16 | 1.06 (0.73 – 1.52) | 0.85 (0.63 – 1.15) | 1.01 (0.95 – 1.07) | 0.88 (0.81 – 0.96)* |
| 17 | 0.99 (0.67 – 1.45) | 0.81 (0.57 – 1.14) | 1.05 (0.98 – 1.13) | 0.95 (0.89 – 1.01) |
| 18 | 1.03 (0.72 – 1.48) | 0.92 (0.70 – 1.19) | 0.97 (0.89 – 1.07) | 1.03 (0.97 – 1.09) |
| 19 | 1.02 (0.71 – 1.45) | 0.86 (0.67 – 1.09) | 0.93 (0.86 – 1.02) | 1.00 (0.91 – 1.09) |
| 20 | 1.18 (0.87 – 1.59) | 1.13 (0.7 – 1.46) | 1.01 (0.91 – 1.11) | 1.04 (0.98 – 1.10) |
| 21 | 0.81 (0.61 – 1.08) | 0.96 (0.66 – 1.38) | 0.96 (0.87 – 1.06) | 0.97 (0.91 – 1.04) |
| 22 | 0.85 (0.57 – 1.27) | 0.85 (0.55 – 1.31) | 1.00 (0.91 – 1.10) | 0.99 (0.91 – 1.08) |
| 23 | 0.86 (0.60 – 1.24) | 0.83 (0.62 – 1.12) | 1.01 (0.94 – 1.09) | 0.96 (0.88 – 1.05) |

*$p < 0.05$

**Supplementary Table 2.** Continued

| ROI | FDmin | FDmax | Anglesmin | Anglesmax |
|---|---|---|---|---|
| 24 | 0.97 (0.67 – 1.40) | 0.83 (0.73 – 1.29) | 0.97 (0.91 – 1.03) | 0.96 (0.88 – 1.05) |
| 25 | 0.93 (0.67 – 1.30) | 0.74 (0.58 – 0.93)* | 0.94 (0.78 – 1.13) | 0.98 (0.87 – 1.11) |
| 26 | 1.22 (0.91 – 1.64) | 1.06 (0.78 – 1.45) | 1.03 (0.97 – 1.10) | 1.07 (0.98 – 1.17) |
| 27 | 0.98 (0.70 – 1.39) | 0.93 (0.73 – 1.41) | 0.96 (0.87 – 1.05) | 1.02 (0.97 – 1.07) |
| 28 | 0.82 (0.60 – 1.12) | 0.69 (0.52 – 0.92)* | 1.00 (0.92 – 1.09) | 1.01 (0.97 – 1.06) |
| 29 | 0.84 (0.64 – 1.10) | 0.80 (0.61 – 1.05) | 0.95 (0.87 – 1.03) | 1.00 (0.92 – 1.09) |
| 30 | 0.85 (0.64 – 1.13) | 0.87 (0.63 – 1.21) | 0.99 (0.92 – 1.06) | 1.01 (0.95 – 1.08) |
| 31 | 0.67 (0.46 – 0.98)* | 0.78 (0.57 – 1.08) | 1.02 (0.95 – 1.10) | 1.01 (0.94 – 1.10) |
| 32 | 0.93 (0.68 – 1.26) | 0.95 (0.69 – 1.30) | 1.01 (0.95 – 1.08) | 1.08 (0.96 – 1.22) |
| 33 | 1.01 (0.74 – 1.39) | 1.06 (0.76 – 1.47) | 0.97 (0.92 – 1.03) | 1.07 (0.90 – 1.27) |
| 34 | 1.01 (0.68 – 1.52) | 1.30 (0.97 – 1.76) | 0.96 (0.91 – 1.01) | 1.16 (0.81 – 1.67) |
| 35 | 0.81 (0.54 – 1.22) | 1.08 (0.74 – 1.58) | 0.97 (0.93 – 1.02) | 0.94 (0.70 – 1.27) |
| 36 | 0.91 (0.61 – 1.35) | 0.99 (0.71 – 1.39) | 1.02 (0.97 – 1.08) | 0.65 (0.42 – 1.01) |
| 37 | 1.07 (0.74 – 1.54) | 1.07 (0.70 – 1.65) | 0.98 (0.93 – 1.03) | 1.10 (0.75 – 1.61) |
| 38 | 1.08 (0.75 – 1.57) | 1.12 (0.75 – 1.68) | 0.98 (0.92 – 1.03) | 0.94 (0.82 – 1.06) |
| 39 | 1.19 (0.85 – 1.66) | 1.32 (0.94 – 1.85) | 0.97 (0.91 – 1.04) | 1.02 (0.90 – 1.15) |
| 40 | 0.81 (0.58 – 1.12) | 0.84 (0.58 – 1.22) | 0.97 (0.92 – 1.03) | 0.99 (0.92 – 1.08) |
| 41 | 0.80 (0.57 – 1.14) | 0.98 (0.67 – 1.43) | 1.00 (0.95 – 1.06) | 1.05 (0.99 – 1.11) |

*$p < 0.05$

**Supplementary Table 3.** Odds ratios (95% confidence interval) of the univariable covariate models to assess incident osteophytes (OST grade ≥ 1 or THR).

| Predictor | Odds ratio |
|---|---|
| Age (years) | 1.00 (0.95 – 1.05) |
| Female gender | 0.71 (0.32 – 1.55) |
| Body mass index (kg/m$^2$) | 1.03 (0.97 – 1.11) |

**Supplementary Table 4.** Odds ratios (95% confidence interval) of univariable texture models to assess incident osteophytes (OST grade ≥ 1 or THR).

| ROI | FDmin | FDmax | Anglesmin | Anglesmax |
|---|---|---|---|---|
| 1 | 1.02 (0.80 – 1.28) | 0.88 (0.66 – 1.15) | 0.97 (0.93 – 1.01) | 1.04 (0.97 – 1.11) |
| 2 | 1.09 (0.88 – 1.36) | 0.81 (0.62 – 1.05) | 1.01 (0.95 – 1.07) | 1.02 (0.97 – 1.08) |
| 3 | 1.06 (0.80 – 1.40) | 0.90 (0.70 – 1.16) | 0.97 (0.90 – 1.04) | 1.00 (0.95 – 1.06) |
| 4 | 1.10 (0.85 – 1.42) | 0.99 (0.77 – 1.27) | 1.00 (0.94 – 1.07) | 1.03 (0.97 – 1.09) |
| 5 | 0.94 (0.74 – 1.21) | 1.00 (0.80 – 1.25) | 0.99 (0.94 – 1.04) | 1.02 (0.94 – 1.10) |
| 6 | 0.84 (0.66 – 1.08) | 0.82 (0.66 – 1.01) | 1.02 (0.98 – 1.06) | 1.03 (0.97 – 1.10) |
| 7 | 0.99 (0.77 – 1.26) | 0.87 (0.58 – 1.11) | 0.97 (0.94 – 1.01) | 1.04 (0.99 – 1.09) |
| 8 | 0.88 (0.69 – 1.11) | 0.87 (0.67 – 1.11) | 1.01 (0.97 – 1.05) | 0.98 (0.92 – 1.05) |
| 9 | 1.00 (0.80 – 1.26) | 1.03 (0.80 – 1.31) | 0.98 (0.93 – 1.03) | 0.99 (0.95 – 1.03) |
| 10 | 0.80 (0.64 – 1.01) | 0.93 (0.75 – 1.15) | 0.99 (0.93 – 1.06) | 0.97 (0.93 – 1.02) |
| 11 | 0.90 (0.70 – 1.16) | 0.89 (0.70 – 1.12) | 1.05 (0.99 – 1.11) | 0.99 (0.95 – 1.03) |
| 12 | 0.96 (0.76 – 1.21) | 0.99 (0.79 – 1.24) | 0.98 (0.93 – 1.03) | 1.00 (0.94 – 1.06) |
| 13 | 0.92 (0.72 – 1.19) | 1.12 (0.90 – 1.40) | 1.00 (0.95 – 1.04) | 0.99 (0.93 – 1.05) |
| 14 | 0.97 (0.78 – 1.22) | 1.00 (0.80 – 1.24) | 1.02 (0.99 – 1.06) | 1.03 (0.95 – 1.11) |
| 15 | 0.92 (0.74 – 1.14) | 0.95 (0.74 – 1.22) | 1.00 (0.96 – 1.03) | 1.07 (0.99 – 1.16) |
| 16 | 0.96 (0.76 – 1.21) | 0.97 (0.75 – 1.24) | 0.99 (0.95 – 1.03) | 1.05 (0.95 – 1.16) |
| 17 | 0.83 (0.66 – 1.04) | 0.84 (0.69 – 1.04) | 1.00 (0.95 – 1.05) | 1.00 (0.94 – 1.06) |
| 18 | 0.97 (0.76 – 1.23) | 0.86 (0.69 – 1.08) | 0.99 (0.92 – 1.07) | 1.02 (0.96 – 1.08) |
| 19 | 0.87 (0.68 – 1.11) | 0.87 (0.69 – 1.09) | 1.08 (1.02 – 1.16)* | 1.01 (0.96 – 1.07) |
| 20 | 0.84 (0.66 – 1.06) | 0.88 (0.69 – 1.10) | 1.00 (0.94 – 1.06) | 1.04 (0.99 – 1.09) |
| 21 | 0.82 (0.66 – 1.04) | 0.95 (0.75 – 1.20) | 0.93 (0.85 – 1.01) | 1.02 (0.96 – 1.08) |
| 22 | 0.85 (0.67 – 1.08) | 0.84 (0.65 – 1.10) | 0.97 (0.89 – 1.06) | 0.98 (0.93 – 1.04) |
| 23 | 0.78 (0.62 – 0.99)* | 0.89 (0.67 – 1.19) | 1.02 (0.97 – 1.08) | 1.03 (0.97 – 1.10) |
| 24 | 0.98 (0.78 – 1.22) | 1.08 (0.85 – 1.38) | 1.00 (0.95 – 1.05) | 1.03 (0.96 – 1.10) |
| 25 | 1.01 (0.81 – 1.25) | 1.04 (0.83 – 1.31) | 0.89 (0.79 – 1.00)* | 0.99 (0.93 – 1.06) |
| 26 | 0.99 (0.78 – 1.25) | 0.90 (0.71 – 1.15) | 1.04 (0.97 – 1.11) | 0.98 (0.92 – 1.05) |
| 27 | 0.93 (0.74 – 1.16) | 1.02 (0.81 – 1.28) | 1.05 (0.97 – 1.14) | 0.98 (0.95 – 1.02) |
| 28 | 0.86 (0.69 – 1.07) | 0.87 (0.70 – 1.10) | 1.01 (0.95 – 1.07) | 1.01 (0.98 – 1.05) |
| 29 | 0.91 (0.73 – 1.15) | 0.95 (0.76 – 1.20) | 0.98 (0.93 – 1.03) | 0.99 (0.95 – 1.04) |
| 30 | 0.95 (0.74 – 1.20) | 1.05 (0.82 – 1.34) | 0.99 (0.94 – 1.04) | 0.99 (0.94 – 1.05) |

*$p < 0.05$

**Supplementary Table 4.** Continued

| ROI | FDmin | FDmax | Anglesmin | Anglesmax |
|---|---|---|---|---|
| 31 | 0.88 (0.71 – 1.11) | 0.83 (0.65 – 1.06) | 1.00 (0.96 – 1.04) | 1.01 (0.95 – 1.06) |
| 32 | 1.05 (0.82 – 1.35) | 0.88 (0.71 – 1.08) | 0.99 (0.94 – 1.04) | 0.99 (0.91 – 1.08) |
| 33 | 1.01 (0.79 – 1.29) | 1.10 (0.86 – 1.42) | 1.02 (0.99 – 1.06) | 0.99 (0.80 – 1.22) |
| 34 | 1.10 (0.84 – 1.45) | 1.10 (0.87 – 1.41) | 0.99 (0.96 – 1.03) | 1.08 (0.92 – 1.27) |
| 35 | 1.00 (0.79 – 1.27) | 1.10 (0.84 – 1.45) | 0.99 (0.96 – 1.03) | 0.96 (0.77 – 1.19) |
| 36 | 1.01 (0.82 – 1.24) | 0.95 (0.75 – 1.21) | 1.06 (1.02 – 1.11)* | 0.91 (0.73 – 1.13) |
| 37 | 1.09 (0.86 – 1.39) | 0.79 (0.62 – 0.99)* | 1.02 (0.99 – 1.06) | 0.85 (0.58– 1.23) |
| 38 | 1.13 (0.86 – 1.48) | 1.05 (0.82 – 1.34) | 1.01 (0.98 – 1.05) | 0.87 (0.71 – 1.05) |
| 39 | 1.37 (0.97 – 1.93) | 1.11 (0.87 – 1.42) | 0.99 (0.95 – 1.04) | 0.92 (0.83 – 1.01) |
| 40 | 1.17 (0.90 – 1.52) | 1.03 (0.82 – 1.28) | 0.97 (0.92 – 1.03) | 1.01 (0.96 – 1.06) |
| 41 | 0.87 (0.67 – 1.13) | 0.89 (0.69 – 1.15) | 1.01 (0.96 – 1.05) | 1.00 (0.95 – 1.04) |

*$p < 0.05$

**Supplementary Table 1.** Variables in the final elastic net model to predict incident rHOA (KL ≥ 2 or THR).

| Variable | Coefficient | Variable | Coefficient |
|---|---|---|---|
| Intercept | -0.47 | Anglesmin ROI7 | 0.00 |
| FDmin ROI1 | -0.04 | Anglesmin ROI15 | -0.01 |
| FDmin ROI2 | -0.04 | Anglesmin ROI16 | -0.01 |
| FDmin ROI6 | -0.05 | Anglesmin ROI18 | -0.01 |
| FDmin ROI10 | -0.03 | Anglesmin ROI22 | -0.01 |
| FDmin ROI15 | 0.01 | Anglesmin ROI26 | 0.00 |
| FDmin ROI19 | -0.08 | Anglesmin ROI28 | 0.01 |
| FDmin ROI24 | 0.01 | Anglesmin ROI29 | 0.00 |
| FDmin ROI27 | 0.00 | Anglesmin ROI36 | 0.00 |
| FDmin ROI28 | -0.03 | Anglesmin ROI38 | 0.00 |
| FDmin ROI29 | -0.05 | Anglesmin ROI41 | 0.00 |
| FDmin ROI32 | 0.03 | Anglesmax ROI1 | -0.02 |
| FDmin ROI33 | 0.01 | Anglesmax ROI3 | 0.01 |
| FDmin ROI40 | -0.02 | Anglesmax ROI4 | 0.00 |
| FDmin ROI41 | -0.08 | Anglesmax ROI6 | -0.01 |
| FDmax ROI2 | -0.05 | Anglesmax ROI9 | 0.00 |
| FDmax ROI3 | -0.04 | Anglesmax ROI10 | -0.01 |
| FDmax ROI6 | -0.03 | Anglesmax ROI11 | -0.01 |
| FDmax ROI14 | 0.05 | Anglesmax ROI12 | -0.02 |
| FDmax ROI15 | 0.01 | Anglesmax ROI14 | -0.01 |
| FDmax ROI19 | -0.03 | Anglesmax ROI15 | 0.00 |
| FDmax ROI21 | -0.01 | Anglesmax ROI17 | -0.02 |
| FDmax ROI22 | -0.09 | Anglesmax ROI20 | 0.00 |
| FDmax ROI23 | -0.04 | Anglesmax ROI23 | -0.01 |
| FDmax ROI24 | 0.04 | Anglesmax ROI24 | -0.01 |
| FDmax ROI30 | 0.01 | Anglesmax ROI27 | -0.01 |
| FDmax ROI33 | 0.05 | Anglesmax ROI28 | 0.00 |
| FDmax ROI34 | 0.06 | Anglesmax ROI29 | 0.00 |
| FDmax ROI35 | 0.04 | Anglesmax ROI31 | 0.00 |
| FDmax ROI37 | 0.04 | Anglesmax ROI32 | 0.00 |
| FDmax ROI38 | 0.07 | Anglesmax ROI40 | 0.01 |
| Anglesmin ROI1 | 0.01 | Age | 0.03 |
| Anglesmin ROI3 | 0.00 | Baseline KL | 0.69 |
| Anglesmin ROI4 | 0.00 | Gender | -0.22 |
| Anglesmin ROI5 | -0.01 | | |

## References


1. Hirvasniemi J, Thevenot J, Kokkonen HT, Finnilä MA, Venäläinen MS, Jämsä T, et al. Correlation of Subchondral Bone Density and Structure from Plain Radiographs with Micro Computed Tomography Ex Vivo. Annals of Biomedical Engineering 2016; 44: 1698-709.
2. Le Corroller T, Pithioux M, Chaari F, Rosa B, Parratte S, Maurel B, et al. Bone texture analysis is correlated with three-dimensional microarchitecture and mechanical properties of trabecular bone in osteoporotic femurs. Journal of Bone and Mineral Metabolism 2013; 31: 82-8.
3. Ollivier M, Le Corroller T, Blanc G, Parratte S, Champsaur P, Chabrand P, et al. Radiographic bone texture analysis is correlated with 3D microarchitecture in the femoral head, and improves the estimation of the femoral neck fracture risk when combined with bone mineral density. European Journal of Radiology 2013; 82: 1494-8.
4. Ranjanomennahary P, Ghalila SS, Malouche D, Marchadier A, Rachidi M, Benhamou CL, et al. Comparison of radiograph-based texture analysis and bone mineral density with three-dimensional microarchitecture of trabecular bone. Medical Physics 2011; 38: 420-8.
5. Steines D, Liew SW, Arnaud C, Vargas-Voracek R, Nazarian A, Müller R, et al. Radiographic trabecular 2D and 3D parameters of proximal femoral bone cores correlate with each other and with yield stress. Osteoporosis International 2009; 20: 1929-38.
6. Bankman IN, Spisz TS, Pavlopoulos S. Two-Dimensional Shape and Texture Quantification. In: Handbook of Medical Image Processing and Analysis (Second Edition), Bankman IN Ed. Burlington: Academic Press 2009:261-77.
7. Pulkkinen P, Saarakkala S, Nieminen MT, Jämsä T. Standard radiography: untapped potential in the assessment of osteoporotic fracture risk. European radiology 2013; 23: 1375-82.
8. Thevenot J, Hirvasniemi J, Pulkkinen P, Maatta M, Korpelainen R, Saarakkala S, et al. Assessment of risk of femoral neck fracture with radiographic texture parameters: a retrospective study. Radiology 2014; 272: 184-91.
9. Buckland-Wright C. Subchondral bone changes in hand and knee osteoarthritis detected by radiography. Osteoarthritis and Cartilage 2004; 12: S10-9.
10. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Physics in Medicine & Biology 1991; 36: 709-22.
11. Lynch JA, Hawkes DJ, Buckland-Wright JC. A robust and accurate method for calculating the fractal signature of texture in macroradiographs of osteoarthritic knees. Medical Informatics 1991; 16: 241-51.
12. Hirvasniemi J, Thevenot J, Guermazi A, Podlipská J, Roemer FW, Nieminen MT, et al. Differences in tibial subchondral bone structure evaluated using plain radiographs between knees with and without cartilage damage or bone marrow lesions-the Oulu Knee Osteoarthritis study. European radiology 2017; 27: 4874-82.

13. Hirvasniemi J, Thevenot J, Multanen J, Haapea M, Heinonen A, Nieminen MT, et al. Association between radiography-based subchondral bone structure and MRI-based cartilage composition in postmenopausal women with mild osteoarthritis. Osteoarthritis and Cartilage 2017; 25: 2039-46.

14. Janvier T, Jennane R, Toumi H, Lespessailles E. Subchondral tibial bone texture predicts the incidence of radiographic knee osteoarthritis: data from the Osteoarthritis Initiative. Osteoarthritis and Cartilage 2017; 25: 2047-54.

15. Janvier T, Jennane R, Valery A, Harrar K, Delplanque M, Lelong C, et al. Subchondral tibial bone texture analysis predicts knee osteoarthritis progression: data from the Osteoarthritis Initiative: tibial bone texture & knee OA progression. Osteoarthritis and Cartilage 2017; 25: 259-66.

16. Kraus VB, Feng S, Wang S, White S, Ainslie M, Brett A, et al. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. Arthritis and Rheumatism 2009; 60: 3711-22.

17. Kraus VB, Feng S, Wang S, White S, Ainslie M, Le Graverand MPH, et al. Subchondral bone trabecular integrity predicts and changes concurrently with radiographic and magnetic resonance imaging-determined knee osteoarthritis progression. Arthritis and Rheumatism 2013; 65: 1812-21.

18. Kraus VB, Collins JE, Charles HC, Pieper CF, Whitley L, Losina E, et al. Predictive validity of radiographic trabecular bone texture in knee osteoarthritis: the Osteoarthritis Research Society International/Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis & Rheumatology 2018; 70: 80-7.

19. detected by plain radiography is associated with an increased risk of knee replacement in patients with osteoarthritis: a 6 year prospective follow up study. Osteoarthritis and Cartilage 2014; 22: 71-5.

20. Podsiadlo P, Nevitt MC, Wolski M, Stachowiak GW, Lynch JA, Tolstykh I, et al. Baseline trabecular bone and its relation to incident radiographic knee osteoarthritis and increase in joint space narrowing score: directional fractal signature analysis in the MOST study. Osteoarthritis and Cartilage 2016; 24: 1736-44.

21. Woloszynski T, Podsiadlo P, Stachowiak GW, Kurzynski M, Lohmander LS, Englund M. Prediction of progression of radiographic knee osteoarthritis using tibial trabecular bone texture. Arthritis and Rheumatism 2012; 64: 688-95.

22. Papaloucas CD, Ward RJ, Tonkin CJ, Buckland-Wright C. Cancellous bone changes in hip osteoarthritis: a short-term longitudinal study using fractal signature analysis. Osteoarthritis and Cartilage 2005; 13: 998-1003.

23. Barbour KE, Murphy LB, Helmick CG, Hootman JM, Renner JB, Jordan JM. Bone Mineral Density and the Risk of Hip and Knee Osteoarthritis: The Johnston County Osteoarthritis Project. Arthritis care & research 2017; 69: 1863-70.

24. Betancourt MCC, Van der Linden JC, Rivadeneira F, Rozendaal RM, Zeinstra SMB, Weinans H, et al. Dual energy x-ray absorptiometry analysis contributes to the prediction of hip osteoarthritis progression. Arthritis Research & Therapy 2009; 11: R162.

25. Hardcastle SA, Dieppe P, Gregson CL, Hunter D, Thomas GE, Arden NK, et al. Prevalence of radiographic hip osteoarthritis is increased in high bone mass. Osteoarthritis and Cartilage 2014; 22: 1120-8.

26. Javaid MK, Lane NE, Mackey DC, Lui LY, Arden NK, Beck TJ, et al. Changes in proximal femoral mineral geometry precede the onset of radiographic hip osteoarthritis: the study of osteoporotic fractures. Arthritis & Rheumatology 2009; 60: 2028-36.

27. Wesseling J, Boers M, Viergever MA, Hilberdink WKHA, Lafeber FPJG, Dekker J, et al. Cohort profile: cohort hip and cohort knee (CHECK) study. International journal of epidemiology 2014; 45: 36-44.

28. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Annals of the Rheumatic Diseases 1957; 16: 494-502.

29. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis and Cartilage 2007; 15: A1-A56.

30. Agricola R, Reijman M, Bierma-Zeinstra SM, Verhaar JA, Weinans H, Waarsing JH. Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: a prospective cohort study (CHECK). Osteoarthritis and Cartilage 2013; 21: 559-64.

31. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of statistical software 2010; 33: 1-22.

32. Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 2005; 67: 301-20.

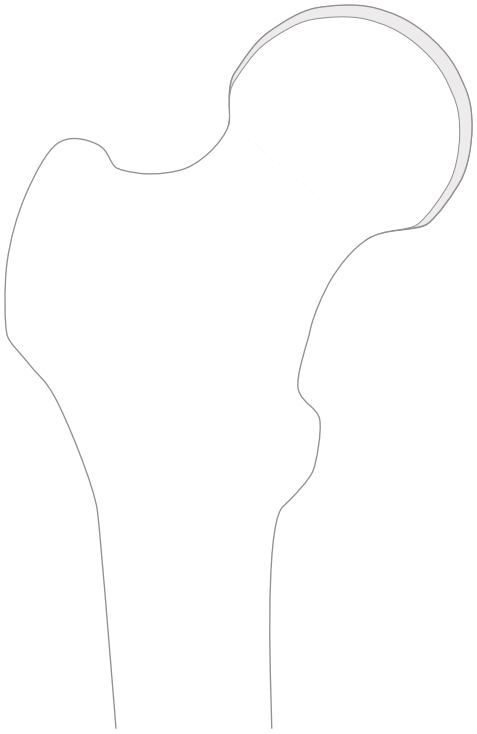
33. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology (Cambridge, Mass.) 2010; 21: 128.

34. Kuhn M. Building predictive models in R using the caret package. Journal of statistical software 2008; 28: 1-26.

35. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S to analyze and compare ROC curves. BMC bioinformatics 2011; 12: 77.

36. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. Journal of clinical epidemiology 2016; 74: 167-76.

37. Boniatis IS, Costaridou LI, Cavouras DA, Panagiotopoulos EC, Panayiotakis GS. Quantitative assessment of hip osteoarthritis based on image texture analysis. The British journal of radiology 2006; 79: 232-8.



# CHAPTER 7

# Is hip morphology a risk factor for or a consequence of hip osteoarthritis? New insights on the chicken and egg dilemma using ten years of follow-up in CHECK

*Annals of Rheumatic Disease*
*Submitted for publication*

Willem Paul Gielis, Roel J. H. Custers, Rintje Agricola, Wouter Foppen, Claudia Lindner, Timothy F. Cootes, Pim A. de Jong, Harrie Weinans

## Abstract

### Background and aim

To investigate if morphological variations of the hip joint and/or a change over time in these morphologies are associated with progression of osteoarthritis (OA).

### Patients and methods

We quantified hip morphology on pelvic radiographs at five time points during ten-year follow-up in CHECK (n=1.002), a cohort focused on hip/knee OA. From a Statistical Shape Model, 12 major hip shape modes (HSMs) were identified. We tested if HSMs and/or changes over time in HSMs (ΔHSMs) predicted a subsequent period of radiographic OA activity; joint space narrowing (JSN (decrease in joint space width)) and osteophytosis (osteophyte growth). These HSMs and ΔHSMs were considered a risk factor. Additionally, we tested which ΔHSMs are associated with simultaneous JSN and/or osteophytosis and were considered a result of OA.

### Results

ΔHSMs describing a decreased size of the lesser trochanter were associated with simultaneous JSN and/or osteophytosis and thus considered a result of OA. HSMs and/or ΔHSMs describing pincer morphology (acetabular overcoverage, head medialization, shorter neck), cam morphology (Aspherical femoral head), decreased neck-shaft angle and small greater trochanter predicted subsequent JSN and/or osteophytosis and are considered a risk factor for OA. These ΔHSMs were also associated with simultaneous JSN and/or osteophytosis. Thus, OA activity may increase morphological features further driving the risk for disease progression.

### Conclusion

We showed that HSMs and ΔHSMs can identify different morphological OA phenotypes. These can be a result of OA disease activity, be dynamic risk factors for OA, but are mainly both and interact with OA to drive further progression as a mediator.

## Introduction

Osteoarthritis (OA) is a cascade of events involving different joint tissues. Currently, no disease-modifying drug for OA exists. In early OA, pathophysiological pathways vary and different phenotypes have been proposed.[1,2] Each phenotype may benefit from different treatments. Clinical trials in OA historically failed partially because OA was regarded as a uniform disease, rather than selecting specific phenotypes for the treatment tested. In the hip, clear morphological phenotypes exist. Some of these, like cam morphology and acetabular dysplasia, are apparent before the onset of hip OA.[3–6] Genetic predisposition may cause an increased risk for OA due to morphological variations.[7] Therefore, some morphological variations are most likely causative risk factors for hip OA. The osteoarthritic process also causes bony morphological changes. Kellgren and Lawrence described radiographic features of OA as osteophyte growth, subchondral sclerosis, and 'bone end deformity'.[8] Osteophytes and sclerosis are clearly defined, but bone end deformities can be subtle and hard to define. Recently, a study quantified these subtle bone end deformities using Statistical Shape Modeling (SSM) and found significant morphological changes in osteoarthritic hips during a one-year follow-up.[9] Therefore, hip morphology can be both a cause and result of OA. To the best of our knowledge, no previous studies aimed to differentiate which morphological variations are a cause and which are a result of OA. The ability to differentiate between them will help to define phenotypes in hip OA with a large morphological aspect. These phenotypes may benefit from specific non-surgical or surgical treatments, including specialized physical therapy, arthroscopic reduction of morphological abnormalities, or acetabular or femoral osteotomies.[10–13] Morphological changes that result from OA could function as sensitive outcome markers for clinical trials focusing on specific morphological phenotypes of hip OA.[9] Therefore, in this study, we aimed to investigate morphological variations of the hip joint and to explore whether these are a cause for or a result of OA.

## Methods

### Participants

We used data from Cohort Hip and Cohort Knee (CHECK), a prospective cohort initiated to study early knee and/or hip OA. CHECK was described in detail previously.[14] In short, 1002 participants, aged between 45 and 65 years, with a first episode of clinically relevant knee and/or hip pain were included between October 2002 and September 2005 and followed for 10 years. Participants had never or

only recently (<6 months) visited a physician for these symptoms. Participants were recruited by general practitioners and/or via local media and included in ten centers throughout the Netherlands. Exclusion criteria included pathology other than OA causing hip/knee symptoms, co-morbidities precluding follow-up or clinical examinations, and inability to understand Dutch. Standardized AP pelvic weight-bearing radiographs, with the feet alongside a wedge in 15° internal rotation were collected at: recruitment (T0), two, five, eight and ten-year follow-up for both hips. For each participant both hips were included. The period between two time points for a single hip was deemed a "hip-period". Each hip with adequate radiographs for all timepoints contributes four hip-periods: from T0 to T2, from T2 to T5, from T5 to T8, and from T8 to T10.

### Assessment of morphology

We used SSM with 75 landmarks to annotate radiographs using BoneFinder® (Manchester, United Kingdom).[15,16] The landmarks do not include evident osteophytes. Radiographs with a prosthesis and/or previous hip fracture, or missing anatomical landmarks (caused by the field of view or an inadequate signal-to-noise-ratio) were excluded from the SSM. We included a total of 7,682 radiographs over five time-points to build the SSM (**supplementary table**). Based on landmark positions, an SSM quantifies all morphological variations present in the studied population sample. Using principle component analysis, all morphological variations are "bundled up" in so-called hip shape modes (HSMs). Each HSM describes distinct morphological variations from the population mean. Each HSM is statistically independent of other HSMs. A value per HSM is defined for each hip at each time-point using ± standard deviations (SD). The value describes the deviation from the mean morphology in that specific HSM. Each HSM describes a part of the total morphological variation, with the percentage described decreasing as HSM numbers increase (e.g. HSM 1 describes the largest part of the total variation). We used all HSMs that described ≥1% of the total morphological variation, resulting in 12 HSMs that together described 86% of the total morphological variation.

For one hip we regarded each period between two time-points as a separate entity, a hip-period, with a baseline and a follow-up moment. The HSM values for the baseline moment define *morphological variations*. We subtracted the HSM value of the follow-up moment from the value at the baseline moment, resulting in ΔHSM between the two time-points defining the *change in morphology*. Thus, for each hip-period both a *morphological variation* (HSM) and a *change in morphology* (ΔHSM) are available. The HSM and ΔHSM could be defined for 5,516 paired radiographs **(Supplementary Table)**.

### Assessment of radiographic hip OA

The CHECK database provides radiographic features of hip OA scored at T0, T2, T5, T8, and T10 using the atlas by Altman et al.[17] Trained readers scored sets of radiographs aware of the chronological order. We selected joint space width (JSW) and osteophytosis to define radiographic disease activity. JSW is a radiographic measure of the space between the femur and acetabulum, reflecting cartilage thickness, and may act as surrogate outcome marker for future hip arthroplasty.[18] Two JSW scores were available: medial and superior. Osteophytes are bone spurs that grow along joint margins, and are also indicative of the future need for hip arthroplasty.[19] Four osteophyte scores were available: medial and superior for both acetabulum and femur. All scores ranged from 0 (normal) to 3 (severe). Reliability for scoring JSW progression and osteophytes cross-sectionally were reported as almost perfect.[20] We defined an increase of ≥1 for JSW score (medial or superior) within a hip-period *as joint space narrowing (JSN)*. We defined an increase of ≥1 for osteophytes score (medial/superior acetabulum or femur) within a hip-period as *osteophytosis*. If a hip had a baseline score 3 for either JSW or osteophytes in any area, or a joint replacement in situ, it was excluded.

### Statistical analyses

Frequencies are given as number (percentage), continuous variables as mean (±SD). We used logistic mixed-effect model regression to test the association between independent variables (HSM and ΔHSM) and dependent variables (*JSN* or *osteophytosis*). Mixed-effect models correct for the correlation between multiple measurements within a participant (left and right, different time-points) and missing data. A model was built using values for *morphological variations* and *change in morphology* described by a single HSM. Corrections were made for sex, age, and BMI. A p-value of $<0.05$ was deemed statistically significant. We performed two analyses **(Figure 1)**:

1) "Morphological variations and changes in morphology predict radiographic disease activity." To test this hypothesis, we used HSM and ΔHSM during a hip-period as predictor variables and *JSN* or *osteophytosis* in the subsequent hip-period as outcome variables. For example, when the morphological parameters are determined for between T2 and T5, the change in radiographic features of OA will be determined between T5 and T8. In this analysis, the ΔHSM can be seen as a risk factor for *JSN* and/or *osteophytosis*.

2) "Radiographic disease activity results in simultaneous changes in morphology." To test this hypothesis, we determined HSM and ΔHSM during each hip-period as predictor variables and *JSN* or *osteophytosis* as outcome variables. In this analysis, the ΔHSM can be seen as a result of radiographic disease activity.

*Morphological variations* described by HSMs are included in both analyses and should be seen as a risk factor for *JSN* and/or *osteophytosis* during follow-up.

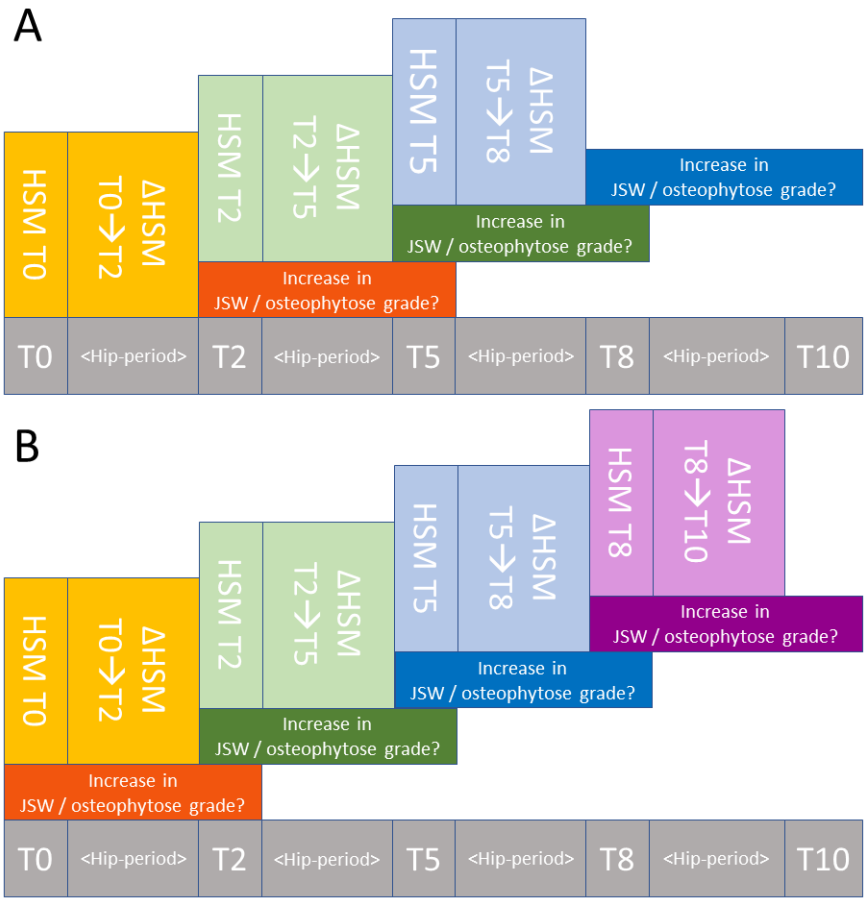



**Figure 1.** Diagram A (analysis 1) and diagram B (analysis 2) show a visualization of how predictors Hip Shape Modes (HSM) and change in Hip Shape Modes (ΔHSM) (morphological parameters) and outcome parameters (*JSN* and *osteophytosis*) are coupled in time within the analyses. Predictors and outcome parameters that are coupled in the analyses are coupled in the figure using colors, light for predictors, dark for outcome parameters. In the first analysis (**A**), the predictors are determined within the hip-period prior to the hip-period in which the outcome parameters are determined. In the second analysis (**B**), the predictors and outcome parameters are determined within the same hip-period.

## Results

Of the 1,002 participants at recruitment, 145 participants dropped out of the study (14%); 31 at two-year; 51 at five-year; 41 at eight-year and 22 at ten-year follow-up. 792 (79%) of the 1,002 participants were female. Mean age and BMI at baseline were 56 (±5.2) years and 26.2 (±4) kg/m$^2$, respectively. An increase of ≥1 score for JSN was present in 1,549 of 6,526 (24%) hip-periods, and for osteophytosis in 2,798 of 6,144 (35%) hip-periods **(supplementary table)**.

### Morphological variations described by HSMs

The HSMs are visualized by depicting hip joints that fit +2.5 SD and -2.5 SD for each HSM **(Table 1 and Figure 2)**. This allows visual interpretation of morphological variations described by each HSM. Below, we describe the morphological variation described by each HSM based on the +2.5 SD HSM value (**blue in Figure 2**).

**Table 1.** Textual description of hip shape modes (HSMs).

| HSM | % of total shape variation | Description |
|---|---|---|
| 1 | 27 | Increased neck-shaft angle in the femur; larger lesser trochanter; narrower pelvis. |
| 2 | 17 | Smaller greater and lesser trochanter with a different shape; smaller frontal aspect of the pelvis |
| 3 | 11 | More spherical femoral head; smaller aspect of the proximal femur (including both trochanter complexes); narrower femoral neck; more medially placed femoral head; smaller aspect foramen obturatorium |
| 4 | 11 | Longer femoral neck; decreased acetabular coverage superiorly |
| 5 | 4 | Shallower acetabulum; less acetabular coverage; proximal position of lesser trochanter; slightly more distal greater trochanter |
| 6 | 4 | Cam-like asphericity of the superior femoral head; decreased Wiberg angle; shorter neck; larger greater trochanter |
| 7 | 2 | Smaller femur, primarily evident in the inferior femoral head, calcar and shaft |
| 8 | 2 | Less pointy aspect of the lateral edge of the acetabular roof |
| 9 | 2 | A lateral position of the acetabulum within the pelvis; decreased coverage superior and posteriorly; narrower aspect of greater and lesser trochanter; shorter neck; valgus neck-shaft angle |
| 10 | 1 | Concave increase of posterosuperior coverage of the acetabulum; a smaller lesser trochanter |
| 11 | 1 | Longer greater trochanter |
| 12 | 1 | Decreased posterior coverage; slimmer inferior head; slightly smaller lesser trochanter |

Each HSM is described based on a +2.5 SD mode value increase (Blue in figure 2)

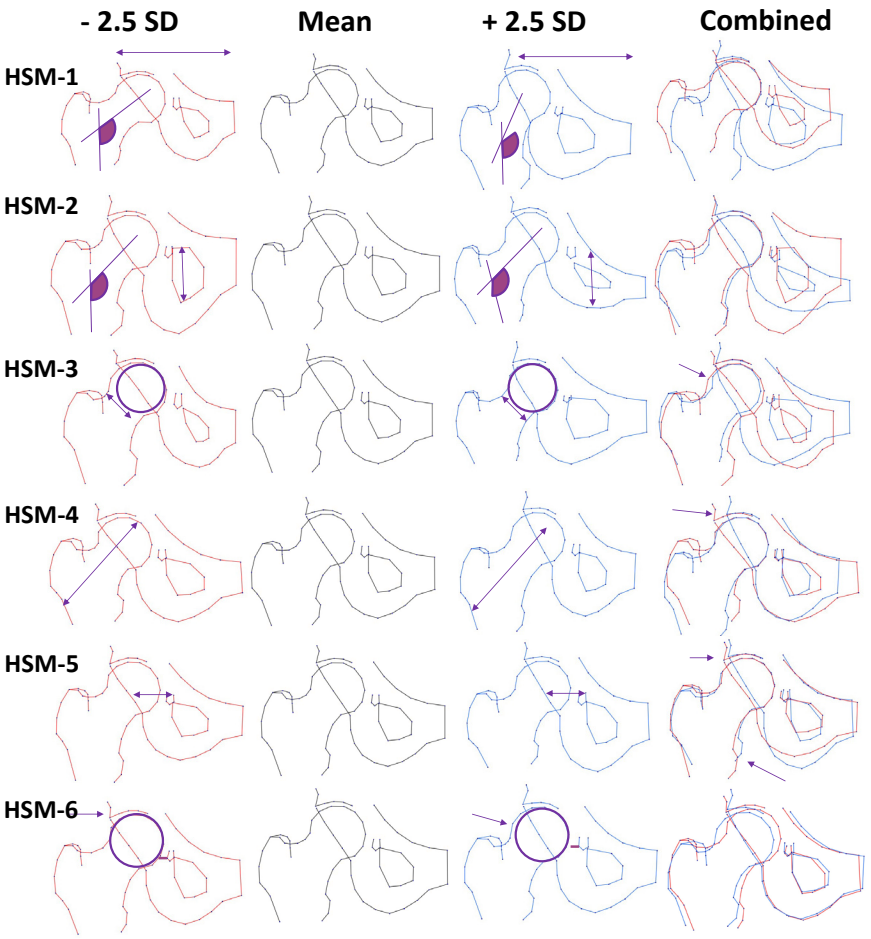



**Figure 2.** The SSM produced 12 hip shape modes (HSMs) that each describes more than 1% of the total shape variation. These HSMs are depicted from left to right as -2.5 SD, mean, +2.5 SD, and the combined +/- 2.5 SD.

### JSN

HSM-3 predicted *JSN* at two/three-year follow-up. HSM-5 and HSM-11 values predicted *JSN* not only at two/three-year follow-up, but also at five-year follow-up **(Table 2)**. ΔHSM-2, ΔHSM-4, ΔHSM-5, and ΔHSM-7 predicted *JSN* in the subsequent hip-period **(Table 3)**. ΔHSM-3, ΔHSM-6, ΔHSM-8, ΔHSM-11, and ΔHSM-12 are associated with *JSN* during the same hip-period.

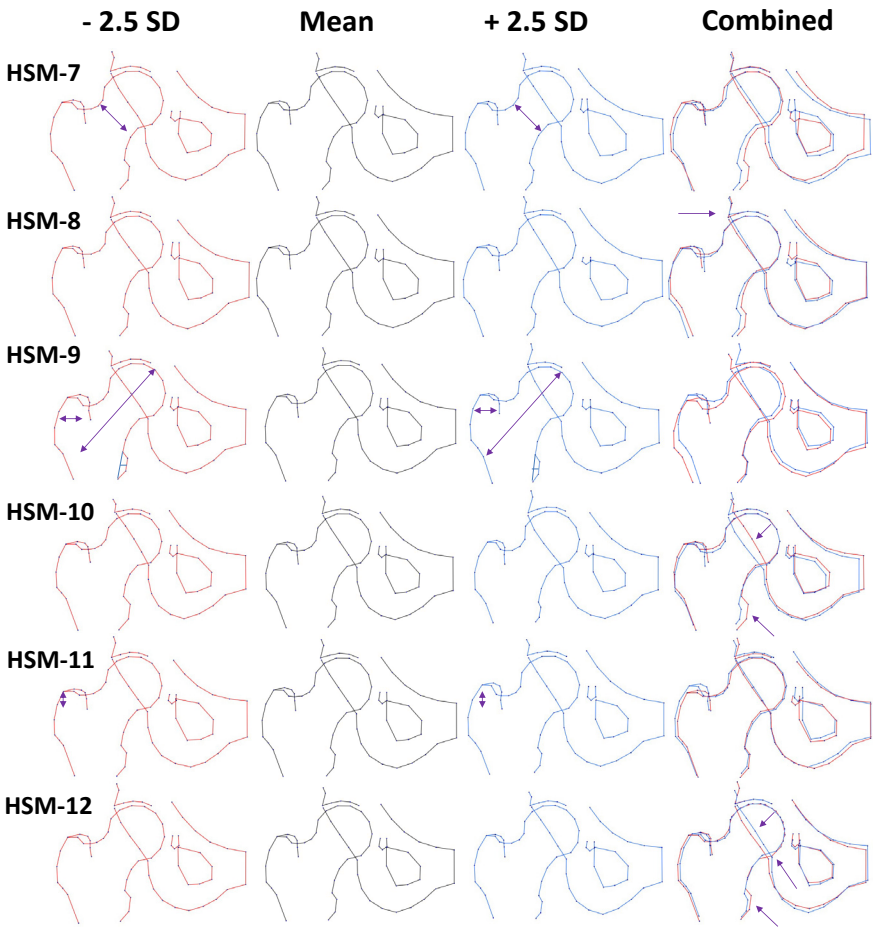



**Figure 2.** Continued

### Osteophytosis

HSM-1 predicted *osteophytosis* at two/three-year follow-up **(Table 2)**. ΔHSM-4, ΔHSM-6, and ΔHSM-9 predicted *osteophytosis* in the subsequent hip-period **(Table 3)**. ΔHSM-1, ΔHSM-2, ΔHSM-3, ΔHSM-5, ΔHSM-7, ΔHSM-8, ΔHSM-10, and ΔHSM-12 are associated with *osteophytosis* during the same hip-period.

**Table 2.** Association between the baseline hip shape modes (HSM) and joint space narrowing (JSN) / osteophytosis

| | OA activity in subsequent period (predictor) | | | OA activity in same period (result) | | |
|---|---|---|---|---|---|---|
| | **OR** | **95%-CI** | **p-value** | **OR** | **95%-CI** | **p-value** |
| **JSN** | | | | | | |
| HSM-3 | | | | 0.91 | 0.83 – 0.99 | 0.032 |
| HSM-5 | 0.90 | 0.82 – 0.98 | 0.020 | 0.88 | 0.81 – 0.94 | 0.001 |
| HSM-11 | 0.90 | 0.82 – 0.99 | 0.027 | 0.90 | 0.83 – 0.98 | 0.010 |
| **Osteophytosis** | | | | | | |
| HSM-1 | | | | 0.93 | 0.86 – 0.99 | 0.031 |

Odd ratios represent the change in odds for JSN and osteophytosis per 1 SD in ΔHSM

**Table 3.** Association between the delta hip shape modes (ΔHSM) and joint space narrowing (JSN) / osteophytosis

| | OA activity in subsequent period (predictor) | | | OA activity in same period (result) | | |
|---|---|---|---|---|---|---|
| | **OR** | **95%-CI** | **p-value** | **OR** | **95%-CI** | **p-value** |
| **JSN** | | | | | | |
| ΔHSM-2 | 1.11 | 1.01 – 1.22 | 0.037 | | | |
| ΔHSM-3 | | | | 0.88 | 0.82 – 0.95 | 0.001 |
| ΔHSM-4 | 0.89 | 0.81 – 0.97 | 0.011 | | | |
| ΔHSM-5 | 0.90 | 0.82 – 0.99 | 0.023 | | | |
| ΔHSM-6 | | | | 1.08 | 1.00 – 1.17 | 0.040 |
| ΔHSM-7 | 1.14 | 1.03 – 1.25 | 0.009 | | | |
| ΔHSM-8 | | | | 0.85 | 0.78 – 0.93 | <0.001 |
| ΔHSM-11 | | | | 0.87 | 0.80 – 0.94 | 0.001 |
| ΔHSM-12 | | | | 0.89 | 0.82 – 0.97 | 0.005 |
| **Osteophytosis** | | | | | | |
| ΔHSM-1 | | | | 0.89 | 0.84 – 0.95 | <0.001 |
| ΔHSM-2 | | | | 1.08 | 1.01 – 1.15 | 0.024 |
| ΔHSM-3 | | | | 0.90 | 0.85 – 0.96 | 0.002 |
| ΔHSM-4 | 0.74 | 0.59 – 0.91 | 0.005 | | | |
| ΔHSM-5 | | | | 0.93 | 0.88 – 0.99 | 0.029 |
| ΔHSM-6 | 1.25 | 1.07 – 1.47 | 0.005 | | | |
| ΔHSM-7 | | | | 1.16 | 1.09 – 1.24 | <0.001 |
| ΔHSM-8 | | | | 0.79 | 0.73 – 0.85 | <0.001 |
| ΔHSM-9 | 0.85 | 0.75 – 0.98 | 0.020 | | | |
| ΔHSM-10 | | | | 0.92 | 0.87 – 0.99 | 0.019 |
| ΔHSM-12 | | | | 0.91 | 0.85 – 0.97 | 0.006 |

Odd ratios represent the change in odds for JSN and osteophytosis per 1 SD in ΔHSM

## Discussion

We investigated morphological variations of the hip and their temporal relation to radiographic hip OA. *Changes in morphology* described by ΔHSM-4 and ΔHSM-9 predicted *JSN* and/or *osteophytosis* in the subsequent period and may be regarded as a risk factor. The *changes in morphology* described by ΔHSM-1, ΔHSM-2, ΔHSM-3, ΔHSM-5, ΔHSM-6, ΔHSM-7, and ΔHSM-11, were associated with simultaneous *JSN* and/or *osteophytosis*, but HSM-1, ΔHSM-2, HSM-3, HSM-5, ΔHSM-5, ΔHSM-6, ΔHSM-7, and HSM-11 also predict *JSN* and/or *osteophytosis*. These HSMs characterize *changes in morphology* that result from osteoarthritic activity. However, these HSMs also describe morphological risk factors for OA. Therefore, aggravation of morphological characteristics described by these HSMs may be a mediator of OA progression. Below we discuss the morphological variations described by the (Δ)HSMs and their temporal relation with hip OA in more detail.

Interpreting the association between the HSMs and OA poses some challenges. The SSM bundles morphological variations in an unsupervised manner. Therefore, parts of previously described and well-known morphological variations, such as the neck-shaft angle, cam morphology or acetabular coverage, may end up in various HSMs. Additionally, the relationship between the HSMs and OA might not be linear. For example, HSM-6 shows increased acetabular coverage (fitting to pincer morphology) in low HSM values and decreased acetabular coverage (fitting to acetabular dysplasia) in high HSM values. Lastly, a perfect differentiation between true morphology and osteophyte growth or positioning artefacts is not possible on radiographs. To try to overcome these issues, we looked for similarity between previously described morphological variations and the HSMs produced by our SSM.

### Neck-shaft angle

In our study, low baseline HSM-1 predicted *osteophytosis* at two/three-year follow-up. A positive ΔHSM-2 (decrease in neck-shaft angle) also predicts *JSN* in the subsequent period. ΔHSM-1 and ΔHSM-2 changed simultaneously with *osteophytosis* in such way that the neck-shaft angle decreased.

Recently, Abdulrahim et al. found an association between high and low neck-shaft angles and hip OA in a case-control study (n=1674).[21] Castaño et al. used baseline HSMs and predefined geometry measures to predict incident hip OA at 6.5 years follow-up in 688 participants.[22] Neck-shaft angle was no linear risk factor for hip OA. Barr et al. studied changes in hip morphology (ΔHSM) resulting from OA

between baseline, 6-month and 12-month follow-up (n=62).[9] They found the neck-shaft angle to decrease in OA hips compared to controls. Combining our findings with previous publications suggests that a decrease in neck-shaft angle is a dynamic risk factor for hip OA, aggravated by the OA process itself. Orthopedic surgeons may account for this when planning to restore natural offset in hip arthroplasty.[23] Future studies may incorporate 3D imaging to elaborate which patients with this morphological phenotype would benefit from osteotomies.[11–13]

### Acetabular coverage

In our study, a negative ΔHSM-4 (increased superior acetabular coverage; shortening of the femoral neck) predicted a period of *JSN* and/or *osteophytosis*. A negative ΔHSM-9 (medialization of the femoral head; shortening of the femoral neck; increase acetabular coverage) predicted *osteophytosis*. A negative ΔHSM-5 (deep acetabulum resulting in more coverage) predicted *JSN* and was associated with simultaneous *osteophytosis*. Additionally, low baseline HSM-5 values predicted *JSN* at 5-year follow-up. A negative ΔHSM-8 (apparent superior acetabular osteophyte growth) and negative ΔHSM-12 (increase of medio-inferior coverage) were associated with simultaneous *JSN* and *osteophytosis*.

It was described previously that acetabular undercoverage (dysplasia) may cause OA at later age.[4,5,21,24,25] However, overcoverage, as seen in pincer impingement, was not a risk factor for subsequent OA in previous studies.[3] Eijer and Hogervorst argue that true pincer morphology is a combination of a larger femoral head, wider and shorter neck, deeper acetabulum and increased posterior coverage. The head migrates posteriorly and medially towards the damaged labrum and cartilage, resulting from repeated impact from the femur onto the acetabular rim.[26] Our data reflects this dynamic theory. As the femoral neck shortens, the femoral head medializes and acetabular coverage increases (negative ΔHSM-4, ΔHSM-5 and ΔHSM-9), the risk for OA activity increases. In turn, the osteoarthritic process aggravates the acetabular overcoverage (negative ΔHSM-5, ΔHSM-8, and ΔHSM-12), which may in turn drive further OA progression (low HSM-5).

### Aspherical femoral head

In our SSM, HSM-3 and HSM-6 include asphericity of the femoral head. Low baseline HSM-3 values (wider femoral neck; lateralization of head; extra bone formation at superior head-neck junction) predicted *JSN* at two/three-year follow-up. A negative ΔHSM-3 was associated with simultaneous *JSN* and/or *osteophytosis*. A positive ΔHSM-6 (growing bone spur at the superior head-neck

junction; lateralization of the femur) was associated with simultaneous *JSN*. A positive ΔHSM-6 also predicted *osteophytosis*.

Previous research showed that cam morphology develops during adolescence and is a major risk factor for future hip OA.[3,27–30] Barr et al. showed that cam morphology, but also osteophytes at the head-neck junction grow faster in hips with OA compared to healthy controls.[9] We add to this knowledge that cam morphology grows during the OA process (negative ΔHSM-3, and positive ΔHSM-6) and that this growth may drive disease progression (positive ΔHSM-6).

As both pincer and cam morphology seem to be evolving morphological variations that drive OA activity as a risk mediator, arthroscopic removal of the cam and pincer lesion may prevent OA progression by disrupting vicious cycles. Longitudinal ideally sham-controlled studies should be performed to characterize the disease-modifying effect of hip arthroscopy in OA.

### Size of the greater and lesser trochanter

Low baseline HSM-11 (smaller greater trochanter) predicted *JSN* at two/three-year and five-year follow-up. A decreasing ΔHSM-11 was associated with simultaneous *JSN*. Lynch et al. and Nelson et al. found a larger greater trochanter as risk factor for incident hip OA.[31,32] Barr et al. found 3 HSMs describing the greater trochanter, but didn't report any interpretation on greater trochanter size changes in OA hips compared to controls. The hip abductors insert at the greater trochanter. Reduced tension stress from the hip abductors may cause bone loss at the greater trochanter.[33] Hip OA tends to change the gait of patients. Duchenne gait is an extreme example, where the trunk is bent laterally to the afflicted hip during the stance phase to relieve the hip adductors.[34] A smaller greater trochanter may indicate muscle weakness and is a risk factor for hip OA. OA activity in its turn may affect gait, decreasing hip abductor activity, further decreasing the greater trochanter bone stock. This process may cycle and drive OA progression.

Negative ΔHSM-12 (growth lesser trochanter) was associated with simultaneous *JSN*. Negative ΔHSM-10 and ΔHSM-12 (growth lesser trochanter) were associated with simultaneous *osteophytosis*. Faber et al. associated a larger lesser trochanter in 2D radiographs with radiographic hip OA in a large cross-sectional study.[35] Furthermore, 3D CT scans confirmed that their findings were based truly on trochanter size, rather than an artefact due to differences in positioning during radiographic acquisition. The OA process may cause a hip to be less stable, requiring increased forces of the musculus iliopsoas for stabilization.[36] The iliopsoas inserts at the lesser trochanter and increased forces may induce bone synthesis. The growth

of the lesser trochanter seems to be a result of OA activity. It would be interesting to combine gait analysis, muscle strength measurement and (subregional) SSM in future research to validate these hypotheses.[37]

## Strengths and limitations

CHECK encompasses a ten-year follow-up with five standardized radiography moments. CHECK included people far before end-stage OA was present, allowing radiographic scores to worsen during ten years. Risk factors for the incidence and progression of OA may vary.[38,39] No clear criteria exist to define the start of pathophysiologic mechanisms in hip OA. As there is no clear differentiation between healthy and OA hips based on Altman grades for radiographic features of OA, we did not differentiate between the incidence and progression of OA. There is a discrepancy between radiographic OA and OA symptoms. As CHECK did not include hip-specific symptom severity at T0 and T2, we did not include clinical OA in the present study. We did use radiographic outcomes related to the risk for hip arthroplasty, a defined clinical end-point in OA.[18,19] Future studies should incorporate clinical OA, as hip morphology may differentiate between hips with radiographic progression vs. clinical progression.[40] We did not use statistical corrections for multiple testing, but rather looked for consistency among morphological variations between HSMs, ΔHSMs and their time-specific association with OA.[41] Uncorrected confidence intervals and p-values are presented to enable the reader to assess the significance of the results. With the current data, we cannot exclude that within a hip-period the morphological variation precedes the OA activity as a fast-acting risk factor. Future studies may incorporate more rapidly changing biomarkers (MRI and biochemical markers) to explore the relationship during shorter intervals.

## Conclusion

In conclusion, our study shows that morphological variations are important in hip OA. Instead of unravelling morphological variations that cause OA from those that result from OA, we found that most morphological variations are both and act as a dynamic mediator of OA risk and activity. Novel insights on patterns of development in morphological variations and their time-specific relation with the osteoarthritic process may help towards phenotyping in hip OA. For the latter to succeed a uniform shape model should be developed and tested on a multitude of hip OA populations. Future studies should consider 3D imaging and biochemical markers to further validate our findings.

### Competing interests

Drs. Cootes and Lindner have a patent US 9928443, EP 2893491 issued.

### Acknowledgements and funding

The CHECK-cohort study is funded by the Dutch Arthritis Foundation. Involved are: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen /Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede /Ziekenhuisgroep Twente Almelo; Reade Center for Rehabilitation and Rheumatology; St.Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen. C. Lindner was funded by the Medical Research Council, UK (MR/S00405X/1). W.P. Gielis and H. Weinans were funded by the APPROACH project. APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement n°115770, resources of which are composed of financial contribution

## Supplementary material

**Supplementary Table.** Radiographs, shape parameter and outcome parameter availability

| | T0 | T2 | T5 | T8 | T10 |
|---|---|---|---|---|---|
| **images available** | | | | | |
| right | 985 | 933 | 882 | 835 | 831 |
| left | 985 | 933 | 882 | 835 | 831 |
| total | 1970 | 1866 | 1764 | 1670 | 1662 |
| | | | | | |
| Shape parameter available | 1355 | 1737 | 1571 | 1503 | 1516 |
| Paired with previous time point | | 1277 | 1486 | 1374 | 1379 |
| images not annotated due inadequate quality or coverage | 615 | 129 | 193 | 167 | 146 |
| | | | | | |
| JSN development/progression | | 201 | 277 | 754 | 735 |
| JSN no development/progression | | 1582 | 1528 | 794 | 739 |
| no data | | 221 | 333 | 456 | 480 |
| | | | | | |
| OST development/progression | | 458 | 647 | 622 | 1071 |
| OST no development/progression | | 1214 | 1258 | 827 | 380 |
| no data | | 332 | 432 | 555 | 553 |

from the European Union's Seventh Framework Programme(FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu.

## References

1. Dell'Isola A, Allan R, Smith SL, et al. Identification of clinical phenotypes in knee osteoarthritis: a systematic review of the literature. BMC Musculoskelet Disord 2016;17:425. doi:10.1186/s12891-016-1286-2

2. Mobasheri A, Saarakkala S, Finnilä M, et al. Recent advances in understanding the phenotypes of osteoarthritis. F1000Research. 2019;8. doi:10.12688/f1000research.20575.1

3. Klij P, Ginai AZ, Heijboer MP, et al. The relationship between cam morphology and hip and groin symptoms and signs in young male football players. Scand J Med Sci Sports 2020;30:1221–31. doi:10.1111/sms.13660

4. Morvan J, Bouttier R, Mazieres B, et al. Relationship between hip dysplasia, pain, and osteoarthritis in a cohort of patients with hip symptoms. J Rheumatol 2013;40:1583–9. doi:10.3899/jrheum.121544

5. Jacobsen S, Sonne-Holm S. Hip dysplasia: a significant risk factor for the development of hip osteoarthritis. A cross-sectional survey. Rheumatology 2005;44:211–8. doi:10.1093/rheumatology/keh436

6. Agricola R, Waarsing JH, Arden NK, et al. Cam impingement of the hip-a risk factor for hip osteoarthritis. Nat Rev Rheumatol 2013;9:630–4. doi:10.1038/nrrheum.2013.114

7. Waarsing JH, Kloppenburg M, Slagboom PE, et al. Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. Arthritis Rheum 2011;63:1349–54. doi:10.1002/art.30288

8. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.http://www.scopus.com/inward/record.url?eid=2-s2.0-62949246527&partnerID=tZOtx3y1

9. Barr RJ, Gregory JS, Yoshida K, et al. Significant morphological change in osteoarthritic hips identified over 6-12 months using statistical shape modelling. Osteoarthr Cartil 2018;26:783–9. doi:10.1016/j.joca.2018.04.004

10. Griffin DR, Dickenson EJ, Wall PDH, et al. Hip arthroscopy versus best conservative care for the treatment of femoroacetabular impingement syndrome (UK FASHIoN): a multicentre randomised controlled trial. Lancet 2018;391:2225–35. doi:10.1016/S0140-6736(18)31202-9

11. Yasunaga Y, Yamasaki T, Ochi M. Patient selection criteria for periacetabular osteotomy or rotational acetabular osteotomy. In: Clinical Orthopaedics and Related Research. Springer New York LLC 2012. 3342–54. doi:10.1007/s11999-012-2516-z

12. Lerch TD, Steppacher SD, Liechti EF, et al. One-third of Hips After Periacetabular Osteotomy Survive 30 Years With Good Clinical Results, No Progression of Arthritis, or Conversion to THA. Clin Orthop Relat Res 2017;475:1154–68. doi:10.1007/s11999-016-5169-5

13. Pogliacomi F, De Filippo M, Costantino C, et al. 2006: The value of pelvic and femoral osteotomies in hip surgery. Acta Biomed l'Ateneo Parm 2007;78:60–70.

14. Wesseling J, Boers M, Viergever MA, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. Int J Epidemiol 2016;45:36–44. doi:10.1093/ije/dyu177

15. Gielis WP, Weinans H, Welsing PMJ, et al. An automated workflow based on hip shape improves personalized risk prediction for hip osteoarthritis in the CHECK study. Osteoarthr Cartil Published Online First: 2019. doi:10.1016/j.joca.2019.09.005

16. Lindner C, Thiagarajah S, Wilkinson JM, et al. Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE Trans Med Imaging 2013;32:1462–72. doi:10.1109/TMI.2013.2258030

17. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15 Suppl A:A1-56. doi:10.1016/j.joca.2006.11.009

18. Altman RD, Abadie E, Avouac B, et al. Total joint replacement of hip or knee as an outcome measure for structure modifying trials in osteoarthritis. Osteoarthr Cartil 2005;13:13–9. doi:10.1016/j.joca.2004.10.012

19. Barr RJ, Gregory JS, Reid DM, et al. Predicting OA progression to total hip replacement: Can we do better than risk factors alone using active shape modelling as an imaging biomarker? Rheumatology 2012;51:562–70. doi:10.1093/rheumatology/ker382

20. Damen J, Schiphof D, Wolde S Ten, et al. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. Osteoarthritis Cartilage 2014;22:969–74. doi:10.1016/j.joca.2014.05.007

21. Abdulrahim H, Jiao Q, Swain S, et al. Constitutional morphological features and risk of hip osteoarthritis: A case-control study using standard radiographs. Ann Rheum Dis 2020;:1–8. doi:10.1136/annrheumdis-2020-218739

22. Castaño-Betancourt MC, Van Meurs JBJ, Bierma-Zeinstra S, et al. The contribution of hip geometry to the prediction of hip osteoarthritis. Osteoarthr Cartil 2013;21:1530–6. doi:10.1016/j.joca.2013.06.012

23. Houcke J Van, Khanduja V, Pattyn C, et al. The history of biomechanics in total hip arthroplasty. In: Indian Journal of Orthopaedics. Wolters Kluwer Medknow Publications 2017. 359–67. doi:10.4103/ortho.IJOrtho_280_17

24. Lane NE, Lin P, Christiansen L, et al. Association of mild acetabular dysplasia with an increased risk of incident hip osteoarthritis in elderly white women: The study of osteoporotic fractures. Arthritis Rheum 2000;43:400. doi:10.1002/1529-0131(200002)43:2<400::AID-ANR21>3.0.CO;2-D

25. Thomas GER, Palmer AJR, Batra RN, et al. Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women. A 20 year longitudinal cohort study. Osteoarthr Cartil 2014;22:1504–10. doi:10.1016/j.joca.2014.06.038

26. Eijer H, Hogervorst T. Femoroacetabular impingement causes osteoarthritis of the hip by migration

and micro-instability of the femoral head. Med Hypotheses 2017;104:93–6. doi:10.1016/j.mehy.2017.05.035

27. Agricola R, Bessems JHJM, Ginai AZ, et al. The Development of Cam-Type Deformity in Adolescent and Young Male Soccer Players. Am J Sports Med 2012;40:1099–106. doi:10.1177/0363546512438381

28. Van Klij P, Heijboer MP, Ginai AZ, et al. Cam morphology in young male football players mostly develops before proximal femoral growth plate closure: A prospective study with 5-yearfollow-up. Br. J. Sports Med. 2019;53:532–8. doi:10.1136/bjsports-2018-099328

29. Thomas GER, Palmer AJR, Batra RN, et al. Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women. A 20 year longitudinal cohort study. Osteoarthr Cartil 2014;22:1504–10. doi:10.1016/j.joca.2014.06.038

30. Agricola R, Leyland KM, Bierma-Zeinstra SMA, et al. Validation of statistical shape modelling to predict hip osteoarthritis in females: Data from two prospective cohort studies (Cohort Hip and Cohort Knee and Chingford). Rheumatol (United Kingdom) 2015;54:2033–41. doi:10.1093/rheumatology/kev232

31. Nelson AE, Liu F, Lynch JA, et al. Association of incident symptomatic hip osteoarthritis with differences in hip shape by active shape modeling: The johnston county osteoarthritis project. Arthritis Care Res 2014;66:74–81. doi:10.1002/acr.22094

32. Lynch JA, Parimi N, Chaganti RK, et al. The association of proximal femoral shape and incident radiographic hip OA in elderly women. Osteoarthr Cartil 2009;17:1313–8. doi:10.1016/j.joca.2009.04.011

33. Burr DB. Muscle strength, bone mass, and age-related bone loss. J. Bone Miner. Res. 1997;12:1547–51. doi:10.1359/jbmr.1997.12.10.1547

34. Reininga IHF, Stevens M, Wagenmakers R, et al. Subjects with hip osteoarthritis show distinctive patterns of trunk movements during gait-a body-fixed-sensor based analysis. J Neuroeng Rehabil 2012;9:3. doi:10.1186/1743-0003-9-3

35. Faber BG, Bredbenner TL, Baird D, et al. Subregional statistical shape modelling identifies lesser trochanter size as a possible risk factor for radiographic hip osteoarthritis, a cross-sectional analysis from the Osteoporotic Fractures in Men Study. Osteoarthr Cartil 2020;28:1071–8. doi:10.1016/j.joca.2020.04.011

36. Yoshio M, Murakami G, Sato T, et al. The function of the psoas major muscle: Passive kinetics and morphological studies using donated cadavers. J Orthop Sci Published Online First: 2002. doi:10.1007/s007760200034

37. Faber BG, Bredbenner TL, Baird D, et al. Subregional statistical shape modelling identifies lesser trochanter size as a possible risk factor for radiographic hip osteoarthritis, a cross-sectional analysis from the Osteoporotic Fractures in Men Study. Osteoarthr Cartil 2020;28:1071–8. doi:10.1016/j.joca.2020.04.011

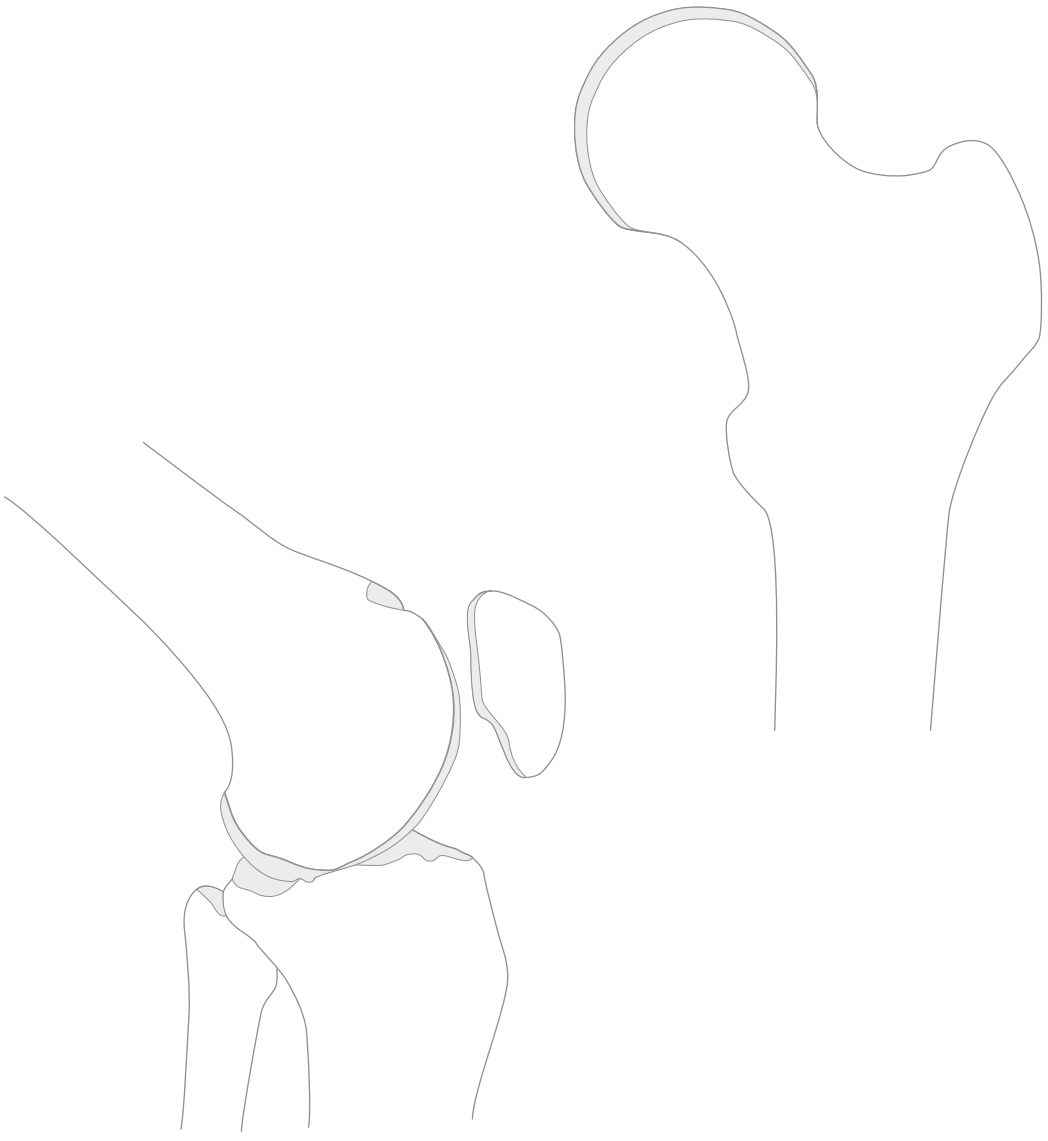
38. Nishimura A, Hasegawa M, Kato K, et al. Risk factors for the incidence and progression of

radiographic osteoarthritis of the knee among Japanese. Int Orthop 2011;35:839–43. doi:10.1007/s00264-010-1073-x

39. Van Spil WE, Welsing PMJ, Bierma-Zeinstra SMA, et al. The ability of systemic biochemical markers to reflect presence, incidence, and progression of early-stage radiographic knee and hip osteoarthritis: data from CHECK. Osteoarthritis Cartilage 2015;23:1388–97. doi:10.1016/j.joca.2015.03.023

40. Waarsing JH, Rozendaal RM, Verhaar JAN, et al. A statistical model of shape and density of the proximal femur in relation to radiological and clinical OA of the hip. Osteoarthr Cartil 2010;18:787–94. doi:10.1016/j.joca.2010.02.003

41. Saville DJ. Multiple Comparison Procedures: The Practical Solution. Am Stat 1990;44:174. doi:10.2307/2684163



# CHAPTER 8

# Unravelling the knee-hip-spine trilemma from the CHECK study: the relation between sagittal pelvic morphology and onset of degenerative knee, hip, and spine disorders

*The Bone & Joint Journal*
*August 2020, doi.org/10.1302/0301-620X.102B9.BJJ-2019-1315.R2*

Joost H.J. van Erp, Willem Paul Gielis, Vahid Arbabi, Arthur de Gast, Harrie Winans, Saeed Arbabi, F. Cumhur Öner, René Castelein, Tom P.C. Schlösser

## Abstract

### Background and aim

The aetiologies of common degenerative spine, hip, and knee pathologies are still not completely understood. Mechanical theories have suggested that those diseases are related to sagittal pelvic morphology and spinopelvic-femoral dynamics. The link between the most widely used parameter for sagittal pelvic morphology, pelvic incidence (PI), and the onset of degenerative lumbar, hip, and knee pathologies has not been studied in a large-scale setting.

### Patients and methods

A total of 421 patients from the Cohort Hip and Cohort Knee (CHECK) database, a population-based observational cohort, with hip and knee complaints < 6 months, aged between 45 and 65 years old, and with lateral lumbar, hip, and knee radiographs available, were included. Sagittal spinopelvic parameters and pathologies (spondylolisthesis and degenerative disc disease (DDD)) were measured at eight-year follow-up and characteristics of hip and knee osteoarthritis (OA) at baseline and eight-year follow-up. Epidemiology of the degenerative disorders and clinical outcome scores (hip and knee pain and Western Ontario and McMaster Universities Osteoarthritis Index) were compared between low PI (< 50°), normal PI (50° to 60°), and high PI (> 60°) using generalized estimating equations.

### Results

Demographic details were not different between the different PI groups. L4 to L5 and L5 to S1 spondylolisthesis were more frequently present in subjects with high PI compared to low PI (L4 to L5, OR 3.717; $p = 0.024$ vs L5 to S1 OR 7.751; $p = 0.001$). L5 to S1 DDD occurred more in patients with low PI compared to high PI (OR 1.889; $p = 0.010$), whereas there were no differences in L4 to L5 DDD among individuals with a different PI. The incidence of hip OA was higher in participants with low PI compared to normal (OR 1.262; $p = 0.414$) or high PI (OR 1.337; $p = 0.274$), but not statistically different. The incidence of knee OA was higher in individuals with a high PI compared to low PI (OR 1.620; $p = 0.034$).

### Conclusion

High PI is a risk factor for development of spondylolisthesis and knee OA. Low pelvic incidence is related to DDD, and may be linked to OA of the hip.

## Introduction

Hip and knee osteoarthritis and low back pain are a major burden for individuals as well as the global health and socioeconomic system. Since they affect a considerable part of the older adult population, they are one of the most expensive medical conditions and a significant burden to society.[1] They lead to pain and loss of function and can present with a variety of symptoms.[2,3] Last decades, the etiopathogenesis of those common degenerative diseases has been studied extensively, but is still not fully understood. It is generally accepted that their multifactorial aetiology includes a mixture of mechanical and biological factors.

In 1992, Duval-Beaupère demonstrated the importance of the pelvis as a key regulator of sagittal spino-pelvic-femoral balance[4]. Since then, sagittal pelvic morphology has been recognized in relation to normal functioning of the human spinopelvic complex as well as in the onset of different degenerative spinal diseases and spinal deformities. Pelvic incidence (PI) was introduced as a parameter for sagittal pelvic morphology and pelvic tilt (i.e. spino-pelvic tilt, PT) and sacral slope (SS) for sagittal pelvic orientation (**Figure 1**).[4] In this, pelvic morphology refers to the orientation of the sacral plate within the pelvic ring, which is not position, but morphology dependent and therefore considered to be fixed during adult life.[5] On the contrary, PT and SS are adaptive and depend on the pelvic positioning in space with respect to the rest of the anatomy and thus change during locomotion. In the sagittal plane, PI is defined as the angle between the line connecting the femoral centre of rotation with the midpoint of the sacral plate and a perpendicular line on the sacral endplate and can be easily measured on lateral pelvic radiographs. In humans, PI has a large range and varies from 33°-85°, increases during childhood and becomes constant when reaching adolescence.[6,7] Measurement of PI on radiographs has shown a variability of 3°-6°, influenced by rotation. PI measured on CT scans is more precise, but assessment of pelvic orientation in the upright position is impossible.[8]

Therefore, the aim of this epidemiological study is to explore the relation between sagittal pelvic morphology and the development of the most common degenerative lumbar, hip and knee pathologies, and the corresponding clinical outcome scores.

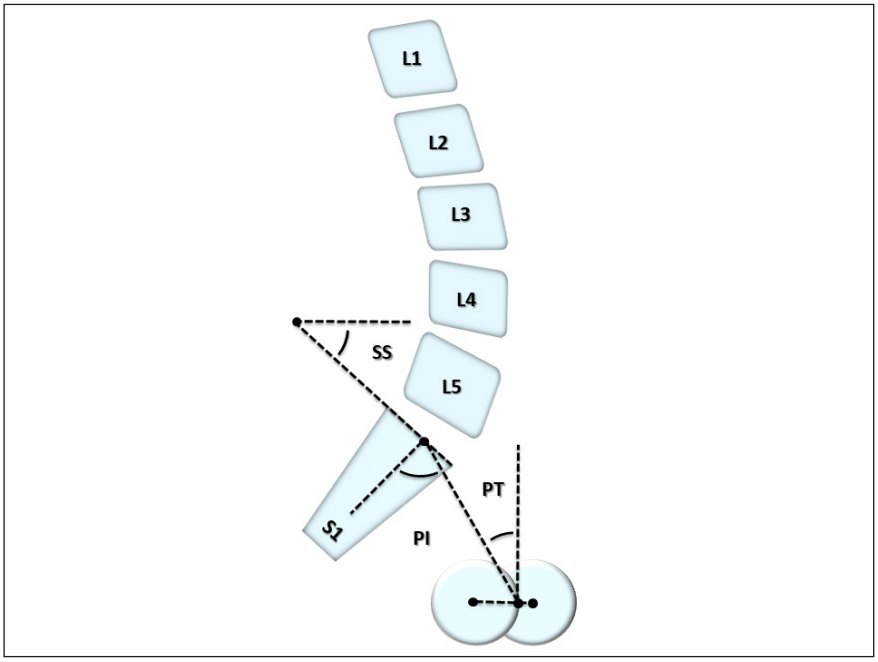



**Figure 1.** Schematic drawing showing pelvic incidence (PI), pelvic tilt (PT), and sacral slope (SS).

## Methods

### Study population

This study utilized data from the CHECK database (Cohort Hip and Cohort Knee), a multicentre population-based prospective observational cohort with 1,002 participants, initiated by the Dutch Arthritis Foundation to study early hip and knee OA. The database collects clinical, radiological, and biochemical data from patients, presenting for the first time with hip or knee pain in the previous six months, and aged between 45 and 65 years old. They underwent clinical and radiological follow-up of their hips and knees over the next ten years. Exclusion criteria were patients having other pathologies (e.g. rheumatic diseases, previous hip and knee arthroplasty, hip dysplasia, Perthes' disease, septic arthritis, osteochondritis dissecans, intra-articular fractures, traumatic ligament or meniscus damage, plica syndrome, or a Bakers cyst). The CHECK study was approved by the local medical ethics committees and all participants gave informed consent before commencement. The protocol of this study, with a detailed description of the cohort, has been published.[9]

A flowchart of the data collection is shown in **Figure 2**. At eight years follow-up, clinical data and radiographs of hip and knee were available on 845 of 1,002 participants (84%). Two participants, who underwent previous lumbar spinal fusion for unknown reasons, were excluded. Out of 843 patients, 421 had lateral radiographs including the femoral heads (for PI and PT measurement) and were included in this analysis. The remaining 422 participants were excluded, since the femoral heads were not visualized. Demographic data were not significantly different between different PI groups (**Table I**). Overall, 97% (n = 409) of the subjects were Caucasian.

### Radiographical parameters

Following the study protocol, lateral spinal radiographs were obtained in the upright position at eight-year follow-up. The independent parameters were PI, PT and SS and L1-S1 lumbar lordosis (LL) and were measured systematically by the method of Legaye et al.[10], using in-house developed software. For comparison of the epidemiology of the various degenerative disorders and clinical outcome scores, three PI subgroups were compared: low PI (<50°), normal PI (50°-60°) and high PI (>60°).[11,12]

The outcome parameters were the presence of L4-5 or L5-S1 degenerative disc disease (DDD) and spondylolisthesis, hip osteoarthritis and knee osteoarthritis. The presence of DDD was scored on the lateral lumbar radiographs acquired at eight year follow-up, by a trained orthopaedic resident using the previously validated Lane classification, which had good (>0.9) intra- and interobserver reliability on lumbar spine radiographs.[13,14] Lane≥2 was defined as the presence of DDD. On the same radiographs, subjects were classified as the radiological presence of L4-5 or L5-S1 degenerative spondylolisthesis if the Meyerding grade was 1 or more.[15] Differentiation between isthmic and degenerative spondylolisthesis was not possible, since no CT imaging was available.

For hip and knee-osteoarthritis, the development of osteoarthritis in eight years in all hips/knees on the anterior-posterior radiographs of the pelvis and knees during the study period had already been measured according to Kellgren and Lawrence (KL) for previous projects, by five observers with good interobserver variability.[16] Both extremities were analyzed separately. The presence of OA (KL≥2) at baseline or development of KL≥2 or a joint arthroplasty during the follow-up was considered as significant radiographic hip or knee OA during follow-up. For each patient, measurements of the spine were performed once, measurements for the hip and knee were performed bilaterally and corrected for in the statistical analyses.

### Clinical parameters

The clinical parameters used were numeric rating scale (NRS) for pain intensity in the hip and knee as well as Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) at eight-year follow-up.[17]

### Statistical analysis

Data were processed and analyzed using Microsoft Excel 2010 and IBM SPSS 23. Demographics were compared between groups using analysis of variance. Epidemiology of the various degenerative disorders and clinical outcome scores were compared between groups using generalized estimating equations, accounting for the correlation between multiple limbs in one patient. PI, PT, SS, age and BMI were all normally distributed. Individuals with early-onset OA (significant hip or knee OA (KL≥2) at baseline) were excluded, to describe the development of hip and knee OA after the onset of hip or knee symptoms. Odd's ratios and 95% confidence intervals were calculated. The relation between the degenerative diseases and SS and PT were analyzed as continuous parameters, using generalized estimated equitation. A correction for potential confounders (age, BMI and sex) was performed by adding them as independent variables to the models. Finally, subanalyses were performed to evaluate whether clinical parameters differ between subjects with significant radiological OA and different degrees of PI. The level of statistical significance was set at 0.05.

## Results

### Demographics

A flowchart of the data collection is shown in **figure 2**. At eight-years follow-up, clinical data and radiographs of hip and knee were available in 845 of 1002 participants (84%). Two participants, who underwent previous lumbar spinal fusion for unknown reason, were excluded. Out of 843 subjects, 421 subjects had lateral radiographs including the femoral heads (for PI and PT measurement) and were included in this analysis. The remaining 422 participants were excluded, since the femoral heads were missing, Demographics did not significantly differ between the different PI groups (**Table 1**). 97% of the subjects were Caucasian.

#### Spondylolisthesis

The prevalence of L4-5 and L5-S1 spondylolisthesis were 6% (n=27) and 9% (n=39), respectively. All spondylolistheses were classified as Meyerding Grade I or II. Statistical comparison showed more L4-5 spondylolisthesis in high PI

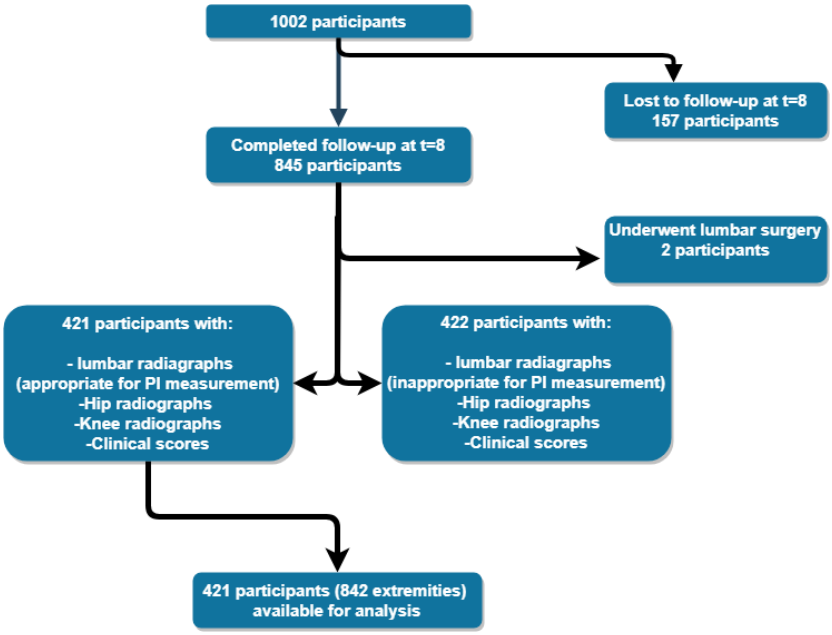



**Figure 2.** Flowchart of the Cohort Hip and Cohort Knee (CHECK) study participants available for the purpose of this study. PI, pelvic incidence.

subjects versus low PI (p=0.024, OR=3.717, 95%-CI 0.086-0.839), but there were no significant differences when compared to the normal PI group (**Table 2**). At L5-S1, spondylolisthesis was more prevalent in high PI versus low PI (p=0.001, OR=7.761, 95%-CI 0.038-0.434) or normal PI (p=0.018, OR=2.732, 95%-CI 0.160-0.840). PT and SS were not related to the prevalence of spondylolisthesis at L4-5, but at L5-S1 there was a significant correlation (p=0.004; **Table 3**).

### Degenerative disc disease

The prevalence of significant DDD on level L5-S1 was higher in participants with a low PI (p=0.010, OR=1.889, 95%-CI 1.163-3.069) or low PT and SS versus high PI, PT and SS, respectively. At L4-L5 the differences did not reach statistical significance (**Table 2 and 3**).

**Table 1.** Characteristics of subjects with low, normal and high pelvic incidence (PI).

| | Low PI (<50°, n=124) | Normal PI (50°-60°, n=120) | High PI (>60°, n=177) | Total (n=421) | p-value |
|---|---|---|---|---|---|
| **Females, n(%)** | 89 (73%) | 88 (73%) | 128 (73%) | 305 (73%) | Low vs. normal: 0.960<br>Low vs. high:0.994<br>Normal vs. high: 0.980 |
| **Age in years at t=0, mean±sd** | 56.4±5 | 56.7±5 | 55.5±5 | 56.1±5 | Low vs. normal: 0.919<br>Low vs. high: 0.288<br>Normal vs. high: 0.135 |
| **Body mass index at t=0** | 26.2±4 | 26.8±4 | 26.7±5 | 26.6±4 | Low vs. normal: 0.524<br>Low vs. high: 0.600<br>Normal vs. high: 0.135 |
| **PI** | 43.0±6 | 55.1±3 | 70.6±8 | 58.1±13 | |
| **PT** | 16.6±9 | 23.8±7 | 30.7±10 | 24.6±11 | |
| **SS** | 26.5±9 | 31.3±8 | 40.9±11 | 33.9±11 | |
| **L1-S1 LL** | 34.0±10 | 39.3±11 | 50.0±13 | 42.3±13 | |

PT=pelvic tilt, SS=sacral slope, LL=lumbar lordosis.

**Table 2:** Comparison of the prevalence of radiographic signs of degenerative lumbar in individuals with different pelvic incidence (PI). N (%)

| | Low PI (<50°, n=124) | Normal PI (50°-60°, n=120) | High PI (>60°, n=177) | Total (n=421) | p-value |
|---|---|---|---|---|---|
| **Spondylolisthesis L4-L5** | 4(3%) | 6(5%) | 17 (10%) | 27 (6%) | Low vs. normal: 0.546<br>Low vs. high: 0.024*<br>Normal vs. high: 0.072 |
| **Spondylolisthesis L5-S1** | 3(2%) | 8(6%) | 28(16%) | 39(9%) | Low vs. normal: 0.129<br>Low vs. high: 0.001*<br>Normal vs. high: 0.018* |
| **L4-5 DDD** | 24 (19%) | 25 (21%) | 28(16%) | 77(18%) | Low vs. normal: 0.835<br>Low vs. high: 0.492<br>Normal vs. high: 0.365 |
| **L5-S1 DDD** | 59(48%) | 53(44%) | 57(32%) | 169(40%) | Low vs. normal: 0.508<br>Low vs. high: 0.010*<br>Normal vs. high: 0.067 |

DDD = Degenerative disk disease

## Hip osteoarthritis

At eight-year follow-up, 137 of the 651 (21%) hips had developed significant radiological OA. The incidence of hip OA during 8 years follow-up was 25% in participants with low PI as compared to 21% and 19% in participants with normal or high PI, respectively. This however, did not reach statistical significance. The same trend, more hip OA in lower PT and SS, was observed (**Table 3 and 4**).

**Table 3.** Relation between pelvic orientation parameters (sacral slope (SS) and pelvic tilt (PT)) and the prevalence of lumbar, hip and knee degenerative diseases is shown for 10°.

| | B | P-value | OR | 95% Confidence interval |
|---|---|---|---|---|
| **Spondylolisthesis L4-L5** | | | | |
| Sacral slope | 0.262 | 0.134 | 1.299 | 0.917-1.825 |
| Pelvic tilt | 0.245 | 0.196 | 1.277 | 0.882-1.857 |
| **Spondylolisthesis L5-S1** | | | | |
| Sacral slope | 0.686 | <0.001* | 1.985 | 1.484-2.701 |
| Pelvic tilt | 0.471 | 0.004* | 1.602 | 1.164-2.238 |
| **DDD L4-L5** | | | | |
| Sacral slope | -0.125 | 0.286 | 0.883 | 0.698-1.107 |
| Pelvic tilt | 0.000 | 0.998 | 1 | 0.784-1.273 |
| **DDD L5-S1** | | | | |
| Sacral slope | -0.335 | 0.001* | 0.715 | 0.587-0.865 |
| Pelvic tilt | -0.118 | 0.246 | 0.889 | 0.727-1.083 |
| **Hip osteoarthritis** | | | | |
| Sacral slope | -0.068 | 0.590 | 0.935 | 0.731-1.195 |
| Pelvic tilt | -0.060 | 0.594 | 0.941 | 0.754-1.175 |
| **Knee osteoarthritis** | | | | |
| Sacral slope | 0.173 | 0.068 | 1.188 | 0.988-1.430 |
| Pelvic tilt | 0.161 | 0.089 | 1.175 | 0.976-1.415 |

DDD = Degenerative disk disease

**Table 4.** Comparison of the prevalence of radiographic signs of knee and hip osteoarthritis and clinical outcomes in individuals with low, normal and high pelvic incidence (PI).

| | Low PI (<50°, n=193) | Normal PI (50°-60°, n=264) | High PI (>60°, n=194) | Total (n=651) | p-value |
|---|---|---|---|---|---|
| **Radiological hip OA (KL≥2), n(%)** | 48(25%) | 40(21%) | 49(19%) | 137(21%) | Low vs. normal: 0.414<br>Low vs. high: 0.274<br>Normal vs. high: 0.837 |
| **Radiological knee OA (KL≥2), n(%)** | 58(30%) | 59(31%) | 108(41%) | 225(35%) | Low vs. normal: 0.843<br>Low vs. high: 0.034*<br>Normal vs. high: 0.025* |
| **WOMAC (0-100), mean±sd** | 24±16 | 21±14 | 24±15 | 23±15 | Low vs. normal: 0.286<br>Low vs. high: 0.818<br>Normal vs. high: 0.577 |
| **NRS for hip pain Intensity (0-10)** | 1.6±2 | 1.7±2 | 1.8±2 | 1.7±2 | Low vs. normal: 0.990<br>Low vs. high: 0.627<br>Normal vs. high: 0.720 |
| **NRS for knee pain intensity (0-10)** | 2.1±2 | 2.1±2 | 2.3±2 | 2.2±2 | Low vs. normal: 0.949<br>Low vs. high: 0.374<br>Normal vs. high: 0.224 |

OA = Osteoarthritis, KL=Kellgren and Lawrence.

At eight-year follow-up, no significant difference in the NRS for hip pain and WOMAC were observed between groups. Subanalysis showed that in the presence of significant radiological hip OA, there were no significant differences in the NRS for hip pain or WOMAC between PI groups.

### Knee osteoarthritis

314 (42%) of the 743 knees had significant radiological OA at eight-year follow-up. A significantly higher incidence of knee OA during 8 years of was observed in subjects with a high PI, compared to subjects with normal PI (p=0.025, OR 1.701, 95%-CI 1.068-2.710) or low PI (p=0.034, OR 1.620, 95%-CI 1.037-2.531). PT and SS were also correlated with radiological knee OA (**Table 4 and 5**). Similar to the clinical scores of the hip, no significant differences were seen in knee pain intensity between individuals with different pelvic morphology. When the knee pain intensity scores were compared between individuals with knee OA KL≥2 and different degree of PI, individuals with high PI had significantly higher scores compared to low PI (p=0.049; **Table 5**).

**Table 5.** the presence of knee and hip pain in individuals with significant radiographic knee or hip OA, with different grades of pelvic incidence (PI).

| | | Low PI (<50°) | Normal PI (50°-60°) | High PI (>60°) | Total |
|---|---|---|---|---|---|
| **Knee Pain** | Yes | 50.5% | 58.0% | 41.7% | 48.6% |
| | No | 49.5% | 42.0% | 58.3% | 51.4% |
| **Hip Pain** | Yes | 70.6% | 70.1% | 75.9% | 72.8% |
| | No | 29.4% | 29.9% | 24.1% | 27.2% |

## Discussion

Human spino-pelvic-femoral alignment is unique and different from any other species.[18] Humans are able to stand in the fully upright position with extended knees and hips. This is a result of an evolutionary lordotic angulation between the ischiac and iliac bones in the pelvis, also known as the pelvic lordosis, in order to keep the centre of gravity in the upright position straight above the pelvis. This sagittal pelvic morphology has been recognized in relation to normal functioning of the human spinopelvic complex as well as in the onset of different degenerative spinal diseases and spinal deformities. Study groups of Roussouly[11], Mac-Thiong[19] and Le Huec[20] demonstrated that, by anterior and posterior pelvic tilt, the pelvis is

a key regulator of the sagittal configuration of the spine in order to keep the centre of mass of the body and the head straight above the femoral heads, and that by posterior pelvic tilt, the pelvis can compensate for delordosing degenerative lumbar deformation.[11,19,20] PI was introduced as a parameter for sagittal pelvic morphology and PT and SS for sagittal pelvic orientation (**Figure 1**).[4] Due to the more horizontal position of the sacrum within the pelvic ring, in a pelvis with high PI, there is an increased ability for posterior pelvic tilt and compensation of kyphogenic spinal pathology.[5]

Previous studies have shown that the PI increases during paediatric growth and remains more or less constant during later life.[6] The only variation at later life can be attributed to body positioning: due to the minor range-of-motion in the sacro-iliac joints there is a 2-3° difference in PI between different body positions.[21] In the present study, PI was measured on lumbar radiographs in the upright position available at 8 years follow-up. Based on the existing literature, however, we assume that an individual's PI has not significantly changed from adolescence until adulthood, and thus also not within the study period. Therefore, this can be considered a pre-existent morphological parameter that can indicate the sagittal biomechanical loading of the spino-pelvic-femoral articulations during later life. Therefore, we believe that the potential etiological conclusions on the role of sagittal pelvic morphology and the onset of certain spinal, hip or knee pathologies hold true. In contrast to PI, the sagittal pelvic orientation parameters (PT and SS) could have changed during the study period as a result of the onset of certain degenerative pathologies.

For degenerative spondylolisthesis, there is already clear evidence that it occurs more in individuals with a relatively high PI and SS, as a result of greater anteriorly directed shear forces at the L4-L5 and L5-S1 level[5]. Our results are in line with the data from previous studies, as the prevalence of L4-L5 and L5-S1 spondylolisthesis was higher in individuals with a high PI and SS.[5]

For DDD, it can be expected that, due to increased axial loading in the anterior spinal elements in low PI and SS, intervertebral disc pathology or degenerative scoliosis is more prevalent in adults with low PI.[5,22] Previously, Barrey et al. observed that patients with DDD on level L4-L5 and L5-S1 had the same PI (52°) as a control group. However, in patients <45 years old, differences in PI were more expressed and demonstrated a significantly lower PI of 48.3° in individuals affected by DDD compared to a control group.[23] Yang et al. described a lower ($P<0.05$) PI and SS in patients with DDD, compared a the control group, and stated that low

PI plays a predisposing role in the pathogenesis of DDD.[22] Furthermore, Strube et al. compared patients with degenerative spondylolisthesis and DDD and found that the PI was higher in spondylolisthesis compared to DDD ($p<0.001$).[24] In our prospective study, DDD was more present in patients with a low PI, but solely on the level L5-S1, there was no effect of low PI on the level L4-L5.

Other, more recent studies found that due to pelvic-femoral dynamics sagittal pelvic morphology cannot only be related to degenerative lumbar pathology, but may also have important biomechanical consequences for the hips and knees.[25,26,27] In a recent systematic review by Saltychev et al., the evidence for a relation between PI and hip OA from ten studies was summarized as inconclusive.[12] No relation was found, but the overall methodology of the included studies and sample sizes for epidemiological data (19–150 subjects) was poor.[12] Furthermore, no earlier study described the relation between PI and knee OA. To the best of our knowledge, the potential link between sagittal pelvic morphology and hip/ knee OA has not yet been studied systematically in a large scale prospective setting.[12]

In the present study, we systematically explored the influence of sagittal pelvic morphology on the development of different degenerative diseases of hip and knee or spine. For this purpose, an existing prospective cohort of 1002 patient presenting with hip and knee pain that underwent extensive radiological follow-up was used. Lateral lumbar radiographs for measurement of PI, and PT and SS, were only available at eight-year follow-up. Based on previous studies, however, it can be assumed that the PI did not change in individuals during the follow-up period, since it is constant after the adolescent growth spurt.[19,28] From the data, we can conclude that low PI was a risk factor for development of L5-S1 DDD, while spondylolisthesis was associated with high PI. Interestingly, we also observed more hip OA in individuals with low PI, whereas knee OA was significantly more prevalent in individuals with a high PI.

From a biomechanical point of view, differences in sagittal pelvic morphology lead to differences in global alignment, spinopelvic-femoral dynamics and compensatory mechanisms, and therefore different mechanical loading of the femoro-acetabular joints and lumbosacral junction. It can be inferred that the hip joints in patients with a low PI are loaded more towards their limit of extension, which could theoretically pose high loads on the cranial and anterior labrum and impingement on the posterior labrum. Conclusions on cause-and-effect of hip and knee OA versus pelvic anatomy, however, cannot be drawn based on the present study.

Potentially, as low PI patients are more prone to development of DDD, and have less ability in the pelvis for compensation of kyphogenic lumbar degenerative pathology, the femoro-acetabular configuration will be not severely affected due to the lack of posterior pelvic tilt. Meanwhile, the consequences of major posterior pelvic tilt in patients with high PI on the biomechanical loading of the hip joint are also unknown.

In contrast to previous theory of Roussouly et al.[5], who expected more hip OA in case of a high PT and PI, we observed a trend of more hip OA in patients with low PI but this did not reach statistical significance. This finding is in line with Bakouny et al.[29], who observed in 91 asymptomatic young adults that individuals with a Roussouly type 2 spinopelvic alignment had a gait pattern with relatively more hip extension and less hip flexion. This altered pelvic orientation may result in different mechanical loading of the coxo-acetebular joints and femoro-acetabular impingement, and with this hip OA.[29] To shed further light on the biomechanical relations between the pelvis and the femoro-acetabular junction, future studies could focus on the anatomical relations between the 3D morphology of the hip joint versus the sagittal pelvic morphology, taking into account the biomechanical loading and orientation of the pelvis.

Clinicians are often confronted with patients suffering from a knee-hip-spine trilemma. In 2003, Murata et al. labelled the knee-spine syndrome, an association between limited extension of the knee and reduced lumbar lordosis.[30] Lee et al. observed that during knee flexion, lumbar lordosis decreases and the sagittal balance is shifted anteriorly, which might give complaints of the lumbar spine[31], whereas PT and SS rarely changed during knee flexion. The influence of PI in the onset of knee OA has not been explored before.

Since the majority of the excluded patients were excluded related due to local imaging differences in participating centres, we don`t believe this process has excluded patients with a different phenotype. This could however, have resulted in under powering of the analysis of the relation between PI and hip OA, which is a weakness of this study. Another weakness is the relative short period of eight-year follow-up to observe the onset of degenerative diseases, since it takes for several decennia to develop most degenerative diseases. This could have resulted in the observation of progression of radiological OA, but this might be too early to correlate this with clinical outcomes, as was seen in this study.

The present study demonstrated that due to the sagittal pelvic morphology and consequent biomechanical loading, individuals may be at higher risk for development of a spectrum of degenerative disorders of the lower extremities or lumbar spine.

## Conclusions

From an etiological perspective, this study provides evidence on the role of sagittal pelvic morphology and the onset of the most common lumbar, hip and knee degenerative pathologies. It can be concluded that high PI is a risk factor for development of degenerative spondylolisthesis and knee osteoarthritis. Contrarily, low PI is a risk factor for degenerative disc disease. Low PI may also be associated with the development of osteoarthritis of the hip. From the observed correlations between the different pathologies and pelvic anatomy, we anticipate that further research on the anatomical and biomechanical knee-hip-spine relations provides more insight into the aetiology and pathogenesis of the most common degenerative spinal, knee and hip disorders.

## Acknowledgements

The CHECK cohort study is funded by the Dutch Arthritis Foundation. Involved are: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen/Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede/Ziekenhuisgroep Twente Almelo; Reade Center for Rehabilitation and Rheumatology; St.Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

## References


1. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003;81(9):646–56.
2. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum* 2008;58(1):26–35.
3. Baar ME van, Dekker J, Lemmens JA, Oostendorp RA, Bijlsma JW. Pain and disability in patients with osteoarthritis of hip or knee: the relationship with articular, kinesiological, and psychological characteristics. *J Rheumatol* 1998;25(1):125–33.
4. Duval-Beaupère G, Schmidt C, Cosson P. A barycentremetric study of the sagittal shape of spine and pelvis: The conditions required for an economic standing position. *Ann Biomed Eng* 1992;20(4):451–462.
5. Roussouly P, Pinheiro-Franco JL. Biomechanical analysis of the spino-pelvic organization and adaptation in pathology. *Eur Spine J* 2011;20 Suppl 5(S5):609–618.
6. Mangione P, Gomez D, Senegas J. Study of the course of the incidence angle during growth. *Eur Spine J* 1997;6(3):163–167.
7. Vaz G, Roussouly P, Berthonnaud E, Dimnet J. Sagittal morphology and equilibrium of pelvis and spine. *Eur Spine J* 2002;11(1):80–87.
8. Brink RC, Vavruch L, Schlösser TPC, Abul-Kasim K, Ohlin A, Tropp H, et al. Three-dimensional pelvic incidence is much higher in (thoraco)lumbar scoliosis than in controls. *Eur Spine J* 2018;
9. Wesseling J, Boers M, Viergever MA, Hilberdink WK, Lafeber FP, Dekker J, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. *Int J Epidemiol* 2016;45(1):36–44.
10. Legaye J, Duval-Beaupère G, Hecquet J, Marty C. Pelvic incidence: A fundamental pelvic parameter for three-dimensional regulation of spinal sagittal curves. *Eur Spine J* 1998;7(2):99–103.
11. Roussouly P, Gollogly S, Berthonnaud E, Dimnet J. Classification of the normal variation in the sagittal alignment of the human lumbar spine and pelvis in the standing position. *Spine (Phila Pa 1976)* 2005;30(3):346–53.
12. Saltychev M, Pernaa K, Seppänen M, Mäkelä K, Laimi K. Pelvic incidence and hip disorders: A systematic review and quantitative analysis. *Acta Orthop* 2018;89(1):66–70.
13. Lane NE, Nevitt MC, Genant HK, Hochberg MC. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol* 1993;20(11):1911–1918.
14. Kettler A, Wilke HJ. Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. *Eur Spine J* 2006;15(6):705–718.
15. Meyerding H. Spondylolisthesis. *Surg,Gynecol, Obs* 1932;

16. Damen J, Schiphof D, Wolde S Ten, Cats HA, Bierma-Zeinstra SMA, Oei EHG. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. *Osteoarthr Cartil* 2014;22(7):969–74.

17. 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(12):1833–1840.

18. Schlösser TPC, Janssen MMA, Hogervorst T, Vrtovec T, Vos J de, Öner FC, et al. The odyssey of sagittal pelvic morphology during human evolution: a perspective on different Hominoidae. *Spine J* Elsevier, 2017;17(8):1202–1206.

19. Mac-Thiong J-M, Labelle H, Berthonnaud E, Betz RR, Roussouly P. Sagittal spinopelvic balance in normal children and adolescents. *Eur Spine J* 2007;16(2):227–234.

20. Huec JC Le, Saddiki R, Franke J, Rigal J, Aunoble S. Equilibrium of the human body and the gravity line: the basics. *Eur Spine J* 2011;20(S5):558–563.

21. Place HM, Hayes AM, Huebner SB, Hayden AM, Israel H, Brechbuhler JL. Pelvic incidence: a fixed value or can you change it? *Spine J* 2017;17(10):1565–1569.

22. Yang X, Kong Q, Song Y, Liu L, Zeng J, Xing R. The characteristics of spinopelvic sagittal alignment in patients with lumbar disc degenerative diseases. *Eur Spine J* 2014;23(3):569–575.

23. Barrey C, Jund J, Noseda O, Roussouly P. Sagittal balance of the pelvis-spine complex and lumbar degenerative diseases. A comparative study about 85 cases. *Eur Spine J* 2007;16(9):1459–1467.

24. Strube P, Pumberger M, Sonnow L, Zippelius T, Nowack D, Zahn RK, et al. Association between lumbar spinal degeneration and anatomic pelvic parameters. *Clin Spine Surg* 2018;31(6):263–267.

25. Yoshimoto H, Sato S, Masuda T, Kanno T, Shundo M, Hyakumachi T, et al. Spinopelvic alignment in patients with osteoarthrosis of the hip: A radiographic comparison to patients with low back pain. *Spine (Phila Pa 1976)* 2005;30(14):1650–1657.

26. Eftekhary N, Buckland A, Schwarzkopf R, Shimmin A, Lazennec JY, Dorr LD, et al. A systematic approach to the hip-spine relationship and its applications to total hip arthroplasty. *Bone Jt J* British Editorial Society of Bone and Joint Surgery, 2019;101 B(7):808–816.

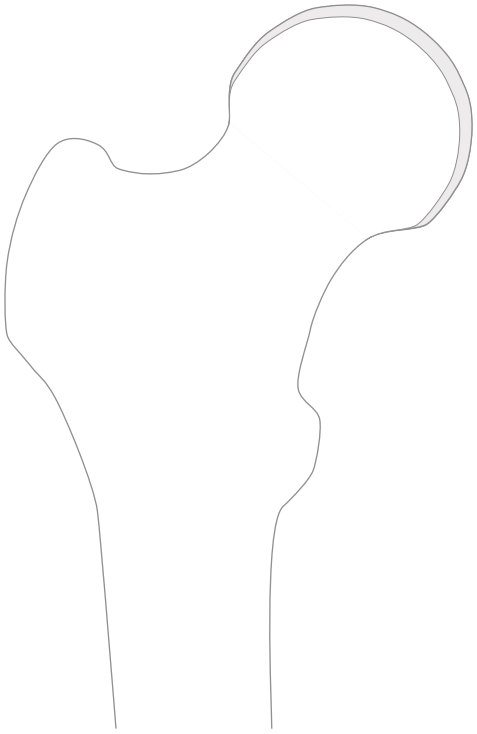
27. Innmann MM, Merle C, Gotterbarm T, Ewerbeck V, Beaulé PE, Grammatopoulos G. Can spinopelvic mobility be predicted in patients awaiting total hip arthroplasty? A prospective, diagnostic study of patients with end-stage hip osteoarthritis. *Bone Jt J* British Editorial Society of Bone and Joint Surgery, 2019;101 B(8):902–909.

28. Schlösser TPC, Janssen MMA, Vrtovec T, Pernuš F, Öner FC, Viergever MA, et al. Evolution of the ischio-iliac lordosis during natural growth and its relation with the pelvic incidence. *Eur Spine J* 2014;23(7):1433–1441.

29. Bakouny Z, Assi A, Massaad A, Saghbini E, Lafage V, Skalli W, et al. Roussouly's sagittal spino-pelvic morphotypes as determinants of gait in asymptomatic adult subjects. *Gait Posture* 2017;54:27–33.

30. Murata Y, Takahashi K, Yamagata M, Hanaoka E, Moriya H. The knee-spine syndrome. Association between lumbar lordosis and extension of the knee. *J Bone Joint Surg Br* 2003;85(1):95–9.

31. Lee CS, Park SJ, Chung SS, Lee KH. The effect of simulated knee flexion on sagittal spinal alignment: novel interpretation of spinopelvic alignment. *Eur Spine J* 2013;22(5):1059–65.



# CHAPTER 9

# Unravelling the hip-spine dilemma from the CHECK-cohort; is sagittal pelvic morphology linked to radiographic signs of femoroacetabular impingement?

*The Bone & Joint Journal*
*Submitted for publication*

Joost H.J. van Erp, Willem Paul Gielis, Vahid Arbabi, Arthur de Gast, Harrie Weinans, Laurens Kaas, René Castelein, Tom P.C. Schlösser

## Abstract

### Background and aim

To date the aetiology of femoroacetabular impingement (FAI) is still not completely understood. There are mechanical theories that suggest that development of symptomatic FAI is linked to sagittal pelvic morphology and spinopelvic-femoral dynamics. The aim of this cohort study is to evaluate the relation of sagittal pelvic morphology and orientation with radiographic signs of FAI. Additionally, we test whether the relation between FAI and spinopelvic parameters differs in osteoarthritic hips.

### Patients and methods

From a prospective observational cohort study, 1002 patients between 45–65 years old with a first episode of knee or hip pain were followed for eight years. All patients that had lateral lumbar radiographs and clinical and radiographic follow-up of the hips during eight years were included in the present study. Range of internal rotation of the hip as well as radiographic parameters for FAI (alpha and Wiberg angle) and the presence of hip osteoarthritis (Kellgren and Lawrence) were systematically measured at baseline. Pelvic incidence (PI), pelvic tilt (PT), sacral slope (SS)) were measured at 8-year follow-up. Associations between PI, PT, SS and FAI parameters was tested using generalized estimating equations.

### Results

421 subjects, 842 hips, were included. No significant relation between PI, PT or SS and alpha or Wiberg angle was found. Comparison of hips with and without radiological sign(s) of FAI showed no differences in PI, PT or SS. There was no relation between range of internal rotation of the hip and the spinopelvic parameters.

### Conclusion

Sagittal pelvic morphology and orientation are not related to the presence of radiological signs of FAI in this study population.

## Introduction

Femoroacetabular impingement (FAI) is a well-known cause of hip pain in young adults and is strongly related to the development of hip osteoarthritis (OA)[1,2]. It is believed to originate from pathologic contact between skeletal prominences of the proximal femur and the acetabular rim or labrum that limits hip range-of-motion, mostly flexion, adduction and internal rotation. Based on the radiographic signs of FAI, the pathoanatomy can be divided in three types: Cam, pincer or combined impingement. Cam deformities are characterized by aspherical deformation of the anterolateral caput-column junction that mostly develops in young men, who practice contact sports during adolescence, like rugby, soccer or hockey[3,4], which can result later on in an symptomatic FAI. Pincer-type deformities are characterized by focal or general over-coverage of the femoral head and can become symptomatic in woman in their 40s.[5]

There is recent evidence that demonstrates that the likelihood of damage to the joint cartilage or labrum in FAI is related to the shape and orientation of the pelvis, and the individual's spino-pelvic-femoral dynamics.[6,7,8] In 1992, Duval-Beaupère introduced the "pelvic incidence" as a key parameter for sagittal pelvic morphology, and pelvic tilt (PT) and sacral slope (SS) for pelvic orientation[9] (**Figure 1**). PI represents the orientation of the sacrum within the pelvic ring in the sagittal plane, and is defined as the angle between the line connecting the hip-axis to the midpoint of the sacrum and a line perpendicular to the sacral endplate. PT and SS represent the pelvic and sacral endplate orientation, relative to the vertical and horizontal, respectively. PI varies enormously within the human population (from 33°-85°) and is highly related to one's sagittal pelvic orientation.[10] By anterior and posterior PT around the hips, the pelvis plays a fundamental role in regulating sagittal spino-pelvic-femoral alignment and the onset the most common lumbar degenerative pathologies.[11,12] Patients with relative low PI or with lumbar degeneration have limited ability for posterior PT when changing from standing to sitting[13]. Anterior and posterior pelvic rotations around the hip axis will have important consequences for the biomechanical loading and range-of-motion of the hips as well. A low PI and PT limits the anterior 'opening' of the acetabular cavities when changing from standing to sitting, with a higher likelihood of impingement. Recently, this hypothesis has been confirmed in two cross-sectional studies that compared the PI of symptomatic FAI patients with a an asymptomatic control group.[14,15] They describe that the PI is on average 4-6° lower in symptomatic FAI patients compared to asymptomatic controls.[14,15] To date, no large-scale prospective cohort study explored the relation between sagittal pelvic

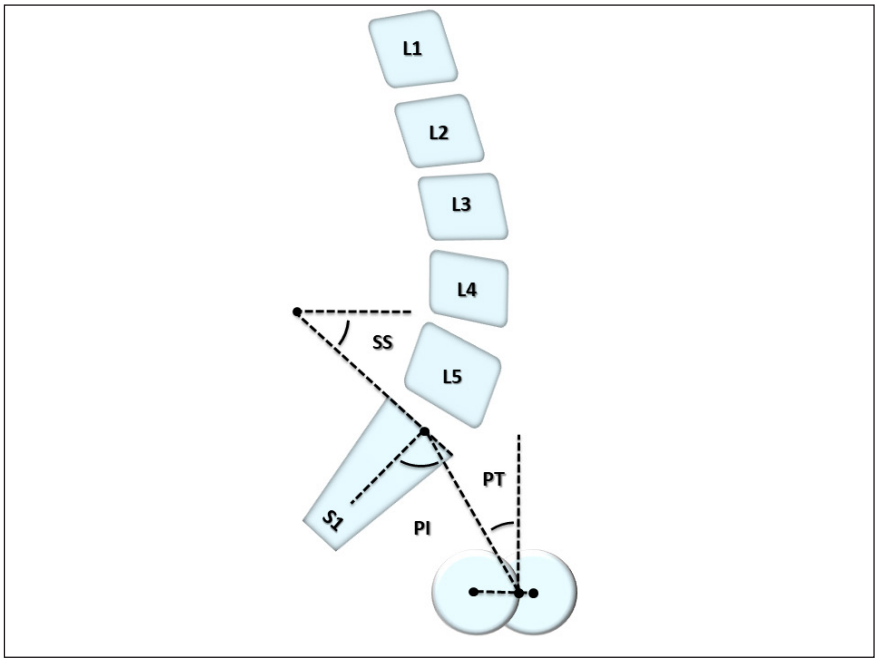



**Figure 1.** Schematic drawing showing pelvic incidence (PI), pelvic tilt (PT), and sacral slope (SS).

parameters, hip range-of-motion and the presence of radiological signs of FAI and/or development of hip OA.

The aim of this study is to evaluate the relation between sagittal pelvic morphology and radiographic parameters of FAI and range-of-internal rotation. Additionally, we test whether the relation between FAI and spinopelvic parameters differs in osteoarthritic hips. For this purpose, data from the prospective CHECK-study (Cohort Hip and Cohort Knee), initiated by the Dutch Arthritis Foundation, was used.[16]

## Methods

### Study population

For this study, all patients in the CHECK-database for which pelvic and lateral lumbar radiographs were available and completed eight-year follow-up, were included. CHECK is a multicentre population based, prospective observational

cohort study on 1002 patients between 45 and 65 years old that presented themselves with <6 months pain of the hip- or knee that were yearly clinically and radiographically followed for ten years[16]. Patients with previous hip or knee surgery, rheumatic diseases, treatment for developmental dysplasia of the hip, osteochondritis dissecans, intra-articular fractures, Perthes' disease, traumatic ligament or meniscus damage, plica syndrome or a Bakers cyst were excluded in the original study. For this study, we also excluded patients of which the pelvic or lateral lumbar radiographs did not include the full pelvis or femoral heads. CHECK was approved by the medical ethic committee and all participants provided informed consent, before inclusion in the study.

### Radiographic analyses

By protocol, lateral standing spine radiographs were obtained at eight-year follow-up. These were used for assessment of PI, PT and SS.[17] Anterior-posterior pelvic radiographs collected at baseline were used in this study to assess the presence of a cam deformity (alpha angle >60°) or pincer deformity (Wiberg angle >40°) and the presence of hip OA of both hips.[18] Both hips were assessed separately, but we corrected for intra-person correlations in the statistical analysis. The following hip parameters were semi-automatically measured in a special developed tool in PSF Python 3.6:

- Alpha angle
- Wiberg angle

The alpha angle is the angle between a centred line through the femoral shaft axis and a line from the centre of the femoral head to the point where the femoral head becomes aspheric. The Wiberg angle is the angle between a perpendicular line between both femoral heads and a line between the centre of the femoral head and the lateral acetabular border. The presence of hip OA was classified according to Kellgren and Lawrence (KL) for previous projects, by five observers with an intraclass correlation coefficient for inter-observer reliability of >0.9.[19,20]

### Clinical parameters

At baseline, all individuals underwent clinical assessment by a trained health professional. The clinical parameters used for this study were the range of internal rotation of the left and right hip. Anterior impingement test (FADIR) was not systematically performed.

### Statistical analysis

Data were processed and analyzed using Microsoft Excel 2010 and IBM SPSS 23. The alpha angle showed a strong bimodal distribution and was therefore dichotomized according to literature (<60°/≥60°).[18] The association between PI and the alpha angle, Wiberg angle, and internal hip rotation was tested using generalized estimating equations for logistic or linear regression, accounting for the correlation between both hips within each patient. All analyses were also adjusted for age and sex, to prevent for confounding. Additionally, the effect of SS and PT on the same outcomes were tested in separate models. An α of 0.05 was used to test for statistical significance.

## Results

In CHECK, 1002 subjects were included. After eight years, clinical data and radiographs of the hip were available for 845 participants (84%), 157 were lost to follow-up. For 421 participants, 842 hips, an appropriate spinal and pelvic radiograph was available for measurement of the PI, PT, SS and FAI parameters and they were included in the present study. Of the majority of the excluded subjects, the lateral lumbar spinal radiographs did not include the femoral heads making PI and PT measurement impossible. 308 (73%) of the included subjects was female. The mean age of the participants was 56±5 years, mean BMI was 27±4. 409 (97%) were Caucasian. Spinopelvic parameters are shown in **(Table 1)**.

**Table 1.** Prevalence of radiological signs of femoroacetabular impingement and hip osteoarthritis and pelvic parameters in the study population.

| | |
|---|---|
| Hip osteoarthritis, percentage (KL≥2) | 89% |
| Alpha angle, mean SD | 54 ± 23 |
| Percentage (≥60°) | 80% |
| Wiberg angle, mean SD | 36 ± 7 |
| Percentage (≥40°) | 30% |
| Hip internal rotation, mean SD | 31 ± 8 |
| Pelvic incidence, mean SD | 58 ± 13 |
| Pelvic tilt, mean SD | 25 ± 11 |
| Sacral slope, mean SD | 34 ± 11 |
| Lumbar lordosis, mean SD | 42 ± 13 |

KL=Kellgren and Lawrence

### Pelvic parameters versus alpha and Wiberg

Analyses between pelvic parameters and the alpha angle and Wiberg angle showed no statistically significant correlation **(Table 2, Figure 2 and 3).**

**Table 2.** Relation between pelvic parameters and femoroacetabular impingement and internal hip rotation.

| | β | p-value | 95% Confidence interval |
|---|---|---|---|
| **Alpha** | | | |
| Pelvic incidence | 0.996 | 0.680 | 0.975-1.017 |
| Pelvic tilt | 0.997 | 0.840 | 0.972-1.023 |
| Sacral slope | 0.999 | 0.930 | 0.972-1.026 |
| **Wiberg angle** | | | |
| Pelvic incidence | -0.023 | 0.455 | -0.085-0.038 |
| Pelvic tilt | -0.023 | 0.579 | -0.105-0.059 |
| Sacral slope | -0.026 | 0.459 | -0.094-0.043 |
| **Hip internal rotation** | | | |
| Pelvic incidence | 0.047 | 0.067 | -0.003-0.097 |
| Pelvic tilt | 0.051 | 0.088 | -0.008-0.110 |
| Sacral slope | 0.038 | 0.253 | -0.027-0.104 |

*OR instead of B due to logistic regression analysis

### Pelvic parameters versus hip range-of-motion

Analyses between pelvic parameters and clinical hip internal rotation showed no correlation between pelvic parameters and internal hip rotation **(Table 2).**

### Pelvic parameters in FAI

The pelvic parameters of patients with or without radiological signs of FAI are shown in **Table 3 and 4**. No correlation was found between pelvic parameters and radiological signs of FAI.

### Pelvic parameters in relation to hip OA

94 (11%) hips had radiological evidence of hip OA (KL≥2). Comparisons of the pelvic parameters between subjects with and without OA and radiological signs of FAI are shown in **Table 4**.

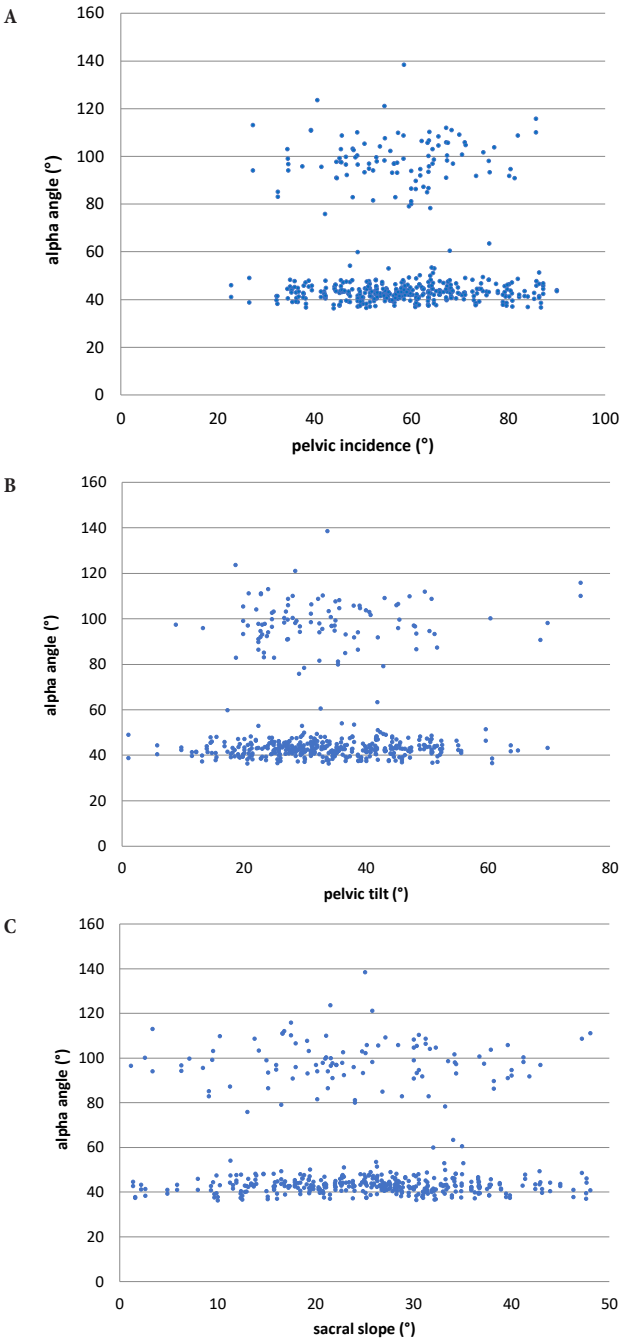



**Figure 2 A)** Pelvic incidence versus alpha angle **B**) Pelvic tilt versus alpha angle **C**) Sacral slope versus alpha angle.

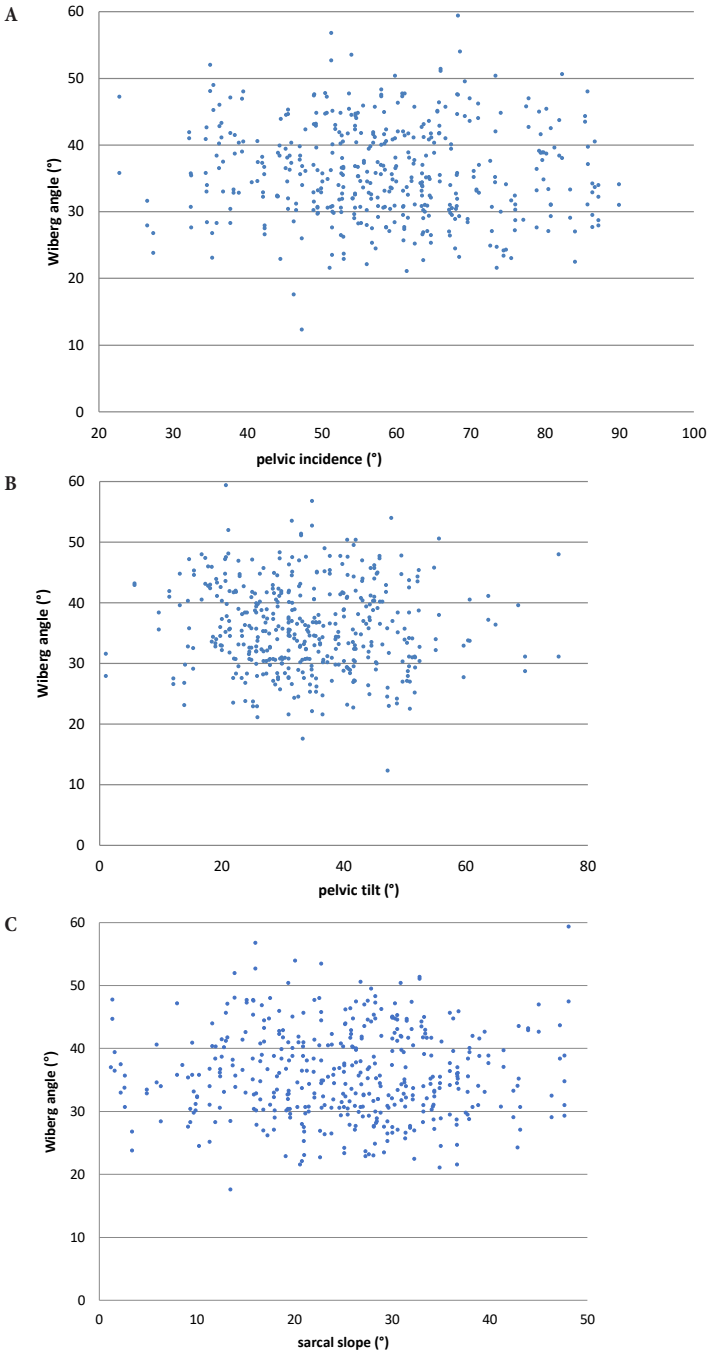



**Figure 3 A)** Pelvic incidence versus Wiberg Angle **B**) Pelvic tilt versus Wiberg Angle **C**) Sacral slope versus Wiberg Angle

**Table 3.** Pelvic parameters for patients with or without a radiographic signs of femoroacetabular impingement (FAI), with exclusion of patients with significant hip osteoarthritis (OA) at baseline (Kellgren and Lawrence ≥2), presented as mean ± SD(°)

| | No FAI (n=263) | Cam deformity only (n=84) | Pincer deformity only (n=84) | Mixed (n=21) |
|---|---|---|---|---|
| Pelvic incidence | 59,1±13,7 | 56,0±13,0 | 56,3±13,4 | 61,7±12,4 |
| Pelvic tilt | 25,7±11,0 | 23,4±10,6 | 23,9±10,6 | 27,2±9,6 |
| Sacral slope | 34,0±11,2 | 33,3±12,2 | 32,5±12,3 | 35,5±13,0 |

**Table 4.** Pelvic parameters for patients with or without hip osteoarthritis (OA) at baseline (Kellgren and Lawrence ≥2) and a radiographic signs of femoroacetabular impingement (FAI), presented as mean ± SD(°)

| | Hip without OA | | | Hip with OA | | |
|---|---|---|---|---|---|---|
| | FAI (n=219) | No FAI (n=263) | P-value | FAI (n=57) | No FAI (n=20) | P-value |
| Pelvic incidence | 56,7±13,1 | 59,1±13,7 | 0,153 | 59,3±13,1 | 55,3±10,0 | 0,340 |
| Pelvic tilt | 24,0±10,5 | 25,7±11,0 | 0,264 | 23,6±11,9 | 23,4±7,7 | 0,811 |
| Sacral slope | 33,1±12,2 | 34,0±11,2 | 0,432 | 36,0±10,7 | 33,1±11,6 | 0,369 |

## Discussion

The role of sagittal pelvic morphology and orientation on the development of FAI was systematically studied in a cohort of 1002 patients between 45 and 65 years old that presented themselves with <6 months pain of the hip or knee. Interestingly, in contrast to previous case-control studies on symptomatic FAI patients, we found no correlation between radiological signs of FAI and sagittal pelvic parameters in this cohort. Furthermore, we found no relation between range of internal rotation and sagittal pelvic parameters. These results suggest that in our cohort, there is limited etiological relevance of sagittal pelvic morphology for development of a symptomatic cam deformity or pincer lesion.

Humans have unique spino-pelvic-femoral alignment, different from any other bipedal mammal. It enables to stand fully upright with extended knees and hips. As a result of a lordosis within the pelvis, humans are able to keep the centre of gravity of the trunk and head straight above the pelvis in the upright position [21]. The pelvis has been recognized in relation to normal functioning of the human spinopelvic complex, because by anterior and posterior pelvic tilt, it is a key

regulator of the sagittal configuration of the spine.[22,12,23] The 'PI' was introduced as a parameter for sagittal pelvic morphology and PT and SS for sagittal pelvic orientation.[9] Sagittal pelvic morphology and orientation varies widely within the human population and plays a significant role in the onset of different degenerative spinal pathologies.[11,21,24,25] Variation in sagittal pelvic morphology and orientation relative to the femoral heads does not only lead to differences in the spinopelvic configuration, but also to different mechanical loading of the femoroacetabular joints. Patients with a low PI as well as patients with degenerative pelvic disorders, have less ability to retrovert the pelvis when changing from standing to sitting. It can be inferred that in those patients the hip joints are loaded more towards their limit of extension in the standing position. The relative anterior over-coverage of the acetabulum could pose higher joint-reaction forces on the cranial and anterior labrum as well as cause anterior impingement.

The pathomechanism of symptomatic FAI remains controversial.[26] Acquired causes as well as genetic predisposition have been implicated in its aetiology.[27] Most of the research on FAI etiopathogenesis has been on the femoral head-neck deformity, acetabular coverage and mechanical impingement. Acquired factors involved in the development of cam-type deformities are repetitive injury of the physis of the femoral head during adolescence. This is supported by the high incidence of cam-type deformities in adolescents participating in high intensity and frequency sports such as soccer, skiing and ice-hockey.[3,28,29] From a genetic perspective, acetabular overcoverage has been associated with certain genotypes. This may also represent the genetic inheritability of certain sagittal spinopelvic alignment.[30]

In the sagittal plane, the motion of the spine has a close relationship with the pelvic-femoral joint. In a study of Grantham et al. individuals, with radiographic signs of FAI had a more rigid spine, but were able to establish more hip flexion compared to controls. This increased amount of required hip flexion, might result in more events of impingement, and development of symptomatic FAI. The asymptomatic individuals with radiological signs of FAI, had a higher ability of spine flexion, but similar hip flexion, so more mechanisms for compensation.[27] Symptomatic patients had more hip flexion in sitting position, due to their lack of compensation in the lumbar spine. More anterior pelvic tilt is required then, which may lead to impingement between the acetabulum and proximal femur. Femoral retroversion might contribute to the development of symptomatic FAI as well, since the anterior neck of a retroverted hip easily collides with the acetabulum labrum in slight degrees of internal rotation. Variations in spinopelvic dynamics between subjects with radiographic FAI may lead to symptoms of FAI.[27]

In the present study, 22% of the study population presented with a first episode of hip or knee pain demonstrated radiographic signs of a cam deformation. In comparison, a systematic review demonstrated a prevalence of 5-75%, varying between 30 studies.[31] In this review, it was unable to demonstrate a higher prevalence in certain subgroups, such as athletes or patients with hip pain. This indicates that radiographic signs of FAI often occur in asymptomatic individuals, so possibly other factors (such as certain acetabular/ proximal femoral morphology or developmental axial rotational deformity of the lower limb are required to develop symptomatic FAI.

Previous studies on sagittal spinopelvic alignment in FAI patients showed that FAI is more prevalent in individuals with a low PI or low SS. This lead to the hypothesis that FAI is a result of different mechanical loading on lumbosacral junctions[6,7]. Hellman et al., recently studied a retrospective cohort of 40 patients with symptomatic FAI, and found a significant lower PI compared to asymptomatic subjects[15]. It was stated that individuals prone for FAI only become symptomatic when the pelvic-femoral compensation mechanism is insufficient. Our study confirms the high prevalence of radiographic signs of FAI (cam and pincer) in the population, instead of the theory that these signs are a result of developing, symptomatic FAI. Differences in spinopelvic morphology in individuals with cam lesions might result in symptomatic FAI. It can be stated that a cam lesion, combined with a low PI, might result in symptomatic FAI.

In earlier studies, based on the same CHECK-cohort, a relation was found between FAI and hip OA. Recently, from the same cohort the possible relation between pelvic morphology and hip OA was explored. There was a trend towards lower PI (and PT) in patients with significant hip OA[11], a higher prevalence of FAI in subjects with a low PI might explain this phenomenon. Pelvises with low PI, may cause increased force on the anterior labrum of the hip during hip flexion, which enables FAI to develop during adolescence, resulting in early OA of the hip.

This study has a few limitations. FADIR/FABER tests of the hip were not systematically performed. These are described as reliable tests with high diagnostic accuracy to assess symptomatic FAI.[32] We explored, however, the relationship between decreased ROM of internal rotation (<20°), another typical finding in patients with FAI[5,32] and PI (**Table 2**), but could not find a relationship. Another limitation is that lateral lumbar radiographs were only available at eight-year follow-up. Based on previous studies, however, it can be assumed that the PI did not change in individuals during the study, since it remains constant after the

adolescent growth spurt, the same period cam deformities develop[4,10]. Before the inclusion, some patients may have developed significant hip OA. Because the development of osteophytes might complicate measurements of the alpha angle and Wiberg angle, we chose to exclude patients with significant OA. This might have introduced selection bias.

In future studies the relation between individual varieties in pelvic morphology and the prevalence of (symptomatic) FAI should be further explored. To better understand the development of FAI, dynamic imaging might be required in individuals with different pelvic morphology, to study the 3D orientation between the pelvis and the proximal femoral in movements which might elicit FAI. Clinical findings of large prospective cohorts of patients should be combined with radiographic signs of FAI, in relation to pelvic parameters on radiographs, or more precise, CT or MR imaging.

## Conclusion

In the present study, no evidence was found for a relation between pelvic parameters (PI, PT and SS) and femoroacetabular impingement of the cam-type FAI nor pincer. Sagittal pelvic morphology and orientation are not directly related to the presence of radiological signs of FAI, which has a high prevalence in the general population. However, individuals with low pelvic tilt may be more at risk for development of symptoms of FAI.



## Acknowledgements

The CHECK-cohort study is funded by the Dutch Arthritis Foundation. Involved are: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen /Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede /Ziekenhuisgroep Twente Almelo; Reade Center for Rehabilitation and Rheumatology; St.Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

## References


1. Ganz R, Parvizi J, Beck M, Leunig M, Nötzli H, Siebenrock KA. Femoroacetabular Impingement: A Cause for Osteoarthritis of the Hip. *Clin Orthop Relat Res* Lippincott Williams and Wilkins, 2003:112–120.

2. Agricola R, Heijboer MP, Bierma-Zeinstra SMA, Verhaar JAN, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: A nationwide prospective cohort study (CHECK). *Ann Rheum Dis* 2013;72(6):918–923.

3. Knapik DM, Gaudiani MA, Camilleri BE, Nho SJ, Voos JE, Salata MJ. Reported Prevalence of Radiographic Cam Deformity Based on Sport: A Systematic Review of the Current Literature. *Orthop J Sport Med* 2019;7(3):2325967119830873.

4. Agricola R, Heijboer MP, Ginai AZ, Roels P, Zadpoor AA, Verhaar JAN, et al. A cam deformity is gradually acquired during skeletal maturation in adolescent and young male soccer players: A prospective study with minimum 2-year follow-up. *Am J Sports Med* SAGE Publications Inc., 2014;42(4):798–806.

5. Pun S, Kumar D, Lane NE. Review: Femoroacetabular impingement. Arthritis Rheumatol. John Wiley and Sons Inc., 2015:17–27.

6. Saltychev M, Pernaa K, Seppänen M, Mäkelä K, Laimi K. Pelvic incidence and hip disorders: A systematic review and quantitative analysis. *Acta Orthop* 2018;89(1):66–70.

7. Pierannunzii L. Pelvic posture and kinematics in femoroacetabular impingement: a systematic review. J. Orthop. Traumatol. Springer-Verlag Italia s.r.l., 2017:187–196.

8. Gebhart JJ, Weinberg DS, Bohl MS, Liu RW. Relationship between pelvic incidence and osteoarthritis of the hip. *Bone Jt Res* 2016;5(2):66–72.

9. Duval-Beaupère G, Schmidt C, Cosson P. A barycentremetric study of the sagittal shape of spine and pelvis: The conditions required for an economic standing position. *Ann Biomed Eng* 1992;20(4):451–462.

10. Schlösser TPC, Janssen MMA, Vrtovec T, Pernuš F, Öner FC, Viergever MA, et al. Evolution of the ischio-iliac lordosis during natural growth and its relation with the pelvic incidence. *Eur Spine J* 2014;23(7):1433–1441.

11. Erp JHJ van, Gielis WP, Arbabi V, Gast A de, Weinans H, Arbabi S, et al. Unravelling the knee-hip-spine trilemma from the CHECK study. *Bone Joint J* NLM (Medline), 2020;102-B(9):1261–1267.

12. Roussouly P, Gollogly S, Berthonnaud E, Dimnet J. Classification of the normal variation in the sagittal alignment of the human lumbar spine and pelvis in the standing position. *Spine (Phila Pa 1976)* 2005;30(3):346–353.

13. Konyves A. Editorial Commentary: Looking Past the Hip Joint—The Role of Pelvic Incidence in Femoroacetabular Impingement. *Arthrosc - J Arthrosc Relat Surg* 2017;33(3):551–552.

14. Lawton CD, Butler BA, Selley RS, Barth KA, Balderama ES, Jenkins TJ, et al. Pelvic incidence in a femoroacetabular impingement population. *J Orthop* Reed Elsevier India Pvt. Ltd., 2020;22:90–94.

15. Hellman MD, Haughom BD, Brown NM, Fillingham YA, Philippon MJ, Nho SJ. Femoroacetabular Impingement and Pelvic Incidence: Radiographic Comparison to an Asymptomatic Control. *Arthrosc - J Arthrosc Relat Surg* W.B. Saunders, 2017;33(3):545–550.

16. Wesseling J, Boers M, Viergever MA, Hilberdink WKHA, Lafeber FPJG, Dekker J, et al. Cohort profile: Cohort Hip and Cohort Knee (CHECK) study. *Int J Epidemiol* 2016;45(1):36–44.

17. Legaye J, Duval-Beaupère G, Hecquet J, Marty C. Pelvic incidence: A fundamental pelvic parameter for three-dimensional regulation of spinal sagittal curves. *Eur Spine J* 1998;7(2):99–103.

18. Klij P van, Reiman MP, Waarsing JH, Reijman M, Bramer WM, Verhaar JAN, et al. Classifying Cam Morphology by the Alpha Angle: A Systematic Review on Threshold Values. *Orthop J Sport Med* SAGE Publications Ltd, 2020;8(8):232596712093831.

19. Damen J, Schiphof D, Wolde S Ten, Cats HA, Bierma-Zeinstra SMA, Oei EHG. Inter-observer reliability for radiographic assessment of early osteoarthritis features: The CHECK (cohort hip and cohort knee) study. *Osteoarthr Cartil* 2014;22(7):969–974.

20. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* Ann Rheum Dis, 1957;16(4):494–502.

21. Schlösser TPC, Janssen MMA, Hogervorst T, Vrtovec T, Vos J de, Öner FC, et al. The odyssey of sagittal pelvic morphology during human evolution: a perspective on different Hominoidae. *Spine J* Elsevier, 2017;17(8):1202–1206.

22. Mac-Thiong JM, Labelle H, Berthonnaud E, Betz RR, Roussouly P. Sagittal spinopelvic balance in normal children and adolescents. *Eur Spine J* 2007;16(2):227–234.

23. Huec JC Le, Saddiki R, Franke J, Rigal J, Aunoble S. Equilibrium of the human body and the gravity line: the basics. *Eur Spine J* 2011;20 Suppl 5(S5):558–563.

24. Barrey C, Jund J, Noseda O, Roussouly P. Sagittal balance of the pelvis-spine complex and lumbar degenerative diseases. A comparative study about 85 cases. *Eur Spine J* 2007;16(9):1459–1467.

25. Labelle H, Roussouly P, Berthonnaud É, Dimnet J, O'Brien M. The importance of spino-pelvic balance in L5-S1 developmental spondylolisthesis: A review of pertinent radiologic measurements. *Spine (Phila Pa 1976)* 2005;30(6 SPEC. ISS.):S27–S34.

26. Leunig M, Beck M, Dora C, Ganz R. Femoroacetabular impingement: Etiology and surgical concept. *Oper Tech Orthop* W.B. Saunders, 2005;15(3):247–255.

27. Grantham WJ, Philippon MJ. Etiology and Pathomechanics of Femoroacetabular Impingement. Curr. Rev. Musculoskelet. Med. Humana Press Inc., 2019:253–259.

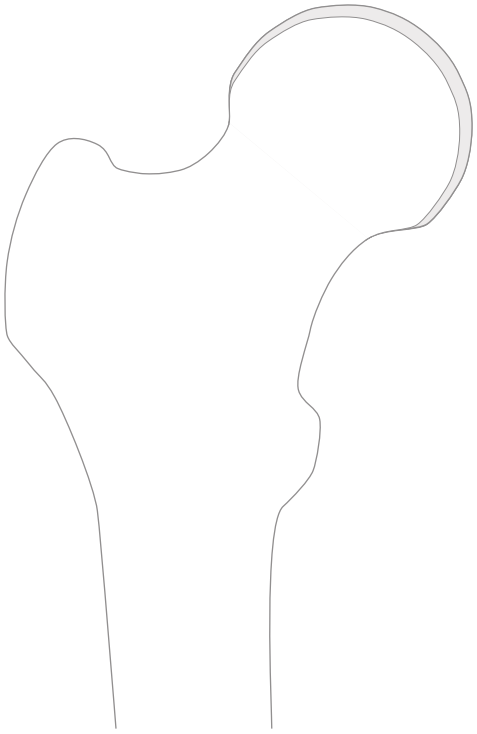
28. Klij P van, Ginai AZ, Heijboer MP, Verhaar JAN, Waarsing JH, Agricola R. The relationship between cam morphology and hip and groin symptoms and signs in young male football players. *Scand J Med Sci Sport* Blackwell Munksgaard, 2020;30(7):1221–1231.

29. Aminoff AS, Agnvall C, Todd C, Jónasson P, Thoreson O, Sansone M, et al. Young elite Alpine and Mogul skiers have a higher prevalence of cam morphology than non-athletes. *Knee Surgery, Sport Traumatol Arthrosc* Springer, 2020;28(4):1262–1269.

30. Sekimoto T, Kurogi S, Funamoto T, Ota T, Watanabe S, Sakamoto T, et al. Possible association of single nucleotide polymorphisms in the 3' untranslated region of HOXB9 with acetabular overcoverage. *Bone Jt Res* British Editorial Society of Bone and Joint Surgery, 2015;4(4):50–55.

31. Dickenson E, Wall PDH, Robinson B, Fernandez M, Parsons H, Buchbinder R, et al. Prevalence of cam hip shape morphology: A systematic review. Osteoarthr. Cartil. W.B. Saunders Ltd, 2016:949–961.

32. Pålsson A, Kostogiannis I, Ageberg E. Combining results from hip impingement and range of motion tests can increase diagnostic accuracy in patients with FAI syndrome. *Knee Surgery, Sport Traumatol Arthrosc* Springer Science and Business Media Deutschland GmbH, 2020;28(10):3382–3392.



# CHAPTER 10

# An automated workflow based on hip shape improves personalized risk prediction for hip osteoarthritis in the CHECK study

*Osteoarthritis and Cartilage*
*January 2020, doi.org/10.1016/j.joca.2019.09.005*

Willem Paul Gielis, Harrie Weinans, Paco M.J. Welsing, Willem E. van Spil, Rintje Agricola, Timothy F. Cootes, Pim A. de Jong, Claudia Lindner

## Abstract

### Background and purpose

To design an automated workflow for hip radiographs focused on joint shape and tests its prognostic value for future hip osteoarthritis.

### Patients and methods

We used baseline and 8-year follow-up data from 1,002 participants of the CHECK-study. The primary outcome was definite radiographic hip osteoarthritis (rHOA) (Kellgren–Lawrence grade ≥2 or joint replacement) at 8-year follow-up. We designed a method to automatically segment the hip joint from radiographs. Subsequently, we applied machine learning algorithms (elastic net with automated parameter optimization) to provide the Shape-Score, a single value describing the risk for future rHOA based solely on joint shape. We built and internally validated prediction models using baseline demographics, physical examination, and radiologists scores and tested the added prognostic value of the Shape-Score using Area-Under-the-Curve (AUC). Missing data was imputed by multiple imputation by chained equations. Only hips with pain in the corresponding leg were included.

### Results

84% were female, mean age was 56 (±5.1) years, mean BMI 26.3 (±4.2). Of 1,044 hips with pain at baseline and complete follow-up, 143 showed radiographic osteoarthritis and 42 were replaced. 91.5% of the hips had follow-up data available. The Shape-Score was a significant predictor of rHOA (odds ratio per decimal increase 5.21, 95%-CI (3.74–7.24)). The prediction model using demographics, physical examination, and radiologists scores demonstrated an AUC of 0.795, 95%-CI (0.757–0.834). After addition of the Shape-Score the AUC increased to 0.864, 95%-CI (0.833–0.895).

### Conclusion

Our Shape-Score, automatically derived from radiographs using a novel machine learning workflow, may strongly improve risk prediction in hip osteoarthritis.

## Introduction

Hip osteoarthritis (HOA) is often diagnosed relatively late in the disease process and currently there are no drugs available to modify disease progression. Therefore, initial HOA treatment is necessarily restricted to education, exercise, weight loss and analgesics. Total hip replacement (THR) often follows when these do not suffice.[1]

To guide current care and develop interventions to modify the disease course, accurate prediction of HOA development in patients presenting with hip complaints is important. Many risk factors for HOA are reported in the literature. However, no established risk prediction tool for HOA is currently available. The rise of automated image processing techniques using artificial intelligence, offers the possibility to extract information from images beyond traditional visual interpretation. For example, deep neural networks can be used on computed tomography scans to classify arterial calcifications or pulmonary peri-fissural nodules. [2,3]

Shape variations in the hip play a role in the development of HOA.[4,5] Geometric measurements for assessing hip dysplasia or cam morphology are used as clinical tools[6], but only describe particular components of the hip shape. Statistical Shape Models (SSMs) have the potential to quantify the overall shape variation of the bone, including more subtle variations.[7–10] However, SSMs require labor-intensive manual input to outline (i.e. segment) shapes with landmark points, hampering their use in large study populations. Therefore, we developed a segmentation software system to automatically extract hip shape from standard pelvic radiographs.[11,12]

This study describes the development and validation of a prediction tool for future HOA in a large cohort of patients with hip pain that had never or only recently (<6 months) consulted a physician for their complaints. Our prediction tool, Shape-Score, utilizes overall hip shape based on SSMs, using our segmentation software system on standard pelvic radiographs. Moreover, we quantified the added predictive value of our tool over clinically available predictors alone.

## Methods

### Participants

Cohort Hip and Cohort Knee (CHECK) aimed to examine the course of early OA in the hip and/or knee.[13] Between October 2002 and September 2005, 1002 participants were enrolled in 10 participating centers throughout the Netherlands. Potential candidates were approached by their general practitioner and/or recruited via local media. Participants were aged 45-65 years at the time of inclusion and

had pain and/or stiffness in at least one knee and/or hip. They had not, or only recently (<6 months) consulted a physician for these complaints. Exclusion criteria were (i) pathology other than OA explaining symptoms, (ii) expected inability to complete 10-year follow-up, and (iii) inability to sufficiently understand Dutch. Radiographic knee OA (defined as KL 2 or higher) was not present in patients at baseline.

## Measurements

### Demographics

Age, gender, BMI and current smoking (yes/no) were registered. Highest education level was scored on a scale from 1 to 8, as a proxy for socio-economic status. The scale is described in **Table 1**.

**Table 1.** Baseline characteristics (n=1140 painful hips)[1]

| | | | (Data based on N=) |
|---|---|---|---|
| **Demographics** | | | |
| Age in years, mean (SD) | 55.9 | (5.1) | 1140 |
| BMI in kg/m², mean (SD) | 26.3 | (4.2) | 1119 |
| Highest education, median (IQR)[2] | 5 | (4 to 7) | 1109 |
| Female sex, N (%) | 954 | (87.3) | 1140 |
| Current smoking, N (%) | 158 | (14.2) | 1115 |
| **Clinical examination** | | | |
| WOMAC total score, mean (SD) | 24.9 | (16.3) | 1113 |
| Pain located around hip, N (%) | 673 | (59.4) | 1133 |
| Pain located around knee, N (%) | 961 | (84.5) | 1137 |
| Analgesic use, N (%) | 387 | (34.5) | 1123 |
| Morning stiffness, N (%) | 423 | (38.3) | 1104 |
| Range of motion hip, mean (SD) | 30.6 | (8.7) | 1134 |
| Pain on internal hip rotation, N (%) | 342 | (30.2) | 1134 |
| **Basic radiographic parameters** | | | |
| Hip Joint space narrowing, N (%) | 367 | (32.8) | 1108 |
| Hip Osteophytosis, N (%) | 374 | (33.8) | 1118 |
| Hip Buttressing, N (%) | 64 | (5.7) | 1118 |
| Hip KL grade 1, N(%)[3] | 434 | (38.6) | 1118 |
| Knee KL grade 1, N(%)[3] | 266 | (23.8) | 1140 |
| **Shape-Score** | | | |
| Shape-Score, mean (SD) | 0.18 | (0.066) | 1140 |

[1] As missing values are drawn from a distribution in multiple imputations, the number of included hip joints (range 1140 to 1143) differed per imputed dataset.
[2] 1= no school, 2= primary school, 3= basic vocational education, 4= secondary education, 5=secondary vocational education, 6= Higher and university preparatory education, 7= higher professional education, 8= university), IQR = interquartile range
[3] Kellgren and Lawrence grade for radiographic osteoarthritis, all hips and knees were graded 0 or 1 at baseline

### Clinical examination

Trained health professionals registered hip pain, when pain was present around the groin/buttock/upper leg. Additionally, knee pain was registered if pain was present around the knee (possible referred pain). The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) total score was used to summarize pain, stiffness, and physical function. Analgesic use and morning stiffness were registered as present or absent. Active internal hip rotation was measured using a goniometer, according to Norkin and White[14]. Pain during internal rotation was registered (yes/no).

### Basic radiographic parameters

Standardized weight-bearing anteroposterior pelvic radiographs, with 15-degree internal hip rotation were made. The presence of joint space narrowing (JSN), osteophytes and thickening of the femoral calcar (buttressing) were scored (yes/no) on baseline radiographs by five trained observers as previously described.[15,16]

### Automatic quantification of subtle shape variations

Statistical shape models (SSMs) provide a global representation of shape rather than reducing shape to a set of geometric measurements, enabling quantification and analysis of complex and subtle shape aspects. Using predefined (anatomical) landmark points, an object, such as the bones of the hip joint, can be outlined and segmented. Based on all landmark points across a set of images, an SSM can be generated by applying principal component analysis to the aligned shapes.[17] The SSM then describes every shape by the combination of a mean shape and a linear combination of a number of shape modes. Each mode describes a distinct shape aspect. The first shape mode explains the highest proportion of variation across the dataset and each additional mode explains a smaller part of the total variation.

We developed a fully automatic segmentation system (FASS) to segment the hip using 75 landmark points **(Supplementary text and supplemenatry figures 1 and 2).**[11,12,18] We used all 1373 baseline pelvic radiographs of sufficient quality with manual segmentations available from previous work, enabling to compare the predictive value of the data produced by the FASS to the data produced by manual segmentations [18]. Below, we give a concise overview of the development of the Shape-Score using FASS/SMM shape modes. More detailed information and a comparison between Shape-Scores from manual versus automatic segmentations is provided in the **Supplementary text**.

We used the first 26 SSM modes, explaining 90% of the overall shape variation across our dataset to develop the Shape-Score. While individual shape modes are independent by the nature of SSMs, the simultaneous effect of two (or more) shape modes on the risk for OA may interact (e.g. be multiplicative instead of additive). For example, a mode describing cam morphology may strengthen the effect of a mode with increased acetabular coverage (i.e. pincer morphology). The theoretical explanation would be that an increased acetabular coverage causes cam morphology to impinge earlier against the acetabulum, which might increase the risk of labral damage and subsequent HOA **(Figure 1)**.[19] Considering interactions between all 26 modes produces 325 combinations (the sum of the arrhythmic row from 1 to 25 with a common difference of 1). Adding the 26 separate modes would produce a total of 351 variables, a number too large for standard regression techniques. Therefore, we used a penalized regression technique (an elastic net) suited for high dimensional data to relate all these variables to the incidence of HOA and produce a single score representing HOA risk based on hip shape **(Supplementary text)**. The resulting Shape-Score ranges between 0 and 1, and contains various clear and subtle aspects of hip geometry. Compared to the low-risk shape, the high-risk shape shows a cam morphology (an aspherical femoral head-neck junction), a narrower superior joint space, decreased acetabular depth in combination with lateralization of the femur. Additionally, the femoral shaft is narrower, while the femoral neck is wider. However, these shape variations do not have to coincide within a patient and a single high risk shape variation may increase the Shape-Score **(Figure 2).**

### Prediction model development and performance testing

To develop the prediction model we only used data of hips with pain around the hip and/or the knee of the respective leg (possible referred pain) as these are the joints where this prediction will most likely be used for in clinical practice. This is in contrast to the development of the shape model itself, in which all available baseline pelvic radiographs with data on hip osteoarthritis on follow-up were used. Furthermore, for a hip to be included in the prediction model a baseline radiograph of sufficient quality had to be available. Depending on these criteria, one or both hips of a participant were included in the analyses. Baseline predictors were used to predict the outcome, rHOA at 8-year follow-up, defined as a Kellgren-Lawrence grade (KL-grade) ≥2, or THR.[16,20] All predictors were measured at baseline and categorized as demographics, clinical examination, basic radiographic parameters, or Shape-Score. To account for missing predictor and outcome data, we imputed 15 datasets using predictive mean matching and logistic regression [21]. We performed a

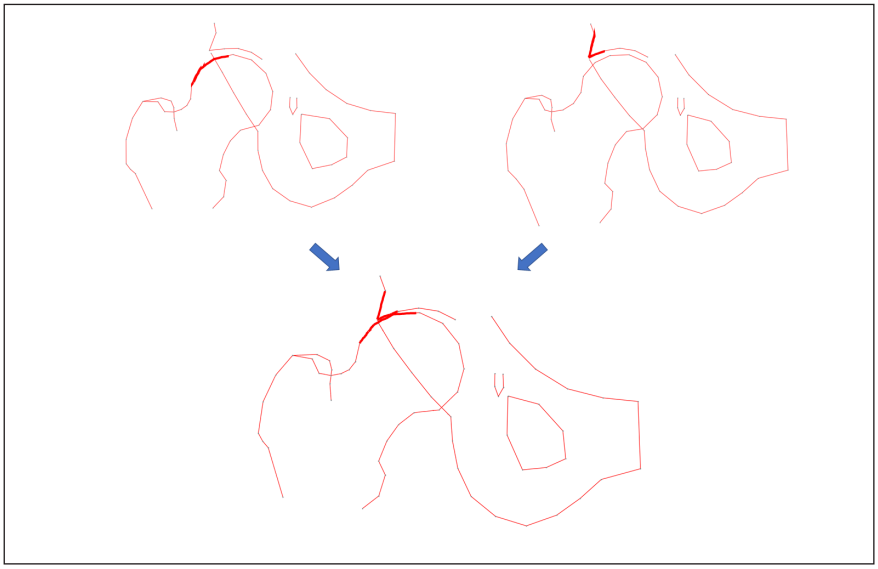


**Figure 1.** Interaction between shape modes. The risk for HOA produced by a shape mode may depend on the presence of other shape features. In this hypothetical example, shape mode A represents cam morphology (Aspherical femoral head-neck junction), the shape mode B represents increased acetabular coverage. Both shape modes have a risk factor for OA. When both features coincide in one hip, the risk for OA may be greater than the sum of two risk factors from shape modes A and B. In C we combine the femur of mode A with the pelvis of mode B and simulate hip motion by applying 15 degrees of abduction to the femur. The risk for femoroacetabular impingement becomes clear and is very plausible. Femoroacetabular impingement may increase the risk of labral damage and subsequent HOA.

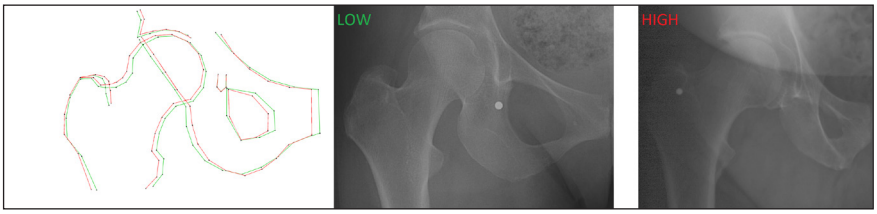



**Figure 2.** High vs low risk Shape-Score. The left shows a schematic representation of the mean shape of the 5% highest Shape-Score (high risk for future HOA, in red) and 5% lowest shape-score (low risk for future HOA, in green). Compared to the low-risk hip, the high-risk hip is characterized by a cam morphology (femoral head-neck asphericity), a narrower superior joint space, and decreased acetabular depth in combination with lateralization of the femur and a higher neck-shaft angle. Additionally, in the high-risk hip the femoral shaft is narrower, while the femoral neck is wider. These shape variations, however, do not have to coincide within a patient and a single high risk shape variation may increase the Shape-Score. In the middle and on the right, a real radiograph of a low risk and high-risk hip are shown.

sensitivity analysis only including hips with complete data using logistic regression and generalized estimating equations (GEE).

To develop the prediction model, logistic regression was used and predictors were added per category. First demographics were added, secondly variables from the clinical examination, thirdly basic radiographic parameters, and finally the Shape-Score. After each step (addition of a category of predictors) we simplified the model by removing redundant predictors from the added category only, using backwards selection with a pooled alpha-level of 0.15.[21,22] To optimize parameter estimates for predictors and avoid overfitting, we used logistic ridge regression on each imputed dataset separately. Optimal penalties were based on corrected Akaike's Information Criteria.[23] When using multiple imputation, the imputed values per imputed dataset may differ, as they are drawn from a distribution. We averaged the intercepts and parameter estimates of the ridge regression models from all imputed datasets, to obtain formulas to calculate individual risks for rHOA or THR in future research or clinical work **(Supplementary text).**[23,24] We calculated predicted risks and stratified all hips into arbitrary risk categories of <20%, 20-50% and >50% risk for OA. We calculated positive and negative predictive value for the low (<20%) and high risk (>50%) categories. For the low-risk category the absence of OA at 8-year follow-up was considered a positive gold standard. For the high-risk category the presence of OA at 8-year follow-up was considered a positive gold standard.

Performance of the model was further assessed in terms of calibration, i.e. the agreement between predictions and observed outcomes, as well as discrimination, i.e. the ability of a model to differentiate hips with developing hip OA from those which will not.[25] To assess calibration of the model, we plotted the percentage of observed OA cases in groups of hips with increasing predicted risk (n= 23 per group to create 50 data points in the plot). Using a lowess smoothing function, we visualized the calibration of the model. We assessed discrimination by calculating Area Under the Curve (AUC) statistics with bootstrapped 95% confidence intervals pooled over the 15 imputed datasets. Any prediction model will perform better in the dataset used to train the model compared to a dataset containing new patients. Therefore, performance measures based on training data will be over-optimistic, also known as overfitted.[26] To internally validate the models and estimate their performance in new patients we used bootstrapping, a resampling method. We drew 1000 bootstrap samples per imputed dataset and pooled the AUC and calibration plots to test for over-optimism of both calibration and discrimination.[26] All data analyses were performed using R v3.3.1. with MICE v2.30, caret v6.0 – 73, rms v5.1-0, and glmnet v2.0-5.

## Results

Baseline characteristics can be found in **Table 1**. Of 1044 hips with data on KL-grade or THR at 8-year follow-up, 143 showed KL-grade 2 or higher and 42 had undergone THR. Among the demographics included as predictor in the initial model, smoking status was non-significant (p-value 0.74) and removed from the model. Among the predictors from clinical examination, morning stiffness (p-value 0.46) and pain on internal rotation (p-value 0.17) were removed. All basic radiographic parameters were significant (at the alpha level of 0.15) and were retained. Predictors that were retained in the models are shown in **Supplementary Table 1** with their respective unpenalized odds ratios. The formulas used to calculate the predicted risks are given in **Supplementary text**.

### Model performance

The discriminative ability of the models improved each time an additional category of predictors was added, meaning that the models' ability to separate cases from non-cases increased **(Figure 3, Table 2)**. A model containing only the Shape-Score discriminated comparable to a model combining demographics, clinical examination and basic radiographic parameters (AUC 0.798 vs 0.795). Adding the Shape-Score to the latter model improved the discriminative ability from an AUC of 0.795 to an AUC of 0.863. Adding the Shape-Score also improved the calibration of the prediction model (**Figure 4**). The calibration curve is very close to the diagonal representing optimal fit, meaning that the predicted risk closely resembles the observed risk for rHOA or THR.

**Table 2.** Discriminative ability

| **Model** | **AUC1 in** development | **AUC1 in** validation | **95%-confidence interval** of AUC in validation |
|---|---|---|---|
| **Demographics** | 0.634 | 0.635 | (0.596 to 0.675) |
| **Demographics and Clinical examination** | 0.710 | 0.710 | (0.668 to 0.751) |
| **Demographics, Clinical examination and Standard radiographic examination** | 0.795 | 0.795 | (0.757 to 0.834) |
| **Demographics, Clinical examination, Standard radiographic examination and Shape-Score** | 0.863 | 0.864 | (0.833 to 0.895) |
| **Shape-Score only** | 0.798 | 0.798 | (0.762 to 0.833) |

[1] Area Under the Curve or C-statistic

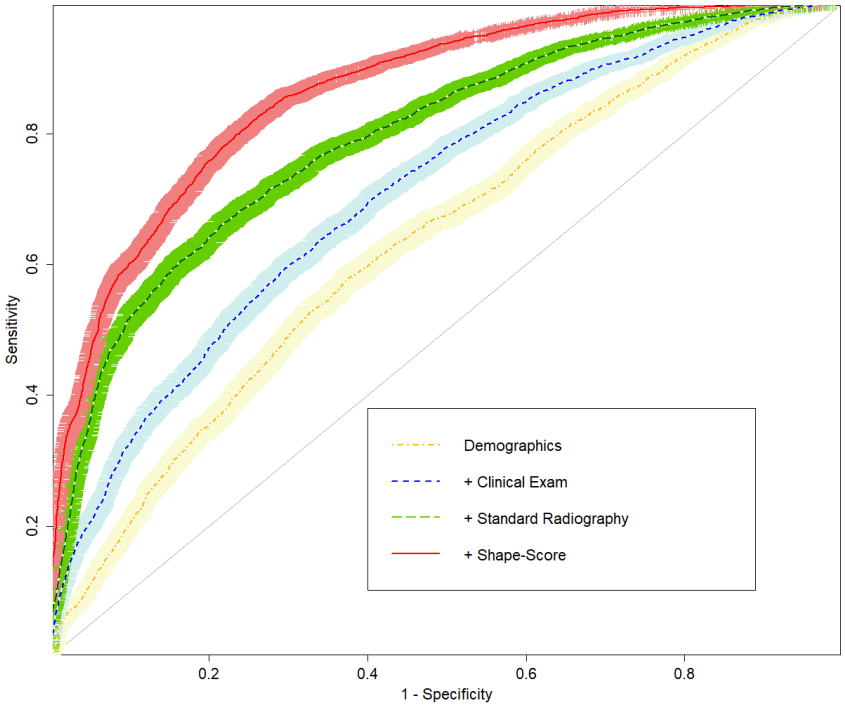



**Figure 3.** Discrimination of the prediction models. The ability to separate cases from non-cases is visualized as area under the curve with the 95%-confidence interval. Sensitivity and specificity of the model improved for all cut-off points after adding the Shape-Score to the model.

Calibration slopes and AUCs in internal-validation based on bootstrapped samples differed minimally from those in development, indicating that the predictive models are not overoptimistic (AUC 0.795 and 0.864 respectively **(Table 2 and Figure 4)**. The distributions of predicted risks resulting from each of the models show that adding the Shape-Score helps to stratify more medium-risk patients into low and high-risk categories **(Table 3)**. In the sensitivity analysis, using both logistic regression and GEE on hips with complete data only, AUC values were within 0.01 of the values found using the imputed datasets, and calibration plots were comparable.

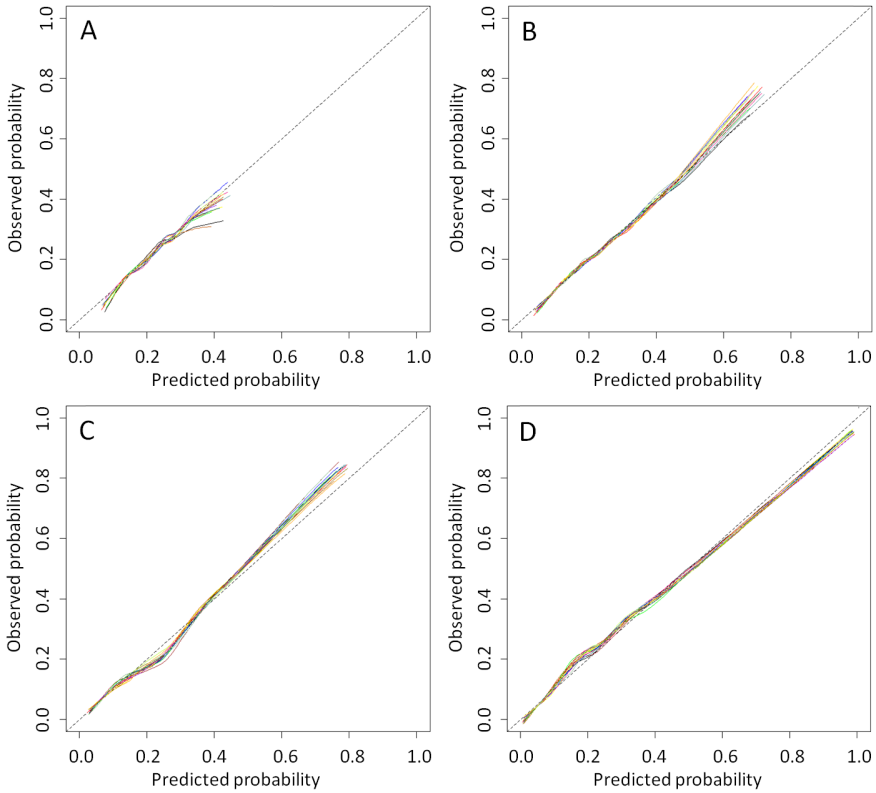



**Figure 4.** Calibration plots in validation. Figure 4 shows the predicted probabilities plotted against the observed outcomes in internal validation. This is used to assess the calibration for five different models. The striped black line represents a perfect match between predicted probabilities and observed outcomes, and thus perfect calibration. The dotted black line represents the calibration in training data. The colored lines each represent a different imputed dataset, and represents the mean calibration in validation, using a 1000 bootstraps. A. Demographics. B. Demographic and clinical examination. C. Demographics, clinical and standard radiographic examination. D. Demographics, clinical, standard radiographic examination and Shape-Score produced using the fully automatic search model.

**Table 3.** Distribution in risk categories with negative (NPV) and positive predictive value (PPV) for low risk (0-20 %) and high risk (50-100%) categories.

| Model | 0 – 20% | Low-Risk | | 20 – 50% | High-Risk | | 50 – 100% |
|---|---|---|---|---|---|---|---|
| | N (%) | NPV | PPV | N (%) | NPV | PPV | N (%) |
| **Demographics** | 682 (60) | 0.27 | 0.86 | 459 (40) | - | - | (0) |
| **Demographics + Clinical examination** | 719 (63) | 0.31 | 0.88 | 395 (34) | 0.82 | 0.66 | 27 (2) |
| **Demographics + Clinical examination + Basic radiographic parameters** | 768 (67) | 0.40 | 0.91 | 288 (25) | 0.85 | 0.67 | 86 (8) |
| **Demographics + Clinical examination + Basic radiographic parameters + Shape-Score** | 800 (70) | 0.48 | 0.93 | 215 (19) | 0.87 | 0.71 | 126 (11) |
| **Shape-Score** | 810 (71) | 0.42 | 0.90 | 243 (21) | 0.85 | 0.68 | 88 (8) |

The numbers given are averages over 15 imputed datasets. For the low-risk category, absence of OA at 8-year follow-up was seen as a positive gold standard. PPV should be interpreted as the probability of not developing OA when being classified as low-risk (< 20 % chance). For the high-risk category, presence of OA at 8-year follow-up was seen as a positive gold standard. PPV should be interpreted as the probability of developing OA when being classified as high-risk (> 50 % chance).

## Discussion

We developed and internally validated models to predict incident rHOA or THR over 8 years in persons with first onset hip pain. Until now, no predictive model for HOA is widely used. We built a prediction model that combines innovative automated analysis of plain radiographs using machine learning, with clinical data that can easily be obtained (**Figure 5**). The discriminative ability of the final model was high (AUC 0.863) given the relatively early stage of possible HOA at baseline and rHOA or THR as outcomes at 8-year follow-up.

In the literature, multiple prediction models are available for HOA. However, most are actually diagnostic models, aiming to diagnose HOA cross-sectionally, situated in a population or end-stage OA cohort. Saberi Hosnijeh et al. recently developed a prediction model for HOA in the Rotterdam cohort.[27] Their model uses demographics, urinary CTX-II levels and radiographic parameters including the Wiberg-angle and alpha-angle (to quantify acetabular coverage and cam morphology, respectively), but no parameters from the physical examination or SSM. Their model showed an AUC of 0.82 in the Rotterdam cohort and 0.71 when validated in CHECK. Furthermore, calibration in CHECK was far off the perfect slope, with observed risks being 2.5 times higher than predicted risks.

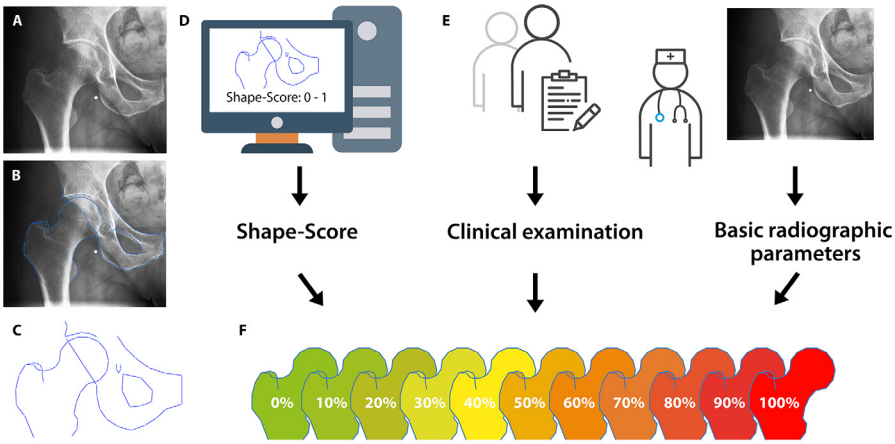



**Figure 5.** Workflow to calculate personalized risk for future hip osteoarthritis. A standard weight-bearing pelvic radiograph is made in the anteroposterior direction with 15-degrees internal rotation. B. The fully automatic segmentation system (FASS) annotates the anatomical landmarks on the radiograph. C. Statistical shape modelling (SSM) quantifies hip shape. D. The machine learning algorithm produces the Shape-Score, a single value representing the risk for incident HOA based on hip shape. E. Demographics, questionnaires, clinical examination, and basic radiographic parameters are assessed by trained personnel. F. Data from demographics, questionnaires, and physical examination and basic radiographic parameters are combined with the Shape-Score in easy to calculate formulas to produce accurate personalized risk for future hip osteoarthritis.

Developing the model in a general population cohort and testing it in a target population with hip pain likely caused this. Our prediction model was developed in CHECK, which represents our target population, and includes parameters from physical examination and the Shape-Score. External validation was not performed as most cohorts focus on OA in later stages and/or do not have pelvic radiographs of sufficient quality available. Nevertheless, internal bootstrap validation suggested that our model is not overoptimistic.[26] In the future, external validation should be performed, preferably in a cohort with symptomatic patients prone to HOA.

Strengths of this study include the use of a large prospective cohort with clinical complaints and inclusion criteria that allude early-stage knee and/or hip OA, with an adequate follow-up time and a sufficient number of incident rHOA or THR after 8 year (185) to test the 16 predictor candidates for the prediction models.[28] We used backward selection on clusters of predictors to mimic the flow of information in clinical care. While this may produce a slight reduction in absolute performance of the models compared to a fully data driven method, it improves the applicability of the models in clinical care and reduces the chance of overfitting. We tested the

association between the Shape-Score and baseline clinical OA characteristics. The Shape-Score was related to hip OA characteristics but not knee OA characteristics (**Supplementary text**). Furthermore, follow-up data were rather complete and we used multiple imputation to reduce bias and increase precision of our analyses. Finally, we used optimism-adjustment methods throughout the development and validation of the models to reduce overfitting and overoptimistic results. By using bootstrap validation instead of multi-fold cross-validation we used the data available more efficiently.[29]

Combining rHOA (KL-grade ≥2) and THR as disease outcome may be debatable. The severity of clinical and radiographic symptoms correspond poorly in HOA, so that rHOA and THR may not always represent similar processes.[30] However, THR most often results from both clinical symptoms and radiographic signs. For a number of participants we included both hips. We used logistic regression analysis, which does not incorporate intra-participant correlation. However, GEE was not applicable in combination with the statistical packages used in the analysis, and mixed models regression had problems to converge when used on the available data. In the sensitivity analysis on the hips with complete data, the results between logistic regression analysis and GEE were very comparable.

The relationship between sex and HOA is less clear. In our models, female sex was initially associated with an increased risk of rHOA or THR, but with a decreased risk after adding the Shape-Score. This suggests that gender differences in OA risk may be related to hip shape. Higher BMI is a well-known risk factor for knee and hand OA, but its relationship with HOA is less clear.[31–33] In our models, a higher BMI had even a mild preventive effect for HOA. Education served as a proxy for socio-economic status in our study. Higher educated persons had a lower risk for HOA or THR, in line with literature.[34] Smoking did not predict rHOA or THR. Although some studies show a protective effect of smoking, this effect may be caused by selection bias.[35]

Pain in the hip area (groin/buttock/upper thigh) increased the risk for HOA, which is in correspondence with literature.[36] Pain around the knee sometimes directs a physician to search for a diagnosis in the knee only. However, hip OA should always be considered as a source of the pain.[37] Limited or painful internal rotation are clinical signs that suggest HOA and may predict THR.[36,38–40] In the present study, pain with/during internal rotation had a significant univariate relation with OA on follow-up (OR 1.7, 95% CI(1.3 – 2.3), but was eliminated from the prediction models as the p-value was 0.17. However, the range of internal rotation

was included in the prediction models, perhaps overrunning the weaker predictive effect of pain on internal rotation. The WOMAC is a tool to measure pain, stiffness and physical functioning in patients with knee and/or hip OA. While it is widely used, its predictive value for incident rHOA is unknown.[41] In this study, baseline WOMAC score was associated with future rHOA or THR. Morning stiffness is included as a diagnostic criterion for HOA in the widely used Altman criteria for HOA, and showed a high sensitivity, but low specificity.[38] Its predictive value as a risk factor for HOA is doubtful and in the present study it did not add predictive value to the models, perhaps because morning stiffness is a non-specific symptom.[27] Surprisingly, the use of analgesics was negatively associated with the future risk for rHOA or THR. Possibly, this is the case because analgesics are used more often in acute pain caused by transient disorders compared to the more elongated pain trajectory in OA.

Joint space narrowing and osteophyte formation may be present before definite rHOA (defined as KL-grade ≥2) can be confirmed. These radiographic signs are known risk factors for progression to definite rHOA and indeed were strong risk factors in the present study.[27]. Buttressing, thickening of the medial femoral neck, is a radiographic sign associated with rHOA.[42] The predictive value of buttressing has not been described before but it showed to be a significant risk factor for future rHOA or THR in this study. The Shape-Score may also include joint space narrowing and osteophyte formation, characteristics used to define KL-grades. However, the current study shows that the Shape-Score has added predictive values on top of traditional radiographic characters alone.

Our Shape-Score incorporates (i) cam morphology (Aspherical femoral head-neck junction) (ii) decreased acetabular depth, and (iii) a higher neck-shaft angle as risk factors for OA. Cam morphology, decreased acetabular depth and a higher neck-shaft angle have been shown to increase risk for OA in large cohorts before.[7,43–48]

Currently, it is challenging for clinicians to predict future hip OA in patients with early-stage joint pain that cannot be explained otherwise, especially for patients who don't have definite signs of OA on the radiograph. Some of these patients will develop OA, while other might have (had) hip pain for other reasons and will not develop OA. The proposed prediction model could help clinicians to optimally inform patients about their personalized future risk for disease and to choose appropriate treatment (intensity) and may boost treatment adherence.

In the future, the proposed FASS and prediction model could be integrated into a software package that can be linked to the electronic patient record (including

PACS for radiographic images). This way, the Shape-Score could be derived fully automatically to assist clinicians in estimating the risk for future hip osteoarthritis. Of note is that the implementation of the proposed prediction model is not dependent on time-consuming visual methods that may be subject to inter-/intra-observer variations. However, variation in positioning during image acquisition may cause differences in Shape-Score values and a standardized acquisition protocol, as used in the current cohort, is necessary. Unfortunately, no data was available to assess the effect of variation in positioning.

Clinical trial inefficiency plays a major role in the current absence of disease-modifying OA drugs (DMOADs). By specifically selecting participants at an increased risk of incident OA, potential DMOADs have more potential to demonstrate detectable effects in a clinical trial. By adding the Shape-score or by using the Shape-Score only, we were able to stratify people for the risk of future rHOA or THR. When the Shape-Score is added to a screening with demographics, physical examination and basic radiographic parameters, 47% more patients could be stratified into the high-risk category (>50% risk), potentially improving screening efficiency.

We have developed an automatic Shape-Score tool, using machine learning algorithms, to optimally predict the risk for incident rHOA or THR based on hip shape as given by a pelvic radiograph. We demonstrated the added value of our Shape-Score in prediction models using easily obtainable parameters in patients with hip pain. Models including the Shape-Score had superior discriminative ability over models without and showed very good calibration. The Shape-Score may therefore prove to be a valuable tool for both patient care and research.

## Acknowledgements

The authors like to thank Chris van Kesteren for his work on the graphic design of **Figure 5**.

## Funding

The CHECK-cohort study is funded by Reuma Nederland. Involved are: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen/Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede /Ziekenhuisgroep Twente Almelo; Reade Center for Rehabilitation and Rheumatology; St.Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

C. Lindner was funded by the Engineering and Physical Sciences Research Council, UK (EP/M012611/1) and by the Medical Research Council, UK (MR/S00405X/1). The current analysis was funded by Reuma Nederland (LLP-22) and the APPROACH project. APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement n°115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme(FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu.

## Data availability

Data from CHECK is available from https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:63523

The FASS and SSM are available via http://bone-finder.com/

FASS and SSM are available via http://bone-finder.com/

## Supplementary data

### Automatic search algorithm

We have followed the approach in [1], [2] to develop a fully automatic segmentation system (FASS) for segmenting the shape of the hip joint. Our FASS identifies 75 landmark points outlining the structures of the hip joint in AP pelvic radiographs (**Supplementary Figure 1**). The FASS was trained using manual ground truth segmentations which were available from previous work for 1373 hips of the participants in this study. [3]

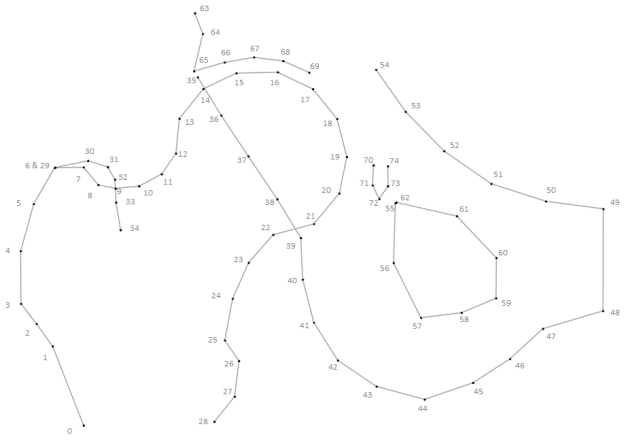
We evaluated the performance of the FASS in two ways: (i) by comparing the automatically obtained segmentations to the manual ground truth segmentations; and (ii) by comparing the performance of the prediction model using an SSM generated from the fully automatic segmentations to using an SSM generated from the manual ground truth segmentations.

Manual vs automatic segmentations

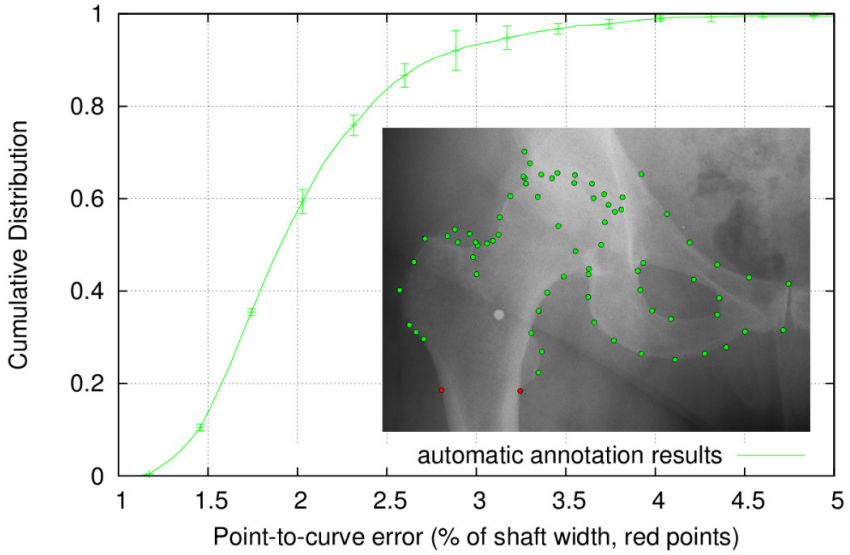
We performed two-fold cross-validation experiments to analyze the difference between manual and fully automatically obtained segmentations. We randomly split the radiographs in two groups and trained a separate FASS in each group. The FASS trained on one group was used to segment radiographs in the other group, so that radiographs were seen for the first time by the FASS. All results were averaged over the performance in both cross-validation groups (**Supplementary Figure 2**). The results are reported as the point-to-curve error which describes the Euclidean distance between an automatically identified landmark point and a curve fitted to the manual ground-truth segmentations. For every image, the point-to-curve error was calculated for each landmark point and then averaged over all points per image. Due to the range of image resolutions and the lack of pixel size information for the images used in this study, we reported the averaged point-to-curve error as a percentage of the shaft width (defined by the two femoral shaft landmark points highlighted in red in **Supplementary Figure 2**), allowing comparability across images.

### Data reduction

Including all 26 or 24 SSM modes, for SSM modes produced by FASS and manual ground truth respectively, and interaction terms between each set of two SSM modes in the prediction model, would have meant to include 351 or 300 predictors. Such model would be prone to overfitting with our data. Therefore, we used a



**Supplementary Figure 1**. 75-point segmentation model. The 75-point model based on anatomical landmarks of the hip joint.




**Supplementary Figure 2.** Point-to-curve error. Performance of the fully automatic segmentation system. The results show the point-to-curve error between the automatically obtained segmentations and the manual ground truth in two-fold cross-validation experiments. All error values were averaged over all 75 points per image and are given as the cumulative distribution over all images. All results were averaged over both cross-validation FASSs, and error bars show the 95% confidence interval.

penalized regression technique that simultaneously fitted a model and reduced the number of predictors, a so called elastic net (penalized regression), to obtain the Shape-Score for each hip. [4] The result can be considered a prediction model for hip OA based only on the optimal combination of shape modes. The data of all participants with both a baseline and eight-year follow-up radiograph of sufficient quality (n= 1262) was used to develop the Shape-Score.

We performed an iterative grid search in which 10000 combinations of alpha and lambda parameters were tested in ten-fold cross-validation to obtain the combination with the highest accuracy for distinguishing cases from non-cases. Alpha resembles the elastic net mixing parameter, where 0 resembles ridge regression and 1 lasso regression. Any value between 0 and 1 represents an elastic net, mixing properties of ridge and lasso regression. Lambda resembles the shrinkage coefficient which refers to the penalty used to prevent overfitting. In our experiments, the optimal alpha was 0.1 and the optimal lambda was 0.0658899. The Shape-Score is calculated using a formula, which can be applied to new patients. The value of the Shape-Score ranges between 0 (lowest risk for OA) and 1 (highest risk for OA).

Formula for calculating the Shape-Score based on the shape mode values produced by the FASS. M_# is the mode value as produced by the SSM.

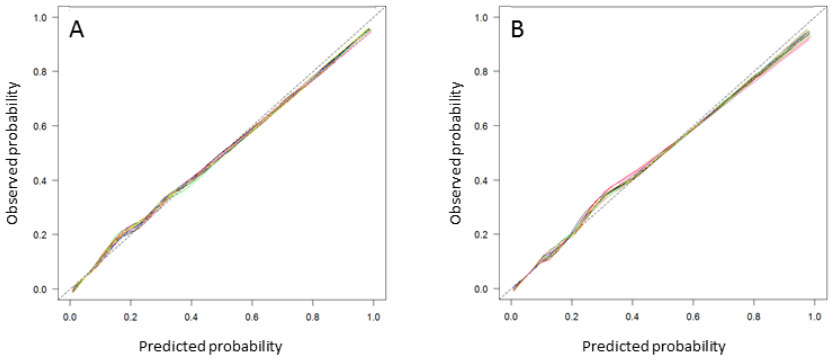
-1.5568902 + M_20*M_25 * -287.959541 + M_13*M_15 * 262.1913582 + M_5*M_21 * 210.1067346 + M_1*M_2 * -21.4053887 + M_21*M_24 * 1078.570449 + M_13*M_16 * 635.6871083 + M_5*M_22 * -33.987232 + M_1*M_23 * -32.1682874 + M_22*M_25 * -86.2056066 + M_13*M_19 * 615.7970254 + M_5*M_24 * -4.4039672 + M_1*M_26 * -217.3489381 + M_23*M_24 * -199.576888 + M_13*M_22 * -536.674347 + M_5*M_25 * -98.1444219 + M_1*M_3 * -12.79741 + M_24 * -1.5482276 + M_14*M_15 * 107.2634382 + M_5*M_8 * -57.970354 + M_1*M_4 * 26.329696 + M_25 * -6.1161671 + M_14*M_17 * 0.1416411 + M_6 * -3.703353 + M_1*M_6 * -49.9387397 + M_3 * -2.7287343 + M_14*M_18 * 1120.833495 + M_6*M_18 * 549.9824169 + M_1*M_7 * -40.6936973 + M_3*M_11 * 203.1342221 + M_14*M_19 * 326.3590625 + M_6*M_19 * -420.5872139 + M_1*M_9 * -6.9882447 + M_3*M_14 * 249.9283829 + M_14*M_20 * 737.4339839 + M_6*M_22 * 176.2472787 + M_10 * -3.6186784 + M_3*M_16 * -61.0959496 + M_15 * 3.6787325 + M_6*M_7 * -87.9910941 + M_10*M_11 * -61.6171982 + M_3*M_18 * 150.1872388 + M_16 * 17.3717681 + M_6*M_8 * 285.226913 + M_10*M_12 * 193.2207746 + M_3*M_22 * -169.8342171 + M_16*M_20 * 743.3275762 + M_6*M_9 * 10.7397399 + M_10*M_17 * 44.4460214 + M_3*M_25 * 181.9545524 + M_16*M_24 * -352.0308105 + M_7*M_11 * -13.3992776 +

M_10*M_18 * 1302.190357 + M_3*M_4 * -1.0007214 + M_16*M_25 * -949.4162743 + M_7*M_21 * 49.987956 + M_10*M_21 * 985.7811976 + M_3*M_6 * 114.9578101 + M_17*M_18 * 509.3953662 + M_7*M_9 * -133.7137365 + M_11*M_13 * 32.8321238 + M_3*M_8 * 44.2613325 + M_17*M_19 * -368.8290594 + M_8 * -4.4170317 + M_11*M_15 * 18.7267828 + M_3*M_9 * 94.2119748 + M_17*M_23 * 37.7878678 + M_8*M_10 * -178.7084949 + M_11*M_18 * -363.9381869 + M_4 * 0.2812895 + M_17*M_24 * 2.7494028 + M_8*M_11 * -44.5584662 + M_11*M_20 * 62.7098343 + M_4*M_10 * 174.195207 + M_18*M_19 * -89.2712441 + M_8*M_13 * -132.8111248 + M_11*M_22 * -63.7140257 + M_4*M_11 * -81.7705555 + M_19 * 8.0303025 + M_8*M_16 * -229.041809 + M_11*M_24 * -408.8633874 + M_4*M_20 * 218.5554976 + M_19*M_23 * 68.4249181 + M_8*M_18 * 166.2689009 + M_11*M_26 * 292.6358193 + M_4*M_26 * 40.673264 + M_19*M_25 * -725.9103928 + M_8*M_21 * 35.4831362 + M_12*M_16 * 190.1231435 + M_4*M_6 * 93.3678123 + M_19*M_26 * -31.2212137 + M_8*M_9 * -315.0005546 + M_12*M_17 * 206.9152064 + M_4*M_7 * -18.6135705 + M_2*M_17 * 9.8960661 + M_9*M_22 * 501.5647865 + M_12*M_18 * 127.9878676 + M_4*M_9 * 63.6571854 + M_2*M_18 * -38.2547196 + M_9*M_24 * -479.4677671 + M_12*M_22 * 253.3392333 + M_5 * -0.7726085 + M_2*M_20 * 166.8343537 + M_9*M_25 * 837.2001561 + M_12*M_24 * 834.4582995 + M_5*M_15 * -236.409445 + M_2*M_23 * 77.5097166 + M_9*M_26 * 448.4389367 + M_13 * 0.4418048 + M_5*M_20 * 296.1666966 + M_2*M_24 * 234.9528434

### Prediction models with manual vs automatic SSM

We built two SSMs, one using the manual ground truth segmentations and another one that used the automatically generated segmentations. The automatically obtained segmentations of both cross-validation groups were combined for generating the automatic SSM. The SSM produced with the automatic segmentations needed 26 SSM modes to explain 90% of the shape variation across our dataset. The SSM produced with the manual ground truth segmentations needed 24 modes to describe 95% of the shape variation. All prediction analyses as described in the paper were mimicked with Shape-Scores derived from the manual ground truth.

The performance of the prediction model using the automatically generated SSM was comparable to using the SSM based on the manual ground truth data. The AUC (95% confidence interval) in our internal validation were 0.86 (95%-CI 0.83 to 0.89) and 0.86 (95%-CI 0.83 to 0.89) for the final prediction models including automatic shape analysis and manual shape analysis, respectively. The calibration of the models including the automatic shape analysis and the manual shape analysis was comparable (**Supplementary Figure 3**).




**Supplementary Figure 3.** Calibration plots in validation. S3 figure shows the predicted probabilities plotted against the observed outcomes in internal validation. The striped black line represents a perfect match between predicted probabilities and observed outcomes, and thus perfect calibration. The dotted black line represents the calibration in training data. The colored lines each represent a different imputed dataset, and represents the mean calibration in validation, using a 1000 bootstraps. A. Demographics, clinical, standard radiographic examination and Shape-Score produced using the fully automatic search model. B. Demographics, clinical, standard radiographic examination and Shape-Score produced using the manual ground-truth segmentations.

### Formulas to calculate personalized risk score for hip osteoarthritis

All formulas calculate the log odds for OA. To obtain the predicted probability, first the log odds should be exponentiated to obtain the odds. Than the probability can be calculated by the formula: probability = odds / (1+ odds)

### Demographics only:

Log odds = -0,730517982 + (Age in years– 45) * 0,052634143 + male sex * 0,716753641 + education level * -0,137110459 + BMI * -0,029671982

### Demographics and Clinical examination:

Log odds = 1,026356634 + (Age in years – 45) * 0,05260851 + male sex * 0,646366906 + education level * -0,154305156 + BMI * -0,037547738 + Painkiller usage * -0,554981525 + Pain located around the hip * 0,403591597 + Pain located around the knee * -0,879734629 + Range of internal rotation in degrees * -0,038447494 + WOMAC total score * 0,012128268

### Demographics, Clinical and standard radiographic examination:

Log odds = -0,488998048 + (Age in years – 45) * 0,030190881 + male sex * 0,244676271 + education level * -0,13939754+ BMI * -0,030841726 + Painkiller

usage * -0,350981376 + Pain located around the hip * 0,325403163 + Pain located around the knee * -0,713907091 + Range of internal rotation in degrees * -0,017635736 + WOMAC total score * 0,0103843 + Osteophytes * 1,266730608 + Joint Space Narrowing * 0,627093864 + Buttressing * 0,626788051

**Demographics, Clinical and standard radiographic examination and Shape-Score:**

Log odds = -2,638948514 + (Age in years – 45) * 0,022390868 + male sex * -0,185755848 + education level * -0,19549462 + BMI * -0,059268856 + Painkiller usage * -0,301113741 + Pain located around the hip * 0,242689642 + Pain located around the knee * -0,668297092 + Range of internal rotation in degrees * -0,00966605 + WOMAC total score * 0,005827585 + Osteophytes * 1,279103437 + Joint Space Narrowing * 0,636889009 + Buttressing * 0,339347062 + Shape-score * 16,36836904

**Associations between the Shape Score and baseline OA characteristics**

The Shape Score, defined on baseline radiographs is used in the current study to predict future rHOA. Additionally, we explored the association between the baseline Shape Score and clinical OA parameters, using a linear regression model with the Shape Score as dependent variable, and clinical parameters as independent variable. We saw a relationship between the shape score and baseline pain, stiffness and limitations (WOMAC), the range of internal hip rotation and the baseline KL grade. However, no relation between the shape score and pain on internal rotation and baseline KL grade for knee OA was present.

**Supplementary Table1.** Association between Shape Score and baseline OA characteristics

| Parameters | Beta | 95%-Confidence Interval | P-value |
|---|---|---|---|
| Female Sex | 0.02797 | (0.01760 to 0.03835) | >0.001 |
| WOMAC score | 0,00027 | (0,00003 to 0.00051) | 0.026 |
| Range of internal rotation hip | -0.00079 | (-0.00124 to -0.00034) | >0.001 |
| Pain on internal rotation hip | -0.00729 | (-0.00114 to 0.01572) | 0.090 |
| Baseline Kellgren Lawrence grade Hip (1 vs 0) | 0.03456 | (0.00463 to 0.02549) | >0.001 |
| Baseline Kellgren Lawrence grade Knee (1 vs 0) | 0.00060 | (0.00396 to -0.00715) | 0.879 |

## References for supplementary data

1. C. Lindner, S. Thiagarajah, J. M. Wilkinson, T. Consortium, G. A. Wallis, and T. F. Cootes, "Fully Automatic Segmentation of the Proximal Femur Using Random Forest Regression Voting," IEEE Transactions on Medical Imaging, vol. 32, no. 8, pp. 1462 - 1472, 2013.

2. C. Lindner, S. Thiagarajah, J. M. Wilkinson, G. a. Wallis, and T. F. Cootes, "Development of a fully automatic shape model matching (FASMM) system to derive statistical shape models from radiographs: Application to the accurate capture and global representation of proximal femur shape," Osteoarthr. Cartil., vol. 21, no. 10, pp. 1537–1544, 2013.

3. R. Agricola, M. Reijman, S. M. a Bierma-Zeinstra, J. a N. Verhaar, H. Weinans, and J. H. Waarsing, "Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: a prospective cohort study (CHECK)," Osteoarthr. Cartil., vol. 21, no. 4, pp. 1–6, 2013.

4. J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," J. Stat. Softw., vol. 33, no. 1, 2010.

## References


1. Gay C, Chabaud A, Guilley E, Coudeyre E. Educating patients about the benefits of physical activity and exercise for their hip and knee osteoarthritis. Systematic literature review. *Ann Phys Rehabil Med*. 2016;59(3):174-183. doi:10.1016/j.rehab.2016.02.005

2. Lessmann N, van Ginneken B, Zreik M, et al. Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions. *IEEE Trans Med Imaging*. 2018;37(2):615-625. doi:10.1109/TMI.2017.2769839

3. Ciompi F, de Hoop B, van Riel SJ, et al. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. *Med Image Anal*. 2015;26(1):195-202. doi:10.1016/j.media.2015.08.001

4. Ganz R, Leunig M, Leunig-Ganz K, Harris WH. The etiology of osteoarthritis of the hip: an integrated mechanical concept. *Clin Orthop Relat Res*. 2008;466(2):264-272. doi:10.1007/s11999-007-0060-z

5. Agricola R, Heijboer MP, Bierma-Zeinstra SM a., Verhaar J a. N, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). *Ann Rheum Dis*. 2013;72(6):918-923. doi:10.1136/annrheumdis-2012-201643

6. Agricola R, Waarsing JH, Arden NK, et al. Cam impingement of the hip-a risk factor for hip osteoarthritis. *Nat Rev Rheumatol*. 2013;9(10):630-634. doi:10.1038/nrrheum.2013.114

7. Agricola R, Leyland KM, Bierma-Zeinstra SMA, et al. Validation of statistical shape modelling to predict hip osteoarthritis in females: data from two prospective cohort studies (Cohort Hip and Cohort Knee and Chingford). *Rheumatology*. 2015;54(11):2033-2041. doi:10.1093/rheumatology/kev232

8. Ahedi HG, Aspden RM, Blizzard LC, et al. Hip shape as a predictor of osteoarthritis progression in a prospective population cohort. *Arthritis Care Res (Hoboken)*. 2016;69(10):1566-1573. doi:10.1002/acr.23166

9. Barr RJ, Gregory JS, Reid DM, et al. Predicting OA progression to total hip replacement: can we do better than risk factors alone using active shape modelling as an imaging biomarker? *Rheumatology (Oxford)*. 2012;51(3):562-570. doi:10.1093/rheumatology/ker382

10. Castaño-Betancourt MC, Van Meurs JBJ, Bierma-Zeinstra S, et al. The contribution of hip geometry to the prediction of hip osteoarthritis. *Osteoarthr Cartil*. 2013;21(10):1530-1536. doi:10.1016/j.joca.2013.06.012

11. Lindner C, Thiagarajah S, Wilkinson JM, Wallis G a., Cootes TF. Accurate bone segmentation in 2D radiographs using fully automatic shape model matching based on regression-voting. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2013;8150 LNCS(PART 2):181-189. doi:10.1007/978-3-642-40763-5_23

12. Lindner C, Thiagarajah S, Wilkinson JM, Wallis G a., Cootes TF. Development of a fully automatic shape model matching (FASMM) system to derive statistical shape models from radiographs: Application to the accurate capture and global representation of proximal femur shape. *Osteoarthr Cartil*. 2013;21(10):1537-1544. doi:10.1016/j.joca.2013.08.008

13. Wesseling J, Boers M, Viergever MA, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. *Int J Epidemiol*. 2016;45(1):36-44. doi:10.1093/ije/dyu177

14. Cynthia C. Norkin PT E, D. Joyce White PT DS. *Measurement of Joint Motion A Guide to Goniometry, 4th Ed.*; 2009. doi:10.2310/6640.2004.00031

15. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage*. 2007;15 Suppl A:A1-56. doi:10.1016/j.joca.2006.11.009

16. Damen J, Schiphof D, Wolde S Ten, Cats HA, Bierma-Zeinstra SMA, Oei EHG. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. *Osteoarthritis Cartilage*. 2014;22(7):969-974. doi:10.1016/j.joca.2014.05.007

17. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models-Their Training and Application. *Comput Vis Image Underst*. 1995;61(1):38-59. doi:10.1006/cviu.1995.1004

18. Agricola R, Reijman M, Bierma-Zeinstra SM a, Verhaar J a N, Weinans H, Waarsing JH. Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: a prospective cohort study (CHECK). *Osteoarthr Cartil*. 2013;21(4):1-6. doi:10.1016/j.joca.2013.01.005

19. Ganz R, Parvizi J, Beck M, Leunig M, Nötzli H, Siebenrock K a. Femoroacetabular impingement: a cause for osteoarthritis of the hip. *Clin Orthop Relat Res*. 2003;(417):112-120. doi:10.1097/01.blo.0000096804.78689.c2

20. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. http://www.scopus.com/inward/record.url?eid=2-s2.0-62949246527&partnerID=tZOtx3y1.

21. Van Buuren S, Groothuis-Oudshoorn K. Multivariate Imputation by Chained Equations. *J Stat Softw*. 2011;45(3):1-67. doi:10.1177/0962280206074463

22. Heymans MW, van Buuren S, Knol DL, van Mechelen W, de Vet HCW. Variable selection under multiple imputation using the bootstrap in a prognostic study. *BMC Med Res Methodol*. 2007;7:33. doi:10.1186/1471-2288-7-33

23. Harrell FE. *Regression Modeling Strategies*. Vol 64.; 2015. doi:10.1007/978-1-4757-3462-1

24. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581-592. doi:10.1093/biomet/63.3.581

25. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138. doi:10.1097/EDE.0b013e3181c30fb2

26. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KGM. Internal and external validation of predictive models: A simulation study of bias and precision in small samples. *J Clin Epidemiol*. 2003;56(5):441-447. doi:10.1016/S0895-4356(03)00047-7

27. Saberi Hosnijeh F, Kavousi M, Boer CG, et al. Development of a prediction model for future risk of radiographic hip osteoarthritis. *Osteoarthr Cartil*. 2018;26(4):540-546. doi:10.1016/j.joca.2018.01.015

28. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstem AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373-1379. doi:10.1016/S0895-4356(96)00236-3

29. Steyerberg EW, Harrell FE, Borsboom GJJ., Eijkemans MJ., Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-781. doi:10.1016/S0895-4356(01)00341-9

30. Kinds MB, Welsing PMJ, Vignon EP, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. *Osteoarthritis Cartilage*. 2011;19(7):768-778. doi:10.1016/j.joca.2011.01.015

31. Reyes C, Leyland KM, Peat G, Cooper C, Arden NK, Prieto-Alhambra D. Association Between Overweight and Obesity and Risk of Clinically Diagnosed Knee, Hip, and Hand Osteoarthritis: A Population-Based Cohort Study. *Arthritis Rheumatol*. 2016;68(8):1869-1875. doi:10.1002/art.39707

32. Jiang L, Rong J, Wang Y, et al. The relationship between body mass index and hip osteoarthritis: A systematic review and meta-analysis. *Jt Bone Spine*. 2011;78(2):150-155. doi:10.1016/J.JBSPIN.2010.04.011

33. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. *BMC Musculoskelet Disord*. 2008;9:132. doi:10.1186/1471-2474-9-132

34. Reyes C, Garcia-Gil M, Elorza JM, et al. Socio-economic status and the risk of developing hand, hip or knee osteoarthritis: a region-wide ecological study. *Osteoarthr Cartil*. 2015;23(8):1323-1329. doi:10.1016/J.JOCA.2015.03.020

35. Hui M, Doherty M, Zhang W. Does smoking protect against osteoarthritis? Meta-analysis of observational studies. *Ann Rheum Dis*. 2011;70(7):1231-1237. doi:10.1136/ard.2010.142323

36. Lievense AM, Koes BW, Verhaar J a N, Bohnen AM, Bierma-Zeinstra SM a. Prognosis of hip pain in general practice: a prospective followup study. *Arthritis Rheum*. 2007;57(8):1368-1374. doi:10.1002/art.23094

37. Dibra FF, Prieto HA, Gray CF, Parvataneni HK. Don't forget the hip! Hip arthritis masquerading as knee pain. *Arthroplast Today*. 2018;4(1):118-124. doi:10.1016/J.ARTD.2017.06.008

38. Sutlive TG, Lopez HP, Schnitker DE, et al. Development of a clinical prediction rule for diagnosing hip osteoarthritis in individuals with unilateral hip pain. *J Orthop Sports Phys Ther*. 2008;38(9):542-550. doi:10.2519/jospt.2008.2753

39. Birrell F, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Predicting radiographic hip osteoarthritis from range of movement. *Rheumatology*. 2001;40(5):506-512. doi:10.1093/rheumatology/40.5.506

40. Holla JFM, Van Der Leeden M, Roorda LD, et al. Diagnostic accuracy of range of motion measurements in early symptomatic hip and/or knee osteoarthritis. *Arthritis Care Res*. 2012;64(1):59-65. doi:10.1002/acr.20645

41. Bellamy N. WOMAC osteoarthritis index user guide. Version V. *Brisbane, Aust*. 2002.

42. Dixon T, Benjamin J, Lund P, Graham A, Krupinski E. Femoral neck buttressing: A radiographic and histologic analysis. *Skeletal Radiol*. 2000;29(10):587-592. doi:10.1007/s002560000260

43. Thomas GER, Palmer AJR, Batra RN, et al. Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women. A 20 year longitudinal cohort study. *Osteoarthr Cartil*. 2014;22(10):1504-1510. doi:10.1016/j.joca.2014.06.038

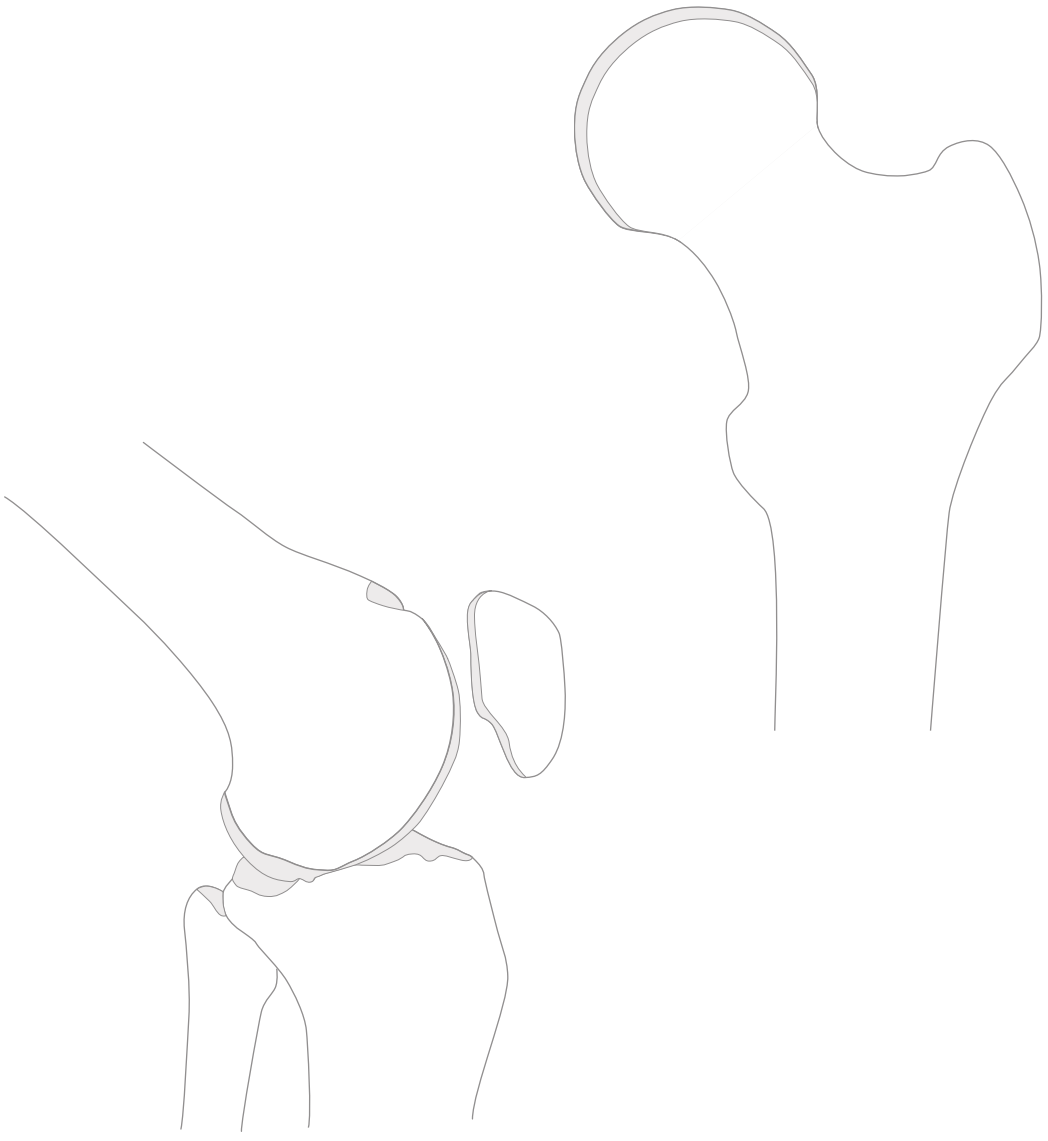
44. Lane NE, Lin P, Christiansen L, et al. Association of mild acetabular dysplasia with an increased risk of incident hip osteoarthritis in elderly white women: The study of osteoporotic fractures. *Arthritis Rheum*. 2000;43(2):400. doi:10.1002/1529-0131(200002)43:2<400::AID-ANR21>3.0.CO;2-D

45. Murphy NJ, Eyles JP, Hunter DJ. Hip Osteoarthritis: Etiopathogenesis and Implications for Management. *Adv Ther*. 2016;33(11):1921-1946. doi:10.1007/s12325-016-0409-3

46. Doherty M, Courtney P, Doherty S, et al. Nonspherical femoral head shape (pistol grip deformity), neck shaft angle, and risk of hip osteoarthritis: A case-control study. *Arthritis Rheum*. 2008;58(10):3172-3182. doi:10.1002/art.23939

47. Ranawat AS, Schulz B, Baumbach SF, Meftah M, Ganz R, Leunig M. Radiographic Predictors of Hip Pain in Femoroacetabular Impingement. *HSS J*. 2011;7(2):115-119. doi:10.1007/s11420-010-9192-x

48. Geoffry Ng KC, Lamontagne M, Beaulé PE. Differences in anatomical parameters between the affected and unaffected hip in patients with bilateral cam-type deformities. *Clin Biomech*. 2016;33:13-19. doi:10.1016/j.clinbiomech.2016.01.007



# CHAPTER 11

# A sex-specific association between incident radiographic osteoarthritis of hip or knee and incident peripheral arterial calcifications: eight-year prospective data from CHECK

*Osteoarthritis and Cartilage*
*July 2017, doi.org/10.1016/j.joca.2017.07.016*

Willem Paul Gielis, Paco M.J. Welsing, Willem E. van Spil,
Jos Runhaar, Harrie Weinans, Pim A. de Jong

## Abstract

### Background and purpose

There is sparse evidence for a relationship between cardiovascular disease and osteoarthritis (OA). We investigated the association between incidence of arterial calcifications and incidence of radiographic knee and/or hip OA.

### Patients and methods

We used baseline and eight-year follow-up data of Cohort Hip and Cohort Knee (CHECK). Knees and hips were either Kellgren-Lawrence grade 0 or 1 at baseline. Arterial calcifications were scored on hip and knee radiographs using a four-grade scale. Scores were summed for patient-level analyses. To investigate incidence, participants with arterial calcifications at baseline or missing follow-up were excluded. Incident OA was defined per joint as KL ≥2 or prosthesis at year eight. The association between incidenct arterial calcifications and incident OA was studied using mixed-effects logistic regression.

### Results

Of 763 participants included, 623 (82%) were women. Mean (sd) age was 56 (5.1) years, mean (sd) BMI 26.2 (4.1) kg/m². Arterial calcifications developed in 174 participants (283 joints). OA developed in 456 participants (778 joints). Sex modified the association between arterial calcification and OA. In women, incident arterial calcification around a joint was positively associated with incident OA in that joint (adjusted OR 2.51 (95%-CI 1.57 to 4.03)). In men, no association was observed on joint-level, but at patient-level the arterial calcification sum score was negatively associated with incident OA (adjusted OR per point increase 0.70 (95%-CI 0.54 to 0.90)) indicating a systemic effect.

### Conclusion

We observed sex-dependent associations between incident arterial calcification and incident radiographic knee and/or hip OA, which differ between joint- and patient-level.

## Introduction

Cardiovascular diseases (CVD) is a major cause of disability and caused nearly one third of the world deaths in 2010. The global economic burden of CVD was estimated at 836 billion US dollar for 2010 and is expected to rise beyond a trillion US dollar in 2025.[1] Calcification in the intimal or medial layer of the arteries can be assessed on radiographs.[2] Presence of arterial calcifications on radiographs strongly predicts cardiovascular events and correlates with the calcification burden as measured by microscopic histology.[2–4]

It is becoming widely accepted that osteoarthritis (OA) is a heterogeneous disease with different phenotypes resulting through different pathways. Amongst possible pathways, multiple epidemiologic studies demonstrate an association between CVD and OA. The diseases co-occur more often than would be expected based on the incidence of each disease, and this increases with age.[5–8] Remarkably, the association between CVD and OA seems to be more evident in women.[6,7,9–11] Both diseases show a peak incidence in women around menopause.[12,13] Establishing an association between OA and arterial calcification may help to discover pathways and related etiology or bring forward new treatment targets for both CVD and OA.

Different working mechanisms for an association between CVD and OA have been suggested. Firstly, the association may be related to chronic inflammation and the metabolic syndrome.[6,8,14] This is supported by similar mediators in CVD and OA, which mainly are drivers of the metabolic syndrome as well as inflammation (e.g. adipokines and cytokines).[15,16] Second, the disability caused by OA may be a driver for CVD.[8] Hoeven et al. suggest that disability and not OA itself may predict the excess of CVD events, probably through physical inactivity.[17] Third, local vascular abnormalities may drive degenerative joint changes. These vascular abnormalities result from repair mechanisms in a joint and lead to poor perfusion of the joint.[18–20] Gosh and Cheras denominate synovitis as a cause for a state of hypercoagulation, hypofibrinolysis, and thrombosis in the subchondral bone vasculature. Such a state may result in subchondral bone necrosis and subsequent OA.[21] Following a large number of post-mortem examinations, Harrison et al. suggest that hypervascularity of the subchondral bone as a reaction to joint damage, may be a main cause of the change in bone metabolism, as seen in the early development of OA.[22]

Depending on the hypothesized working mechanism the association between CVD and OA would be expected to be mainly present on either patient-level (first and second mechanisms) and/or joint-level (third mechanism). The present research aims at studying the association between incidence of arterial calcifications as a marker of CVD development and incidence of knee and/or hip OA, distinguishing between both patient- and joint-level.

## Methods

### Study participants

We used data and radiographs from the Cohort Hip and Cohort Knee (CHECK). CHECK is a prospective cohort initiated to study early OA in the knee and/or hip described in detail elsewhere.[23] From October 2002 untill September 2005, 1002 participants were approached by their general practitioner and/or recruited via local media and included in one of 10 participating centers throughout the Netherlands.

Participants aged 45-65 years at the time of inclusion, were included when they had pain and/or stiffness in at least one knee and/or hip. They had never or only recently (<6 months) visited a physician for these symptoms. Participants were excluded if pathology other than OA could explain these symptoms, or if they suffered from co-morbidities precluding follow-up or precluding clinical examinations, or were unable to understand Dutch.

### Scoring of arterial calcifications

A board-certified radiologist (PJ), blinded for OA scores, assessed arterial calcifications on baseline and eight year follow-up radiographs. Standing antero-posterior pelvic radiographs were used to asses iliac and femoral arteries. Standing semi-flexed antero-posterior and lateral knee radiographs were used to asses femoral, popliteal and crural arteries. A detailed description of radiographic acquisition can be found elsewhere.[24] Calcifications were scored using a four grade scale representing absence of (grade 0), mild (grade 1), moderate (grade 2) and severe (grade 3) calcifications (**Figure 1**). To reflect involvement at joint-level, the areas surrounding the right and left hip and right and left knee were scored separately. To reflect the total load of peripheral arterial calcifications (i.e. patient-level) we summed the score of the four joints to obtain the arterial calcification sum score (range 0-12).

We tested inter- and intra-observer agreement of calcification grading in 30 patients (60 hips and 60 knees). Intra-observer (PJ) agreement was tested by scoring the sample with a one-year interval. Quadratic weighted kappa's were 0.78 and 0.94, for hip and knee radiographs respectively. A trained reader (WPG) scored the same sample to assess inter-observer agreement. Quadratic weighted kappa's were 0.78 and 0.88, for hip and knee radiographs, respectively.

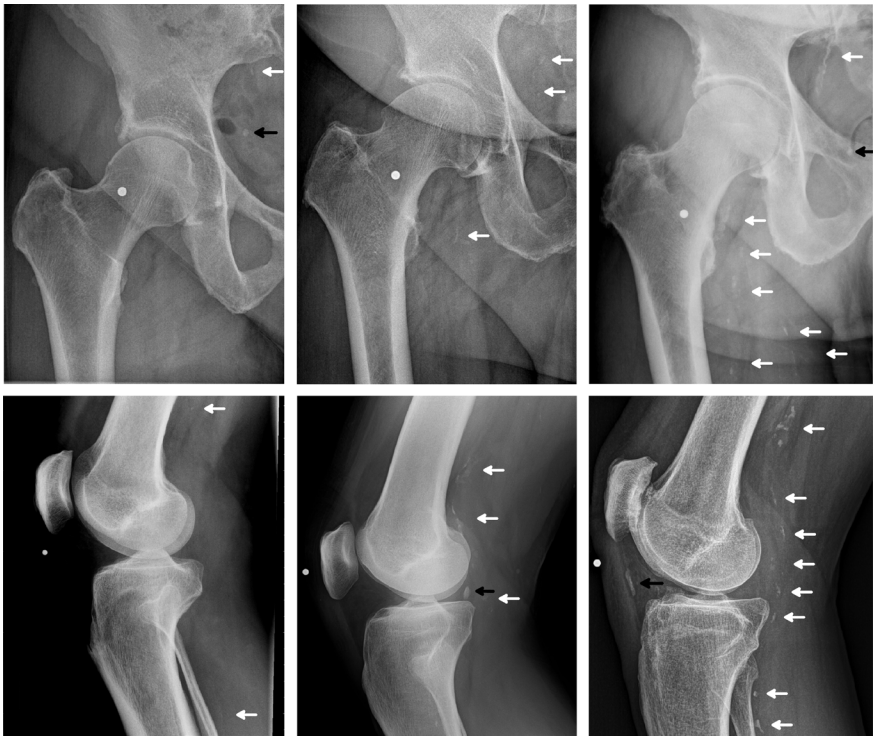


**Figure 1.** Radiographs of the right hip (top) and knees (bottom) which show mild (left), moderate (middle), and severe (right) calcifications. White arrows indicate arterial calcifications, black arrows indicate phlebolites (top), fabella (bottom middle), and calcified tendinitis (bottom right).

## Scoring of osteoarthritis

Kellgren and Lawrence grades (KL) were scored by five trained observers, unaware of the present research question, to assess structural OA features.[25] Methods and inter-observer agreement for scoring of OA in CHECK were reported earlier.[24] CHECK aimed to include very early OA and all knees of included patients had KL grades of 0 and 1 for hips and knees (no or doubtfull OA) at baseline. In the present study incident radiographic OA was defined per joint as a KL ≥2 at eight-year follow-up. Joint replacements were also considerd incident OA, as structural damage visualised by radiographs plays a major role in the decision for joint replacement.

### Other variables

Data on multiple characteristics of CHECK participants were collected. Length in centimetres (self-reported) and weight in kilograms were measured. Information on education level, smoking (current/non-current), and comorbidities were collected via questionnaires.[23]

### Statistical analysis

Participant characteristics are presented using numbers (proportions) and means (sd). As discussed, both arterial calcifications and OA can be seen as a either local processes or systemic diseases. Therefore, we analysed the association between incidence of arterial calcification and OA on patient- as well as joint-level.

To identify incidence of arterial calcification we used the scores at eight years and excluded participants that already had calcifications present at baseline around any joint. To account for interdependence between observations on joints within patients we used a mixed-effects logistic regression model. Using this model, we examined the association between the arterial calcification sum score (i.e. on patient-level) at follow-up (independent variable) and incident OA (dependent variable) with joint as the unit of observation. For the analysis on joint-level we used the same analysis method, but now with the occurrence of arterial calcification defined on joint-level. As moderate and severe arterial calcifications occurred infrequently we used the occurrence of any arterial calcifications (grade ≥1 yes/no) as definition for incident arterial calcification in this analysis.

We investigated modification of the association between incident arterial calcifications and OA by sex, joint type, and the presence of pain at baseline, by adding interaction terms to the models. Possible confounders of the association between incident arterial calcification and OA (age, sex, body mass index (BMI) at baseline, difference in BMI between baseline and follow-up baseline KL, smoking status (baseline), education level, hypertension (at baseline and follow-up), and diabetes (at baseline and follow-up), presence of pain (at baseline), menopausal status for women (at baseline), and joint type for the analysis on joint-level) were investigated in both models, by adding these variables as covariates. We used a manual backward elimination strategy removing covariates based on their association with the outcome and change in the regression coefficient for arterial calcification (when Akaike's Information Criteria lowers and/or the parameter estimate for association changes > 5%). As both diseases are highly dependent on age and sex, these were never eliminated from the model. P-values under 0.05 were

considered statistically significant. All available data was used, missing data was not imputed. We performed a sensitivity analysis to evaluate the effect of including joint replacements as incident radiographic OA by excluding these cases from the analysis. As KL 1 may already represent osteoarthritic changes, we performed an additional sensitivity analysis by excluding joints with baseline KL 1. Statistical analyses were performed using R (version 3.2.2).

## Results

### Participants and outcomes

Of the 1002 participants included in CHECK, 763 participants (3052 hip and knee joints) could be included in the present study (**Figure 2**). Relatively more men than women were excluded because of arterial calcifications at baseline (62 (8% of women) vs. 49 (24% of men)). The incidence of arterial calcifications was higher in men compared to women (**Table 1 and 2**). Men and women were comparable in terms of age and BMI, but women were less likely to smoke (71 (12%) vs 23 (18%)).

778 out of 2941 joints developed radiographic OA (**Table 2 and supplementary table 1**). In men incident OA occurred equally frequent in hips (68 joints (27%)) and knees (68 joints (27%)). Women developed OA more often in their knees (406 joints (36%)) as compared to their hips (185 joints (17%)).

### The association between arterial calcifications and osteoarthritis on patient-level

Modification of the association between OA and calcification by sex was observed (OR for interaction 0.56, 95%-CI (0.41 to 0.77)) and results are therefore presented separately for men and women (**Table 3**). Modification by joint type was not observed.

In men, the sum score of arterial calcification showed a negative association with the development of OA. This effect became slightly more evident after adjustment for confounders (age, sex, body mass index (BMI), menopausal status at baseline, difference in BMI between baseline and follow-up baseline KL, smoking status (baseline), education levelhypertension (baseline and follow-up), and diabetes (baseline and follow-up), presence of pain at baseline,). In women, the sum score of arterial calcifications was positively associated with developing OA in the crude analysis. However, this effect was no longer statistically significant after adjustment for confounders (**Table 3**). Higher age and a higher baseline KL grade increased the odds of incident OA.

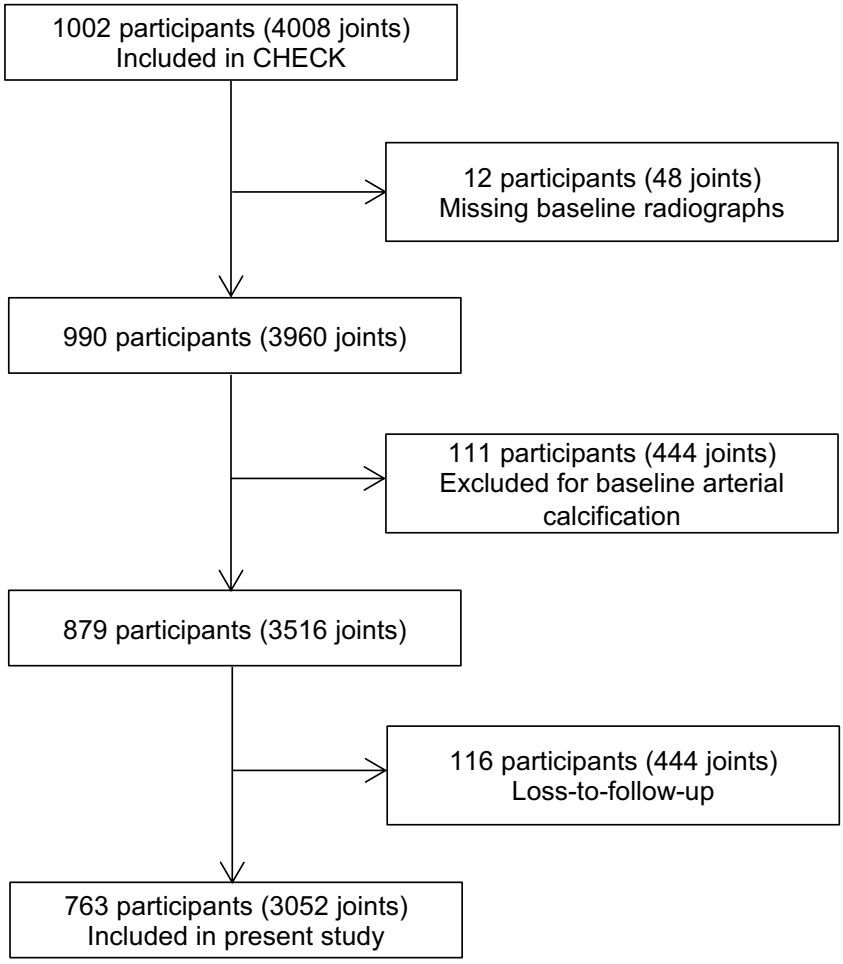



**Figure 2.** Flowchart of participants included in the present study.

## The association between osteoarthritis and arterial calcifications on joint-level

Also, for the analysis on joint-level, modification by sex was observed (OR for interaction 0.34, 95%-CI (0.14 to 0.82)) and results are therefore presented separately for men and women. Modification by joint type was not observed. In men, incident arterial calcification was not associated with incident OA in either crude or adjusted analysis (**Table 4**). In women, local incident arterial calcification was associated with incident OA, also after adjustment for confounders (**Table 4**).

**Table 1.** Characteristics

| **Baseline** | **Total** | **N[1]** | **Men** | **N[1]** | **Women** | **N[1]** |
|---|---|---|---|---|---|---|
| Age in years, mean (SD) | 56 (5.1) | 763 | 55 (5.3) | 140 | 56 (5.0) | 623 |
| BMI in kg/m$^2$, mean (SD) | 26.2 (4.1) | 749 | 26.5 (3.3) | 136 | 26.2 (4.3) | 613 |
| Currently smoking, N (%) | 97 (13.0) | 747 | 23 (16.9) | 136 | 74 (12.1) | 611 |
| Hypertension, N(%) | 148 (19.7) | 750 | 23 (16.8) | 137 | 125 (20.4) | 613 |
| Diabetes, N(%) | 18 (2.4) | 750 | 3 (2.1) | 137 | 15 (2.4) | 613 |
| **Menopausal status** | | | | | | 620 |
| Pre-menopausal, N(%) | | | | | 128 (20.6) | |
| Post-menopausal, N(%) | | | | | 397 (64.0) | |
| Hormone usage, N(%) | | | | | 14 (2.3) | |
| Unknown, N (%) | | | | | 81 (13.1) | |
| **KL for osteoarthritis[2]** | | 740 | | 134 | | 606 |
| KL-1 in 1 joint, N (%) | 204 (27.6) | | 38 (28.4) | | 166 (27.4) | |
| KL-1 in 2 joints, N (%) | 167 (22.6) | | 28 (20.9) | | 139 (22.9) | |
| KL-1 in 3 joints, N (%) | 63 (8.5) | | 17 (12.7) | | 46 (7.6) | |
| KL-1 in 4 joints, N (%) | 39 (5.3) | | 9 (6.7) | | 30 (5.0) | |
| **8-year follow-up** | **Total** | **N[1]** | **Men** | **N[1]** | **Women** | **N[1]** |
| Change in BMI, mean (SD) | 0.24 (2.1) | 749 | 0.07 (1.9) | 135 | 0.28 (2.1) | 608 |
| Currently smoking, N (%) | 66 (8.8) | 748 | 18 (12.9) | 139 | 48 (7.7) | 609 |
| Hypertension, N(%) | 208 (27.5) | 756 | 32 (23.0) | 139 | 176 (28.3) | 617 |
| Diabetes, N(%) | 43 (5.6) | 756 | 10 (7.2) | 139 | 33 (5.3) | 617 |

**1** Number of the in total 763 included participants, for whom data on the specific variable was available. **2** All joints were graded 0 or 1 as described by Kellgren and Lawrence at baseline. Here the number of KL 1 joints in one patient is given.

A higher age and higher baseline KL were associated with a higher incidence of OA. Current smoking was associated with a lower incidence of OA. In women, joint type was a significant confounder. Hips were less likely to develop OA compared to knees.

## Sensitivity analysis

Results were not significantly altered (<10% change in parameters and same conclusions) by excluding joint replacements or baseline KL 1 joints from the analysis.

**Table 2.** Incidence of osteoarthritis and arterial calcifications

| | Total | Men | Women |
|---|---|---|---|
| **Osteoarthritis** | 2941 joints | 541 joints | 2400 joints |
| No osteoarthritis, N (%) | 2163 (73.5) | 391 (73.0) | 1772 (73.8) |
| Incident osteoarthritis, N (%) | 778 (26.5) | 150 (27.0) | 628 (26.2) |
| **Joint-level arterial calcifications** | 2958 joints | 542 joints | 2416 joints |
| None, N (%) | 2675 (90.4) | 454 (82.9) | 2221 (91.9) |
| Mild, N (%) | 219 (7.4) | 65 (12.5) | 154 (6.4) |
| Moderate, N (%) | 58 (2.0) | 21 (4.2) | 37 (1.5) |
| Severe, N (%) | 6 (0.2) | 2 (0.4) | 4 (0.2) |
| **Patient-level arterial calcifications** | 731 participants | 134 participants | 597 participants |
| Sum score =0, N (%) | 557 (76.2) | 85 (63.4) | 472 (79.1) |
| Sum score ≥1, N (%) | 174 (23.8) | 49 (36.6) | 125 (20.9) |
| Sum score ≥3, N (%) | 34 (4.7) | 15 (11.2) | 19 (3.2) |
| Sum score ≥6, N (%) | 8 (1.1) | 2 (1.5) | 6 (1.0) |

The category score ≥1 includes all participants with an arterial calcification sum score of one or more. Therefore, this category includes partly the same participants as categories ≥3 and ≥6. For some patients the arterial calcification score could not be calculated due to missing information for one or more joints. However, the data of available joints of these patients were used in the analysis on joint-level.

**Table 3.** Association between incident arterial calcification on patient-level and incident osteoarthritis.

| | Crude model | | Adjusted model | |
|---|---|---|---|---|
| | OR | 95% CI | OR | 95% CI |
| **Men** | | | | |
| Arterial calcification sum score (Points) | 0.74 | (0.58 to 0.95) | 0.71 | (0.55 to 0.92) |
| Age (Years) | | | 1.06 | (1.00 to 1.13) |
| **Women** | | | | |
| Arterial calcification sum score (Points) | 1.14 | (1.00 to 1.29) | 1.09 | (0.96 to 1.24) |
| Age (Years) | | | 1.05 | (1.02 to 1.08) |

[1]Kellgren-Lawrence grade for osteoarthritis

**Table 4.** Association between incident arterial calcification on joint-level and incident osteoarthritis

| | Crude model | | Adjusted model | |
|---|---|---|---|---|
| | OR | 95% CI | OR | 95% CI |
| **Men** | | | | |
| Arterial calcification (Present) | 0.91 | (0.46 to 1.81) | 0.83 | (0.41 to 1.68) |
| Age (Years) | | | 1.03 | (0.97 to 1.09) |
| Smoking (Current) | | | 0.47 | (0.19 to 1.09) |
| Baseline[1] KL (1 vs 0) | | | 3.25 | (1.96 to 5.40) |
| **Women** | | | | |
| Arterial calcification (Present) | 2.00 | (1.36 to 2.93) | 2.51 | (1.57 to 4.03) |
| Age (Years) | | | 1.05 | (1.01 to 1.09) |
| Smoking (Current) | | | 0.61 | (0.35 to 1.07) |
| Joint type (Hip) | | | 0.31 | (0.25 to 0.39) |
| Baseline KL [1] (1 vs 0) | | | 6.96 | (5.27 to 9.18) |

[1]Kellgren and Lawrence grade for osteoarthritis

## Discussion

Within CHECK, a large prospective cohort of middle-aged Dutch participants with complaints of early hip and/or knee OA, we found sex-dependent associations between incidence of peripheral arterial calcification and hip and/or knee OA. Arterial calcification and OA were measured using radiographic criteria, both showing good inter and intra-observer reliability. In men, the extent of incident arterial calcifications was negatively associated with the incidence of OA (i.e. on patient-level), but there was no association between local arterial calcifications and OA (i.e. joint-level). In contrast, in women incident arterial calcifications were not associated with OA on patient-level, but a positive association between incident arterial calcifications in the proximity of a joint and incidence of OA in that joint was observed (i.e. joint-level).

The effect of local arterial calcification on OA as we found in women, is in line with the hypothesis that local vascular pathology is a main driver of the degenerative joint changes leading to OA.[19–22,26] Remarkably, we found this effect in women only. In men, we found a surprising negative association between incidence of arterial calcifications and incidence of OA on patient-level. Such a negative association has not been described previously.

To further study the different results between analyses at joint- and patient-level, we evaluated whether the arterial calcification sum score at patient-level had an effect above the local (i.e. joint-level) effect of arterial calcifications. We added the arterial calcification sum score of the other three joints (0-9) as a covariate in the joint-level model. This did not significantly affect the local effect and again the sum score was not related to OA in women (data not shown). In men, the effect of local calcification diminished (OR 0.93, 95% – CI 0.46 to 1.91)) and the effect of the sum score of the other joints showed an OR of 0.64, 95% – CI 0.46 to 0.91).

As discussed below, most papers describe a positive association between markers of CVD and hip and/or knee OA in women, but no association in men. Hoeven et al. studied 5650 middle-aged participants and found an independent association between carotid intima media thickness and prevalence of knee OA in women.[6] In the same cohort, the authors found an association between blood levels of vascular cell adhesion molecule 1 and CD40L (markers of atherosclerosis) and prevalence, but not progression of knee OA, only in women. [9] In a recent cohort study (n=2158) Veronese et al. found a higher incidence of atherosclerotic CVD (coronary artery disease, stroke or peripheral artery disease (PAD, ankle brachial index <0.9)) in women with baseline hip or knee OA, but not in men.[27] In a cohort of 142 asymptomatic middle-aged women, Wang et al. found that higher baseline popliteal arterial wall thickness was associated with increased loss of medial tibial cartilage during two-year follow-up.[10] In a related study, they found that baseline popliteal arterial wall thickness was associated with lower medial tibial cartilage volume cross-sectionally and higher medial tibial cartilage loss over 2 years in 278 asymptomatic adults.[11] They did not mention any sex-dependent effects in either paper. We are the first to describe not only sex-specific, but also location specific (i.e. patient- and joint-level) associations between CVD and hip and knee OA.

The relationship between arterial calcification and OA has been studied sparsely and results are conflicting. Karasik et al. studied lateral lumbar and hand radiographs of 777 men and 1241 women. They found that a higher prevalence of abdominal aortal calcification was associated with a higher prevalence of anterior lumbar osteophytes, but not hand osteophytes. The strength of the association did not differ between men (OR 1.20) and women (OR 1.25).[28] Jonsson et al. found that in women (n=3078) both carotid plaque severity and coronary calcifications were associated with hand OA. No associations were found in men (n=2264).[7] In a longitudinal cohort study (n=1669), Hoeven et al. found an association between neither presence nor progression of knee OA and baseline coronary artery calcification, in men nor women.[9]

Several explanations for our results in women can be hypothesized. Decreasing estrogen levels as a consequence of menopause have a pro-inflammatory effect, giving momentum to the atherosclerotic process.[29] Of the 525 women for whom menopausal status was determinable at baseline, 397 (76%) were postmenopausal and 128 (24%) are expected to have passed menopause during the eight-year follow-up (mean age of 50 years at baseline). As atherosclerosis in women is more diffusely spread in coronary vessels as compared to the focal pattern in men [30], it is plausible that peripheral atherosclerosis is also more widespread.[30] In addition, in women atherosclerosis takes place in smaller diameter vessels and calcification starts in a later stage of the atherosclerotic process.[31] Microvascular dysfunction and thrombus formation via shear stress are main causes of disrupted blood flow in women.[29,30,32] Thrombus formation and disrupted blood flow in the supplying vessel of a joint may cause subchondral ischemia and initiate or undertow the osteoarthritic process.[18] Physical impairment caused by OA is a larger problem in women, when compared to men.[33,34] Disability is a risk factor for CVD and development of arterial calcification, however we would expect this relationship to act on the patient-level.[17] Additionally, factors other than ischemia may have a smaller role in the development of OA in women (e.g. trauma or joint shape).[35–37]

We also hypothesize about the unexpected negative association between arterial calcification and osteoarthritis on the patient-level in men. Atherosclerosis in men is mostly distributed in focal plaques. Calcification may stabilize these plaques and thus prevent rupture and arterial occlusion with subsequent subchondral ischemia and OA.[26,38] Another explanation may lay in CVD risk awareness and profiling. Participants might be medicalized because of complains related to OA (CHECK inclusion criteria), which may in turn trigger CVD risk evaluation and treatment. Resulting treatments and/or lifestyle changes may decrease the risk of atherosclerosis and calcification. This effect may be stronger in men, as awareness for CVD risk is still suboptimal in women.[32,39,40] There is ambivalent evidence on a role for statins in the treatment of OA.[41] However, we do not have detailed information on drug use in this cohort.

We studied the co-occurrence of incident peripheral arterial calcification and OA on hip and knee radiographs over eight years. This limits the causal interpretation of the relationship and results may not be extrapolated to clinically relevant artery disease or OA. Standard radiography is very specific, but relatively insensitive to detect arterial calcifications.[2] At baseline only 184 of 3946 joints (4.6%) had arterial calcification present on the same radiographs. We believe this number would not give us sufficient statistical power to relate baseline arterial calcifications to the

presence of OA at baseline or progression to OA during follow-up. Future research aimed at studying these relationships may include tools more sensitive to detect arterial calcification such as CT-scans.

Arterial calcifications were scored blinded for KL-grade. However, blinding for joint appearance (joint space/osteophytes/sclerosis) was considered infeasible. Missing data was sparse in our data set. Although no detailed information about the causes of missing data was available, we assumed reasons for missing data were independent of the association between arterial calcification and OA. Therefore, we assumed data to be missing at random in which case mixed-effects models will produce valid parameter estimates.[42,43]

## Conclusion

The association between incident arterial calcification and OA in this population-based cohort study of middle-aged participants with complaints of early hip and/or knee OA (CHECK), is sex-dependent and differs between patient- and joint-level. In men, the incidence of radiographic OA is negatively associated with arterial calcifications on patient-level, indicating a systemic effect. In women, local arterial calcification around a joint is positively associated with incidence of radiographic OA in that joint. The current results point to an underlying pathway between CVD and OA that might differ between men and women. Elucidating this pathway and knowledge about the shared features of CVD and OA is essential for finding common targets that may lead to disease modifying treatments for both diseases.

## Acknowledgements

The authors would like to thank the CHECK-cohort, initiated by the Dutch Arthritis Association and performed within; Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen/ Allied Health Care Center for Rheum. and Rehabilitation Groningen; Medical Spectrum Twente Enschede/ Ziekenhuisgroep Twente Almelo; Reade, formerly Jan van Breemen Institute/VU Medical Center Amsterdam; St. Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen; and especially Gerard van Hoorn for organizing all radiographs.

## Funding

The Cohort Hip and Cohort Knee was funded by the Dutch Arthritis Association. The research leading to these results was undertaken as part of the APPROACH

consortium (www.approachproject.eu). APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement n°115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme(FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu.

Contributions to this work by JR have been partly funded by a program grant of the Dutch Arthritis Foundation for their centre of excellence "Osteoarthritis in primary care".

## Supplementary material

**Supplementary Table 1**. Kellgren-Lawrence grades vs Arterial calcification score at follow-up

| | Absent N (%) | Mild N (%) | Moderate N (%) | Severe N (%) | Total N (%) |
|---|---|---|---|---|---|
| **Hip** | | | | | |
| KL 0 | 566 (91.6) | 44 (7.1) | 7 (1.1) | 1 (0.2) | 618 (100) |
| KL 1 | 518 (89.8) | 49 (8.5) | 9 (1.6) | 1 (0.2) | 577 (100) |
| KL 2 | 178 (88.6) | 20 (10.0) | 3 (1.5) | 0 (0) | 201 (100) |
| KL 3 | 10 (90.9) | 1 (9.1) | 0 (0) | 0 (0) | 11 (100) |
| Prosthesis | 50 (86.2) | 8 (13.8) | 0 (0) | 0 (0) | 58 (100) |
| **Total** | 1322 (90.2) | 122 (8.3) | 19 (1.3) | 2 (0.1) | 1465 (100) |
| **Knee** | | | | | |
| KL 0 | 227 (94.2) | 9 (3.7) | 5 (2.1) | 0 (0) | 241 (100) |
| KL 1 | 659 (90.8) | 46 (6.3) | 20 (2.8) | 1 (0.1) | 726 (100) |
| KL 2 | 389 (89.0) | 34 (7.8) | 11 (2.5) | 3 (0.7) | 437 (100) |
| KL 3 | 38 (84.4) | 4 (8.9) | 3 (6.7) | 0 (0) | 45 (100) |
| Prosthesis | 19 (90.5) | 2 (9.5) | 0 (0) | 0 (0) | 21 (100) |
| **Total** | 1332 (90.6) | 95 (6.5) | 39 (2.7) | 4 (0.3) | 1470 (100) |

KL=Kellgren and Lawrence grade for osteoarthritis.

## References


1. Bloom DE, Cafiero E, Jané-Llopis E, et al. The Global Economic Burden of Noncommunicable Diseases. *World Econ Forum*. 2011;(September):1-46. http://ideas.repec.org/p/gdm/wpaper/8712.html.

2. Orr DP, Myerowitz RL, Herbert DL, Friday P. Correlation of radiographic and histologic findings in arterial calcification. *Invest Radiol*. 1978;13(2):110-114. doi:10.1097/00004424-197803000-00002

3. Peters SAE, den Ruijter HM, Bots ML, Moons KGM. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. *Heart*. 2012;98(3):177-184. doi:10.1136/heartjnl-2011-300747

4. McClelland RL, Jorgensen NW, Budoff M, et al. 10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors: Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) With Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Stu. *J Am Coll Cardiol*. 2015;66(15):1643-1653. doi:10.1016/j.jacc.2015.08.035

5. Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. *Ann Rheum Dis*. 2004;63(4):408-414. doi:10.1136/ard.2003.007526

6. Hoeven T a., Kavousi M, Clockaerts S, et al. Association of atherosclerosis with presence and progression of osteoarthritis of the knee and hand: the Rotterdam study. *Ann Rheum Dis*. 2013;72:646-651. doi:10.1016/j.joca.2012.02.386

7. Jonsson H, Helgadottir GP, Aspelund T, et al. Hand osteoarthritis in older women is associated with carotid and coronary atherosclerosis: the AGES Reykjavik study. *Ann Rheum Dis*. 2009;68(11):1696-1700. doi:10.1136/ard.2008.096289

8. Nüesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Jüni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ*. 2011;342(mar08_2):d1165. doi:10.1136/bmj.d1165

9. Hoeven TA, Kavousi M, Ikram MA, et al. Markers of atherosclerosis in relation to presence and progression of knee osteoarthritis: a population-based cohort study. *Rheumatology (Oxford)*. May 2015:kev106-. doi:10.1093/rheumatology/kev106

10. Wang Y, Dawson C, Hanna F, et al. Association between popliteal artery wall thickness and knee cartilage volume loss in community-based middle-aged women without clinical knee disease. *Maturitas*. 2015;82(2):222-227. doi:10.1016/j.maturitas.2015.07.010

11. Wang Y, Novera D, Wluka AE, et al. Association between popliteal artery wall thickness and knee structure in adults without clinical disease of the knee: a prospective cohort study. *Arthritis Rheumatol (Hoboken, NJ)*. 2015;67(2):414-422. doi:10.1002/art.38922

12. Kannel WB, Hjortland MC, McNamara PM, Gordon T. Menopause and risk of cardiovascular disease: the Framingham study. *Ann Intern Med*. 1976;85(4):447-452. doi:10.7326/0003-4819-85-4-447

13. Nadkar MY, Samant RS, Vaidya SS, Borges NE. Relationship between osteoarthritis of knee and menopause. *J Assoc Physicians India*. 1999;47(12):1161-1163. http://www.ncbi.nlm.nih.gov/pubmed/11225216. Accessed October 6, 2016.

14. Kornaat PR, Sharma R, van der Geest RJ, et al. Positive association between increased popliteal artery vessel wall thickness and generalized osteoarthritis: is OA also part of the metabolic syndrome? *Skeletal Radiol*. 2009;38(12):1147-1151. doi:10.1007/s00256-009-0741-7

15. Scotece M, Conde J, Gómez R, et al. Role of Adipokines in Atherosclerosis: Interferences with Cardiovascular Complications in Rheumatic Diseases. *Mediators Inflamm*. 2012;2012:1-14. doi:10.1155/2012/125458

16. Fernandes GS, Valdes AM. Cardiovascular Disease and Osteoarthritis: Common Pathways and Patient Outcomes. *Eur J Clin Invest*. 2015;45:n/a-n/a. doi:10.1111/eci.12413

17. Hoeven T a, Leening MJG, Bindels PJ, et al. Disability and not osteoarthritis predicts cardiovascular disease; a prospective population-based cohort study. *Osteoarthr Cartil*. 2013;21(2013):S46. doi:10.1136/annrheumdis-2013-204388

18. Findlay DM. Vascular pathology and osteoarthritis. *Rheumatology (Oxford)*. 2007;46(12):1763-1768. doi:10.1093/rheumatology/kem191

19. Conaghan PG, Vanharanta H, Dieppe PA. Is progressive osteoarthritis an atheromatous vascular disease? *Ann Rheum Dis*. 2005;64(11):1539-1541. doi:10.1136/ard.2005.039263

20. Binks DA, Gravallese EM, Bergin D, et al. Role of vascular channels as a novel mechanism for subchondral bone damage at cruciate ligament entheses in osteoarthritis and inflammatory arthritis. *Ann Rheum Dis*. 2015;74(1):196-203. doi:10.1136/annrheumdis-2013-203972

21. Ghosh P, Cheras P a. Vascular mechanisms in osteoarthritis. *Best Pract Res Clin Rheumatol*. 2001;15(5):693-709. doi:10.1053/berh.2001.0188

22. Harrison MH, Schajowicz F, Trueta J. Osteoarthritis of the hip: a study of the nature and evolution of the disease. *J Bone Joint Surg Br*. 1953;35-B(4):598-626. http://www.ncbi.nlm.nih.gov/pubmed/13108925. Accessed June 23, 2016.

23. Wesseling J, Boers M, Viergever MA, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. *Int J Epidemiol*. 2016;45(1):36-44. doi:10.1093/ije/dyu177

24. Damen J, Schiphof D, Wolde S Ten, Cats HA, Bierma-Zeinstra SMA, Oei EHG. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. *Osteoarthritis Cartilage*. 2014;22(7):969-974. doi:10.1016/j.joca.2014.05.007

25. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. http://www.scopus.com/inward/record.url?eid=2-s2.0-62949246527&partnerID=tZOtx3y1.

26. Findlay DM. Vascular pathology and osteoarthritis. *Rheumatology (Oxford)*. 2007;46(12):1763-1768. doi:10.1093/rheumatology/kem191

27. Veronese N, Trevisan C, De Rui M, et al. Association of Osteoarthritis With Increased Risk of Cardiovascular Diseases in the Elderly: Findings From the Progetto Veneto Anziano Study Cohort. *Arthritis Rheumatol (Hoboken, NJ)*. 2016;68(5):1136-1144. doi:10.1002/art.39564

28. Karasik D, Kiel DP, Kiely DK, et al. Abdominal aortic calcification and exostoses at the hand and lumbar spine: The Framingham study. *Calcif Tissue Int*. 2006;78(1):1-8. doi:10.1007/s00223-005-0054-3

29. Fairweather D. Sex differences in inflammation during atherosclerosis. *Clin Med Insights Cardiol*. 2014;8(Suppl 3):49-59. doi:10.4137/CMC.S17068

30. Shaw LJ, Bugiardini R, Merz CNB. Women and Ischemic Heart Disease. Evolving Knowledge. *J Am Coll Cardiol*. 2009;54(17):1561-1575. doi:10.1016/j.jacc.2009.04.098

31. Hirsch AT, Allison M a., Gomes AS, et al. A call to action: Women and peripheral artery disease: A scientific statement from the american heart association. *Circulation*. 2012;125(11):1449-1472. doi:10.1161/CIR.0b013e31824c39ba

32. EUGenMed, Cardiovascular Clinical Study Group, Regitz-Zagrosek V, et al. Gender in cardiovascular diseases: impact on clinical manifestations, management, and outcomes. *Eur Heart J*. 2016;37(1):24-34. doi:10.1093/eurheartj/ehv598

33. Keefe FJ, Lefebvre JC, Egert JR, Affleck G, Sullivan MJ, Caldwell DS. The relationship of gender to pain, pain behavior, and disability in osteoarthritis patients: the role of catastrophizing. *Pain*. 2000;87(3):325-334. doi:S0304395900002967 [pii]

34. D.S. L, J. Z, L. S-M, Logerstedt DS, Zeni Jr J, Snyder-Mackler L. Sex differences in patients with different stages of knee osteoarthritis. *Arch Phys Med Rehabil*. 2014;95(12):2376-2381. doi:10.1016/j.apmr.2014.07.414 [doi]

35. Majewski M, Susanne H, Klaus S. Epidemiology of athletic knee injuries: A 10-year study. *Knee*. 2006;13(3):184-188. doi:10.1016/j.knee.2006.01.005

36. Agricola R, Heijboer MP, Bierma-Zeinstra SM a., Verhaar J a. N, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). *Ann Rheum Dis*. 2013;72(6):918-923. doi:10.1136/annrheumdis-2012-201643

37. Hack K, Di Primio G, Rakhra K, Beaulé PE. Prevalence of Cam-Type Femoroacetabular Impingement Morphology in Asymptomatic Volunteers. *J Bone Jt Surg*. 2010;92(14).

38. Wong KK, Thavornpattanapong P, Cheung SC, et al. Effect of calcification on the mechanical stability of plaque based on a three-dimensional carotid bifurcation model. *BMC Cardiovasc Disord*. 2012;12(1):7. doi:10.1186/1471-2261-12-7

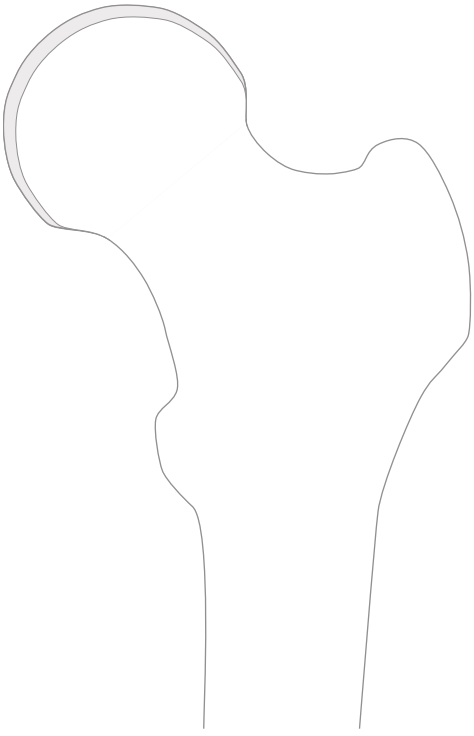
39. Zhang H, Plutzky J, Shubina M, Turchin A. Drivers of the Sex Disparity in Statin Therapy in Patients with Coronary Artery Disease: A Cohort Study. *PLoS One*. 2016;11(5):e0155228. doi:10.1371/journal.pone.0155228

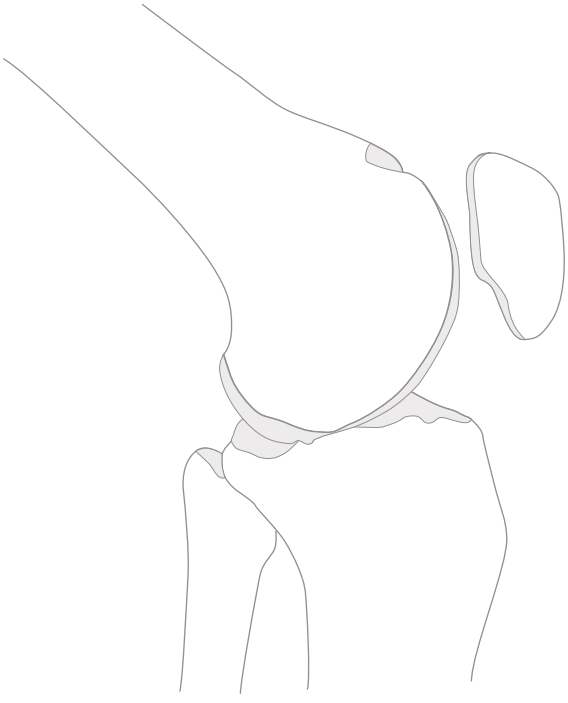
40. Maas AHEM, Leiner T. Gender and age-specific focus needed for cardiovascular outcome measures to improve life-time prevention in high risk women. *Maturitas*. 2016;86:74-76. doi:10.1016/j.maturitas.2016.01.016

41. de Munter W, van der Kraan PM, van den Berg WB, van Lent PLEM. High systemic levels of low-density lipoprotein cholesterol: fuel to the flames in inflammatory osteoarthritis? *Rheumatology (Oxford)*. July 2015:kev270-. doi:10.1093/rheumatology/kev270

42. Molenberghs G, Kenward MG. *Missing Data in Clinical Studies*. John Wiley & Sons, Ltd.; 2007.

43. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581-592. doi:10.1093/biomet/63.3.581





# CHAPTER 12

# Osteoarthritis in pseudoxanthoma elasticum patients: an explorative imaging study

*Journal of Clinical Medicine,*
*special Issue: Pseudoxanthoma Elasticum Pathophysiology, from Clinic to Bench Side Dec 2020,*
*doi.org/10.3390/jcm9123898*

Willem Paul Gielis, , Pim A. de Jong , Jonas W. Bartstra,
Wouter Foppen, Wilko Spiering, Annemarie M. den Harder

## Abstract

### Background and aim

Pseudoxanthoma elasticum (PXE), is a systemic disease affecting the skin, eyes, and cardiovascular system of patients. Cardiovascular disease is associated with osteoarthritis (OA), the most common cause of joint pain. Systematic investigations on joint manifestations in PXE lack from literature. In this explorative study, we aimed to investigate whether patients with PXE are more at risk for developing osseous signs of OA.

### Patients and methods

Patients with PXE and hospital controls with whole-body low dose CT examinations available were included. OA was assessed using the OACT-score, a 4-point Likert scale, in the acromioclavicular (AC), glenohumeral (GH), facet, hip, knee, and ankle joints. Additionally, intervertebral disc degeneration was scored. Data were analyzed using ordinal logistic regression adjusted for age, body mass index (BMI) and smoking status.

### Results

In total 106 PXE patients (age 56 [48 – 64], 42% males, BMI 25.3 [22.7 – 28.2]) and 87 hospital controls (age 55 [43 – 67], 46% males, BMI 26.0 [22.5 – 29.2]) were included. PXE patients were more likely to have a higher OA score for the AC joints (OR 2.00 [1.12 – 3.61]), tibiofemoral joint (OR 2.63 [1.40 – 5.07]), and patellofemoral joint (2.22 [1.18 – 4.24]). For the other joints, the prevalence and severity of OA did not differ significantly.

### Conclusion

This study suggests that patients with PXE are more likely to have structural OA of the knee and AC-joints, which needs clinical confirmation in larger groups and further investigation into the mechanism.

## Introduction

Pseudoxanthoma elasticum (PXE), is a systemic disease affecting the skin, eyes, and cardiovascular system of patients. It is caused by mutations in the *ABCC6* gene and it is associated with low levels of inorganic pyrophosphate (PPi). [1] PPi is found in bones and inhibits the precipitation and dissolution of hydroxyapatite (HA). It is therefore thought to regulate the entry and exit of calcium and phosphate in mineralized tissues and stabilize already formed calcifications. [2,3] In soft tissues, PPi is an important inhibitor of calcification. A deficiency of PPi is characterized by extensive arterial calcifications. The consequences of a deficiency in the PPi homeostasis are shown in several monogenetic disorders. In Generalized Arterial Calcification of Infancy syndrome *(GACI, OMIM #208000)*, the complete lack of PPi results in extensive arterial calcification at birth, but calcification of the joints has also been described in these children.[4] Arterial calcification due to a deficiency in *CD73 (ACDC, OMIM #211800)* typically results in both periarticular and arterial calcification due to the increased conversion of PPi into calcification promoting phosphate.[5–7] The evaluation of the joints in PXE is limited to a recent study on peri-articular calcifications around the shoulder.[8]

In our PXE practice, we noticed that several PXE patients complained about painful joints. In some case reports PXE is co-existing with Rheumatoid arthritis, cervical arthritis, and Still's disease, but we did not find systematic investigations into structural joint disease in PXE.[9–11] Osteoarthritis (OA) is the main cause of disability and joint pain in the general population.[12] Furthermore, OA is related to cardiovascular disease.[13] The most important risk factors are age, body mass index (BMI), gender, and occupation.[14,15] OA is a multifactorial disease with a clear genetic component but 'treatable' phenotypes have not been discovered.[16] A promising potential treatment target is the remodeling of the subchondral bone.[17,18]

The goal of the current study was to investigate whether PXE is associated with a higher prevalence of OA related structural bone disease compared to hospital controls. This line of research could give insight into the origin of the joint pain in PXE patients.

## Methods

### Patients

To determine the prevalence of OA in PXE patients, we used a cohort of consecutive patients with a confirmed PXE diagnosis[19] from the UMC Utrecht in

the Netherlands. At least two of the three diagnostic criteria should be fulfilled: skin lesions; peau d'orange or angioid streaks; pathogenic variants on both *ABCC6* alleles.[20] Sanger sequencing was performed to identify single nucleotide polymorphisms (SNPs) and small deletions and insertions, and multiplex ligation-dependent probe amplification (MLPA) was performed to screen for larger deletions in the *ABCC6* gene (reference sequence NM_001171.5, MLPA kit P092B (https://www.mrcholland.com/). All patients with PXE received a non-contrast-enhanced whole-body CT as part of the routine clinical workup to evaluate the amount of vascular calcification. CT-acquisitions were low-dose (effective dose <3 mSv in a 70 kg adult male) and performed on a 64-slice CT system (Brilliance 64, Philips). PXE Patients did not receive bisphosphonate or anti-vitamin K treatment at the time of or before the scan.

We used a cohort of hospital controls to compare the prevalence of OA. These patients received a whole-body low-dose CT as part of a fluorodeoxyglucose (FDG) positron emission tomography (PET) CT examination between June 2011 and November 2015 for various medical indications. These scans were performed on a Siemens Biograph 40 scanner (Siemens Healthcare, Erlangen, Germany). Patients with suspicion of endocarditis, vasculitis, osteomyelitis, arthritis, or infected osteosynthesis material were excluded. Of all patients, age, gender, BMI, and smoking status within 6 months of the CT acquisition were extracted from the electronic patient file. Previously, Kranenburg et al. investigated the prevalence and severity of arterial calcifications in the same cohort, which is described in detail elsewhere.[21]

### Ethical approval

The need for informed consent was waived by the local institutional review board of the UMC Utrecht, (protocol number 15/446-C) since this concerned a retrospective analysis of data acquired in routine clinical care where clinical and radiological data as acquired in routine clinical care were provided in an anonymized fashion to the researchers.

### Scoring

OA was assessed on the whole-body CTs using the OACT-score.[22] One experienced observer (WPG), scored all scans in random order. His intra-observer and inter-observer reliability compared to two other trained readers are reported previously.[23] Only CT images were assessed, without PET data or DICOM tags. The acromioclavicular (AC) joints, glenohumeral (GH), hip, knee and ankle joints were scored using a 4-grade scale with a score of 0 meaning no OA and a score

of 3 the most advanced stage of OA (osteophytes, marked joint space narrowing, and subchondral sclerosis/cysts). Joints with a prosthesis received the highest score possible (score 3). Degenerative disc disease and facet joint OA were scored separately for the cervical, thoracic and lumbar levels. For each section, only the scores for the two most degenerated levels were given a 4-grade scale based on disc/joint space narrowing, osteophytes, sclerosis and for degenerative disc disease specifically; endplate irregularity. An example of the various grades in the tibiofemoral joint is presented in **Figure 1**.

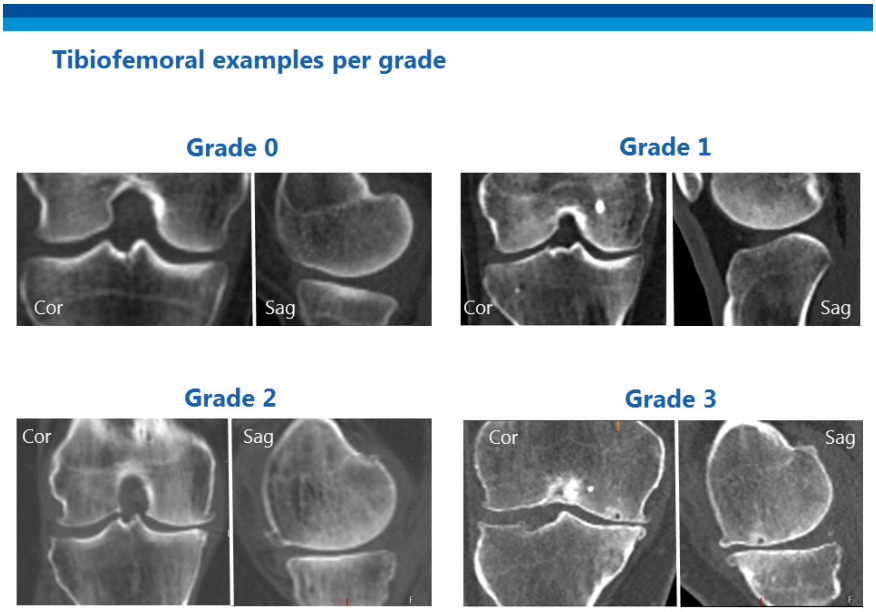



**Figure 1.** Examples of the different scores of osteoarthritis for the tibiofemoral joint.

## Statistical analysis

Statistical analysis was performed with RStudio version 1.1.414 (RStudio Team, Boston, USA). Normally distributed continuous variables are provided as mean ± standard deviation and non-normally distributed continuous variables as median [interquartiles]. Categorical variables are provided as n (percentage). Multiple imputation (number of imputations = 25) was used for missing data (R package 'mice', based on classification and regression trees). Overall, 2% of the data was missing. The imputations of the missing variables were merged into a single variable

by computing the mean of all imputed values (R package 'sjmisc'). Therefore, the analysis was performed on a single dataset, which is a suitable method when the proportion of missing data is limited.[24] Ordinal logistic regression was applied (R package 'MASS') with the OA score for each joint as outcome variable. Unadjusted and adjusted ordinal logistic regression was performed with age, gender, BMI, smoking status (current smoker yes or no), and group (PXE or hospital control) as dependent variables. The chi-squared score test for the proportional odds assumption was used to assess whether the main model assumption was violated or not (R package 'VGAM'). Reported are the crude and adjusted proportional Odds Ratios (OR's) with 95% confidence intervals (CI). Sensitivity analysis was performed for the hip and the knee after excluding patients with a joint prosthesis. In joints that showed a higher prevalence of OA in the PXE group, a subgroup analysis was performed to test the association between the genotype and the prevalence of OA. The PXE group was stratified into the number of truncating variants in the *ABBC6* gene: 2 truncating; 1 truncating and 1 non- truncating; 2 truncating variants. A similar ordinal logistic regression model was used to test the association between the number of truncating mutations and the prevalence of OA.

## Results

In total 106 patients with PXE and 87 hospital controls were included. Baseline characteristics are provided in **Table 1**. The median age was 56 [48 – 64] years in PXE patients and 55 [43 – 67] years in hospital controls. Almost half of the patients were male (42% of PXE patients; 46% of hospital controls) and the median BMI was 25.3 [22.7 – 28.2] in PXE patients and 26.0 [22.5 – 29.2] in hospital controls. The indication for the FDG PET-CT in the hospital controls was suspicion of infection (n=53, 61%), malignancy (n=33, 38%) or lymphadenopathy (n=1, 1%). Overall, 2% of the data was missing with a maximum of 11 (12%) missing scores per variable for the ankle. The number of missings per variable are provided in **Supplementary Table 1**. Missing scores were caused by poor image quality or the lack of coverage of a joint in the field of view. Several patients in the control group had a joint prosthesis of the hip (5 prostheses in 4 patients) or knee (3 prostheses in 2 patients). We could stratify 98 patients in the PXE group based on genetic information. Five patients had 2 non-truncating gene variants, 32 patients had one truncating and one non-truncating variants, and 61 patients had two truncating variants in the *ABCC6* gene. The scores for each joint are provided in **Figure 2**. The results of the ordinal logistic regression are shown in **Table 2**.

**Table 1.** Baseline characteristics. Data are provided as median [interquartiles] and n (%). BMI = body mass index. * missing in 7 patients, ** missing in 9 patients.

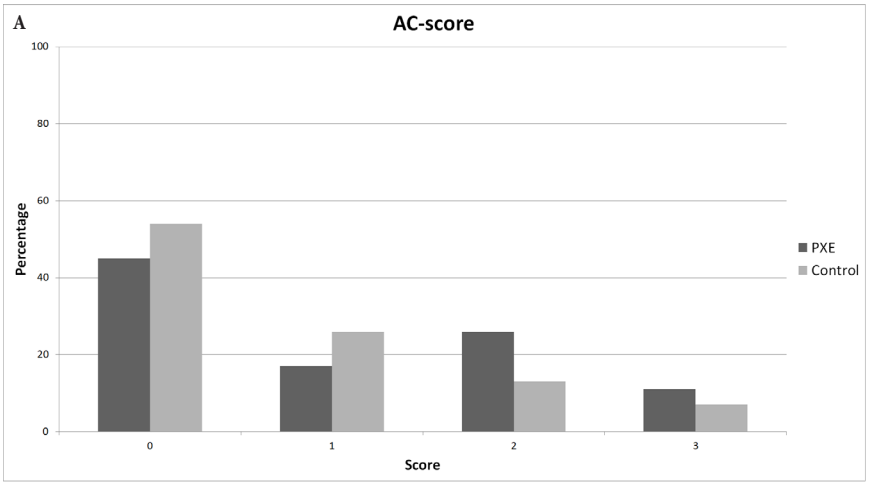
| | Control group (N = 87) | PXE group (N=106) |
|---|---|---|
| Age (years) | 55 [43 - 67] | 56 [48 - 64] |
| Gender (male) | 40 (46%) | 44 (42%) |
| BMI (kg/m2)* | 26.0 [22.5 – 29.2] | 25.3 [22.7 – 28.2] |
| Current smoking** | 16 (20%) | 17 (16%) |
| Joint prosthesis hip | 4 (5%) | 0 (0%) |
| Joint prosthesis knee | 2 (2%) | 0 (0%) |

**Table 2.** Results of ordinal logistic regression analysis with each joint as outcome variable.

| | Crude OR (95% CI) | p-value | Adjusted OR (95% CI) | p-value |
|---|---|---|---|---|
| AC-score | 1.69 [0.99 – 2.90] | 0.054 | 2.00 [1.12 - 3.61] | 0.020 |
| GH-score | 0.83 [0.41 – 1.70] | 0.606 | 1.19 [0.49 - 2.54] | 0.822 |
| Intervertebral disc score | 1.34 [0.80 – 2.26] | 0.265 | 1.61 [0.91 - 2.87] | 0.102 |
| Facet joint score | 1.08 ]0.65 – 1.83] | 0.782 | 1.31 [0.73 - 2.38] | 0.369 |
| Hip score | 0.51 [0.26 – 1.00] | 0.050 | 0.56 [0.27 - 1.15] | 0.117 |
| Tibiofemoral score | 1.83 [1.06 – 3.22] | 0.033 | 2.63 [1.40 - 5.07] | 0.003 |
| Patellofemoral score | 1.70 [0.97 – 3.02] | 0.066 | 2.22 [1.18 - 4.24] | 0.014 |
| Ankle score | 0.68 [0.34 – 1.35] | 0.267 | 0.78 [0.37 - 1.66] | 0.522 |

Crude analysis and analysis adjusted for age, gender, BMI, smoking status (current smoker yes or no) were performed. The hospital controls were used as a reference, therefore an OR > 1 implies that PXE patients have a higher OA score compared to hospital controls. Patients with a joint prosthesis received the highest score possible.
OR Oddsratio, CI Confidence interval, AC Acromioclavicular; GH glenohumeral; OR Odds Ratio; CI Confidence Interval

The test for the proportional odds assumption was always non-significant and therefore the assumptions for performing ordinal logistic regression were not violated. After adjustment, PXE patients were more likely to have a higher OA score for the AC joints (OR 2.00 [1.12 – 3.61]), tibiofemoral joint (OR 2.63 [1.40 – 5.07]), and patellofemoral joint (OR 2.22 [1.18 – 4.24]). Sensitivity analysis after excluding joints with a prosthesis showed the same direction and effect size (Data not shown). No differences were found between males and females. Compared with patients with 2 truncating variants, OA scores for the AC (OR 0.154 [0.033 – 0.712]) and patellofemoral joint (OR 0.137 [0.025 – 0.739]) were lower in patients with 2 non-truncating variants (**supplementary Table 2**). Compared with patients with 2 truncating variants, OA scores for the tibiofemoral joint (OR 0.407 [0.208 – 0.797]) were lower in patients with one non-truncating and one truncating variants.



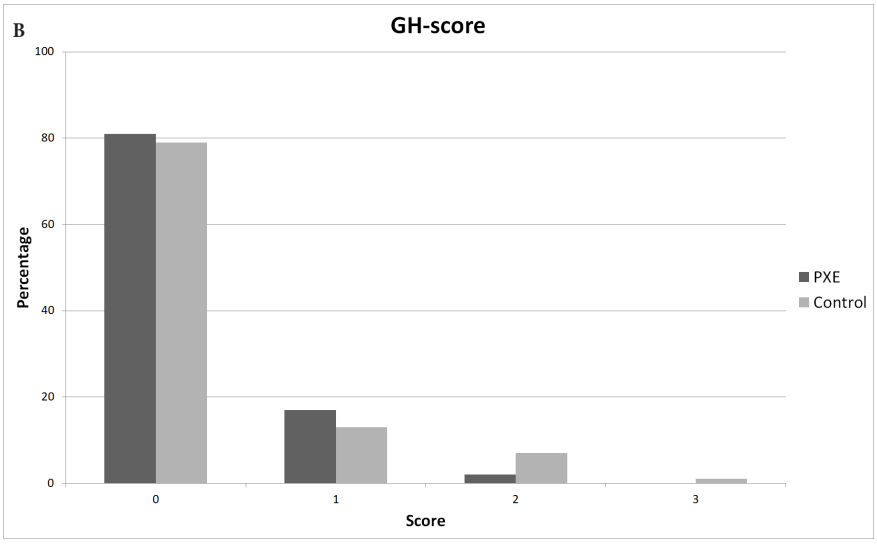
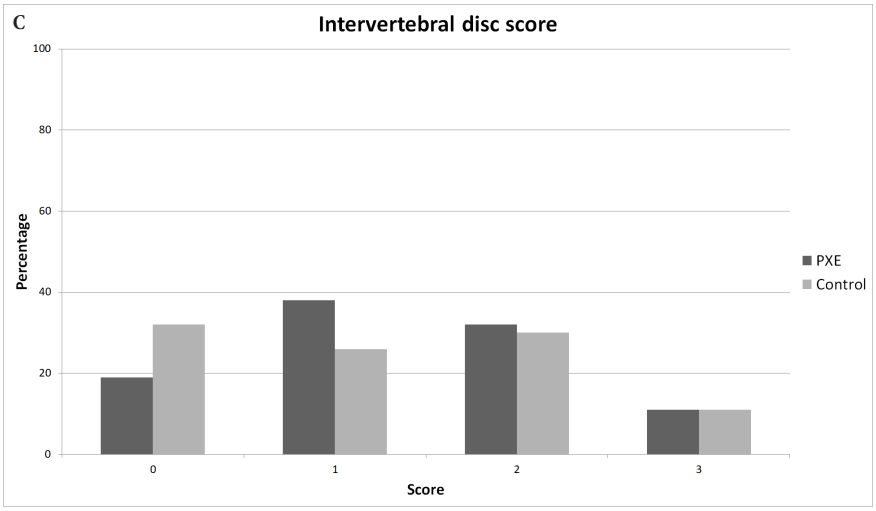




**Figure 2.** Frequencies of osteoarthritis scores for each joint for the PXE and controls. AC Acromioclavicular; GH glenohumeral.



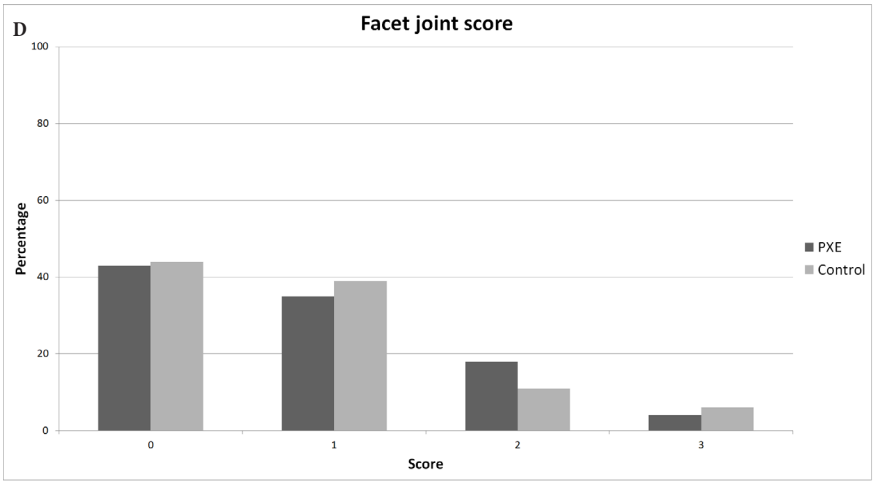
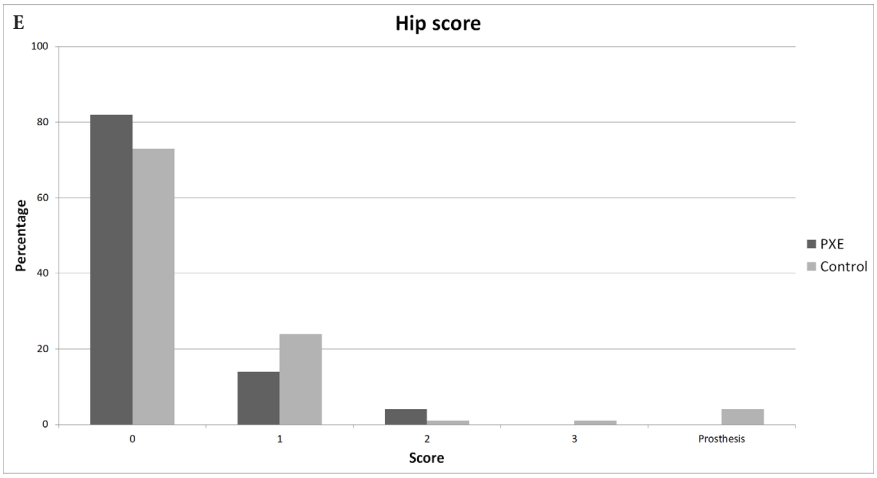




**Figure 2.** Continued I



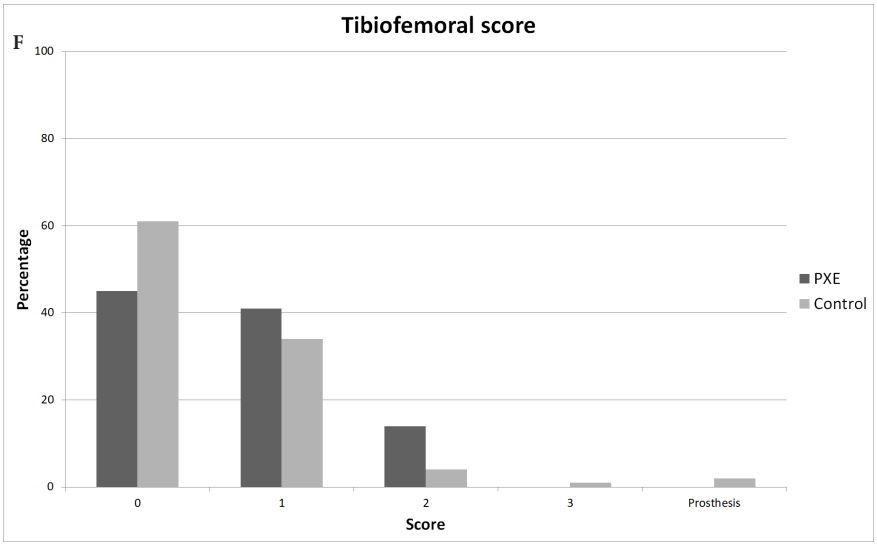
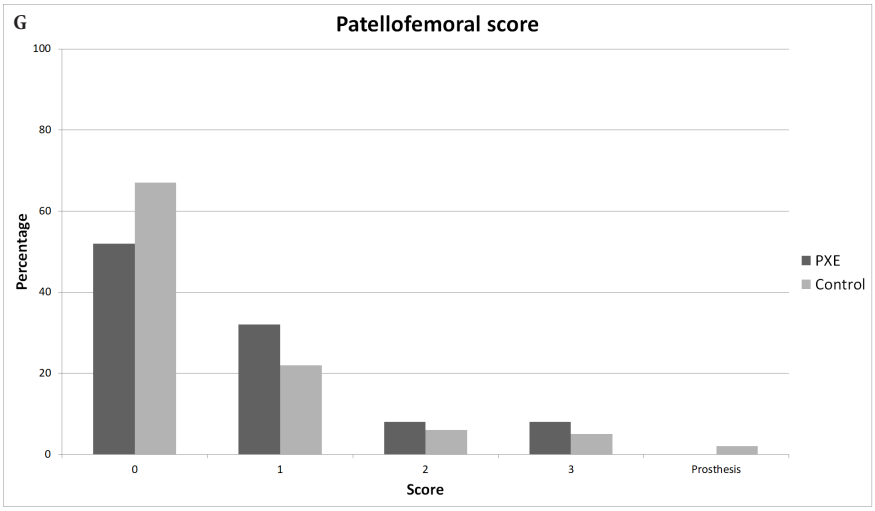




**Figure 2.** Continued II



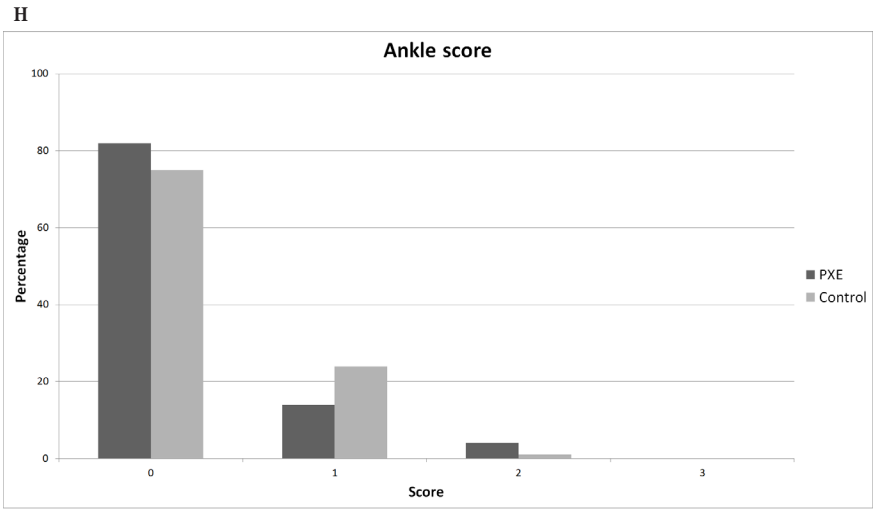




**Figure 2.** Continued III

## Discussion

In this case-control study, we found that patients with PXE, were more likely to have osseous signs of OA in the knee and AC joints compared to hospital controls. Our clinicians noted that PXE patients complained of painful joints more often than expected. This explorative study suggests that the patients may develop OA earlier than expected. In addition, 2 truncating mutations appeared to be associated with a higher prevalence of OA in knees and acromioclavicular joints. However, our observations require prospective validation preferable by clinical assessment and (magnetic resonance) imaging. If confirmed, further investigation into the pathogenesis and possible treatment is needed.

What could be a possible mechanism? PPi is present both intracellular and extracellular. At physiological levels, PPi suppresses HA crystal formation and is a precursor of inorganic phosphate (Pi). In normal physiology, the Pi/PPi ratio is precisely balanced to prevent pathologic calcification. A disbalance in this ratio results in a calcification promoting environment, where low PPi levels result in HA formation, whereas excess PPi results in the formation of calcium pyrophosphate dehydrate (CPPD) crystal deposition.[25] Studies in *ENPP1* null mice and *ank/ank* mice, resulting in low PPi levels, show extensive HA crystal deposits in articular cartilage.[26–28] Although patients with PXE have a different mutation, namely in the *ABCC6* gene, it also results in low PPi blood levels. Bisphosphonates are nonhydrolyzable PPi analogs. A recent randomized controlled trial showed that a bisphosphonate can limit the development of arterial calcification in patients with PXE.[29] Some studies suggest that bisphosphonates are effective in subgroups of patients with OA, specifically in patients with active remodeling of the subchondral bone.[17,30] Whether this effect is mediated through PPi levels or due to other factors such as suppression of bone turnover is unclear. In future research, it should be determined whether the increased OA seen in the current study is attributable to HA crystal formation, and it needs to be confirmed that PXE patients have low PPi levels in synovial fluid.

Another hypothetical mechanism could be related to vascular disease. The incidence of arterial calcifications around knees and hips is related to the incidence of OA in the same joint, although this relationship is only described in women.[31] Previous research in the present population demonstrated that PXE patients have an increased prevalence of vascular calcifications in arm (20% vs 3%), femoro-popliteal (74% vs 44%) and sub-popliteal arteries (84% vs 38%), but not in the vertebral (17% vs 10%) and external iliac arteries (16% vs 30%), when compared to

hospital controls.[21] This would explain that the relationship between PXE and OA in the AC and knee joints, but not in the hip and spine. As glenohumeral and ankle OA have a low prevalence we had little power to detect a difference between groups in the present study. Possibly, vascular disease at a local level is pathophysiological related to OA in PXE patients. For future research, it would be interesting to study this effect using mediation analysis.

Other reasons are also considered. Calcific enthesophytes are related to tendinopathy, but also seen in abundance in PXE patients.[8] Tendinopathy is associated with OA and may be an important mediator for OA in PXE patients.[32] Additionally, it has been suggested that patients with PXE have lower levels of vitamin K.[33] While low vitamin K levels are also associated with an increased prevalence of OA of the hand and knee.[34] Future research should show whether arterial calcifications, enthesophytes, and low vitamin K levels are a risk factor for OA development in PXE patients.

This study has several limitations. First, due to the retrospective nature of the study we did not have clinical data on the symptoms, such as pain and swelling. Future studies should confirm if patients with PXE have more symptomatic OA compared to controls. Second, the patients included in the current study were relatively young with a median age of 56 years, while the incidence of OA is strongly age-related. Therefore, most patients presented with a low OA score. Furthermore, the age range in the hospital control group was larger than in the PXE group, although we corrected for age in the regression analysis. Third, we used CT scans from 2 different vendors, but the resolution and image quality were similar. We believe this did not influence the OA scores. The disease status was not visible for the reader and PET data and DICOM tags were not reviewed. However, as CT scanner brands differed for the PXE and control group, full prevention of observer bias was not guaranteed. Fourth, a population-based sample would be the preferable control group. However, we used hospital controls, as whole body CTs were readily available and harmful radiation for new study participants was thus avoided. Data on comorbidities was lacking, due to limitations posed by the ethical committee to protect the privacy of the participants. hospital controls most likely have more comorbidities compared to population-based controls.[35] As comorbidities are more frequent in people suffering from OA, our control group might have a higher prevalence of OA compared to population-based controls.[36] This would, however, result in a bias towards the null hypothesis, meaning a stronger association between PXE and OA may be observed when PXE patients are compared with population-based controls. Finally, we did not perform a sample size calculation for the main

analyses, as this is a first explorative study. PXE is a rare genetic disorder with an estimated prevalence between 1:25,000 to 1:50,000. This means that there are 350 to 700 PXE patients in the Netherlands.[37] In our opinion, the inclusion and analysis of 106 PXE patients is a substantial number, given the rarity of the disease. When stratifying the PXE patients based on the number of truncating gene variants, only five patients had two non-truncating variants. Therefore, we used the group with two truncating variants as reference group in our analysis. Differences with the group with two non-truncating variants should be interpreted with care. A larger (multicenter) study including a clinical symptoms score is deemed necessary to truly investigate the OA burden in PXE.

## Conclusions

This study suggests that patients with PXE are more likely to have structural OA of the knee and AC-joints. Several explanations for this phenomenon were discussed that may provide new insights for future research in patients with PXE and patients with OA in general.

## Funding

This research was funded by the APPROACH project. APPROACH has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement n°115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme(FP7/2007-2013) and EFPIA companies' in kind contribution. See www.imi.europa.eu.

## Acknowledgments

We like to thank the entire team of the Dutch National health expertise centre for PXE, without their efforts the current research would not have been possible.

## References


1. Jansen RS, Duijst S, Mahakena S, et al. ABCC6-mediated ATP secretion by the liver is the main source of the mineralization inhibitor inorganic pyrophosphate in the systemic circulation - Brief report. Arterioscler Thromb Vasc Biol. 2014;34(9):1985-1989. doi:10.1161/ATVBAHA.114.304017

2. Mebarek S, Hamade E, Thouverey C, et al. Ankylosing spondylitis, late osteoarthritis, vascular calcification, chondrocalcinosis and pseudo gout: toward a possible drug therapy. Curr Med Chem. 2011;18(14):2196-2203. doi:10.2174/092986711795656153

3. Perkins HR, Walker PG. The occurence of pyrophosphate in bone. J Bone Jt Surg Br. 1958;40(B2):333-339.

4. Nitschke Y, Baujat G, Botschen U, et al. Generalized arterial calcification of infancy and pseudoxanthoma elasticum can be caused by mutations in either ENPP1 or ABCC6. Am J Hum Genet. 2012;90(1):25-39. doi:10.1016/j.ajhg.2011.11.020

5. Ichikawa N, Taniguchi A, Kaneko H, et al. Arterial calcification due to deficiency of CD73 (ACDC) as one of rheumatic diseases associated with periarticular calcification. J Clin Rheumatol. 2015;21(4):216-220. doi:10.1097/RHU.0000000000000245

6. Gutierrez LB, Link T, Chaganti K, Motamedi D. Arterial calcification due to CD73 deficiency (ACDC): imaging manifestations of ectopic mineralization. Skeletal Radiol. 2016;45(11):1583-1587. doi:10.1007/s00256-016-2465-9

7. Hilaire CS, Ziegler SG, Markello TC, et al. NT5E mutations and arterial calcifications. N Engl J Med. 2011;364(5):432-442. doi:10.1056/NEJMoa0912923

8. Hoppé E, Etienne M, Legrand E, Rutsch F, Leftheriotis G, Martin L. Prevalence of shoulder calcification in pseudoxanthoma elasticum patients. Jt Bone Spine. 2018;85(6):777-778. doi:10.1016/j.jbspin.2018.03.008

9. Gold SC. Still's Disease with Pseudoxanthoma Elasticum. Proc R Soc Med. 1957;50:473.

10. Sairanen E, Itkonen A, Kangas S. Pseudoxanthoma elasticum (pxe) and joint manifestations. Acta Rheumatol Scand. 1970;16(42008):130-135. doi:10.3109/rhe1.1970.16.issue-1-4.16

11. Biot J, Desbordes P. A. A propos d'un nouveau cas d'elastorrhexie systématisée. Association avec une cervicarthrose et un syndrome de Barré Liéou [Apropos of a new case of pseudoxanthoma elasticum. Association with cervical arthritis and Barre Lieou syndrome]. Med Trop. 1968;28(1):75-84.

12. Van Saase JLCM, Van Romunde LKJ, Cats A, VanDenBroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis. 1989;48(4):271-280. doi:10.1136/ard.48.4.271

13. Bierma-Zeinstra SMA, Waarsing JH. The role of atherosclerosis in osteoarthritis. Best Pract Res Clin Rheumatol. September 2018. doi:10.1016/J.BERH.2018.08.006

14. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthr Cartil. 2005;13(9):769-781. doi:10.1016/j.joca.2005.04.014

15. Corti M, Rigon C. Epidemiology of Osteoarthritis: Prevalence, Risk Factors and Functional Impact. Aging Clin Exp Res. 2003;15(5). doi:10.1007/BF03327356

16. Sandell LJ. Etiology of osteoarthritis: Genetics and synovial joint development. Nat Rev Rheumatol. 2012;8(2):77-89. doi:10.1038/nrrheum.2011.199

17. Laslett LL, Dore D a., Quinn SJ, et al. Zoledronic acid reduces knee pain and bone marrow lesions over 1 year: a randomised controlled trial. Ann Rheum Dis. 2012;71(8):1322-1328. doi:10.1136/annrheumdis-2011-200970

18. Pelletier J-P, Roubille C, Raynauld J-P, et al. Disease-modifying effect of strontium ranelate in a subset of patients from the Phase III knee osteoarthritis study SEKOIA using quantitative MRI: reduction in bone marrow lesions protects against cartilage loss. Ann Rheum Dis. 2015;74(2):422-429. doi:10.1136/annrheumdis-2013-203989

19. Plomp AS, Toonstra J, Bergen a. a B, Van Dijk MR, De Jong PTVM. Proposal for updating the pseudoxanthoma elasticum classification system and a review of the clinical findings. Am J Med Genet Part A. 2010;152(4):1049-1058. doi:10.1002/ajmg.a.33329

20. Plomp AS, Toonstra J, Bergen AAB, Van Dijk MR, De Jong PTVM. Proposal for updating the pseudoxanthoma elasticum classification system and a review of the clinical findings. Am J Med Genet Part A. 2010;152(4):1049-1058. doi:10.1002/ajmg.a.33329

21. Kranenburg G, de Jong PA, Mali WP, Attrach M, Visseren FLJ, Spiering W. Prevalence and severity of arterial calcifications in pseudoxanthoma elasticum (PXE) compared to hospital controls. Novel insights into the vascular phenotype of PXE. Atherosclerosis. 2017;256:7-14. doi:10.1016/J.ATHEROSCLEROSIS.2016.11.012

22. Gielis WP, Weinans H, Nap FJ, Roemer FW, Foppen W. Scoring Osteoarthritis Reliably in Large Joints and the Spine Using Whole-Body CT: OsteoArthritis Computed Tomography-Score (OACT-Score). J Pers Med. 2020;11(1):5. doi:10.3390/jpm11010005

23. Gielis WP, Weinans H, Nap FJ, Roewer FW, de Jong PA, Foppen W. Manuscript currently under review at Insights into Imaging. Please enquire if details on the score are needed.

24. Burns RA, Butterworth P, Kiely KM, et al. Multiple imputation was an efficient method for harmonizing the Mini-Mental State Examination with missing item-level data. J Clin Epidemiol. 2011;64(7):787-793. doi:10.1016/j.jclinepi.2010.10.011

25. Hamilton E, Pattrick M, Doherty M. Inorganic pyrophosphate, nucleoside triphosphate pyrophosphatase, and cartilage fragment in normul human synovial fluid. Rheumatology. 1991;30(4):260-264. doi:10.1093/rheumatology/30.4.260

26. Johnson K, Goding J, Van Etten D, et al. Linked Deficiencies in Extracellular PPi and Osteopontin Mediate Pathologic Calcification Associated With Defective PC-1 and ANK Expression. J Bone Miner Res. 2003;18(6):994-1004. doi:10.1359/jbmr.2003.18.6.994

27. Johnson K, Terkeltaub R. Inorganic pyrophosphate (PPI) in pathologic calcification of articular cartilage. Front Biosci. 2005;10:988-997. doi:10.2741/1593

28. Zhang Y, Brown MA, Peach C, Russell G, Wordsworth BP. Investigation of the role of ENPP1 and TNAP genes in chondrocalcinosis. Rheumatology. 2007;46(4):586-589. doi:10.1093/rheumatology/kel338

29. Kranenburg G, de Jong PA, Bartstra JW, et al. Etidronate for Prevention of Ectopic Mineralization in Patients With Pseudoxanthoma Elasticum. J Am Coll Cardiol. 2018;71(10):1117-1126. doi:10.1016/j.jacc.2017.12.062

30. Xing RL, Zhao LR, Wang PM. Bisphosphonates therapy for osteoarthritis: a meta-analysis of randomized controlled trials. Springerplus. 2016;5(1). doi:10.1186/s40064-016-3359-y

31. Gielis WP, Welsing PMJ, van Spil WE, Runhaar J, Weinans H, de Jong PA. A sex-specific association between incident radiographic osteoarthritis of hip or knee and incident peripheral arterial calcifications: 8-year prospective data from Cohort Hip and Cohort Knee (CHECK). Osteoarthr Cartil. 2017. doi:10.1016/j.joca.2017.07.016

32. Meknas K, Johansen O, Steigen SE, Olsen R, Jørgensen L, Kartus J. Could tendinosis be involved in osteoarthritis? Scand J Med Sci Sports. 2012;22(5):627-634. doi:10.1111/j.1600-0838.2010.01287.x

33. Vanakker OM, Martin L, Schurgers LJ, et al. Low serum vitamin K in PXE results in defective carboxylation of mineralization inhibitors similar to the GGCX mutations in the PXE-like syndrome. Lab Investig. 2010;90(6):895-905. doi:10.1038/labinvest.2010.68

34. Neogi T, Booth SL, Zhang YQ, et al. Low vitamin K status is associated with osteoarthritis in the hand and knee. Arthritis Rheum. 2006;54(4):1255-1261. doi:10.1002/art.21735

35. Neupane B, Walter SD, Krueger P, Loeb M. Community controls were preferred to hospital controls in a case-control study where the cases are derived from the hospital. J Clin Epidemiol. 2010;63(8):926-931. doi:10.1016/j.jclinepi.2009.11.006

36. Suri P, Morgenroth DC, Hunter DJ. Epidemiology of Osteoarthritis and Associated Comorbidities. PM R. 2012;4(5 SUPPL.). doi:10.1016/j.pmrj.2012.01.007

37. Kranenburg G, Baas AF, de Jong PA, Asselbergs FW, Visseren FLJ, Spiering W. The prevalence of pseudoxanthoma elasticum: Revised estimations based on genotyping in a high vascular risk cohort. Eur J Med Genet. 2019;62(2):90-92. doi:10.1016/j.ejmg.2018.05.020

# CHAPTER 13

## Summary and general discussion

## Summary

Osteoarthritis (OA) is a debilitating disease that causes a giant socioeconomic burden. Until now, there is no drug to limit disease progression. There are important steps towards a drug for OA. First, we need to be able to predict in which patients the disease will progress and at what time using accurate prediction models. We may use this knowledge to better identify "early OA" patients and include patients who are expected to progress in clinical trials. Second, we need to select the right patient for the right treatment. This can be achieved by defining robust pheno- or endotypes of OA and tailoring treatments towards specific pathomechanisms within pheno/endotypes. Third, we need a better understanding of pathologic mechanisms in OA to identify treatment targets. Fourth, we need to develop sensitive outcome markers for follow-up. In this thesis, small steps brought us closer to reaching these four goals (**Table 1**).

### Part I Developing new imaging techniques to analyze OA in patients

Malalignment of the lower extremity plays a key role in knee OA. Patients with malalignment of the lower extremity and unicompartmental knee OA may be treated with a correctional osteotomy. Under correction is an important cause for failed outcomes after correctional osteotomies. Alignment is best measured using the Hip-Knee-Ankle angle (HKAA) determined from Whole Leg Radiographs, which is also used to plan correctional osteotomies. Only a small study with 8 participants studied the test-retest reliability of this examination and thus the precision of pre-operative planning is precarious. In **Chapter 2**, we performed an *in vitro* experiment using a sawbone, to test the effect of knee flexion, leg rotation, and X-ray beam height on the measured HKAA. 20 degrees of external leg rotation alone caused an underestimation of one degree in measured HKAA. Flexion and X-ray beam height alone did not alter the measured HKAA. A combination of flexion and rotations strongly effects the HKAA and makes the error difficult to predict. We implemented these findings into a new Whole Leg Radiography protocol, that is easy to apply and aims to minimize errors. In **Chapter 3**, we tested the test-retest reliability of the proposed Whole Leg Radiography protocol in 30 patients. Each patient underwent two radiographs on the same day. Three observers measured the HKAA. Inter-observer reliability for measurements on the same radiograph was excellent with an ICC of 0.982. Reliability between the two separate radiographs was also excellent with an ICC of 0.985. The mean absolute error between two separate radiographs of the same knee was 0.442 degrees, which can be considered clinically acceptable.

**Table 1.** Conclusions

| | |
|---|---|
| | **Part I**<br>*Developing new imaging techniques to analyze osteoarthritis in patients* |
| Chapter 2 | Leg rotation alone alters the measured Hip-Knee-Ankle. Flexion and X-ray beam height alone does not. A combination of flexion and rotations strongly effects the measured Hip-Knee-Ankle angle and makes the error difficult to predict. We implemented these findings into a new Whole Leg Radiography protocol. |
| Chapter 3 | The test-retest reliability of the newly implemented Whole Leg Radiography protocol was excellent with an ICC of 0.985 for the Hip-Knee-Ankle angle. |
| Chapter 4 | We developed an automated image analysis pipeline to measure the Femoral-Tibial angle from a standard knee radiograph. Using this pipeline, we could predict the Hip-Knee-Ankle angle a Pearson correlation of 0.90 and mean absolute error of 1.8 °. |
| Chapter 5 | We developed a reliable scoring system for assessing structural osteoarthritis burden in large joints and the spine and demonstrated its' inter-observer reliability. Furthermore, we presented an atlas to train new readers. |
| | **Part II**<br>*The multifactorial pathways to OA* |
| Chapter 6 | Radiography-based bone texture variables in proximal femur and acetabulum help to predict incident radiographic hip osteoarthritis over a ten-year period. |
| Chapter 7 | We could differentiate between morphological risk factors for OA and morphological variations that result from OA. Morphological variations as cam and pincer lesions may initiate OA. However, the osteoarthritic process may aggravate these lesions, further increasing the risk for OA progression. |
| Chapter 8 | High pelvic incidence is related to incident knee OA and the prevalence of spondylolisthesis. Low pelvic incidence is related to the prevalence of degenerative disc disease and possibly to incident hip OA. |
| Chapter 9 | Sagittal pelvic morphology was not related to hip-joint anatomy and range-of-motion, and the onset of radiographic signs of FAI and hip OA. |
| Chapter 10 | We developed an automated workflow and defined the Shape-Score, a risk score to predict hip OA based on morphology. The Shape-Score demonstrated added predictive value on top of classic predictors for hip OA. |
| Chapter 11 | In women, the incidence of arterial calcifications and incidence of radiographic knee and hip osteoarthritis are related on a local level. In men, the incidence of arterial calcifications and incidence of radiographic knee and hip osteoarthritis are negatively related on a patient-level. |
| Chapter 12 | Patients with Pseudoxanthoma Elasticum are more at risk for developing osseous signs of acromioclavicular and knee OA. |

Many OA cohorts and trials do not incorporate Whole Leg Radiography as it is expensive, requires specialized equipment, and entails additional radiation exposure. Being able to predict the HKAA from standard radiographs would enable researchers to study malalignment in these studies. To achieve this, the Femoral-Tibial angle (FTA) may be used. Many definitions of the FTA exist and it is unknown which definition is best to predict true alignment. In **Chapter 4**, we developed an automated image analysis pipeline to measure the FTA using 9 different definitions. We predicted the HKAA using the automatically derived FTA on 110 pairs of standard knee and Whole Leg radiographs in a cross-validation experiment. The best-performing FTA definition used a femoral axis between the mid-shaft of the femur (approximately 10 cm above the joint line) and the femoral notch, and a tibial axis running through 2 points in the mid-shaft of the tibia (approximately 4 cm and 10 cm beneath the tibial plateau). The Hip-Knee-Ankle angle could be predicted using this pipeline with a Pearson correlation of 0.90 and mean absolute error of 1.8 °. The automated image analysis pipeline may be used in existing and future studies to study alignment.

In OA research, many biomarkers are not joint specific. Examples are biochemical marker levels in blood and urine, performance tests (e.g. timed up and go test), and quality of life measures. The OA status of all joints in the body will have a collective impact on these biomarkers. Until now, no standardized method is available to assess OA throughout the body. In **Chapter 5**, we developed a reliable scoring system for assessing the structural OA burden in large joints and the spine, the OsteoArthritisComputedTomography (OACT) score. We developed the score and an atlas to train new readers using 196 CT scans. The score incorporates the intervertebral discs, glenohumeral, acromioclavicular, facet, hip, tibiofemoral, patellofemoral and ankle joints. Each joint is graded on a scale from no OA (0) to severe OA (3). A total burden of OA was calculated by summing all grades. To test the intra and inter-observer reliability, three readers scored a random set of 25 scans, containing 600 joints. The intraclass correlation coefficient for intra and inter-observer reliability of the total burden of OA was 0.97 (95%-CI, 0.94 to 0.99) and 0.95 (95%-CI, 0.90 to 0.98), respectively. Square weighted kappa's for intra and inter-observer reliability of the individual joint scores ranged from 0.79 to 0.95 and from 0.48 to 0.95, respectively. The OACT-score can be used to define the structural burden of OA in joints throughout the body. The relationship with biomarkers and symptomatic OA should be established in future research, including the APPROACH study. The calculation of the OACT-score using the individual joint scores should be adapted to the purpose of each study.

## Part II The multifactorial pathways to OA

Standard radiography plays a key part in the diagnosis and follow-up of OA. Computational analysis of radiographs may help us to better predict OA. Bone texture analysis was already used to predict incident knee OA and follow the progression of knee OA. In **Chapter 6**, we developed an automated image analysis pipeline to extract bone texture parameters from pelvic radiographs to predict hip OA over 10 years. We used fractal dimension from 41 regions of interest to predict radiographic OA at follow-up, defined as a Kellgren and Lawrence grade (KL) of 2 or higher (definite moderate OA or worse) or joint replacement, in 987 hips with baseline KL 0 and 1 (no or possible OA). The area under the curve (AUC) of the prediction model containing baseline demographics (Sex, age, BMI) and KL showed an AUC of 0.69 in validation. A model with the texture parameters alone, showed an AUC of 0.68. The full model including demographics, KL and texture parameters had an AUC of 0.73 in validation. We concluded that bone texture parameters contain additional information on the prediction of hip OA and may be used in prediction models. However, external validation of our findings should be performed.

The morphology of the hip joint is known to influence the risk for OA. However, the osteoarthritic process may also influence the morphology of the hip. In **Chapter 7**, we aimed to elucidate which morphological variations are pre-existent and a risk factor for OA, and which are a result of OA. We used a 75 landmark statistical shape model (SSM) to annotate pelvic radiographs from five time points during the ten-year follow-up of 1,002 participants. We included shape modes that explained more than 1% of the total shape variance. This led to a total of 12 shape modes. We used the grades for radiographic features of OA (joint space narrowing (JSN) and osteophytes) as described by Altman et al. to define worsening of structural OA. We tested which morphological variations developed during radiographic JSN and osteophytosis and can thus be seen as a result of the disease. Furthermore, we tested which morphological variations develop that predict a period of radiographic JSN and osteophytosis and can thus be seen as a risk factor for the disease. Finally, we included baseline morphology as a predictor in our models. We found that the severity of the cam lesion increased over time, further increasing the risk for OA progression. Additionally, we demonstrated that Pincer morphology has a dynamic nature, as the femoral neck shortens, the head medializes and the acetabular coverage increases, the risk for OA increases and osteophytes grow superiorly and inferiorly as a result. Coxa vara tends to be the result of the OA process opposed to a risk factor for OA. The greater trochanter

seems to shrink as a result of OA, presumably due to disuse of the hip abductor muscles. Furthermore, the lesser trochanter seems to grow as a result of OA, most likely due to psoas activity to compensate for instability These novel insights may help in phenotyping hip OA patients.

The alignment of the spine and pelvis impacts the shear forces in the spine and load distribution through the lower extremity. Mechanical theories suggest that sagittal pelvic morphology is related to common degenerative spine, hip and knee diseases. The pelvic incidence (PI) is a commonly used measure for sagittal spine morphology. The effect of PI on the probability of degenerative lumbar, hip and knee disease was not studied in a large cohort. In **Chapter 8**, we measured the PI from lateral lumbar radiographs of 421 participants and tested it as a predictor for the onset of hip and knee OA and the prevalence of degenerative disc disease (DDD) and spondylolisthesis at eight-year follow-up. The incidence of hip OA showed a trend towards more OA in lower PI groups, but this was not statistically significant. The incidence of knee OA was higher in participants with a high PI compared to a low PI (Odds ratio 1.6, $p = 0.03$). Spondylolisthesis in L4L5 and L5S1 was more frequent in participants with a low PI compared to a high PI (Odds ratios 3.7 and 7.7; $p = 0.02$ and 0.00, respectively). Low PI was a predictor of DDD in L5S1 (Odds ratio 1.9, $p = 0.01$), but not for L4L5. We concluded that high PI is a risk factor for the onset of knee OA and the prevalence spondylolisthesis. Low PI is a risk factor for the presence of DDD and might be linked to the onset of hip OA. In Chapter 7 we demonstrated that cam and pincer morphology are a risk factor for OA. As a result of OA development, the morphology changes and the chance on impingement and subsequent OA progression further increases. The acetabulum is part of the pelvis and the sagittal spinopelvic alignment might be related to the acetabular coverage, which in turn plays a major role in FAI. In **Chapter 9**, we used the same population as in Chapter 8 to analyze the relationship between spinopelvic alignment and radiographic and clinical signs of FAI (alpha, Wiberg angle, and range of internal rotation of the hip). We did not find a significant relation between PI, pelvic tilt, sacral slope and FAI parameters and concluded that sagittal pelvic morphology is not related to FAI.

It is very hard to predict which patients progress to end-stage OA, especially patients with early symptoms. For hip OA, no widely implemented model to predict end-stage hip OA in patients with a first episode of complaints existed. Hip morphology is an important risk factor for hip OA and can be quantified on standard pelvic radiographs, commonly used in the clinical workup of hip symptoms. In **Chapter 10**, we developed an automated workflow to predict hip OA at eight-year follow-

up in patients with a first episode of symptoms around hip and/or knee. We used an SSM similar to Chapter 7 to quantify the baseline morphology of the hip. We used machine learning algorithms to provide a single value, the Shape-Score, that describes the risk for hip OA at eight-year follow-up based on morphology alone. Subsequently, we built and internally validated prediction models by adding a category of variables to the model, starting with baseline demographics, physical examination, radiologist scores, and finally the Shape-Score. The added value of each category was tested using AUC. The prediction model based on demographics, physical examination and radiologist scores predicted incident hip OA, defined as a KL 2+ or joint replacement at follow-up with an AUC of 0.80. When we added the Shape-Score, the AUC rose to 0.86 and we concluded that the Shape-Score may strongly improve predictions for hip OA. However, this model should be externally validated. The generalizability of the model may greatly increase if training data is added from versatile cohort studies.

The risk factors and pathologic mechanisms of atherosclerosis and OA have major overlap. In **Chapter 11**, we studied the simultaneous incidence of arterial calcifications and hip and knee OA. We used baseline and eight-year follow-up radiographs from 763 CHECK cohort participants. We define arterial calcifications on a four-scale grade (no, mild, moderate, severe) and the presence of radiographic OA (KL ≥ 2 or prosthesis). We performed two mixed-effect model analyses. One on joint-level, where the incidence of arterial calcifications was used, and a second on patient-level, where the calcification scores were summed. As the effect of calcifications differed heavily between male and female sex, we stratified the analyses per sex. In women, a significant relationship between the incidence of arterial calcification and OA (Odds ratio 2.51 (1.57–4.03)). On patient-level, the relationship was not significant. In men, the relationship on the local level was not significant. Surprisingly, we found a negative association between arterial calcifications on patient-level and OA in men (Odds ratio 0.71 (0.55–0.92)). The association at the local level in women suggests that local vascular pathology may be a driver of knee and hip OA in women. The association found in men indicates a systemic effect. This might be a true pathophysiologic pathway, for example, a stabilizing effect of calcification on atherosclerotic plaques. Alternatively, this might reflect a health care effect. Men with OA complaints seek medical help more frequently. This potentially triggers secondary CVD risk evaluation and prevention that in turn prevents arterial calcifications.

Pseudoxanthoma Elasticum (PXE) is a systemic disease that affects the skin, eyes, and cardiovascular system. In our hospital, clinicians noted that people with PXE complain about joint pain, more often than expected. As cardiovascular pathology may drive OA, we studied whether PXE patients are more at risk for developing osseous features of OA in **Chapter 12**. Whole-body Low Dose CT scans are part of the routine workup of PXE patients. We used a cohort of hospital controls that underwent PET-CT scans for various indications. We graded DDD and OA in the acromioclavicular, glenohumeral, facet, hip, knee and ankle joints using the score presented in Chapter 5. Group differences were analyzed using ordinal logistic regression adjusted for age, BMI, and smoking status. PXE patients were more likely to have higher OA scores in the acromioclavicular (Odds ratio 2.00 (1.12 – 3.61)), tibiofemoral (Odds ratio 2.63 (1.40 – 5.07)), and patellofemoral joint (Odds ratio 2.22 (1.18 – 4.24)). For other joints, OA scores did not differ significantly. We conclude that PXE patients are more likely to have structural OA of the knee and acromioclavicular joints. Future studies in larger groups are needed to confirm this finding and may further investigate the pathophysiologic mechanism between PXE and OA.

## General discussion and future perspectives

OA is a debilitating disease that poses a major socioeconomic burden on society. No disease modifying drug exists for OA. In the introduction four underlying problems were posed.

### Heterogeneous progression

OA symptoms and structural deterioration may progress slowly or remain stable over a period of years. In such periods, it is not possible to show a treatment effect. Novel biomarkers may help us to better predict methods. In chapter 5 we developed and tested the reliability of the OACT-score. The OACT-score will be an important tool in the search for novel biochemical markers related to the progression of knee OA in the APPROACH cohort. Most research in biochemical markers uses systemic marker levels in either urine or blood. These marker levels are then related to OA status in a single joint or joint group. In all probability, these marker levels will be affected by all joints throughout the body and major confounding lurks. The OACT-score can be used to correct for such confounding and will give future research towards the predictive value of biochemical markers more power. However, as the hands are not included in the score, it would be prudent to add an established radiographic score for hand OA.[1] Furthermore, the correlation between the OACT-score and clinical symptoms should be investigated.

The morphology of the hip joint and pelvis affect the probability for OA as was described in chapters 7, 8, and 10. The morphology of the hip joint may also be affected the osteoarthritic process as we saw in chapter 7. The morphology, but also changes in this morphology, may be used to predict periods of disease progression. In chapter 10, we managed accurate prediction for hip OA using baseline characteristics, physical examination, standard radiologist scores and the newly developed Shape-Score. This finding needs external validation, particularly because we included mainly Caucasian women. Furthermore, shape modes defined by the SSM are population specific and comparison to other studies is difficult. The World COACH initiative is striving to overcome this problem by incorporating multiple cohorts over the world and form a standardized SSM. This would provide similar and thus comparable shape modes over the cohorts and a possibility to validate models using different population distributions over the world. In chapter 10, we predicted hip OA at eight-year follow-up as we needed enough cases while using a fairly insensitive outcome measure, KL ≥2. An eight-year medical trial would be very expensive and ensures a long time between drug development and market launch. This brings us to the following problem.

**Insensitive outcome markers**

At the moment there is no efficient outcome marker to test the treatment effect of potential drugs. Joint arthroplasty can be seen as a hard outcome, as the structural degradation and symptoms have reached such a level that the joint should be replaced. However, the choice to replace a joint is heavily reliant on both patient and surgeon. Life events or job commitments may postpone surgery from the patients' side. Comorbidities or lifestyle choices such as smoking may drive a surgeon to postpone surgery. The FDA and EMA only accepted joint space narrowing measured from standard radiographs, which is notoriously insensitive to change. As a consequence, the FDA and EMA request a trial duration of at least two years. While they acknowledge MRI and biochemical markers to be promising, they request further research on the correlation with symptoms and radiographic progression. The APPROACH project includes many novel and established biomarkers determined on baseline and six-month follow-up. Changes in these biomarkers will be correlated to joint space narrowing over a two-year follow-up. Hopefully this brings forward new surrogate outcome markers, leading to shorter and powerful trials.

## Knowledge gap

It became clear that OA is not just wear-and-tear. OA is a heterogeneous disease with many pathological pathways and stages. By broadening our knowledge, we may discover new treatment targets. We may better stratify patients and disease stages into pheno- and endotypes, helping us to select the right time and patient for the right treatment. Pathologic pathways and risk factors for cardiovascular disease and OA heavily overlap. Oddly enough, research in this area is sparse. In chapter 11, we studied the coincidence of arterial calcification and hip and knee OA. This relation was heavily sex dependent. In women, the incidence of calcifications around a joint increased the risk for OA in the same joint. In men, the severity of calcifications on a patient-level was negatively correlated to the risk for OA in hips and knees. As the study population was 45-65 years old during inclusion, this difference may be explained by perimenopausal hormone level changes.[2] Furthermore, the female vasculature generally has a smaller diameter and atherosclerosis is more widespread.[3,4] When the microvasculature fails or thrombi detach from the vascular wall, the joint may become hypovascularized causing tissue damage and subsequent OA. In men, the vasculature has a larger diameter and atherosclerosis forms larger focal plaques. Calcification is a sign of plaque formation, but may also stabilize the plaques, creating a protective effect against thrombi.[5] Finally, the protective effect in men may be explained by a healthcare effect. Men with OA complaints are more prone to cardiovascular risk profiling and prevention, compared to men without OA symptoms, but also women with OA symptoms. This may prevent atherosclerosis to progress to calcifications in men that develop symptomatic end-stage OA.

As stated earlier, the risk factors and pathophysiologic mechanisms of cardiovascular disease and OA strongly overlap. Furthermore, both diseases progress slowly and simultaneously. Furthermore, innovative research methods are needed that deal with confounding and mediation. PXE is a systemic disease caused by mutations in the *ABCC6* gene, affecting the skin, eyes and cardiovascular system. The main cardiovascular symptom is calcifications in the arterial wall caused by dysregulations in inorganic pyrophosphate (PPi) levels. In the UMC Utrecht, many PXE patients complained about joint pain. As we found a relation between arterial calcifications and OA and as OA is the main cause for joint pain, we performed an explorative study to test whether PXE patients truly had more structural features of OA compared to hospital controls. We found that PXE patients had OA more often in the acromioclavicular and knee joints. Future research should validate this finding, but also focus on pathological pathways such as PPi and vitamin-K levels, cytokines affecting bone remodeling, subchondral bone integrity and

arterial calcifications. This is just one example of a research line that may guide us towards new endotypes in OA. With novel techniques arising in rapid tempo, the possibilities to study pathological pathways and potential treatment targets emerge. These novel techniques should be combined with established techniques in large cohorts that focus on including patients that are likely to progress.

### One-size-fits-all

Many trials that test the effect of a drug for OA have uniform inclusion criteria, for example the ACR criteria for knee OA. However, due to a lack of definitions and selection criteria for both pheno- and endotypes, it is difficult for researchers to select the right patient for their treatment. Again, large cohorts that combine established with novel biomarkers and contain a decent amount of progressors are needed to stratify patients into these phenotypes. Finding phenotypes of OA is also one of the main goals of the APPROACH project. However, we should not stop testing potential treatments for OA until phenotypical definitions are clear cut. Researchers should already adjust their inclusion criteria for trials towards the hypothesized mechanism of action. An example of this is the ZOledronic acid as DIseAse-modifying drug in Knee OA (ZODIAK) study. This study hypothesizes that bisphosphonates may counter the loss of integrity of the subchondral bone. OA causes the porosity of the subchondral bone to increase, nerves and bloodvessels may grow towards the cartilage, providing a route for pain stimuli and inflammatory cytokines. Counteracting this process by disrupting osteoclast activity may prevent pain and structural degradation of cartilage. To select patients with active subchondral bone remodeling, the ZODIAK study screens patients for bone marrow lesions (BML), as they are related to bone remodeling.[6] By excluding patients without BML the researchers hope to exclude patients that lack the intended treatment target, leading to a more efficient trial design.

### The future of OA

A disease modifying drug for OA will not unemploy orthopedic surgeons. The first goal is to postpone joint arthroplasty and delay or even prevent revision surgery. Future research in OA should focus on getting the right treatment to the right patient at the right time. This requires large cohort studies to define endotypes of OA, selection criteria for these endotypes and criteria to define successful treatments within a specific endotype. Until we reach that point, potential treatments may be tested, but researchers should select their patients carefully. Furthermore, when blueprints emerge for endotype specific trials, we should consider retesting drugs that failed due to bad trial design.

## References

1. Visser AW, Bøyesen P, Haugen IK, Schoones JW, van der Heijde DM, Rosendaal FR, et al. Radiographic scoring methods in hand osteoarthritis - a systematic literature search and descriptive review. Osteoarthr Cartil. 2014;22(10):1710–23.

2. Fairweather D. Sex differences in inflammation during atherosclerosis. Clin Med Insights Cardiol [Internet]. 2014 [cited 2016 Sep 19];8(Suppl 3):49–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25983559

3. Shaw LJ, Bugiardini R, Merz CNB. Women and Ischemic Heart Disease. Evolving Knowledge. J Am Coll Cardiol [Internet]. 2009;54(17):1561–75. Available from: http://dx.doi.org/10.1016/j.jacc.2009.04.098

4. Hirsch AT, Allison M a., Gomes AS, Corriere M a., Duval S, Ershow AG, et al. A call to action: Women and peripheral artery disease: A scientific statement from the american heart association. Circulation. 2012;125(11):1449–72.

5. Wong KK, Thavornpattanapong P, Cheung SC, Sun Z, Tu J, O'Rourke R, et al. Effect of calcification on the mechanical stability of plaque based on a three-dimensional carotid bifurcation model. BMC Cardiovasc Disord [Internet]. 2012 Dec 15 [cited 2016 Sep 19];12(1):7. Available from: http://bmccardiovascdisord.biomedcentral.com/articles/10.1186/1471-2261-12-7

6. Donell S. Subchondral bone remodelling in osteoarthritis. EFORT Open Rev [Internet]. 2019 Jun 1 [cited 2020 Sep 4];4(6):221–9. Available from: /pmc/articles/PMC6549114/?report=abstract

# CHAPTER 13

## Nederlandse samenvatting

## Artrose, het plaatje snappen

Artrose maakt patiënten minder mobiel, geeft pijn, verminderde kwaliteit van leven en veroorzaakt een grote sociaal-economische last. Er is momenteel geen medicijn dat de ziekteprogressie kan remmen. Om tot zo'n medicijn te komen moeten we een aantal drempels overbruggen. Allereerst moeten we de juiste patiënt kunnen identificeren. Aangezien artrose in een vergaand stadium vermoedelijk niet meer terug te draaien is moeten we het liefst al in een vroeg stadium kunnen voorspellen of iemand eindstadium artrose zal krijgen. Nauwkeurige voorspelmodellen kunnen hier een grote rol in spelen, door bijvoorbeeld 'vroege artrose'-patiënten te identificeren. Daarnaast moeten we de juiste patiënt selecteren voor de juiste behandeling. Hiervoor dienen feno- en/of endotyperingen van artrose te worden gedefinieerd. Per patiënt moet de behandeling worden afgestemd op de specifieke pathomechanismen binnen zijn/haar feno- en/of endotype. Tenslotte moeten we pathologische mechanismen binnen artrose beter in kaart brengen om behandeldoelen te vast te stellen. Om al deze doelen te bereiken moeten er gevoelige uitkomstmarkers worden ontwikkelen voor follow-up van artrose. In dit proefschrift komen we met kleine stappen dichter bij deze doelen.

### Deel I De ontwikkeling van nieuwe beeldvormingstechnieken voor artrose

Het alignement van een been speelt een belangrijke rol in knieartrose. Mensen bij wie de knie uit het lood staat hebben vaak last van artrose in één kniecompartiment. Deze patiënten kunnen behandeld worden middels een correctie-osteotomie. Een te kleine correctie is de belangrijkste oorzaak voor suboptimale resultaten na zo'n ingreep. De heup-knie-enkelhoek (HKAA), gemeten van lange beenas-opnames, is de belangrijkste maat om het alignement van een knie te bepalen en wordt gebruikt om correctie-osteotomieën te plannen. De betrouwbaarheid van de HKAA is slecht onderzocht. Tot nu toe is er slechts één kleine studie met 8 mensen is hiervoor uitgevoerd.[1] In **Hoofdstuk 2** hebben we een experiment uitgevoerd met een kunstbeen om de meetfout veroorzaakt door flexie, rotatie en röntgenbundelhoogte op de gemeten HKAA te testen. Bij 20 graden exorotatie wordt een onderschatting van één graad in gemeten HKAA veroorzaakt. Flexie en röntgenbundelhoogte op zichzelf zorgden niet voor een meetfout. Echter, een combinatie van flexie en rotaties zorgt voor een onvoorspelbare meetfout. De bevindingen van deze studie hebben we geïmplementeerd in een nieuw protocol voor lange beenas-opnames, dat eenvoudig is toe te passen en zich richt op het minimaliseren van meetfouten. In **Hoofdstuk 3** hebben we de betrouwbaarheid van dit nieuwe protocol bij 30 patiënten getest. Elke patiënt onderging op dezelfde dag twee lange beenas-

opnames. De HKAA werd door drie waarnemers gemeten. De betrouwbaarheid tussen waarnemers was uitstekend met een 'Intra Class Correlatiecoëfficiënt (ICC) van 0,982. De betrouwbaarheid tussen twee afzonderlijke lange beenas-opnames was ook uitstekend met een ICC van 0,985. De gemiddelde absolute fout tussen twee aparte lange beenas-opnames van dezelfde knie was 0,442 graden, wat klinisch aanvaardbaar is.

Binnen artroseonderzoek worden lange beenas-opnames vaak niet uitgevoerd omdat ze duur zijn, speciale apparatuur vereisen en extra blootstelling geven aan schadelijke röntgenstraling. Als de HKAA op basis van standaard knie röntgenfoto's voorspeld kan worden, zouden onderzoekers de rol van alignement binnen artrose makkelijker kunnen onderzoeken. De femoro-tibiale hoek (FTA) kan gebruikt worden om de HKAA te voorspellen. Er bestaan echter veel verschillende definities van de FTA en het is niet goed onderzocht welke definitie de meest accurate voorspelling geeft. In **Hoofdstuk 4** hebben we de FTA gemeten met behulp van zelfontworpen automatische beeldanalyse software. We testten de voorspellende waarde van 9 verschillende definities van de FTA op een set van 110 paar röntgenfoto's (standaard knie- en lange beenasfoto's) in een kruisvalidatie-experiment. De best presterende FTA werd gedefinieerd door een femorale as vanuit de schacht van het femur (ongeveer 10 cm boven de gewrichtsspleet) naar de femorale inkeping, en een tibiale as die midden door de schacht van de tibia loopt (tussen twee punten middenin het tibia op ongeveer 4 cm en 10 cm onder de gewrichtsspleet). Met behulp van deze FTA konden we de HKAA voorspellen met een Pearson-correlatie coëfficiënt van 0,90 en een gemiddelde absolute fout van 1,8°. De geautomatiseerde beeldanalysesoftware kan worden gebruikt in bestaande en toekomstige cohorten om de rol van alignement binnen artrose te bestuderen.

Onderzoek naar artrose maakt vaak gebruikt van biomarkers die niet specifiek gekoppeld zijn aan veranderingen binnen één gewricht. Voorbeelden zijn biochemische markers in bloed en urine, prestatietests (zoals de timed up and go-test) en kwaliteit van leven. Deze biomarkers worden beïnvloed door de artrose-status van alle gewrichten in het lichaam. Tot nu toe is er geen gestandaardiseerde methode beschikbaar om de artrose-status door het hele lichaam te meten. In **Hoofdstuk 5** hebben we een betrouwbaar scoresysteem ontwikkeld voor structurele artroseveranderingen in grote gewrichten en de wervelkolom, de zogenaamde OsteoArthritisComputedTomography (OACT)-score. De score omvat de tussenwervelschijven, glenohumerale, acromioclaviculaire, facet-, heup-, tibiofemorale, patellofemorale en enkelgewrichten. Elk gewricht wordt gescoord op een schaal van geen artrose (0) tot ernstige artrose (3). We hebben een atlas

ontwikkeld, op basis van 196 CT-scans, om nieuwe lezers te trainen. Om de intra- en interobserver betrouwbaarheid te testen, scoorden drie beoordelaars een willekeurig gekozen set van 25 scans met totaal 600 gewrichten. De totale artrosebelasting werd berekend door de score van alle gewrichten bij elkaar op te tellen. De ICC voor intra-observer betrouwbaarheid van de totale artrosebelasting was 0,97 (95% -CI, 0,94 tot 0,99), de inter-observer betrouwbaarheid was 0,95 (95% -CI, 0,90 tot 0,98). Kwadraat gewogen kappa's voor intra-observer betrouwbaarheid van de individuele gewrichtsscores varieerde tussen 0,79 en 0,95. De inter-observer betrouwbaarheid varieerde tussen 0,48 en 0,95. Met deze bevindingen kunnen we stellen dat de OACT-score kan worden gebruikt om de artrose-status van grote gewrichten door het hele lichaam te meten. De relatie met biomarkers en symptomatische artrose moet worden vastgesteld in toekomstig onderzoek, waaronder de APPROACH-studie. Aangezien deze laatste studie zich richt op de knie, is het hierbij van belang te weten of de systemische biochemische markers een representatie zijn van artrose in de knie, of wellicht meer iets zeggen over artrose in andere gewrichten.

### Deel II Er zijn meerdere wegen die naar artrose leiden

Röntgenfoto's staan aan de basis van de diagnostiek en follow-up onderzoek van artrose. Computergestuurde analyse van deze röntgenfoto's zou ons kunnen helpen om artrose beter te volgen en mogelijke progressie voorspellen. De bottextuur is een parameter die eerder gebruikt is voor het voorspellen en vervolgen van knieartrose. In **Hoofdstuk 6**, ontwikkelden we geautomatiseerd software om bottextuur parameters uit röntgenfoto's van de heup te halen. We gebruikten fractale dimensies uit 41 regio's om incidentele heupartrose na tien jaar te voorspellen. Heupartrose definieerden we als een Kellgren en Lawrence score (KL) van 2 of hoger, of een heupvervanging met een prothese. We gebruikten 987 heupen met een KL van 0 of 1 op baseline. De "area-under-the-curve"(AUC) van een model met alleen demografische informatie en de baseline KL was 0.69. Een model met alleen bottextuur parameters toonde een AUC van 0.68. Een model dat de bottextuur parameters combineerde met het eerste model haalde een AUC van 0.73. We concludeerden daarom dat bot textuurparameters enige extra informatie bieden bij het voorspellen van heupartrose. Deze bevinding dient echter in een externe validatie bevestigd te worden.

De morfologie van de heup kan het risico op heupartrose beïnvloeden. Op zijn beurt kan het artrotisch proces ook de vorm van de heup aantasten. In **Hoofdstuk 7** poogden wij op te helderen welke morfologische varianten pre-existent zijn aan

heupartrose en een risico factor vormen, en welke morfologische variantie juist resultaat zijn van de ziekte. We gebruikten een statistical shape model (SSM) op basis van 75 anatomische oriëntatiepunten op bekkenfoto's en hadden hiervoor de data van 1,002 mensen, van vijf tijdspunten over een periode van tien jaar verspreid beschikbaar. Vanuit het SSM gebruikten we alle vormvariabelen die meer dan 1% van de totale vormvariatie verklaren. Het voortschrijden van artrose definieerden we op basis van de criteria voor gewrichtsspleetversmalling en osteofytose zoals gesteld door Altman et al. We onderzochten welke vormvariaties ontstonden tijdens het voortschrijden van de ziekte en dus als resultaat van artrose bestempeld konden worden. Daarnaast testten we welke vormvariaties aanwezig waren voorafgaand aan een periode van voortschrijdende ziekte, en dus een risicofactor voor artrose vormen. Als laatste voegden we de baseline vormen toe aan de modellen als voorspeller voor artrose. We vonden dat zogenaamde cam-laesies over tijd verergerden, en een zichzelf versterkende risicofactor is voor artrose. Daarnaast is de zogenaamde pincer morfologie een dynamisch proces, waarbij de femur nek korter wordt, de femur kop medialiseert en de acetabulaire overkapping toeneemt. Daarmee neemt het risico op artrose toe en groeien acetabulaire osteofyten. Verder lijkt Coxa vara eerder het resultaat van artrose, dan een risico factor. De trochanter major krimpt als een gevolg van artrose, waarschijnlijk door het ontlasten van de heup abductoren. De trochanter minor lijkt juist te groeien door artrose, waarschijnlijk door toegenomen acitivieit van de psoas major spier als compensatie voor instabiliteit veroorzaakt door het artrotisch proces. Deze inzichten in de relatie tussen morfologie en artrose kunnen ons helpen bij de phenotypering van patiënten met heupartrose.

De uitlijning tussen de rug en het bekken beïnvloedt hoe belasting door de rug en naar de benen vloeit. Mechanische theorieën relateren de sagittale bekkenmorfologie aan veelvoorkomende rug-, heup- en knieziekten. De pelvic incidence (PI) is een maat voor sagittale bekkenmorfologie. Het effect van PI op het risico voor degeneratieve ziekten van rug, knie en heup werd nog niet eerder beschreven. In **Hoofdstuk 8**, hebben we de PI gemeten op laterale rugfoto's van 421 patiënten en gebruikten dit als voorspeller voor de prevalentie van een degeneratieve discus (DDD) en spondylolisthesis en als voorspeller voor de incidentie van heup- en knieartrose na acht jaar follow-up. De incidentie van heupartrose leek hoger in de groep met een lage PI, maar dat bleek niet statistisch significant. De incidentie van knieartrose was hoger in deelnemers met een hoge PI vergeleken met een lage PI (Odds ratio 1.6, p = 0.03). Spondylolisthesis in L4L5 en L5S1 werd meer gezien in deelnemers met een lage PI vergeleken met een hoge PI (respectievelijke odds

ratios 3.7 en 7.7; p = 0.02 en 0.00). Lage PI was een voorspeller van DDD in L5S1 (odds ratio 1.9, p = 0.001), maar niet voor L4L5. We concludeerden dat hoge PI een risicofactor is voor het ontwikkelen van knieartrose en de prevalentie van spondylolisthesis. Lage PI was een risicofactor voor de aanwezigheid van DDD en is mogelijk gerelateerd aan heupartrose. In hoofdstuk 7 toonden we aan dat cam- en pincermorfologie een risicofactor vormen voor artrose. Het acetabulum is onderdeel van het bekken en de sagittale uitlijning van de rug en het bekken speelt mogelijk een rol in de acetabulaire overkapping, die op zijn beurt weer een rol speelt in heupimpingement (Femoral Acetabular Impingement, FAI). In **Hoofdstuk 9**, gebruikten we dezelfde populatie als in hoofdstuk 8, om de relatie tussen de sagittale uitlijning van de rug en het bekken enerzijds en radiografische en klinische kenmerken van FAI anderzijds, te onderzoeken. We vonden geen significant verband tussen PI, kanteling van het bekken, de lumbo-sacrale hoek en parameters die de FAI beschrijven. We concludeerden dat de sagittale bekkenmorfologie geen verband houdt met FAI.

Het is zeer moeilijk te voorspellen welke patiënten eindstadium artrose ontwikkelen, zeker bij patiënten met alleen vroege symptomen. Er is momenteel geen breed geïmplementeerd predictiemodel voor heupartrose in mensen met vroege symptomen. De morfologie van de heup speelt een grote rol in de ontwikkeling van heupartrose. Morfologie is te kwantificeren op röntgenfoto's, die onderdeel maken van de standaard work-up van heupartrose. In **Hoofdstuk 10**, ontwikkelden we een geautomatiseerde methode om, op basis van morfologische kenmerken, het ontstaan van heupartrose na acht jaar te voorspellen in patiënten met een eerste episode van klachten aan heup en/of knie. We gebruikten wederom een SSM, overeenkomend met die in hoofdstuk 7, om de 'baseline' morfologie van de heup te kwantificeren. Met machine learning algoritmes konden we een zogenaamde Shape-Score berekenen, één enkele waarde die het risico voor heupartrose beschrijft op basis van morfologie bepaald d.m.v de vormparameters in het SSM. Vervolgens testten we predictiemodellen op basis van interne validatie, d.w.z. validatie op basis van experimenten in dezelfde dataset. Allereerst maakten een predictiemodel op basis van alleen demografische variabelen. Achtervolgens voegden we het lichamelijk onderzoek, radiologische scores en als laatste de Shape-Score toe aan het predictiemodel. De toegevoegde waarde van iedere groep variabelen werd getoetst middels de AUC. Het predictiemodel gebaseerd op demografie, lichamelijk onderzoek en radiologische scores voorspelde artrose, gedefinieerd als een KL 2 of hoger of een heupprothese na acht jaar, met een AUC van 0.80. Na het toevoegen van de Shape-Score steeg de AUC naar 0.86. We

concluderen derhalve dat de Shape-Score een sterk toegevoegde waarde heeft in het voorspellen van heupartrose. Deze bevinding dient echter extern gevalideerd te worden. Daarnaast zal de generaliseerbaarheid van het model sterk verbeteren als data van meerdere cohorten wordt toegevoegd.

De risicofactoren en pathofysiologie binnen atherosclerose en artrose overlappen sterk. In **Hoofdstuk 11**, bestudeerden we de simultane incidentie van arteriële verkalking en artrose van de heup en knie. We gebruikten hiervoor 'baseline' en acht jaar follow-up röntgenfoto's van 763 deelnemers van CHECK. Per gewricht scoorden we arteriële verkalking in vier gradaties (geen, mild, matig, ernstig) en de aanwezigheid van artrose (KL 2 of hoger of een prothese). Er werden twee 'mixed-effects' modellen gemaakt. Eén op gewrichtsniveau, waarbij de incidentie van verkalking in een gewricht werd gebruikt en één op patiëntniveau waarbij de somscore van de verkalkingen bij beide heupen en knieën werd gebruikt. De relatie tussen de mate van verkalking en het al dan niet ontwikkelen van artrose verschilden sterk per geslacht, waarop we de analyses stratificeerden. Bij vrouwen werd een significante relatie gevonden tussen de incidentie van vaatverkalking rond een gewricht en het ontstaan van artrose in datzelfde gewricht (Odds ratio 2.51 (1.57-4.03)). Op patiëntniveau was de relatie niet significant. Bij mannen, was de relatie op gewrichtsniveau niet significant. Echter, op lichaamsniveau vonden we een negatieve associatie tussen de mate van de somscore voor vaatverkalking en artrose (Odds ratio 0.71 (0.55-0.92)). De associatie op gewrichtsniveau gevonden in vrouwen suggereert dat vasculaire pathologie op lokaal niveau, heup- en knieartrose in de hand kan werken. De associatie bij mannen, suggereert een systemisch effect. Dit kan berusten op een waar pathofysiologisch mechanisme, zoals het stabiliserende effect van verkalking op atherosclerotische plaques. Echter kan het ook een zorgeffect weerspiegelen. Mannen met artroseklachten komen wellicht eerder bij de dokter, welke cardiovasculaire risicoanalyse kan inzetten en in een vroeg stadium preventieve medicatie voorschrijven, waardoor vaatverkalking voorkomen wordt.

Pseudoxanthoma Elasticum (PXE) is een systemische genetische ziekte waarbij kalk zich opstapelt in bindweefsel waardoor er uiteindelijk er schade optreedt aan o.a. huid, ogen, en het cardiovasculair systeem. In ons ziekenhuis merkten doctoren op de dat mensen met PXE vaak klagen over pijnlijke gewrichten. Gezien cardiovasculaire pathologie en artrose aan elkaar gerelateerd zijn, bestudeerden we in **Hoofdstuk 12** of mensen met PXE meer risico lopen op het ontwikkelen van benige kernmerken van artrose. Whole-body Low Dose CT scans zijn deel van de standaard work-up van PXE-patiënten. We gebruikten een cohort met

controle-patiënten die voor wisselende indicaties PET-CTs ondergingen in het ziekenhuis. We gebruikten de OACT-score beschreven in hoofdstuk 5 om DDD en artrose in de acromioclaviculaire, glenohumerale, facet, heup, knie en enkelgewrichten te scoren. Verschillen tussen PXE en controle-patiënten werden getest middels ordinale logistische regressie, gecorrigeerd voor leeftijd, geslacht, BMI en roken. PXE-patiënten hadden meer kans op hogere artrosescores in de acromioclaviculaire (Odds ratio 2.00 (1.12 – 3.61)), tibiofemorale (Odds ratio 2.63 (1.40 – 5.07)) en patellofemorale gewrichten (Odds ratio 2.22 (1.18 – 4.24). In andere gewrichten verschilden artrosescores niet significant. We concludeerden dat PXE-patiënten een grotere kans hebben op artrose in de acromioclaviculaire en kniegewrichten vergeleken met controle-patiënten. Deze bevindingen moeten in de toekomst gevalideerd worden. Daarnaast moeten de pathofysiologische relaties tussen PXE en artrose verder onderzocht worden.

## Discussie en toekomstperspectief

Artrose is een slopende ziekte voor patiënten. Voor de samenleving brengt het een enorme socio-economische last met zich mee. Op dit moment is er geen medicijn om ziekteprogressie te remmen, laat staan te stoppen. Historisch gezien falen onderzoeken naar een behandeling voor artrose door vier problemen. In dit proefschrift hebben we aan de weg naar een oplossing voor deze vier problemen gewerkt.

### Heterogene progressie

Het eerste probleem is het wisselvallige beloop van de progressie van artrose. Symptomen en structurele kenmerken van de ziekte verergeren langzaam over de jaren. Sommige patiënten maken een korte periode van versnelde progressie mee, maar het is met de huidige stand van kennis niet mogelijk zo'n periode te voorspellen. Omdat een periode van progressie meestal niet wordt 'gevangen' binnen een gekaderd onderzoek, is het lastig een potentieel behandel effect aan te tonen. In hoofdstuk 5 ontwikkelden we de OACT-score en testen de betrouwbaarheid ervan. De OACT-score zal binnen de APPROACH studie een belangrijke rol spelen in de zoektocht naar nieuwe biochemische markers die zijn gerelateerd aan de progressie van knieartrose. Veel biochemische markers worden namelijk gemeten in urine of bloed en vervolgens geassocieerd met artrose in één gewricht(sgroep). Echter het markerniveau wordt hoogstwaarschijnlijk beïnvloed door alle gewrichten in het lichaam van de patiënt, wat een bron kan zijn voor 'confounding'. De OACT-score kan corrigeren voor artrosebelasting door het hele

lichaam en zal onderzoek naar biochemische markers daarmee preciezer en sterker maken. De handen zijn helaas niet in de OACT-score meegenomen en zullen via één van de beschikbare gevalideerde methoden voor het scoren van handartrose moeten worden toegevoegd.[2] Daarnaast moet de relatie tussen de OACT-score en symptomen nog onderzocht worden.

In hoofdstukken 7, 8 en 10 beschreven we dat de morfologie van de heup en het bekken de kans op artrose beïnvloedt. In hoofdstuk 7 lieten we zien dat het artrotisch proces op zijn beurt ook weer de morfologie beïnvloedt. De morfologie, maar ook veranderingen in die morfologie kunnen gebruikt worden om artroseprogressie te voorspellen. In hoofdstuk 10 bereikten we een accurate voorspelling van heupartrose op basis van demografie, lichamelijk onderzoek, bestaande radiologische scores en de nieuwe ontwikkelde 'Shape-Score'. Onze bevindingen dienen nog extern gevalideerd te worden, zeker gezien het feit dat de studiepopulatie voornamelijk bestond uit Kaukasische vrouwen. Daarnaast is het vergelijken van 'shape modes' tussen verschillende onderzoeksresultaten niet mogelijk, aangezien deze modes afhankelijk zijn van het gebruikte model en dus studie-specifiek zijn. Het World COACH initiatief streeft er naar dit probleem op te lossen door meerdere internationale cohorten samen te brengen en een gestandaardiseerd model over alle data te laten draaien. Dit zou zorgen dat de data tussen verschillende cohorten vergelijkbaar wordt en biedt de mogelijkheid modellen te valideren in verschillende populaties over de hele wereld. In hoofdstuk 10 voorspelden we artrose gedefinieerd als KL ≥2, wat een uitkomstmaat is die een lange follow-up nodig heeft om sensitief te worden. Acht jaar is veel te lang voor een medisch onderzoek. Het zou leiden tot dure klinische studies en een veel te lange tijd tussen uitvinding en implementatie van behandelingen. Dit brengt ons naar het volgende probleem.

### Ongevoelige uitkomstmaten

Op dit moment zijn er geen efficiënte uitkomstmaten om het behandeleffect van een potentieel medicijn voor artrose te testen. Gewrichtsvervanging kan als harde uitkomstmaat gezien worden voor artrose, gezien de structurele achteruitgang en symptomen dusdanig zijn dat het gewricht vervangen moet worden. Echter is het vervangen van een gewricht een keuze die berust bij de patiënt en zijn behandelaar. Gebeurtenissen in iemands leven, of arbeidsovereenkomsten kunnen voor uitstel zorgen vanuit de patiënt. Comorbiditeiten en levensstijl, waaronder roken, kunnen zorgen voor uitstel of afstel vanuit de orthopedisch chirurg. De FDA en EMA accepteren momenteel alleen gewrichtsspleetversmalling zoals gemeten

op röntgenfoto's als erkende uitkomstenmaat voor de structurele progressie van artrose. Als gevolg eisen beide organisaties ook dat medicijnonderzoeken minimaal twee jaar moeten duren. Hoewel biochemische uitkomstmaten en metingen van MRI veelbelovend zijn, vinden de FDA en EMA dat er eerst meer onderzoek moet plaatsvinden naar de correlatie met symptomen en uiteindelijke röntgenologische progressie. De APPROACH studie meet bestaande en nieuwe biomarkers op baseline en na zes maanden, met de hoop nieuwe surrogaatmarkers te vinden om zo kortere en krachtiger geneesmiddelenonderzoeken voor artrose op te kunnen zetten.

### Kennis lacunes

Het is duidelijk geworden dat artrose niet louter slijtage omvat. Het is een ziekte met een heterogene pathofysiologie. Door onze kennis te vergroten kunnen we nieuwe behandeldoelen ontdekken. We kunnen patiënten beter in feno- en endotypes indelen, wat ons helpt de juiste behandeling te kiezen op het juiste moment, bij de juiste patiënt. De pathofysiologie van artrose en cardiovasculaire ziektes overlapt fors. Gek genoeg is er maar weinig onderzoek naar de relatie tussen beiden. In hoofdstuk 11 bestudeerden we de coïncidentie van arteriële verkalkingen en heup- en knieartrose. De relatie bleek zeer geslachtsafhankelijk. Bij vrouwen verhoogde het ontstaan van arteriële verkalking rondom een gewricht de kans op heup/knie artrose binnen datzelfde gewricht. Bij mannen was de ernst van de calcificaties op lichaamsniveau negatief gecorreleerd aan de kans op artrose in heupen en knieën. Gezien de studiepopulatie tussen 45 en 65 jaar oud was bij inclusie, kan het verschil eventueel verklaard worden door hormonale veranderingen die gepaard gaan met de menopause.[3] Daarnaast heeft het vaatstelsel van vrouwen over het algemeen een kleinere diameter en komt artherosclerose meer wijdverspreid voor.[4,5] Als de microvasculatuur faalt of thrombi loskomen van de vaatwand, kan er te weinig bloed richting het gewricht stromen. Dit kan weefselschade en daaropvolgende artrose in de hand werken. In mannen is het vaatstelsel over het algemeen wijder en vormt athersclerose focale plaques. Calcificatie is een teken van plaquevorming, maar kalk stabiliseert de plaque ook, waardoor het een beschermend effect kan hebben tegen loskomende thrombi.[6] Als laatste kan de associatie in mannen een zorgeffect zijn. Mannen met artrose komen sneller bij de huisarts. Bij mannen zal eerder het cardiovasculair risico ingeschat worden dan bij vrouwen. Uiteindelijk kan het zijn dat preventieve maatregelen ten aanzien van atherosclerose eerder worden ingezet bij mannen met artroseklachten dan bij mannen zonder artroseklachten.

Zoals hierboven vermeld is er een sterke overlap tussen de risicofactoren en pathofysiologie van cardiovasculaire ziekten en artrose. Daarnaast ontwikkelen beide ziekten zich traag. Daarom is het nodig om innovatieve onderzoeksstrategieën in te zetten om met 'confounding' en 'mediation' om te gaan. PXE is een systemische ziekte veroorzaakt door mutaties in het ABCC6 gen die leidt tot mineralisaties in fibreus weefsel en daarmee afwijkingen van de huid, ogen en het cardiovasculair systeem veroorzaakt. Cardiovasculair staan calcificaties in de vaatwand, veroorzaakt door ontregeling van inorganisch pyrofosfaat (PPi) niveaus, voorop. In het UMC Utrecht klaagden PXE-patiënten vaker dan verwacht over gewrichtspijn. Gezien we een relatie tussen arteriële calcificaties en artrose vonden in hoofdstuk 11 en artrose de nummer één oorzaak is van pijnlijke gewrichten, voerden we een exploratieve studie uit om te testen of PXE-patiënten meer kenmerken van structurele artrose hadden dan controle patiënten. PXE-patiënten bleken vaker structurele tekenen van artrose in de acromioclaviculaire en kniegewrichten te hebben. Deze bevinding dient in vervolgstudies gevalideerd te worden. Daarnaast moet er onderzoek plaatsvinden naar verklarende pathofysiologische mechanismes zoals PPi en vitamine-K niveaus, cytokines met invloed op botmetabolisme, subchondrale botintegriteit en arteriële verkalkingen. Dit is slechts één voorbeeld van een onderzoekslijn die ons kan leiden naar nieuwe endotypes binnen artrose. De snelle ontwikkeling van nieuwe technieken biedt in rap tempo nieuwe kansen om de pathofysiologie en behandeldoelen van artrose te onderzoeken. Deze nieuwe technieken dienen gecombineerd te worden met gevalideerde bestaande technieken in grote cohorten die zich richten op artrosepatiënten met hoge kans op snelle progressie.

### One-size-fits-all

Veel medicijnstudies binnen artrose gebruiken uniforme inclusiecriteria, zoals de ACR-criteria voor knieartrose. Door een gebrek aan definities en criteria voor feno- en endotypes binnen artrose is het lastig voor onderzoekers om de juiste patiënten voor hun behandeling te selecteren. Grote cohorten met genoeg ziekteprogressie en een combinatie van innovatieve en bestaande biomarkers zijn nodig om patiënten in fenotypes te kunnen stratificeren. Het lopende APPROACH project heeft als één van haar doelen gesteld om deze fenotypes te ontdekken. Tot deze fenotypes volledig gedefinieerd zijn moeten we echter niet wachten met het testen van potentiële behandelingen. Onderzoekers moeten nadenken over hun verwachte behandelmechanisme en hun inclusiecriteria hierop aanpassen. Een voorbeeld daarvan is de ZOleDroninezuur als behandelingen In Artrotische Knieën

(ZODIAK) studie. Deze studie vaart op de hypothese dat bisfosfonaten de integriteit van het subchondrale bot kunnen beschermen. Artrose veroorzaakt een toename in porositeit van het subchondrale bot, waardoor zenuwen en bloedvaten richting het kraakbeen groeien. Dit biedt een route voor pijnstimuli en inflammatoire cytokines. Dit proces tegengaan, door het remmen van osteoclastactiviteit, kan pijn en structurele achteruitgang van het kraakbeen in artrose tegengaan. Om patiënten met activiteit in het subchondrale bot te includeren, maakt de ZODIAK-studie gebruikt van een screenende MRI. Hierbij wordt gekeken naar BeenMergLaesies (BML), gezien deze gerelateerd zijn aan botremodellering.[7] Door patiënten zonder BML te excluderen hopen de onderzoekers patiënten zonder het beoogde behandeldoel te excluderen. Dit zou moeten leiden tot een efficiënter onderzoek met meer kans op een positief effect van de zoledroninezuur voor het behandelen van artrose.

## De toekomst van artrose

Een medicijn tegen artrose zal orthopedisch chirurgen niet werkloos stellen. Het eerste doel is om gewrichtsvervangende operaties uit te stellen en daarmee ook protheserevisies uit te stellen of te voorkomen. Onderzoek naar artrose moet zich richten op de juiste behandeling bij de juiste patiënt krijgen, op het juiste moment. Dit vergt grote cohortstudies naar definities van endotypes binnen artrose met goed omschreven selectiecriteria voor die endotypes en uitkomstmaten om een behandeleffect binnen een specifiek endotype aan te tonen. Tot we dat bereikt hebben moeten potentiële behandeling wel getest worden, maar onderzoekers moet kritisch kijken naar de gekozen selectiecriteria. Als blauwdrukken voor medicijnstudies binnen endotypes van artrose uitgerold worden, moeten we overwegen om medicijnen die eerder gefaald zijn wegens slechte studieopzetten, opnieuw te testen.

## Referenties


1. Odenbring S, Berggren AM, Peil L. Roentgenographic assessment of the hip-knee-ankle axis in medial gonarthrosis. A study of reproducibility. *Natl Med J India*. 1993.

2. Visser AW, Bøyesen P, Haugen IK, et al. Radiographic scoring methods in hand osteoarthritis - a systematic literature search and descriptive review. *Osteoarthr Cartil*. 2014;22(10):1710-1723. doi:10.1016/j.joca.2014.05.026

3. Fairweather D. Sex differences in inflammation during atherosclerosis. *Clin Med Insights Cardiol*. 2014;8(Suppl 3):49-59. doi:10.4137/CMC.S17068

4. Shaw LJ, Bugiardini R, Merz CNB. Women and Ischemic Heart Disease. Evolving Knowledge. *J Am Coll Cardiol*. 2009;54(17):1561-1575. doi:10.1016/j.jacc.2009.04.098

5. Hirsch AT, Allison M a., Gomes AS, et al. A call to action: Women and peripheral artery disease: A scientific statement from the american heart association. *Circulation*. 2012;125(11):1449-1472. doi:10.1161/CIR.0b013e31824c39ba

6. Wong KK, Thavornpattanapong P, Cheung SC, et al. Effect of calcification on the mechanical stability of plaque based on a three-dimensional carotid bifurcation model. *BMC Cardiovasc Disord*. 2012;12(1):7. doi:10.1186/1471-2261-12-7

7. Donell S. Subchondral bone remodelling in osteoarthritis. *EFORT Open Rev*. 2019;4(6):221-229. doi:10.1302/2058-5241.4.180102

# CHAPTER 13

## Curriculum vitae

Willem Paul Gielis was born in Breda in 1990 and grew up in Eindhoven. During his youth, he experienced the impact of orthopedic surgery. His dad became a new person after receiving hip replacement surgery. This motivated him to study medicine at the university of Amsterdam. He started rowing at the A.S.R. Nereus and became a member M.A.R.N.I.X. He organized a major event for medical students at the A.S.C. and was board member of O.P.I.U.M., a student association for A.S.C. members with medical interest.

During the final years of university Willem Paul focused on gaining experience in the field of orthopedics. The first step was a research internship under Rudolf Poolman and Jacob van Oldenrijk at the OLVG, resulting in his first paper: "Increased persistent mid-thigh pain after short-stem compared with wedge-shaped straight-stem uncemented total hip arthroplasty at medium-term follow-up: a randomized double-blinded cross-sectional study." Thereafter, Willem Paul went through internships at the department of orthopedic surgery of the OLVG and AMC. In 2015 he graduated and began the journey as PhD candidate at the University Medical Center Utrecht under supervision of Prof. dr. ir. Harrie Weinans and Prof. dr. Pim A. de Jong. During this period Dr. Wouter Foppen and Dr. Roel J. H. Custers were of invaluable contribution. Willem Paul supervised masters' students, became outstanding reviewer for Osteoarthritis and Cartilage, graduated cum laude from the postgraduated master epidemiology, and won an award for one of nine highest rated abstracts by young investigators at the OARSI 2017. In 2019 and 2020 Willem Paul worked as an orthopedic surgey resident at the Bergman Clinics | Bewegen Delft and Spaarne Gasthuis. In 2021 he started the specialist medical training for orthopedic surgery in the Utrecht region.

Manuscripten buiten het proefschrift

# CHAPTER 13

## Dankwoord

Een promotietraject voelt soms aan als een individuele marathon waarbij af en toe iemand je een glas water aanreikt. Zonder water bereik je de eindstreep niet. Zonder steun en liefde van collega's, vrienden en familie was dit boekje niet tot stand gekomen.

Harrie en Pim, ik had me geen fijnere promotoren kunnen wensen. De hoeveelheid tijd en aandacht die jullie mij geschonken hebben laat menig promovendus groen aanlopen. Elke week kon ik bij jullie terecht. Was het niet voor mijn onderzoeksvragen, dan kon ik veel van jullie leren over management, geldstromen en alle valkuilen van leiderschap in het ziekenhuis. In jullie zag ik de ideale combinatie van een creatieve ideeëngenerator en een pragmaticus. Ik hoop in de toekomst nog veel van jullie te mogen leren.

Roel en Wouter, jullie waren een ongelofelijke bron aan ervaring die ik kon aanboren. Altijd kon ik bij jullie aankloppen om mijn vragen te stellen en altijd kwam ik met betere plannen naar buiten. Wouter, zonder jouw adviezen was de OACT-score waarschijnlijk een puzzel geworden voor elke radioloog. Jouw adviezen hebben een warboel van ideeën tot een bruikbaar product gemaakt. Roel, je kon mij altijd helpen om de klinische waarde van onderzoek te wegen. Ik heb al veel van je mogen leren op de polikliniek van de Mobility clinic en hoop dat dit de komende opleidingsjaren een vlucht zal nemen.

Maarten en Richard, zonder jullie waren de laatste loodjes te zwaar geweest. We zijn nu bijna dertien jaar vrienden en uit jullie inzet de afgelopen weken blijkt de vriendschap onvoorwaardelijk. Liever had ik meer tijd gehad om samen met jullie te genieten van het proces, maar samen genieten zullen we nog genoeg. Ik waardeer het enorm jullie zoveel tijd bij me te hebben.

Roy, je hebt me enorm geholpen met de eindsprint. De laatste weken waren erg druk voor mij, maar je hebt me er doorheen begeleid. Met jouw hulp is dit geen werkstuk, maar een mooi boek geworden. Bedankt voor je inzet en doorzettingsvermogen.

Niet achter, maar naast elke succesvolle man staat een sterke vrouw. Natali, jij staat altijd voor mij klaar. Als geen ander kan je mij kalmeren als ik in de stress schiet. Je ondersteunde mij in tijden dat ik druk was, maar zat ook achter me aan als ik aan het fluimen was. We kennen elkaar al sinds het eerste moment van de opleiding en groeiden langzaam naar elkaar toe. Van samen in de UB blokken voor de bloktoetsen, naar samen wonen in een doorleefde antikraak woning. We hebben veel gezien, ontdekt en geleerd, we zijn samen volwassen geworden. De bezegeling van onze liefde raakte door COVID in het verdring, maar hoe mooi

was het alternatief. Zo klein, zo zoet, zo lief. Het schrijven van dit boekje duurde even, wij drie zijn voor het leven. Lieve Julia, jij hebt het afmaken van dit boekje niet persé eenvoudiger gemaakt. Toch ben ik erg dankbaar voor je vrolijkheid en energie. Ik geniet elke dag van de groei en ontwikkeling die je doormaakt. Ooit hoop ik je ook te kunnen interesseren voor de wetenschap, maar eerst zal ik je uit De Kleine Prins voorlezen.

Zonder jullie steun in mijn studententijd was ik wellicht nooit zo enthousiast geworden voor de wetenschap. Mijn ouders gaven mij de tijd om mijn interesses te ontwikkelen. Tijd is onvervangbaar en daarom ben ik ze enorm dankbaar. Papa, mama, zonder jullie opvoeding had ik nooit de discipline gehad om dit af te maken. Dit is ook jullie prestatie.

Anne Lize en Joery, Sven en Jonne, opa Cas en oma Corrie, Frederique, Lian, Djoko, Shanomi, Mace, Jayh, Marcella, Jeremy, Tatjana, Alexander, Ilia en Annemijn, je familie wordt voor je gekozen. Toch zou ik jullie allemaal vaker willen zien. Meer tijd hebben om samen te zijn. De drukke levens en de afstand maken de drempel vaak hoog om met elkaar af te spreken. Ik wil dat jullie weten dat jullie altijd in mijn hart zijn en ik aan jullie denk. Oom Hennie, aan jou was de wetenschap niet helemaal besteed. Je had een sterke eigen mening. Je was een bijzonder persoon en deed je eigen ding. Ook daar heb ik van geleerd. Ik denk aan je.

Pepijn, Stefan, Wouter, Robert en Pim, ik ben jullie onwijs dankbaar voor het feit dat er in het zuiden altijd een warm nest is om naar terug te keren. Tentje ploffen in Renesse, pizza met hete kolen in de tent op hockeytoernooi, een eigen villa in Chersonissos, schoenen ruilen op Dominator, samen naar Scooter en de zwarte piste op Solar. We hebben hele mooie dingen meegemaakt, te veel om op te noemen. We staan voor elkaar klaar in mindere tijden, dat onderscheid kennissen van echte vrienden. Ik waardeer het enorm dat jullie altijd de moeite nemen om een klein biertje met me te doen als ik weer in het zonnige zuiden ben. Ik ken de meesten van jullie vanaf mijn 4e en kijk vol goesting uit naar de komende kinderfeestjes.

Louis, Danitscha, Florus, Thijs, Dick, Jeroen en Mingus, jullie vormen mijn warme nest in Amsterdam. Bizar wat we met z'n allen hebben meegemaakt. De eerste keer gele lasers in je ogen. Een bierfontijn. Uitvinden dat er naast de partyboat ook een boozecruise bestaat. Samen verdwaald raken in de bossen van Amsterdam Noord. Het resultaat van 60 jaar hoarden op een zolder van een sociale huurwoning bekijken. Een liefde voor kaas. Ook bij jullie is het altijd vanouds gezellig en gaat de vriendschap diep. Ik hoop dat ik daar nog lang van mag genieten.

Arthur, Dolf en Thomas, we hebben veel samen in een boot gezeten. Weinig mensen kunnen dat zeggen. Als je samen in een boot zit bouw je een band op, als je samen naar een Skadi feest gaat verbetert die band. Ik hoop dat we in de toekomst nog een keer samen roeien.

Wouter, Merle, Eva en Anne, we hebben elkaar beter leren kennen tijdens een onstuimig weekend in Renesse. Het werd een vriendschap vol pieken en dalen. Van epidemiologische of anatomische vraagstukken naar de vroege uurtjes in de Armadillo. Juist nu dat laatste toekomstmuziek is, koester ik die momenten misschien wel het meest.

Rintje, Chien, Joost, Jukka, Vahid, Floris, Bruce, Chella, Jonneke, Koen, Jasmijn, Rob, Jelle, Mattie, Huub, Eefje, Nick, Justin, Sebastiaan, Steven, Lorenzo, Mechteld, Pauline en Hilde, naast veel serieus werk, cijfertjes en lange teksten heb ik ook jullie persoonlijke kant gezien. Schreeuwend in een taxi voor een broodje döner. Lachend over een gefrituurde sparerib na een congres. Een traantje wegpinken voor het anthem van het Q mini-tafeltennistoernooi. Als eerste naar Taverne vanuit Brabanthallen, omdat je je daar wel kan misdragen. Met je collega's breng je meer tijd door dan met je naasten. Het is fijn als die tijd dan vol avonturen en gezelligheid zit. Daarvoor wil ik jullie bedanken.