Unraveling hip morphology
from development to osteoarthritis

Fleur Boel

Unraveling Hip Morphology

from development to osteoarthritis

Fleur Boel

ISBN: 978-94-6510-971-8
Cover design: Fleur Boel
Layout and design: Optima
Printing: Optima

The printing of this thesis was generously supported by the Department of Orthopaedic Surgery & Sports Medicine at Erasmus MC, University Medical Center Rotterdam, The Netherlands; Erasmus University Rotterdam; and ChipSoft.

Het ontrafelen van heupmorfologie van ontwikkeling tot artrose

Unraveling Hip Morphology

from development to osteoarthritis

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr.ir. A.J. Schuit

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op
dinsdag 4 november 2025 om 10.30 uur

door

Fleur Désirée Emile Maria Boel
geboren te Breda.

Erasmus University Rotterdam

Erasmus

PROMOTIECOMMISSIE:

Promotor:	Prof.dr. S.M.A. Bierma - Zeinstra
Overige leden:	Dr. M. van Middelkoop Dr. J.F. Felix Prof.dr. J.N. Doornberg
Copromotoren:	Dr. R. Agricola Dr. J. Runhaar

CONTENTS

Chapter 1

General introduction

DEVELOPMENT OF THE HIP JOINT

The hip joint is an articulating joint formed by the acetabulum, the femoral head, and the surrounding soft tissue (Figure 1). Hip joint development begins in utero and continues throughout fetal development, infancy, and childhood until skeletal maturity is reached. The morphology of both the acetabulum and femoral head at skeletal maturity are interconnected and depend on their interaction during the development of the hip joint.

Prenatal development

Prenatal development consists of the embryonic period, the first 8 weeks of gestation, and the fetal period. During the embryonic period, the limb buds form and rapidly differentiate into the infantile extremities[1,2]. By approximately 7 weeks of gestation, the cartilaginous femur and acetabulum models are complete. The ligament teres, transverse acetabular ligament, acetabular labrum, joint capsule and synovium are microscopically identifiable by 8 weeks of gestation. Additionally, the primary ossification center of the femur is formed.

During fetal development, the focus shifts from differentiation to growth, ossification, vascularization and maturation of the hip joint[2]. By 11 weeks of gestation, the infantile configuration of the hip joint is reached. The proximal femur consists of a fully formed femoral head with a spherical contour, a short femoral neck and a primitive greater trochanter. The joint capsule, acetabular labrum and transverse ligament are well-defined. By 16 weeks of gestation, the femur has ossified to the level of the lesser trochanter, and the primary ossification centers of the ilium, ischium, and pubis have appeared. The hip joint space is complete, with mature hyaline cartilage covering the articular surfaces, and the muscle structures have matured. At this stage of development, active motion of the legs can be observed. Hip differentiation will continue until approximately 20 weeks of gestation.

Postnatal development

Acetabulum

At birth, the acetabulum primarily consists of a cartilaginous ring around the femoral head. This cartilage ring will develop further during growth, forming the load-bearing, articular surface[2].

The acetabulum develops during infancy and childhood through interstitial growth in the triradiate cartilage[3], see Figure 2. The triradiate cartilage, the growth plate of the

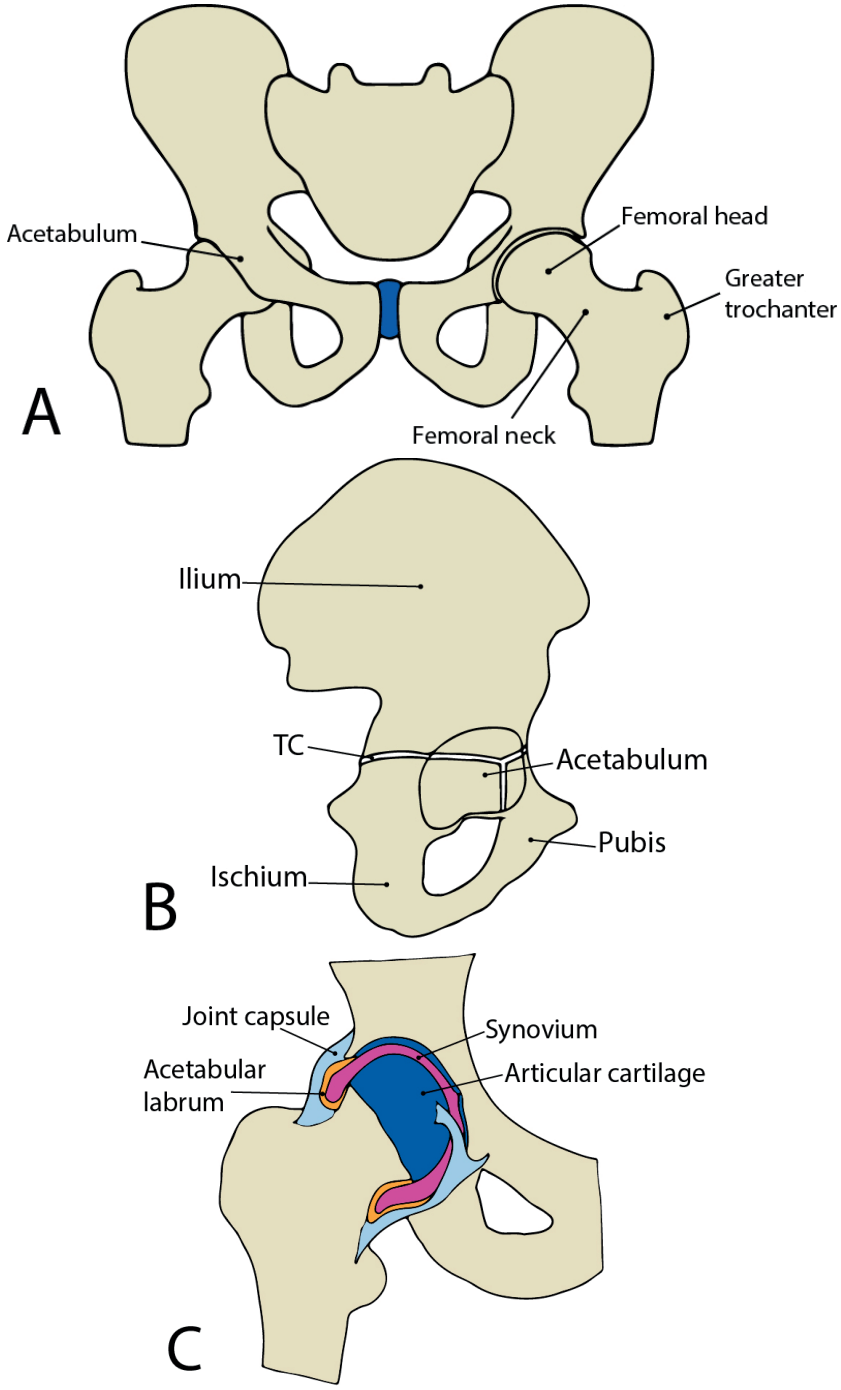


Figure 1. Anatomy of the hip joint. **A**. Anterior view of the pelvis bones. **B**. Lateral aspect of the hip bone. TC: triradiate cartilage. **C**. Soft tissues of the hip joint.

acetabulum, is present at the center of the acetabulum between the ilium, ischium and pubis, see Figure 1B. The ossification center from the pubis will ossify to form the anterior acetabular wall. The ossification center from the ilium will form the superior acetabular dome, and the ischial center will form the posterior wall of the acetabulum. The triradiate cartilage ossification will start around age 8 in girls and 10 in boys and will be closed around age 12 in girls and 14 in boys[4].

An additional ossification center, the os acetabili, will form on the acetabular rim, influencing the final bony shape of the acetabulum[5]. The ossification of the os acetabuli will start around age 9 in girls and 11 in boys and the ossification center will be entirely fused by age 11 in girls and 13 in boys[4].

During skeletal maturation of the hip joint, the acetabular inclination orientation becomes increasingly horizontal and the coverage of the femoral head by the acetabulum increases. The neck-shaft angle decreases during the same period, improving the centric alignment of the femoral head and the acetabulum. This helps maintain joint stability, allows for a good range of motion, and even distribution of the forces across the joint surface[6].

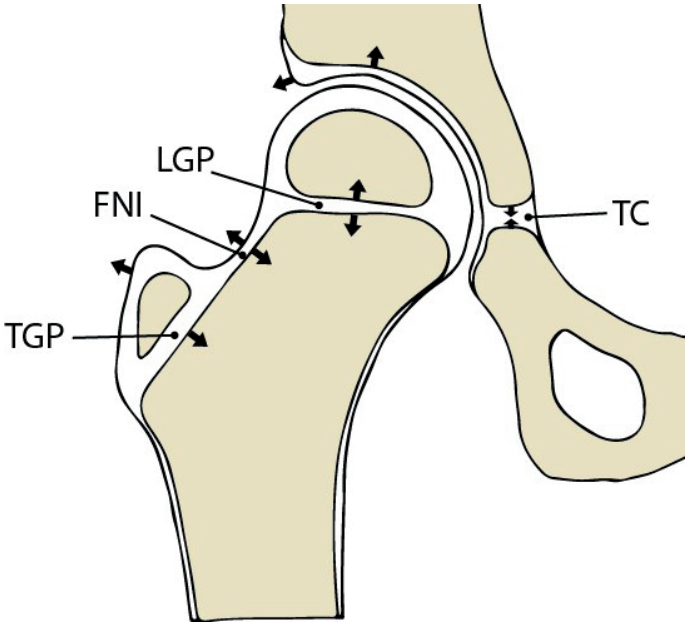


Figure 2. The development of the hip joint during growth occurs trough interstitial and appositional proliferation, indicated by the black arrows. TC: triradiate cartilage, LPG: longitudinal growth plate, FNI: femoral neck isthmus, TGP: greater trochanter growth plate.

Femur

At birth, the proximal femur remains cartilaginous and gradually ossifies during infancy and childhood. The femoral head enlarges through appositional cartilage cell proliferation followed by ossification[3,7]. Three growth plates further contribute to the morphology and growth of the proximal femur: the longitudinal growth plate (LGP) of the femoral neck, the greater trochanter growth plate (TGP) and the femoral neck isthmus (FNI) that connects the LPG and TGP on the lateral neck (Figure 2). The LGP and TGP are responsible for the longitudinal growth of the femur and the lateral width of the femoral head. Additionally, the LGP helps maintain the sphericity of the femoral head. Normal development of the proximal femur depends on the growth balance from these growth plates. A disruption of the ossification of these growth plates can lead to deformities of the proximal femur, such as coxa vara, a decreased neck-shaft angle, or coxa valga, an increased neck-shaft angle.

ABNORMAL DEVELOPMENT

Abnormal hip joint development can arise from altered forces on the growth zones or changes at the growth plates. Furthermore, as the development of the femur and acetabulum are interdependent, alterations in the morphology of one can lead to changes in the morphology of the other. This thesis focuses on the abnormal development of the acetabulum, namely acetabular dysplasia and pincer morphology.

Acetabular dysplasia

Acetabular dysplasia is typically characterized by a shallow acetabulum with undercoverage of the femoral head. Acetabular dysplasia can lead to hip instability and increased labrum and articular cartilage stress. Consequently, acetabular dysplasia is associated with hip pain, decreased function, and early-onset hip osteoarthritis[8-13]. It is crucial to distinguish between developmental dysplasia of the hip (DDH), which develops during the perinatal period and infancy, and acetabular dysplasia, which develops later in life.

DDH is a spectrum that includes hip dislocation, hip subluxation, hip instability and a stable hip with insufficient acetabular coverage of the femoral head[14,15]. DDH develops during the perinatal period and infancy. According to a recent meta-analysis, the prevalence of DDH in infants is 1.4 % (95% CI 0.86 – 2.3)[16]. Risk factors for DDH include female sex assigned at birth, breech position, breech presentation, a positive family history of DDH, cesarean section and firstborn status[14,17]. Conversely, low birth weight is generally considered protective against DDH[14].

Various screening programs have been implemented worldwide to detect and treat DDH in infants[18,19]. In the Netherlands, this screening is performed in child healthcare centers, where parents can take their children for regular check-ups, vaccinations, growth monitoring and developmental assessments. Screening involves medical history questions and physical examination at 1, 3 and 7 months. Referral for diagnostic evaluation of DDH happens if the child has a positive family history of DDH or early-onset hip osteoarthritis (<50 years old), a breech position after 32 weeks of gestation, a breech presentation at birth, or abnormal findings on physical examination such as limited abduction < 70°, an abduction difference ≥ 20°, or a difference in leg length or knee height[20]. The positive predictive value of the Dutch screening program is 15.6% (95% CI 12.0-19.2). while the negative predictive value is 99.4% (95% CI 99.0-99.8). Diagnostic evaluation is performed using ultrasound Graf classification (under the age of 6 months) or pelvic radiographs (after the age of 6 months)[21]. If DDH is diagnosed early, the first treatment is usually a Pavlik harness, which is a non-surgical treatment with a hip abduction device stimulating normal development of the acetabulum. If the Pavlik harness treatment is not successful or not indicated, closed or open reduction of the hip joint can be considered, with the possibility of surgery on the surrounding ligaments and/or tendons. Following surgery, a hip abduction device and regular monitoring are required. If the acetabular development is unsatisfactory after the age of 1 year, additional surgeries, such as a pelvic osteotomy and/or a proximal femur osteotomy, may be necessary[22].

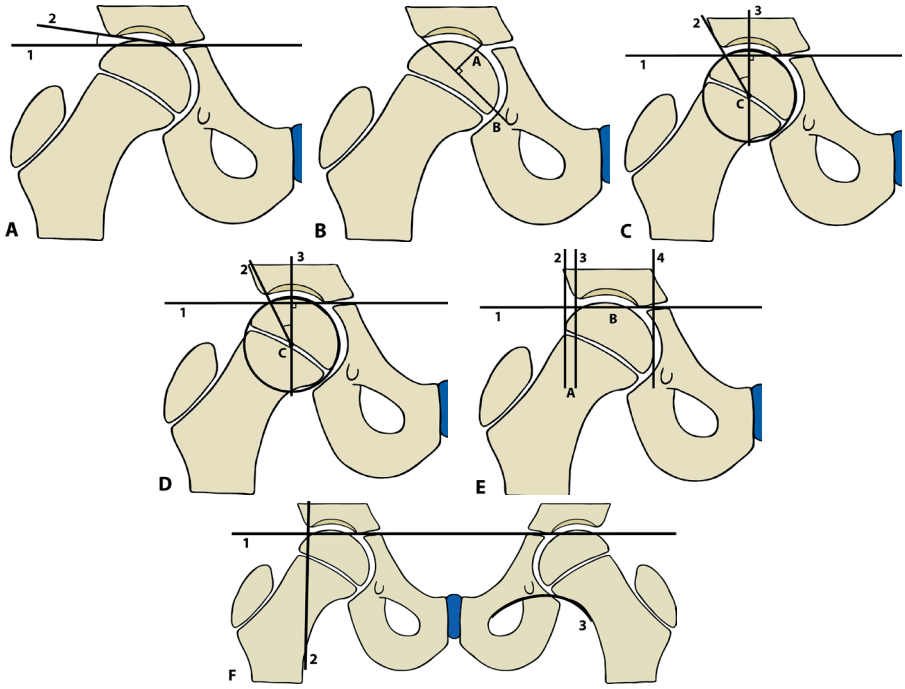


Figure 3. Radiographic measurements of acetabular dysplasia. **A: The acetabular index (AI)** describes the angle between the horizontal reference line of the pelvis (line 1) and line 2 through the most lateral bony part of the acetabulum and the most lateral point of the triradiate cartilage. **B: The acetabular depth-width ratio (ADR)**, ADR = (A/B)*1000, is the ratio between the acetabular width (B) measured from the most lateral bony part of the acetabulum to the most inferior point of the teardrop and the acetabular depth (A) measured from the most medial point of the sourcil perpendicular to line B. **C: The lateral center edge angle (LCEA)** is the angle between line 2 from the most lateral bony part of the acetabulum to the femoral head center and line 3, a line through the femoral head center perpendicular to the horizontal reference line of the pelvis (line 1). **D: The center edge angle of Wiberg (WCEA)** is the angle between a line from the most lateral part of the sourcil to the femoral head center (line 2) and line 3, a line through the femoral head center perpendicular to the horizontal reference line of the pelvis (line 1). **E: The extrusion index (EI)**, EI = A/(A+B) * 100%, is the percentage of the part of the femoral head not covered by the acetabulum (A) compared to the entire width of the femoral head (A+B). **F:** Line 1 is Hilgenreiner's line, which is a horizontal line connecting both triradiate cartilages. Line 2 is Perkin's line, which is perpendicular to Hilgenreiner's line, intersecting the most lateral edge of the acetabular roof. Line 3 is Shenton's line, which is a curved line drawn along the superior border of the obturator foramen and along the medial border of the proximal femoral neck.

Acetabular dysplasia can also be diagnosed later in life. Individuals are often diagnosed after experiencing symptoms such as groin or gluteal pain and difficulty standing or walking for an extended period, which are related to damage to the soft tissues and articular cartilage. The development timeline of these cases of late-diagnosed acetabular dysplasia remains unclear. The prevalence of acetabular dysplasia in adults from the general population is higher than in infants, ranging between 3.3 – 9.4% compared to 1.4% in infants[23-25]. This higher prevalence in adults may indicate that acetabular dysplasia can also develop later in life. Other possible explanations for late diagnosed acetabular dysplasia can be missed diagnoses of DDH in infancy, or residual acetabular dysplasia after DDH treatment[26-28].

Radiographs are used for the detection and characterization of acetabular dysplasia, as well as for monitoring the hip over time. The measurements of acetabular dysplasia that can be performed depend on the age of the patient and the stage of hip development, which is determined by factors such as the ossification of the femoral head and the closure of growth plates. Common radiographic measurements used to assess acetabular dysplasia on anteroposterior pelvic radiographs include the acetabular index, the acetabular depth-width ratio, the center edge angle, the extrusion index, Shenton's line, and Hilgenreiner's and Perkin's lines (Figure 3). Using multiple measurements in conjunction can provide a more comprehensive assessment of the hip joint. While reference values are available for some of these measurements, this is often limited to adults or specific age ranges. Interpretation of the measurements can, therefore, be difficult in childhood and adolescent populations. This difficulty is enhanced by the changing morphology of the hip joint throughout development and the different rates at which males and females develop. This emphasizes the need for age and sex-specific reference values.

The acetabular index, also known as the Tönnis angle, measures the acetabular roof inclination. The acetabular index can be determined until the triradiate cartilage closes. During development, the acetabular roof inclination, as determined on radiographs, will decrease over time. The Tönnis table is commonly used as a reference for normal and dysplastic hips, with biological sex and age-specific reference values. The acetabular depth-width ratio is a measure of acetabular depth, calculated as the ratio of the acetabular depth to the entire length of the acetabular. In adults, an acetabular depth-width ratio ≤ 250 is often used to indicate acetabular dysplasia.

The center edge angle measures the coverage of the femoral head by the acetabulum. Once the acetabular lip is fully developed, a distinction can be made between the bony and the weight-bearing coverage of the femoral head by the acetabulum. The bony coverage can be assessed using the lateral center edge angle, while the weight-bearing coverage can be assessed using the Wiberg center edge angle. The lateral center edge angle will increase during maturation of the hip. While clear cut-off values to determine acetabular dysplasia in children are lacking, in adults a Wiberg center edge angle ≤ 20° and ≤ 25° are often used to indicate severe and mild acetabular dysplasia, respectively. The extrusion index, also known as the migration index, is another measure of femoral head coverage by the acetabulum. An extrusion index ≥ 25% in adults is commonly used to quantify acetabular dysplasia.

Shenton's line is a curved line drawn along the superior border of the obturator foramen and along the medial border of the proximal femoral neck. Usually, the line is continuous,

but this line can be interrupted in dysplastic hips. Hilgenreiner's line is a horizontal line connecting both triradiate cartilages, and Perkin's line is perpendicular to Hilgenreiner's line, intersecting the most lateral edge of the acetabular roof. Usually, the femoral head is located in the inferior of Hilgenreiner's line and medial of the Perkin's line.

Pincer morphology

Pincer morphology is characterized by an overcoverage of the femoral head by the acetabulum. The overcoverage can be focal, caused by retroversion of the acetabulum, or global, resulting from bony overcoverage by the entire acetabulum or a deep acetabular socket[29]. The reported prevalence of pincer morphology in the literature varies widely, from 3% to 74%, along with the definition of pincer morphology used[30]. Focal overcoverage is often defined by indirect measures of acetabular retroversion, such as the crossover sign, posterior wall sign, and ischial spine sign (Figure 4A-C). However, these measurements generally have poor reliability and validity when defining true retroversion or pincer morphology[31]. Global overcoverage is frequently defined by the lateral center edge angle (Figure 3C), which assesses bony overcoverage, or coxa profunda and protrusio acetabuli, which characterize a deep acetabular socket (Figure 4D). However, there seems to be no association between the presence of pincer morphology and coxa profunda and protrusio acetabuli[32]. These variations in definition and lack of reliability for some of these measurements lead to inconsistencies in found prevalence.

Although pincer morphology has been studied extensively in adults, there is a lack of knowledge on the development of pincer morphology. The precise timing of pincer morphology development remains unclear, but it has been suggested that the development of pincer morphology starts around age 12 years[33].

HIP OSTEOARTHRITIS

Osteoarthritis is a complex chronic disease of the whole joint, impacting the articular cartilage subchondral bone, synovial membrane, capsule, ligaments and periarticular musculature. Osteoarthritis is common and debilitating, with a significant impact on quality of life, as well as a high healthcare and societal burden due to the required resources and associated costs[34]. Any joint can be affected by osteoarthritis, including the hip joint.

The dominant symptom of hip osteoarthritis is pain, but it can also result in stiffness and reduced range of motion. These symptoms can significantly impact an individual's mo-

bility, daily activities, and overall quality of life. The development of hip osteoarthritis is a result of genetic, environmental and lifestyle factors.

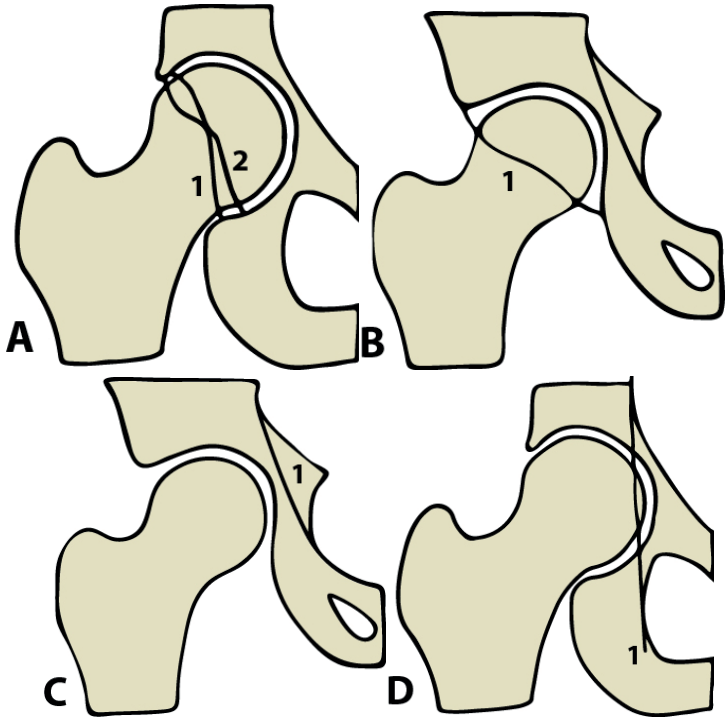


Figure 4. Radiographic measurements of pincer morphology. **A: The crossover sign** is present when the posterior acetabular wall (line 1) intersects with the anterior acetabular wall (line 2). **B: The posterior wall sign** is positive when the posterior acetabular wall runs lateral to the femoral head center. **C: The ischial spline sign** is positive when the ischial spline (part 1) projects medially to the pelvic rim. **D: Coxa profunda** is present if the acetabular fossa extends medially to the ilioischial line (line 1). Protrusio acetabuli is present when the femoral head extends medially to the ilioischial line (line 1).

The prevalence of hip osteoarthritis in the general population ranges between 0 and 47%, depending on the definition of osteoarthritis, age, country of origin, and biological sex distribution of the population studied[35]. Overall, the prevalence is higher for radiographic osteoarthritis than for symptomatic osteoarthritis. Additionally, the prevalence has been increasing worldwide in the last 30 years and is expected to continue increasing[36]. It is predicted that one in four people may develop symptomatic osteoarthritis in their lifetime[37].

In the Netherlands, the clinical diagnosis of hip osteoarthritis can be made without additional testing if the patient is 45 years or older, presents with activity-related hip pain, and has no or short (< 30 minutes) morning stiffness in the hip. Otherwise, additional

testing, such as radiographic imaging, should be considered[38]. Radiographic hip osteoarthritis is generally classified by several structural changes, including the development of osteophytes, narrowing of the joint space, increased density of the subchondral bone, the appearance of subchondral cysts, and deformity of the femoral head and acetabulum[39].

Typical management of hip osteoarthritis is focused on conservative treatment of the symptoms rather than proactive or preventative action[34]. No curative treatment for hip osteoarthritis exists, and joint replacement surgery is considered a last resort when conservative treatment fails[37]. Given the increasing incidence and burden of hip osteoarthritis, management should shift towards prevention and individualized treatment. Therefore, identifying modifiable risk factors and early detection of individuals at risk is paramount. This could aid in both developing and implementing preventative measures and provide the opportunity to slow disease progression or delay the need for joint replacement surgery.

Abnormal hip morphology as a risk factor for hip osteoarthritis

Numerous risk factors for hip osteoarthritis have been identified, including older age, female sex, various genetic variations, high physical workload, participation in high-impact sports, and abnormal hip morphology[34]. Abnormal hip morphology is an important risk factor for the development of hip osteoarthritis, including acetabular dysplasia and pincer morphology[11].

Acetabular dysplasia has been hypothesized to lead to osteoarthritis due to increased mechanical stress on the articular cartilage caused by the global undercoverage of the femoral head by the acetabulum[40]. This increased loading on the small contact area between the femoral head and dysplastic acetabulum results in articular cartilage and labrum damage[41-43].

The association between acetabular dysplasia and hip osteoarthritis has been studied in literature, and the results vary. A meta-analysis found that dysplastic hips had 2.38 (95% CI 1.84-3.07) higher odds of developing radiographic hip osteoarthritis than control hips[11]. However, the resulting associations can vary widely in size, and some studies find no association at all[11,12]. This variability is likely due to the difference in the definition of acetabular dysplasia and hip osteoarthritis, as well as the age of the study population and follow-up time. Additionally, acetabular dysplasia is thought to be a risk factor for the early and rapid development of hip osteoarthritis[12]. The contribution of acetabular dysplasia to the development of hip osteoarthritis might diminish over more extended periods because non-dysplastic hips also develop hip osteoarthritis over time.

Therefore, the true impact of acetabular dysplasia on the development of hip osteoarthritis remains unclear.

Pincer morphology has been hypothesized to lead to osteoarthritis due to impingement of the labrum between the femoral neck and the acetabular bone during movement[44]. During impingement, the labrum is compressed and the force is transmitted to the acetabular cartilage along the acetabular rim. Repeated trauma due to impingement causes degeneration and ossification of the labrum.

However, the association between pincer morphology and the development of hip osteoarthritis remains unclear. A recent systematic review with meta-analysis of prospective studies showed no higher likelihood of developing hip osteoarthritis in hips with pincer morphology, defined by a lateral center edge angle ≥ 40°, compared to control hips over a median follow-up of 9.2 years. On the contrary, in cross-sectional studies, hips with osteoarthritis were 3.7 times more likely to have a lateral center edge angle ≥40° than control hips. Several hypotheses may explain these disparate findings. Firstly, the progression of hip osteoarthritis associated with pincer morphology is thought to be slow due to the gradual degeneration of the acetabular cartilage[11,45]. Thus, the association between pincer morphology and hip osteoarthritis might remain undetected with insufficient follow-up times. Second, the use of the lateral center edge angle alone to determine pincer morphology may underestimate its prevalence. Lastly, pain could be an effect modifier of the association between pincer morphology and hip osteoarthritis[46]. Further research is therefore warranted to better understand the relationship between pincer morphology and hip osteoarthritis.

AIMS AND OUTLINE OF THIS THESIS

This thesis investigates acetabular dysplasia and pincer morphology to better understand their prevalence, development, and contribution to hip osteoarthritis risk. Two large cohorts will be used in this thesis: the Generation R study and the Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) consortium. The Generation R study, a prospective, population-based cohort from Rotterdam, the Netherlands, was established to investigate children's growth, development and health from fetal life onwards[47]. In this cohort, high-resolution dual-energy x-ray absorptiometry (DXA) images of the right hip are made at multiple time points throughout childhood and adolescence, allowing us to study how the hip develops over time. Secondly, the World COACH consortium is a global collaboration of all available prospective cohort studies

with prospective pelvic or hip imaging[48]. This consortium will allow us to further study the relationship between hip morphology and the development of hip osteoarthritis.

The thesis is structured in three parts:

Part I focuses on developing and validating an open-access automated method for quantifying hip morphology on two-dimensional imaging. This method aims to address the need for a consistent definition of these measurements, paving the way for more reliable research and better comparison between studies.

Part II investigates the prevalence and etiology of acetabular dysplasia and pincer morphology during childhood and early adolescence, a critical, yet understudied period for hip development. This part will shed light on the early development of these features and potential risk factors.

Part III explores the relationship between acetabular dysplasia and pincer morphology and the development of radiographic hip osteoarthritis. This part leverages the established methods and knowledge from previous parts to contribute valuable insights into the long-term consequences of these hip morphologies.

Part I covers developing and validating the automated methods used to quantify hip morphology. In **Chapter 2**, we present our in-house developed automated method to determine radiographic hip morphology measurements on DXA images and evaluate the results. **Chapter 3** is an external validation of this method in the adult population from the World COACH consortium. Traditionally, these radiographic hip morphology measurements are determined on pelvic radiographs. However, within the Generation R cohort, we make use of DXA images to determine hip morphology. In **Chapter 4**, we compare hip morphology measurements performed on DXA images and pelvic radiographs to see if these imaging modalities can be used interchangeably.

Utilizing the validated methods from Part I, Part II investigates the prevalence and risk factors associated with hip morphology in a general population of children and early adolescents. In **Chapter 5**, we perform a systematic review of the literature on the prevalence of acetabular dysplasia between the ages of two and eighteen. Additionally, we describe the radiological measurements used to diagnose acetabular dysplasia on the imaging. **Chapter 6** describes the prevalence of acetabular dysplasia in 3,986 early adolescents of the general population. **Chapter 7** mirrors chapter 6 and describes the prevalence of pincer morphology within this same group of early adolescents of the general population. In **Chapter 8**, we aim to further our understanding of acetabular

dysplasia in early adolescents. We investigate whether known risk factors for DDH are also associated with acetabular dysplasia in early adolescents from the general population. In **Chapter 9**, we continue the work of chapter 8 and investigate the development of the acetabulum during childhood. We investigate how acetabular coverage changes over time and its association with birth-assigned sex, weight status, triradiate cartilage orientation, and proximal femur shape, providing insights into the developmental trajectory of the acetabulum.

Part III focusses on the association between hip morphology and the development of incident radiographic hip OA (RHOA). We make use of the automated methods described and validated in Part I to determine acetabular dysplasia and pincer morphology within the World COACH consortium. In **Chapter 10**, we describe the association between acetabular dysplasia and the development of RHOA. Acetabular dysplasia will be defined using multiple hip morphology measurements to gain insight into the different aspects of acetabular dysplasia and their relationship with RHOA. **Chapter 11** mirrors chapter 10 and describes the association between pincer morphology and the development of RHOA.

Part 1

Quantifying Hip Morphology

Chapter 2

Automated radiographic hip morphology measurements: An open-access method

F. Boel, S. de Vos-Jakobs, N.S. Riedstra, C. Lindner, J. Runhaar,
S.M.A. Bierma-Zeinstra, R. Agricola.

Osteoarthritis imaging. 2024 Jun;4(2):100181

ABSTRACT

Objective: The aim of this study is to present a newly developed automated method to determine radiographic measurements of hip morphology on dual-energy x-ray absorptiometry (DXA) images. The secondary aim was to compare the performance of the automated and manual measurements.

Design: 30 DXA scans from 13-year-olds of the prospective population-based cohort study Generation R were randomly selected. The hip shape was outlined automatically using radiographic landmarks from which the acetabular depth-width ratio (ADR), acetabular index (AI), alpha angle (AA), Wiberg and lateral center edge angle (WCEA) (LCEA), extrusion index (EI), neck-shaft angle (NSA), and the triangular index (TI) were determined. Manual assessments were performed twice by two orthopedic surgeons. The agreement within and between observers and methods was visualized using Bland-Altman plots, and the reliability was studied using the intraclass correlation coefficient (ICC) with 95 % confidence intervals (CI).

Results: The automated method was able to perform all radiographic hip morphology measurements. The intermethod reliability between the automated and manual measurements ranged from 0.57 to 0.96 and was comparable to or better than the manual interobserver reliability, except for the AI.

Conclusion: This open-access, automated method allows fast and reproducible calculation of radiographic measurements of hip morphology on right hip DXA images. It is a promising tool for performing automated radiographic measurements of hip morphology in large population studies and clinical practice.

INTRODUCTION

Osteoarthritis (OA) of the hip leads to pain, disability, and poor quality of life[49]. Risk factors for developing hip OA include age, genetics, trauma, physical workload, and hip morphology[8-11,50-52]. Hip morphology can be quantified using radiographic measurements, which are most often performed manually on anterior-posterior (AP) pelvic radiographs[11]. However, manual measurements are time-consuming, especially when multiple measurements are performed, and the accuracy is highly observer dependent. Moreover, the definitions used for most radiographic measurements of hip morphology are inconsistent, and there is often no clear description of how the measurements are performed, making it challenging to compare between studies[53-55]. A recent international consensus statement on hip pain in young and middle-aged adults therefore recommended detailed and consistent definitions, measurements, and statistical reporting of radiographic measurements of hip morphology[54].

Automation of radiographic measurements of hip morphology has the potential to increase their reproducibility. It would also allow rapid calculation of multiple measurements for each patient. Automated calculation of radiographic measurements would not only aid clinical practice but also enable these measurements to be taken in large population studies. However, automated methods are scarce, poorly described and generally not published as open access.

Hip or pelvic radiographs have traditionally been used to define and classify hip OA and hip morphology. Alternatively, modern dual-energy X-ray absorptiometry (DXA) scanners are increasingly used to assess hip morphology[56-60] and are also suitable for hip OA grading[61]. DXA images of the hip expose study participants to a much lower radiation burden (0.36-70 µSv) than hip or pelvic radiographs (600-700 µSv)[62,63]. The reduced radiation exposure makes DXA images suitable to study hip morphology, growth plates and hip joint development in a pediatric population.

This study provides a detailed description of a newly developed and open-access, automated method of determining radiographic measurements of hip morphology on DXA images of 13-year-olds from the general population. The secondary aims were to compare the performance of the automated method with manual determination of the radiographic measurements and to assess the reliability of the automated measurements, using intraclass correlation coefficients (ICC).

METHODS

Participants

Generation R is a population-based prospective cohort study that follows participants from fetal life until young adulthood in the multi-ethnic urban population of Rotterdam, the Netherlands[47]. Generation R is designed to identify early environmental and genetic factors and causal pathways causing normal and abnormal growth, development and health. The Medical Ethical Committee of the Erasmus MC approved the study (MEC-2015–749). All participants provided written informed consent. The 4625 participants who underwent a DXA scan at around age 13 years formed the population of interest. We randomly selected two training sets of 500 participants each and a different test set of 30 participants around age 13 for validation of the automated measurements.

Image acquisition

DXA scans were obtained by the GE-Lunar iDXA densitometer (GE Healthcare, Madison, WI, USA) and enCORE software (enCORE 2010; GE Healthcare). The participants were scanned in supine position with legs slightly apart and big toes touching, with their feet secured in this position. A unilateral anteroposterior (AP) right hip DXA image and an AP full-body DXA image were acquired consecutively for each patient. We extracted the full-body and right hip images from the enCORE software in BMP format.

Definition and calculation of radiographic measurements of hip morphology

We developed methods in-house to automatically calculate radiographic measurements of hip morphology based on radiographic landmarks. The proximal femur and acetabulum were outlined with 80 radiographic landmarks using the BoneFinder® software (www.bone- finder.com; The University of Manchester, UK). The protocol for radiographic landmark definition can be found in Supplementary material 1. To automate the radiographic landmark placement, an automatic search model (ASM), a random forest-based machine learning algorithm, was trained on the first training set of 500 right hip DXA images[64]. Of these 500 images, 20 images were incomplete, with one or more landmarks missing from the image. This resulted in a training set of 480 images suitable for training of the ASM. This ASM was used to annotate the images of the other training set for development of the automated methods and the 30 images of the separate validation set.

To assess the influence of radiographic landmark placement, a second set of radiographic landmarks was created using the ASM. The radiographic landmark placement for all 30 hips from the validation set were manually assessed and minor corrections were performed by the same researcher who developed the protocol for radiographic

landmark placement. This manual correction was performed to remove the influence of incorrect placement of the radiographic landmarks on the performance of automated measurements. The landmarks that most often needed adjustments were the most lateral bony point of the acetabulum, the most lateral and most medial point of the sourcil, the triradiate cartilage points, the most caudal point of the teardrop, and the landmarks at the start and end of the best fitting circle. Any points equally spaced between these landmarks were also influenced by the changed landmark position. Adjustments were needed in 20–50 % of images. The radiographic measurements were performed twice: once based on the uncorrected landmarks and again based on the manually adjusted landmarks.

The automated method for radiographic measurements was created in Python v3.9.13 [65]. We implemented the following radiographic measurements: the acetabular depth-width ratio (ADR), the acetabular index (AI), the alpha angle (AA), the center edge angle of Wiberg (WCEA), the lateral center edge angle (LCEA), the extrusion index (EI), the neck-shaft angle (NSA), and the triangular index (TI), see Figure 1. Some of the measurements are based on similar features, such as the femoral head center, the femoral neck axis, and the horizontal reference line of the pelvis. Additionally, similar mathematical concepts are used in different measurements, such as the angle between two vectors. These general concepts will be reported before describing each radiographic measurement in detail. The output of the algorithm is the performed measurement and a visualization of the measurement. This visualization is created upon request by the user to provide insight into how the measurement was calculated. The workflow to calculate the radiographic measurements can be found in Figure 2.

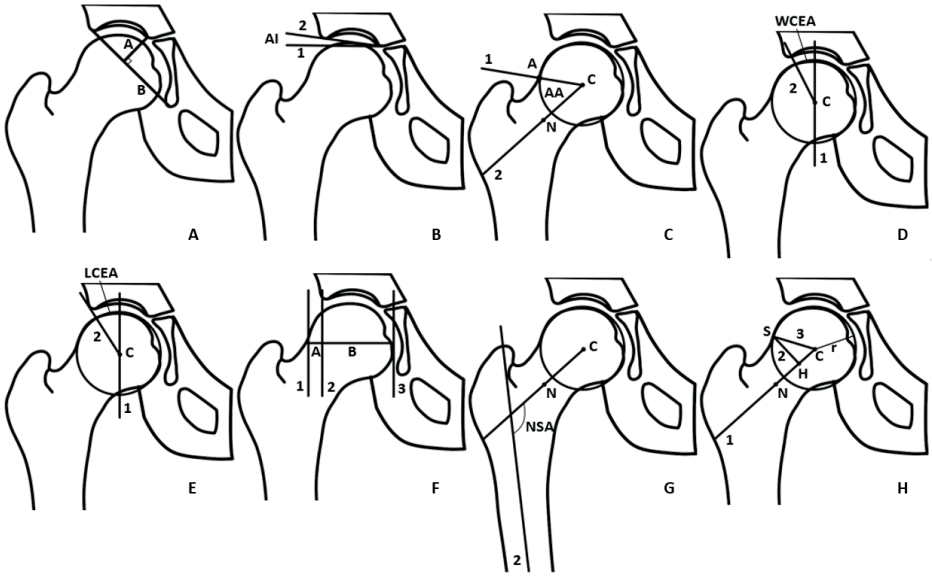


Figure 1. Definition of radiographic measurements of hip morphology implemented in this study. **A: The acetabular depth-width ratio (ADR)**, ADR = (A/B)*1000, is the ratio between the acetabular width (B) measured from the most lateral bony part of the acetabulum to the most inferior point of the teardrop and the acetabular depth (A) measured from the most medial point of the sourcil perpendicular to line B. **B: The acetabular index (AI)** describes the angle between the horizontal reference line of the pelvis (line 1) and line 2 through the most lateral bony part of the acetabulum and the most lateral point of the triradiate cartilage. **C: The alpha angle (AA)** is the angle between line 1 through the alpha point and the femoral head center, and the femoral neck axis (line 2). **D: The center edge angle of Wiberg (WCEA)** is the angle between line 1, a line through the femoral head center perpendicular to the horizontal reference line of the pelvis, and a line from the most lateral part of the sourcil to the femoral head center (line 2). **E: The lateral center edge angle (LCEA)** is the angle between line 1, a line through the femoral head center perpendicular to the horizontal reference line of the pelvis, and line 2 from the most lateral bony part of the acetabulum to the femoral head center. **F: The extrusion index (EI)**, EI = A/(A+B) * 100%, is the percentage of the part of the femoral head not covered by the acetabulum (A) compared to the entire width of the femoral head (A+B). **G: The neck-shaft angle (NSA)** is the angle between the femoral neck axis (line 1) and the femoral shaft axis (line 2). **H: The triangular index (TI)** is the length of line 3, the line between point S, a point at the intersection of the cortex of the femoral head and line 2, and the femoral head center (C).

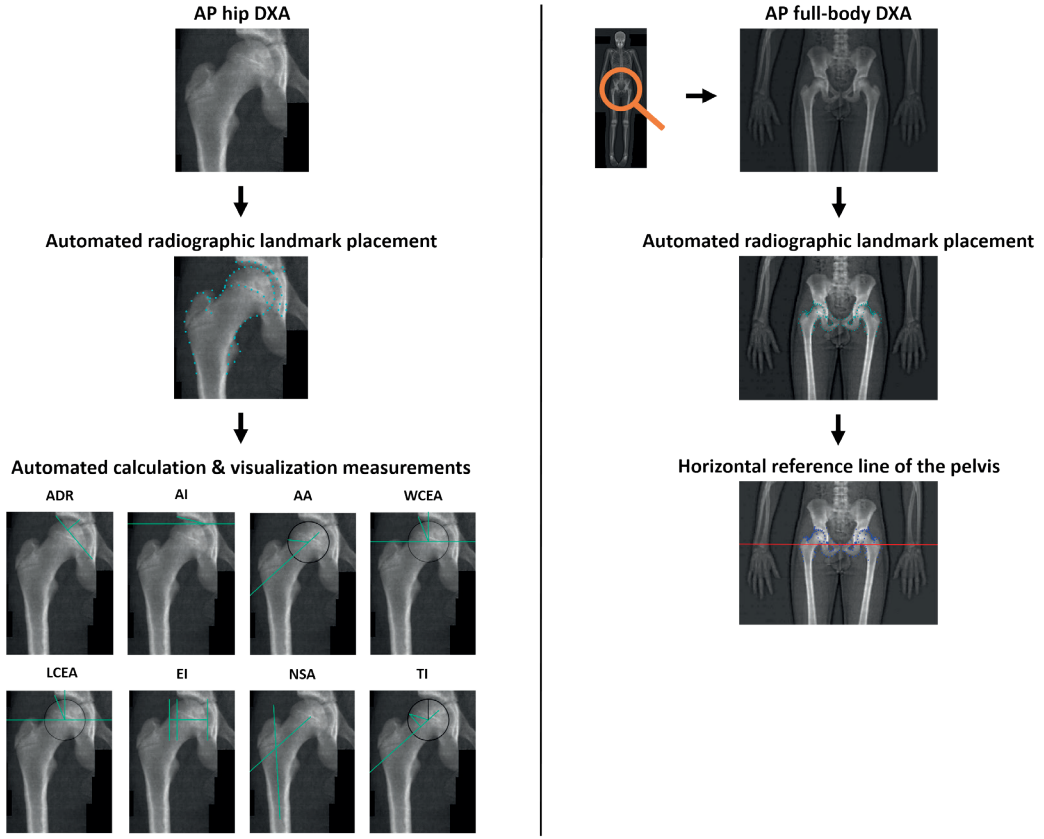
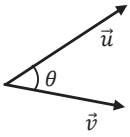


Figure 2. Overview of the workflow to obtain the automated radiographic measurements of hip morphology on the AP hip DXA, as well as the horizontal reference line of the pelvis on the AP full-body DXA. AP: anteroposterior. DXA: dual-energy x-ray absorptiometry. ADR: acetabular depth-width ratio. AI: acetabular index. AA: alpha angle. WCEA: center edge angle of Wiberg. LCEA: lateral center edge angle. EI: extrusion index. NSA: neck shaft angle. TI: triangular index.

General concepts

The angle between two vectors – We used Walker's[66] method for calculating the angle between two vectors to determine the angle between landmarks. Each vector was defined by two landmarks or originated in a landmark and pointed in the direction of one of the image's axes. The angle between the vectors can be calculated using Eq. (1).

(1)



$$\theta = \tan^{-1}\left(\frac{\|\vec{u}\times\vec{v}\|}{\vec{u}\cdot\vec{v}}\right),$$

The horizontal reference line of the pelvis – The horizontal reference line of the pelvis is determined automatically on the full-body DXA image to correct for potential pelvic obliquity. The correction of potential pelvic obliquity is applied to the AI, the WCEA, and the LCEA. Two landmarks on both hips were used: the most inferior point of the ischial tuberosity and the most superior point of the obturator foramen.

The slope of the line through each set of landmarks on both hips relative to the horizontal axis of the image was determined as the angle between two vectors, Eq. (1). The horizontal reference line of the pelvis was then determined to be the mean of these measurements. A negative slope indicates that the right hip is positioned more cranial than the left hip.

Femoral head center – The femoral head center is defined as the center of the best-fitting circle around the femoral head. We selected the hyper fit to determine the best-fitting circle since it offers fast calculation owing to its non-iterative nature[67]. Additionally, it has no essential bias and outperforms the geometric circle fit[67], which is viewed as the golden standard.

The circle fit was optimized for each hip by performing the calculations with nine different combinations of points (Fig. 3), to obtain the circle fit with the smallest root mean square error of the distances between the points and the circle with the smallest radius. A trade-off was made between these two features if this was not the same circle. This optimization was performed to prevent the best-fitting circle becoming too large and influenced by possible erroneous points.

Femoral neck axis – The femoral neck axis is the axis through the femoral head center and the femoral neck center. To determine the femoral neck center, the distance between all femoral neck landmarks was calculated. The femoral neck center was then determined as the center of all distances.

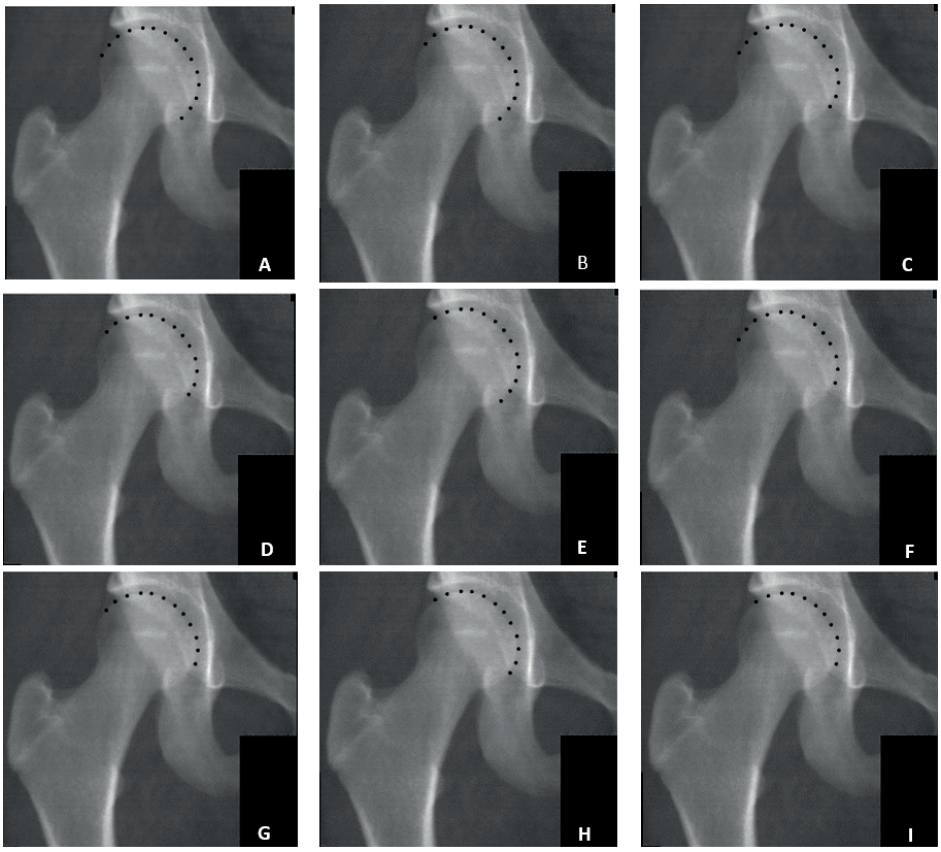


Figure 3. All possible combinations of radiographic landmark points used to define the best-fitting circle. **A**: All femoral head points. **B**: One point less on the lateral side of the femoral head. **C**: One point less on the medial side of the femoral head. **D**: One point less on both the lateral and the medial side of the femoral head. **E**: Two points less on the lateral side of the femoral head. **F**: Two points less on the medial side of the femoral head. **G**: One point less on the lateral and two points less on the medial side of the femoral head. **H**: Two points less on the lateral and one point less on the medial side of the femoral head. **I**: Two points less on both the lateral and the medial side of the femoral head.

Radiographic measurements of hip morphology

Acetabular depth-width ratio – The acetabular depth-width ratio (ADR) measures the acetabular depth. The ADR is the ratio of the length of the acetabular depth (A) to the entire length of the acetabular opening (B); see Eqs. (2-4). The length of the acetabular opening was measured as the distance between the most lateral point of the bony acetabulum (LA) and the most inferior point of the teardrop (TD). The acetabular depth was measured from the most medial point of the sourcil (MS), the weight-bearing part of the acetabulum, to line B (see Fig. 1A).

(2) $$ADR = \frac{A}{B} \bullet 1000,$$

$$A = \frac{|(x_{MS} - x_{LA})(y_{LA} - y_{TD}) + (x_{LA} - x_{TD})(y_{MS} - y_{LA})|}{\sqrt{(x_{LA} - x_{TD})^2 + (y_{LA} - y_{TD})^2}}, \quad (3)$$

$$B = \sqrt{(x_{LA} - x_{TD})^2 + (y_{LA} - y_{TD})^2} \quad (4)$$

Acetabular index – The acetabular index (AI), also known as the Tönnis angle, describes the acetabular roof inclination. The AI is the angle between the horizontal reference line of the pelvis, line 1, and the line bisecting the most lateral point of the triradiate cartilage and the most lateral bony point of the acetabulum, line 2 (Fig. 1B)[68]. The AI is determined by calculating the angle between the vector representing line 2 and a vector originating in the most lateral point of the triradiate cartilage parallel to the horizontal axis of the image, Eq. (1). The found AI is corrected for any potential pelvic obliquity using the horizontal reference line of the pelvis (HRLP) using Eq. (5) and (6) for the left and right hip respectively.

$$AI_{left\_hip} = AI_{uncorrected} - HRLP, \quad (5)$$

$$AI_{right\_hip} = AI_{uncorrected} + HRLP \quad (6)$$

Alpha angle – The alpha angle (AA) is used to detect asphericity (cam morphology) of the femoral head. The AA is the angle between the alpha point and the femoral neck axis (see Fig. 1C). The alpha point is defined as the point where the femoral head or femoral neck leaves the best-fitting circle. To simulate clinical practice, if only a small bony protrusion leaves the best-fitting circle but returns inside the best-fitting circle around the head-neck junction, the alpha point is indicated at the first point on the femoral neck that leaves the best-fitting circle (see Fig. 4). First, the best-fitting circle is calculated as described previously. To find the alpha point, all radiographic landmarks in the femoral head-neck area were investigated to see if any of them were outside of the best-fitting circle. Previous studies have reported a more oval-like shape of the femoral head in children aged 13 years, which might result in higher alpha angles while the femoral head seems spherical[69]. Therefore, an error margin of 4 % of the radius of the best-fitting circle is used to avoid false positives. An error margin in relation to the radius was chosen so that the error margin was not affected by the size of the hip joint. Error margins of 2–7 % were evaluated in the second training set of 500 images (containing 472 complete images), to optimize the detection of cam deformity without creating too many false positives. The AA measurement was assessed visually by an orthopedic surgeon (RA). The error margin of 4 % was chosen since it prevented under-detection of a cam deformity, with a false negative rate of 1 and false positive rate of 7 in the 472 images. To refine the alpha point detection, an interpolating B-spline is fitted through the lateral femoral head and neck points using the interpolate function from the Python SciPy-package[70]. This will allow for identification of the alpha point around the radiographic landmark,

which is the first point outside of the best-fitting circle. Lastly, the intersection of the best-fitting circle and the B-spline around this radiographic landmark is defined as the alpha point. Once the alpha point is identified, the AA is calculated using Eq. (1).

Center edge angle – The CEA is the angle between the vertical line through the femoral head center, perpendicular to the horizontal reference line of the pelvis, and the line tangential to the lateral margin of the acetabulum. On an AP hip DXA image, two types of CEA can be measured: the CEA of Wiberg (WCEA) and the lateral CEA (LCEA). The WCEA is measured from the most lateral point of the sourcil (weight-bearing part of the acetabulum) and represents the anterosuperior coverage of the femoral head (see Fig. 1D). The LCEA is measured from the acetabulum's most lateral bony point and represents the femoral head's superolateral coverage (see Fig. 1E).

To calculate the CEA, the center of the femoral head is determined as described previously. Next, we calculate the CEA as the angle between the vector from the center of the femoral head parallel to the vertical axis of the image and the vector from the center of the femoral head to the most lateral point of the bony acetabulum or sourcil (Eq. (1)). The found CEA is corrected for any potential pelvic obliquity using the horizontal reference line of the pelvis (HRLP) using Eqs. (7) and (8) for the left and right hip respectively.

(7) $$CEA_{left\_hip} = CEA_{uncorrected} + HRLP,$$

(8) $$CEA_{right\_hip} = CEA_{uncorrected} - HRLP$$

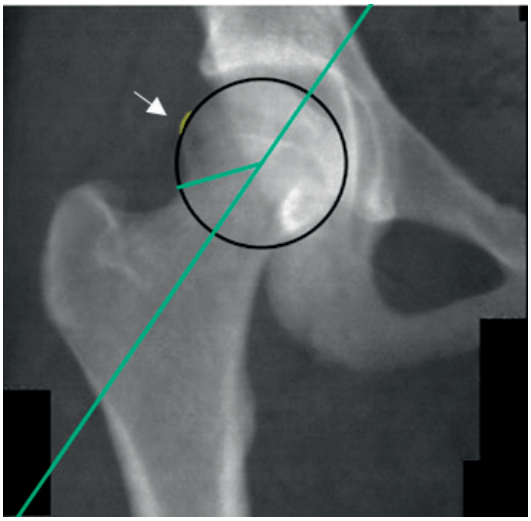


Figure 4. Example of small bony protrusion outside of the best-fitting circle (indicated by the arrow and highlighted in yellow), which returns inside the best-fitting circle around the head-neck junction. The alpha angle is indicated using the light green lines.

Extrusion index – The extrusion index (EI) describes the femoral head coverage by the acetabulum. The EI is the percentage of the total femoral head width (A + B) not covered by the bony acetabulum (A), (see Fig. 1F). The width of the femoral head is calculated as the difference in x-coordinate between the most lateral and most medial points of the femoral head. The uncovered part of the femoral head is calculated as the difference in x-coordinate between the most lateral point of the femoral head and the most lateral bony point of the acetabulum. Lastly, the EI can be determined using Eq. (9).

(9) $$EI = \frac{A}{A+B} \bullet 100\%$$

Neck shaft angle – The neck shaft angle (NSA) is a measure for coxa vara and valga and is the angle between the neck axis and the shaft axis (Fig. 1G). To determine the shaft axis, first, the image is cropped to below the minor trochanter to increase the accuracy of the shaft axis determination and to make the calculation faster. Next, the femoral shaft is segmented in the cropped image using multi-otsu thresholding with three levels for the Python Skicit-image package[71]. Next, we cleaned up the segmentation using morphological closing to remove noise and small artifacts, and smooth the edges of the segmentation result[72]. The shaft axis was defined as the center of the lateral and medial cortices. The shaft axis will only be determined if the shaft is visible for a length of at least half of the radius of the best-fitting circle below the minor trochanter. The NSA was calculated as the angle between two lines using the slope of the shaft axis (m_1) and the neck shaft axis (m_2) with Eq. (10).

(10) $$NSA = \tan^{-1}\left|\frac{m_2 - m_1}{1 + m_1 m_2}\right|$$

Triangular index – The triangular index (TI) measures the sphericity of the femoral head, see Figure 1H. Several structures need to be identified to calculate the TI:

1. The best-fitting circle of the femoral head, the femoral head center (point C), and the radius of the circle (r),
2. The neck axis,
3. Point H at a distance of half the head circle radius from the femoral head center, along the neck axis,
4. Point S at the cortex of the femoral head, on the line through point H perpendicular to the neck axis,
5. The TI is the distance between point S and the femoral head center.

A unit vector ($\vec{u}$) in the direction of the neck axis was used to find the coordinates of point H (Eq. (11)). The unit vector ($\vec{u}$) was calculated using Eq. (12), where vector $\vec{v}$ is the vector from the femoral head center to the femoral neck center. Next, the line perpendicular to the neck axis through point H was determined. This line could then be used to identify point S, using similar methods as finding the alpha point, namely finding the intersection between this perpendicular line and the B-spline through all femoral head neck points. Lastly, the TI was calculated as the distance between the femoral head center and point S.

(11) $$H = C + 0.5 \bullet r \bullet \vec{u}$$

(12) $$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$$

Manual determination of the radiographic measurements of hip morphology

On 30 right hip DXA scans, two experienced orthopedic surgeons performed a manual assessment of all the above described radiographic measurements of hip morphology at two different time points, at least one month apart. Both surgeons had access to a protocol providing precise descriptions of the measurements, (see Supplementary material 2). The measurements were performed using a DICOM viewer (Synedra View, Version 21.0.0, Synedra Information Technologies). The mean of all four measurements was used as the reference standard to which the automated method was compared.

Statistical analysis

The agreement within and between observers, between automated measurements on the automated landmarks and the manually adjusted landmarks, and between methods was investigated using Bland-Altman plots with limits of agreement. The Bland-Altman

plot shows the differences between the reference standard and the automated measurements over the average between the measurements for each individual hip. The limits of agreement were estimated using the 95 % CI of the differences between the reference standard and the automated measurements. Reliability was tested through intraclass correlation coefficients (ICCs) and reported with 95 % confidence intervals. Intraobserver reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. Interobserver reliability between both manual observer and between the automated and manually adjusted landmarks was tested with a 2-way random-effects model, single rater, absolute agreement ICC. Lastly, intermethod reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. ICCs were rated as poor (<0.50), moderate (0.50–0.75), good (0.76–0.90), or excellent (>0.90). All statistical analyses were performed using R Statistical Software (v4.1.0; R Core Team 2021). The ICCs were calculated using the irr-package[73] and the Bland-Altman plots were created using the ggplot2-package[74].

RESULTS

The mean age of the participants was 13.6 ± 0.2 years, and 18 (60 %) were female. The automated method was able to perform almost all radiographic hip measurements for all 30 hips. Except, however, for the AI measurement which could only be performed in 8 of the 30 hips, since only those 8 still had open triradiate cartilage.

Agreement

The Bland-Altman plots for agreement between the manual and automated measurements can be found in Figure 5. Bland-Altman inter-observer and intermethod agreement are presented in Table 1. The manual measurements were consistently higher than the automated measurements for the NSA based on the automated landmarks, and for the AI, the AA, the EI, the NSA, and the TI based on the manually adjusted landmarks. For the ADR, the WCEA, and the LCEA, there was almost no overall difference between the measurements. However, most measurements showed proportion errors, where the difference between the automated and manual measurements was dependent on the measurement size. The limits of agreement were mostly similar or smaller in the comparison of the manual and automated method, as well as in the comparison of the automated measurements on automated and manually adjusted landmarks, than in the comparison of the manual measurements of both observers. Only for the ADR and the AI the intermethod limits of agreement were larger than the interobserver.

Reliability

The intra- and interobserver and intermethod reliability for all measurements is shown in Table 2, as well as the mean absolute difference between the reference standard and the automated measurements. The intermethod reliability is comparable to or better than the interobserver reliability. Only for the acetabular index are manual measurements more reliable than the automated measurements. The interobserver ICCs from the automated measurements using the manually adjusted landmarks and the fully automated landmarks showed that the AI and AA were mostly influenced by the manual corrections. For all the other measurements, the ICCs ranged from good to excellent.

Table 1. Interobserver and intermethod agreement.

	Manual		Automated		Manual vs Automated			
	Observer 1 vs observer 2		Automated vs manual adjusted landmarks		Automated landmarks		Manual adjusted landmarks	
Measurement	Interobserver bias (mean ± SD)	Interobserver limits of agreement	Interobserver bias (mean ± SD)	Interobserver limits of agreement	Intermethod bias (mean ± SD)	Intermethod limits of agreement	Intermethod bias (mean ± SD)	Intermethod limits of agreement
Acetabular depth-width ratio	-27.1 ± 18.5	-63.4 to 9.2	1.2 ± 11.9	-22.0 to 24.5	-0.1 ± 21.8	-42.9 to 42.8	-1.3 ± 17.4	-35.5 to 32.9
Acetabular index*	-0.4° ± 2.4°	-5.2° to 4.4°	-3.3° ± 3.3°	-9.8° to 3.1°	-1.1° ± 4.1°	-9.1° to 6.9°	2.3° ± 2.5°	-2.5° to 7.1°
Alpha angle	0.4° ± 4.9°	-9.3° to 10.0°	-2.2° ± 6.0°	-14.0° to 9.6°	0.3° ± 5.5°	-10.4° to 11.0°	2.5° ± 3.5°	-4.3° to 9.3°
Center edge angle of Wiberg	-2.5° ± 2.6°	-7.6° to 2.6°	-0.9° ± 2.3°	-5.4° to 3.5°	-0.6° ± 2.1°	-4.7° to 3.5°	0.3° ± 2.4°	-4.4° to 5.1°
Lateral center edge angle	-3.5° ± 2.3°	-8.0° to 1.0°	-0.2° ± 0.8°	-1.7° to 1.4°	-0.7° ± 1.6°	-3.8° to 2.4°	-0.6° ± 1.3°	-3.1° to 1.9°
Extrusion index	3.0 ± 2.5	-1.8 to 7.8	-0.6 ± 1.5	-3.5 to 2.3	0.5 ± 1.8	-3.1 to 4.0	1.1 ± 1.7	-2.1 to 4.3
Neck shaft angle	4.4° ± 3.0°	-1.4° to 10.3°	0.3° ± 1.8°	-3.3° to 3.9°	5.2° ± 2.3°	0.6° to 9.7°	4.9° ± 2.4°	0.2° to 9.5°
Triangular index	0.2 ± 3.3	-6.4 to 6.8	-1.0 ± 1.3	-3.6 to 1.6	0.1 ± 2.7	-5.2 to 5.3	1.0 ± 2.5	-4.0 to 6.0

Bland-Altman interobserver and intermethod bias (mean and standard deviation) and limits of agreement, n=30. *The acetabular index was only measured in hips with an open triradiate cartilage (n=8).

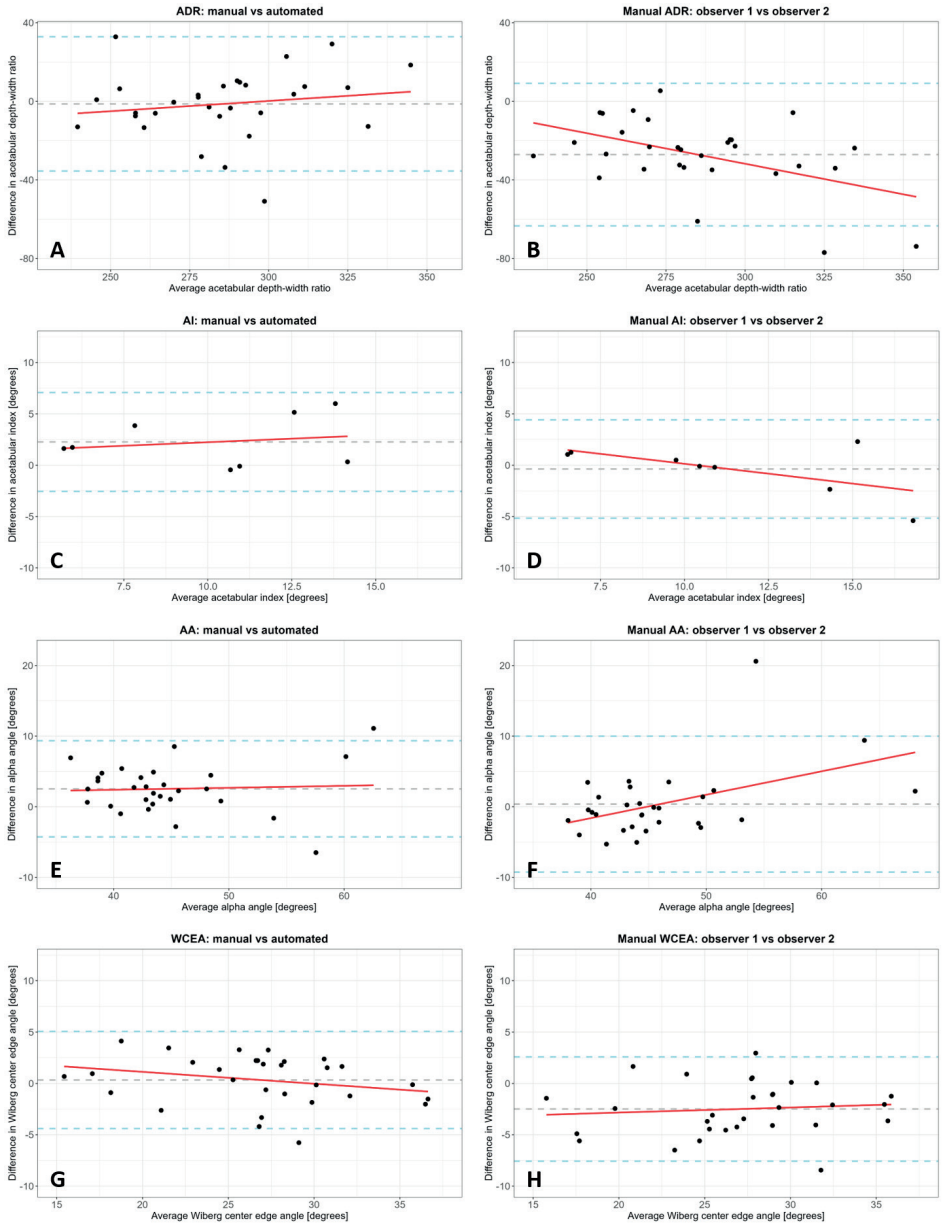


Figure 5. Bland-Altman plots of the mean manual vs. automated hip morphology measurements and the manual measurement as performed by observer 1 vs. observer 2. **A: The acetabular depth-width ratio (ADR)** – manual vs automated. **B: ADR** – observer 1 vs observer 2. **C: The acetabular index (AI)** – manual vs automated. The AI was only measured in hips with an open triradiate cartilage (n=8). **D: AI** – observer 1 vs observer 2. The AI was only measured in hips with an open triradiate cartilage (n=8). **E: The alpha angle (AA)** – manual vs automated. **F: AA** – observer 1 vs observer 2. **G: The center edge angle of Wiberg (WCEA)** – manual vs automated. **H: WCEA** – observer 1 vs observer 2.

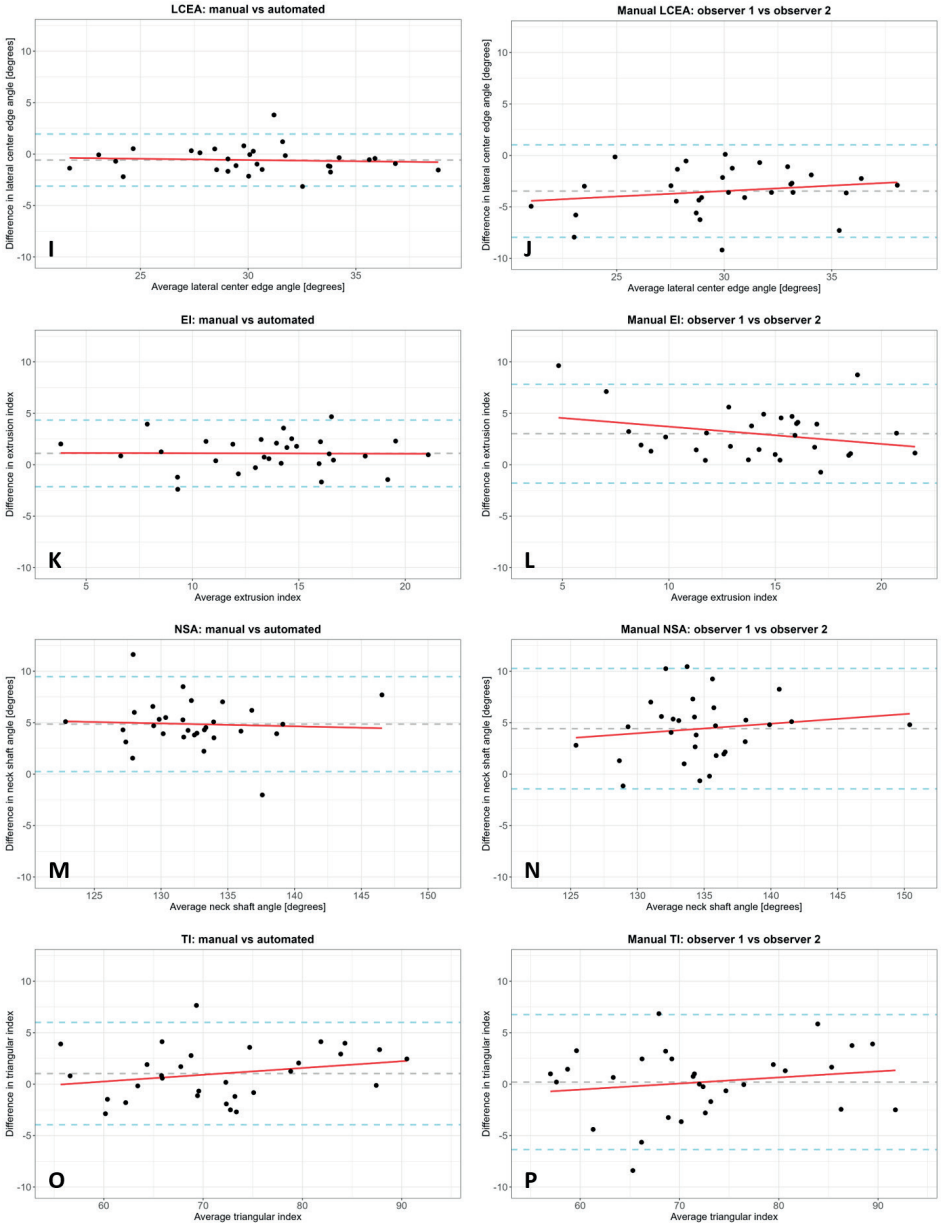


Figure 5. - continued. **I: The lateral center edge angle (LCEA)** – manual vs automated. **J: LCEA** – observer 1 vs observer 2. **K: The extrusion index (EI)** – manual vs automated. **L: EI** – observer 1 vs observer 2. **M: The neck-shaft angle (NSA)** – manual vs automated. **N: NSA** – observer 1 vs observer 2. **O: The triangular index (TI)** – manual vs automated. **P: TI** – observer 1 vs observer 2.

Table 2. Comparing the performance of two manual raters and the automated measurements.

	Manual		Automated		Manual vs Automated			
	Observer 1	Observer 2	Observer 1 vs observer 2	Automated vs manual adjusted landmarks	Automated landmarks		Manual adjusted landmarks	
Measurement	Intraobserver ICC	Intraobserver ICC	Interobserver ICC	Interobserver ICC	Intermethod ICC	Mean absolute difference [range]	Intermethod ICC	Mean absolute difference [range]
Acetabular depth-width ratio	0.72 (0.50 – 0.86)	0.57 (0.26 – 0.77)	0.58 (0 – 0.85)	0.90 (0.80 – 0.95)	0.69 (0.44 – 0.84)	17.2 [1.7 – 47.9]	0.80 (0.62 – 0.90)	12.6 [0.44 – 50.9]
Acetabular index*	0.60 (0 – 0.92)	0.99 (0.66 – 0.998)	0.83 (0.36 – 0.96)	0.58 (0 – 0.90)	0.62 (0 – 0.91)	2.9° [0.08° – 8.0°]	0.65 (0 – 0.92)	2.4° [0.1° – 6.0°]
Alpha angle	0.89 (0.78 – 0.95)	0.74 (0.52 – 0.87)	0.77 (0.58 – 0.89)	0.68 (0.43 – 0.84)	0.77 (0.56 – 0.88)	3.7° [0.05° – 18.2°]	0.81 (0.46 – 0.92)	3.4° [0.08° – 11.1°]
Center edge angle of Wiberg	0.85 (0.70 – 0.92)	0.83 (0.66 – 0.91)	0.80 (0.25 – 0.93)	0.91 (0.81 – 0.96)	0.92 (0.84 – 0.96)	1.6° [0.05° – 6.2°]	0.91 (0.81 – 0.95)	2.0° [0.1° – 5.8°]
Lateral center edge angle	0.79 (0.60 – 0.89)	0.86 (0.73 – 0.93)	0.65 (0 – 0.89)	0.98 (0.96 – 0.99)	0.92 (0.81 – 0.96)	1.4° [0.03° – 4.5°]	0.95 (0.87 – 0.98)	1.1° [0.05° – 3.8°]
Extrusion index	0.71 (0.47 – 0.85)	0.86 (0.72 – 0.93)	0.66 (0 – 0.88)	0.93 (0.85 – 0.97)	0.91 (0.81 – 0.95)	1.6 [0.29 – 3.4]	0.89 (0.66 – 0.95)	1.6 [0.1 – 4.7]
Neck shaft angle	0.69 (0.19 – 0.88)	0.87 (0.75 – 0.94)	0.58 (0 – 0.85)	0.94 (0.87 – 0.97)	0.60 (0 – 0.88)	5.3° [0.20° – 9.5°]	0.57 (0 – 0.86)	5.0° [1.6° – 11.6°]
Triangular index	0.81 (0.61 – 0.91)	0.95 (0.90 – 0.98)	0.94 (0.89 – 0.97)	0.98 (0.93 – 0.99)	0.96 (0.92 – 0.98)	2.2 [0.15 – 5.2]	0.96 (0.91 – 0.98)	2.2 [0.13 – 7.7]

Intraclass correlation coefficients (ICC) of intra- and interobserver, and intermethod reliability of the radiographic measurements of hip morphology, n=30. ICCs are presented with 95% CI. Intraobserver reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. Interobserver reliability was tested with a 2-way random-effects model, single rater, absolute agreement ICC. Intermethod reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. Interpretation: poor (<0.50), moderate (0.50-0.75), good (0.76-0.90), or excellent (>0.90). *The acetabular index was only measured in hips with an open triradiate cartilage (n=8).

DISCUSSION

We presented an open-access, automated method for determining radiographic measurements of hip morphology on hip DXA images, and evaluated the agreement and reliability compared to manual assessments. We automatically calculated the ADR, AI, AA, CEA, EI, NSA, and TI for all 30 images assessed. The automated measurements were calculated based on a set of radiographic landmarks describing the shape of the acetabulum and proximal femur. The agreement between the automated and manual measurements was similar to or better than the agreement between two manual observers, with respect to the width of the 95 % CI of the intermethod differences. For most measurements, there was a bias towards smaller values for the automated method compared to the manual measurements. The intermethod reliability was comparable to or better than the manual interobserver reliability.

The AA showed both a larger 95 % CI for the agreement as well as erratic behavior in the higher values. This is likely the result of the difficulty in the correct identification of the alpha point. This is especially true if some asphericity seems to be present and thus a higher alpha angle is found. In clinical practice this could lead to an under- or overestimation of the cam deformity. The ADR was the measurement with the largest 95 % CI around the difference for both the intermethod and the interobserver analyses. This is likely because the ADR is a ratio that is tenfold bigger than the other ratio measurements. Further, the AI was only measured in a limited number of hips, since only 8 hips in our study population had an open triradiate cartilage. This makes it difficult to judge the performance of the automated AI measurement. The modified AI, which is measured from the most medial part of the weight-bearing part of the acetabulum instead of the most lateral part of the triradiate cartilage, will allow for measurement of the acetabular roof inclination in hips with a closed triradiate cartilage. Lastly, interpreting the measurement performance of TI can be difficult as it is a length measurement which is also dependent on the size of the hip of the individual. An alternative is the TI ratio, also known as the Gosvig ratio[75], which is the ratio of the TI to the radius of the best-fitting circle around the femoral head.

Some methods are presently available for the automatic determination of radiographic morphology measurements of the hip[76-80]. However, most of these algorithms are not open-source and do not have clear descriptions of how the measurements are calculated. This makes reproducibility and implementation in clinical research and practice impossible. Faber et al.[76] developed an open-access method determining the AA on DXA images based on radiographic landmarks and reported a concordance correlation coefficient between the automated method and the manual measurement of 0.88 (95% CI

0.84–0.92). Three articles used the AI software HIPPO and validated the software for the WCEA, NSA, EI, and AI against manual measurements[77-79]. The ICCs were comparable to those reported in the present study. Lastly, Jensen et al.[80] investigated the reliability and agreement of the RBhip software, which was able to measure the WCEA and the AI, also showing a bias in the WCEA measurement.

The proposed automated methods have some limitations. There is no gold standard for determination of radiographic measurements of hip morphology. Therefore, we needed to create a reference standard to assess the software's performance. We believe that the created reference standard reflects current clinical practice since it was created by two orthopedic surgeons, who are the specialists who ultimately work with these measurements and use them to decide on diagnoses and treatment. Moreover, the proposed automated calculations need radiographic landmarks describing the shape of the hip as the input.

The creation of these landmark sets can be a very time-consuming process. However, the radiographic landmark set for this study was created automatically using the BoneFinder® software. Additionally, the algorithm's performance depends on the quality of the landmark set provided as input. It should be taken into account that correct landmark placement influences the performance of the ADR, AI and AA the most. Before implementation of the fully automated pipeline in clinical practice, the ASM needs to be improved by additional training using images of different databases. Another limitation for the AA needs to be noted. The selection of the error margin needed for the determination of the alpha point was based on subjective visual inspection and should be confirmed in populations with a higher prevalence of cam morphology. Lastly, the presented method was created and validated on only 30 right hip DXA images in 13-year-old subjects. The generalizability of the automated method in other populations is therefore limited. Nevertheless, we believe that following validation this automated method could also be applied to radiographs.

We think that this algorithm is a promising tool for performing automated radiographic measurements of hip morphology. The fact that it is automated, but still very insightful due to the use of radiographic landmarks, makes it feasible to analyze vast amounts of data. This makes the software highly applicable for large population studies. Additionally, it can be used in clinical practice, where the user can see how the measurement is performed based on the output image if desired, making it more insightful. Another application of these automated measurements could be more standardized recruitment for trials pertaining to certain hip morphologies. The use of automated measurements could reduce selection bias. DXA images allow for x-ray-like imaging with a lower radia-

tion burden. A full-body DXA, however, is also needed when using hip DXA images to be able to correct for the potential obliquity of the pelvis.

In conclusion, the proposed algorithms allow for fast and reproducible calculation of radiographic measurements of hip morphology on right hip DXA images. Furthermore, by providing open access to the algorithms, we aimed at transparency and provided the opportunity for better inter-research comparison. This can help advance insights into the morphology and development of the hip, as well as provide information on the development and risk of diseases such as hip OA.

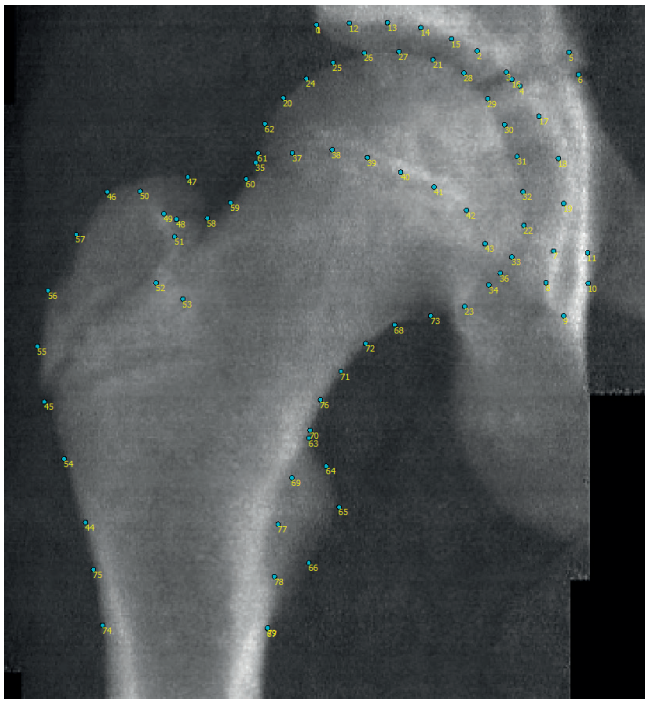
Software availability

The in-house developed methods can be accessed here: https://github.com/FleurBoel/Automated-Hip-Morphology-Measurements/tree/main. License: Apache 2.0 license. BoneFinder® and the model for automatic point placement are freely available from the website (www.bone-finder.com; The University of Manchester, UK)[64].

Acknowledgements

This research is supported by the Dutch Arthritis Association [grant number 19-2-202, 2019]. The Generation R Study is conducted by the Erasmus MC, University Medical Center Rotterdam in close collaboration with the Erasmus University Rotterdam and the city of Rotterdam. We gratefully acknowledge the contribution of children and parents. The general design of Generation R Study is made possible by long-term financial support from Erasmus MC, University Medical Center Rotterdam, the Netherlands, Organization for Health Research and Development (ZonMw) and the Ministry of Health, Welfare and Sport. CL is funded by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (223267/Z/21/Z). This research was funded in whole, or in part, by the Wellcome Trust [Grant number 223267/Z/21/Z]. For the purpose of open access, the authors have applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

SUPPLEMENTARY MATERIAL 1: SHAPE PROTOCOL



Acetabulum

Note that the middle of the points needs to be on the lower border of the acetabulum.

Point(0): lateral edge acetabulum
Point(1): lateral edge sourcil
Point(2): medial edge sourcil, the medial edge of the weight baring part.
Point(3-4): 2 points at the lateral side of the triradiate cartilage, at the superior respectively inferior edge. If not visible put them in the same location as each other where you perceive that the triradiate used to be.
Point(5-6): 2 points at the medial side of the triradiate cartilage, at the superior respectively inferior edge. If not visible put them in the same location as each other where you perceive that the triradiate used to be.

Teardrop

Point(7): On the superolateral corner of the visible teardrop (on the wall of the acetabular fossa). On the sharp concavity where the teardrop starts.
Point(8): On the most lateral point of the teardrop or evenly spaced between previous and next point.
Point(9): On the most caudal point of the teardrop.
Point(10-11): 2 points mirrored to the opposite side, parallel to the horizontal axis of the image.
Point(12-15): 4 Points between point(1) and point(2)
Point(16-19): 4 Points between point(2) and point(7).

Femoral head

Note that the middle of the points needs to be on the most outer line of the femoral head.

Point(20): On the superolateral side of the femoral head, where the "best fitting circle" around the convexity of the femoral head seems to start. In case of an irregularity: place point(20) right after this bump ends, and the circle begins.

Point (23): On the inferomedial side of the femoral head, where the convexity of the femoral head seems to end. (The neck bends off after this point).

Points(24-27, 21, 28-32, 22, 33-34): 13 points evenly spaced between point(20) and points(23), following the convexity of the femoral head.

Femoral head growth plate

If visible:

- Note that these points should be placed on the lower edge of the sclerotic/irregular line representing the growth plate.
- Point(35): lateral point growth plate.
- Point(36): medial point growth plate.
- Point(37-43): 7 points between point(38) and point(39)

If not visible:

- Note in sheet, and place the 7 points along the estimated growth plate.

Trochanters

Point (44): Most proximal point of the lateral cortex.

Trochanter major

Point(45): On the lower lateral corner of the greater trochanter, if visible on the middle of the growth plate
Point(46): On the upper lateral corner of the posterior greater trochanter
Point(47): On the medial upper corner of the posterior greater trochanter. If not visible, equally spaced between (46) and (48).
Point (48): Where the greater trochanter joins the femoral neck (usually at an angle and at a sclerotic corner).
** If the posterior greater trochanter is somehow not visible: overlap points with anterior trochanter.
Point (49): On the upper medial corner of the anterior greater trochanter.
Point (50): Between (46) and (49), following the contour. If there is a clear angle, put it there.
Point (51): On the medial corner of the posterior greater trochanter, where it starts to drop downwards (caudal). This is independent of the femoral neck, so it can be before or after it dips behind the femoral neck, depending on the rotation of the proximal femur.
Point (52): Where the posterior greater trochanter is dropping straight down, right before it bends medially
Point (53): On the end of the sclerotic line right after the medial bend, following the contour of the posterior greater trochanter.
Point(54): Between point(44) and point(45).
Point(55-57): 3 Between point(45) and point(46).

Lateral edge femoral neck

Point(58-62): 5 Between point(48) and point(20).

Trochanter minor

If visible:

- Point(63): Where the lesser trochanter seems to start proximally, can be behind the shaft.
- Point(64): Between point(63) and point(65)
- Point(65): Most medial point trochanter minor.
- Point(66): Between point(65) and point(67)
- Point(67): Where the lesser trochanter seems to end distally, can be behind the shaft.

If not visible:

- Point(63-67): Same spacing as points (68-76) to distal from point(76).

Medial edge femoral neck

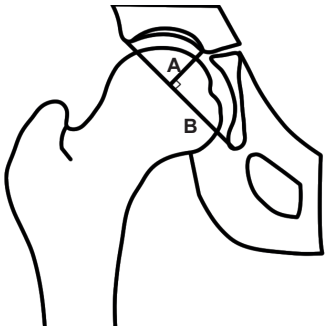
Point (68): At the deepest point of the inferomedial concavity of the femoral neck, so that (68) will follow the medial cortex of the femoral neck as closely as possible.
Point(71-72): 2 Between point(76) and point(68).
Point(73): Between point(68) and point(23).

Shaft

Point(74): Mirrored to point(67) along the shaft axis.
Point(75): Equally spaced between point(44) and point(74)
Point(76): Mirrored to point(45) along the shaft axis.
Point(70, 69, 77, 78): 4 Equally spaced between point(76) and point(79)
Point(79): On the shaft, in line with the shaft axis through point(67)

SUPPLEMENTARY MATERIAL 2: PROTOCOL FOR MANUAL MEASUREMENTS

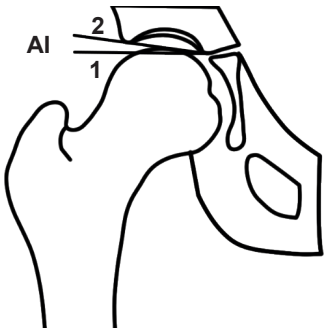
Acetabular depth-width ratio (ADR)



Ratio of the length of the acetabular depth to the entire length of the lateral acetabular opening.

1. Draw line B connecting the most lateral point of the acetabulum and the most inferior point of the teardrop.
2. Draw line A perpendicular to line B through the most medial point of the sourcil.
3. $ADR = \frac{A}{B} \bullet 1000$

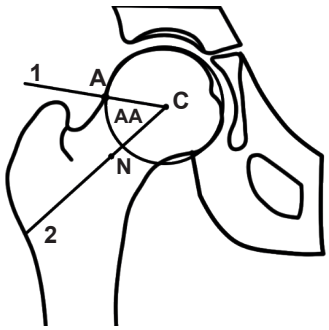
Acetabular index (AI)



Only determine the AI if triradiate cartilage is still open.

1. Draw line 1 horizontal to the image axis through the most lateral point of the triradiate cartilage.
2. Draw line 2 bisecting the most lateral point of the triradiate cartilage and the most lateral bony point of the acetabulum.
3. The AI is the angle between these lines.

Alpha angle

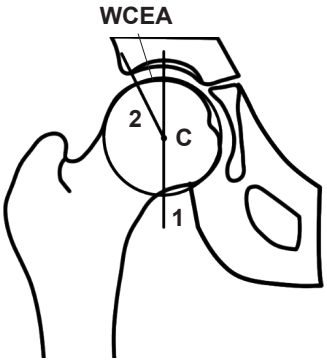


The angle between the longitudinal axis of the femoral neck and the line between the alpha point and the femoral head center.

1. Draw the best fitting circle of the femoral head. The **femoral head center** (point C) is defined as the center of this circle.
2. Define the alpha point (point A). This is the point where the femoral head first exits the best fitting circle on the lateral side of the femoral head.
3. Define the femoral neck center (point N). This is the mid- point on a line drawn across the narrowest part of the femoral neck.
4. Draw line 1 connecting the alpha point and the femoral head center.
5. Draw line 2 connecting the femoral neck center and the femoral head center.
6. The alpha angle (AA) is the angle between line 1 and line 2.

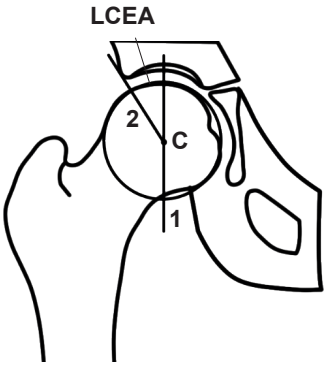
Center edge angle (CEA)

Center edge angle of Wiberg



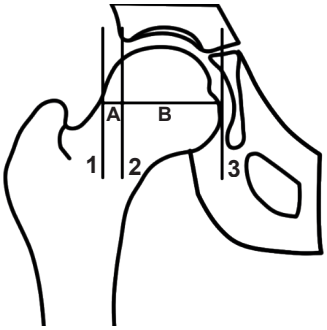
1. Draw the best fitting circle of the femoral head, and the define the femoral head center (point C).
2. Draw line 1 vertically through the femoral head center, parallel to the image.
3. Draw line 2 connecting the femoral head center and the most lateral point of the sourcil (weight bearing part of the acetabulum).
4. The CEA of Wiberg (WCEA) is the angle between line 1 and 2.

Lateral center edge angle



1. Draw the best fitting circle of the femoral head, and the define the femoral head center (point C).
2. Draw line 1 vertically through the femoral head center, parallel to the image.
3. Draw line 2 connecting the femoral head center and the most lateral bony point of the acetabulum.
4. The lateral CEA (LCEA) is the angle between line 1 and 2.

Extrusion index

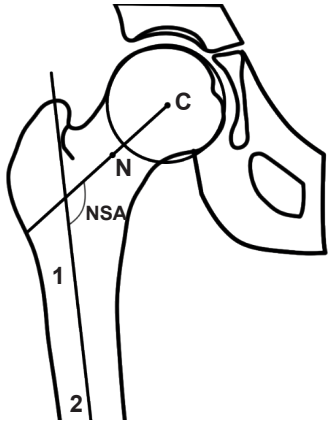


The extrusion index (EI) is a radiographic measurement of femoral head coverage by the acetabulum.

1. Draw line 1 vertical through the most lateral point of the femoral head.
2. Draw line 2 vertical through the most lateral bony point of the acetabulum.
3. Draw line 3 vertical through the most medial point of the femoral head.
4. A: distance between line 1 and line 2
5. B: distance between line 2 and line 3

$EI = \frac{A}{A+B} \bullet 100\%$

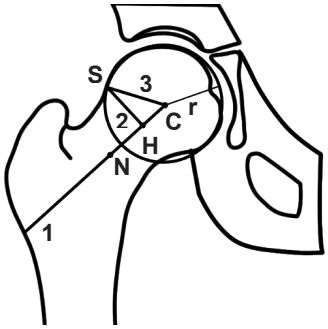
Neck shaft angle



Angle between the line through the femoral head center and the femoral neck center, and the line drawn along the femoral shaft axis.

1. Draw the best fitting circle of the femoral head, and the define the femoral head center (point C).
2. Define the femoral neck center (point N). This is the mid- point on a line drawn across the narrowest part of the femoral neck.
3. Draw line 1 connecting the femoral head center and the femoral neck center.
4. Draw line 2 along the femoral shaft axis.
5. The neck shaft angle is the angle between line 1 and 2.

Triangular index (TI)



1. Draw the best fitting circle of the femoral head, and the define the femoral head center (point C).
2. Define the femoral neck center (point N). This is the mid- point on a line drawn across the narrowest part of the femoral neck.
3. Draw line 1 connecting the femoral head center and the femoral neck center.
4. Measure the radius (r) of the best fitting circle.
5. Define point H along line 1 at a distance of half of the radius (r) from the femoral head center (point C).
6. Draw line 2 from point H to the superior cortex of the femoral head, point S, perpendicular to line 1.
7. Draw line 3 from the center of the femoral head (point C) to point S, and measure the length of line 3.
8. The triangular index is the length of line 3.

Chapter 3

Reliability and agreement of manual and automated morphological radiographic hip measurements

F. Boel*, N.S. Riedstra*, J. Tang, D.F. Hanff, H. Ahedi, V. Arbabi, N.K. Arden, S.M.A. Bierma-Zeinstra, M.M.A. van Buuren, F.M. Cicuttini, T.F. Cootes, K. Crossley, D. Eygendaal, D.T. Felson, W.P. Gielis J. Heerey, G. Jones, S. Kluzek, N.E. Lane, C. Lindner, J. Lynch, J.B.J. van Meurs, A.E. Nelson, A. Mosler, M.C. Nevitt, E.H. Oei, J. Runhaar, H. Weinans, R. Agricola.

*Shared first authorship

Osteoarthritis and Cartilage Open. 2024 Sept;6(3):100510.

ABSTRACT

Objective: To determine the reliability and agreement of manual and automated morphological measurements, and agreement in morphological diagnoses.

Methods: Thirty pelvic radiographs were randomly selected from the World COACH consortium. Manual and automated measurements of acetabular depth-width ratio (ADR), modified acetabular index (mAI), alpha angle (AA), Wiberg center edge angle (WCEA), lateral center edge angle (LCEA), extrusion index (EI), neck-shaft angle (NSA), and triangular index ratio (TIR) were performed. Bland-Altman plots and intraclass correlation coefficients (ICCs) were used to test reliability. Agreement in diagnosing acetabular dysplasia, pincer and cam morphology by manual and automated measurements was assessed using percentage agreement. Visualizations of all measurements were scored by a radiologist.

Results: The Bland-Altman plots showed no to small mean differences between automated and manual measurements for all measurements except for ADR. Intraobserver ICCs of manual measurements ranged from 0.26 (95%-CI 0-0.57) for TIR to 0.95 (95%-CI 0.87-0.98) for LCEA. Interobserver ICCs of manual measurements ranged from 0.43 (95%-CI 0.10-0.68) for AA to 0.95 (95%-CI 0.86-0.98) for LCEA. Intermethod ICCs ranged from 0.46 (95%-CI 0.12-0.70) for AA to 0.89 (95%-CI 0.78-0.94) for LCEA. Radiographic diagnostic agreement ranged from 47% to 100% for the manual observers and 63%-96% for the automated method as assessed by the radiologist.

Conclusion: The automated algorithm performed equally well compared to manual measurement by trained observers, attesting to its reliability and efficiency in rapidly computing morphological measurements. This validated method can aid clinical practice and accelerate hip osteoarthritis research.

INTRODUCTION

There is evidence that hip morphology is a leading contributing factor to the development of hip osteoarthritis (OA)[11]. Furthermore, studies have shown that specific hip morphologies, such as acetabular dysplasia (undercoverage of the femoral head by the acetabulum), pincer morphology (excessive coverage of the femoral head by the acetabulum) and cam morphology (aspherical femoral head) are associated with radiographic hip OA[11,30,44,81-83].

In order to quantify hip morphology, morphological measurements can be performed on pelvic anteroposterior (AP) radiographs, which are inexpensive and routinely obtained in clinical practice. Manual morphological measurements, however, are time-consuming and can be unreliable when performed by different observers[55]. Additionally, a lack of consistency exists in the current definitions for some morphological measurements[84].

Automated morphological measurements could enhance reproducibility while facilitating rapid assessment of multiple measurements per radiograph. Automation, therefore, has the potential to aid clinical practice and allows for the quantification of hip morphology in large cohort studies. There are currently few open-access, publicly available algorithms, and those that are available are sometimes poorly described[76-78].

We aim to study the reliability and agreement of manual and our inhouse developed, open-access, automated morphological hip measurements through quantitative and qualitative assessment of both methods. This ensures that results from future studies where this automated method is applied are clinically relevant. The secondary aim was to assess the agreement in making radiographic morphological diagnoses based on manual and automated measurements.

METHODS

Participants

The Worldwide Collaboration of OsteoArthritis prediCtion of the Hip (World COACH) consortium is a global collaboration of all prospective cohort studies with available sequential pelvic or hip imaging. The included cohorts are Cohort Hip and Cohort Knee, the Multi-center OSteoarthritis sTudy, the OsteoArthritis Initiative, the Rotterdam Study-I, the Rotterdam Study-II, the Rotterdam Study-III, the Chingford Study, the Johnston County Project, the Study of Osteoporotic Fractures, and the Tasmanian Older Adults Cohort. The World COACH consortium currently counts 37,732 participants aged 42–100

(mean 65.72 years) at baseline, and 71.33% are female individuals. The consortium profile and protocol have previously been published in detail[48]. From the consortium,

30 baseline radiographs were selected proportionate to the cohort size in the consortium for qualitative and quantitative assessment of the manual and automated morphological measurements. A power analysis was performed assuming type I errors of 0.05, type II errors of 0.20, two replications, a minimally acceptable level of reliability of 0.75 and an expected level of reliability between 0.8 and 0.9, a minimum of 27 inclusions was needed. Therefore, we selected a total of 30 random radiographs for inclusion[85]. A flowchart of the radiograph selection is shown in Figure 1. The baseline characteristics were: 18 females (60%), the mean age was 62.5 ± 8.6 years (range 47–78), and the mean BMI was 26.5 ± 3.9 kg/m2. All included hips had no definite RHOA as defined by Kellgren and Lawrence classification, modified Croft classification or modified OA score of 0 or 1.

Radiographs

The AP pelvic radiographs were obtained according to a protocol previously decided on by each cohort, and details on cohort-specific radiographic protocols can be found in the World COACH description paper[48]. Seven cohorts (CHECK, MOST, OAI, RS-I, RS-II, RS-III, TASOAC) contained weight-bearing AP pelvic radiographs. In contrast, three cohorts (the Chingford Study, JoCo, and SOF) contained supine AP pelvic radiographs.

Hip morphology and morphological measurements

Morphological measures used in this manuscript to determine acetabular dysplasia include the acetabular depth-width ratio (ADR), the modified acetabular index (mAI), the Wiberg center edge angle (WCEA), and the extrusion index (EI)[23,86,87]. The lateral center edge (LCEA) angle determined pincer morphology[8,10,58]. Cam morphology was defined by the alpha angle (AA) and the triangular index ratio (TIR)[30,88,89]. The neck-shaft angle (NSA) is used to determine coxa valga and vara[90]. All measurements are shown in Figure 2 and are explained in detail elsewhere[91]; a brief overview, including radiological thresholds for radiographic diagnosis, is provided below.

Acetabular depth-width ratio

The acetabular depth-width ratio (ADR) quantifies the depth of the acetabulum. The acetabular width was defined by a line from the lateral bony edge of the acetabulum to the pelvic teardrop to measure the acetabular opening. Next, the acetabular depth was defined by a line perpendicular to the acetabular width, extending from the most medial point of the sourcil (Fig. 2B). The ADR is the depth ratio to the width multiplied by 1000. Acetabular dysplasia is diagnosed by an ADR ≤ 250 [92].

Modified Acetabular Index

The mAI measures the acetabular roof's inclination. The original acetabular index is applied to hips with an open triradiate cartilage; a modified version was created to obtain this measurement in adults. The mAI measures the angle between the line from the medial sourcil to the lateral bony edge of the acetabulum and the horizontal reference line of the pelvis (Fig. 2C). Acetabular dysplasia is defined by mAI ≥ 13°, acetabular overcoverage is defined by mAI ≤ 3° [92,93].

Wiberg Center Edge Angle

The degrees of weight-bearing coverage of the femoral head by the acetabulum is measured by the WCEA[92]. The WCEA is formed by a vertical line through the center of the femoral head, perpendicular to the horizontal reference line of the pelvis, and a second line from the center of the femoral head to the most lateral weight-bearing part of the sourcil (Fig. 2E). Although the threshold has been debated, acetabular dysplasia is generally defined by a WCEA ≤ 25° in prospective studies[8,9,11,52].

Lateral Center Edge Angle

The degrees of bony coverage of the femoral head by the acetabulum is measured by the LCEA[11,30,94]. The LCEA is formed by a vertical line through the center of the femoral head, perpendicular to the horizontal reference line of the pelvis, and a second line from the center of the femoral head to the most lateral bony part of the acetabulum (Fig.2F). Pincer morphology is generally defined by an LCEA ≥ 40° in prospective studies[10,11].

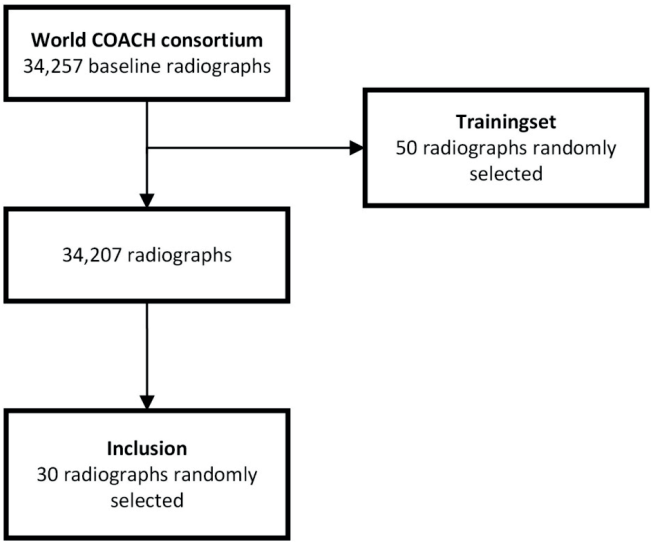


Figure 1. Flowchart of the radiograph selection.

Extrusion Index

The EI quantifies bony femoral head coverage by the acetabulum. The EI is obtained by dividing the horizontal distance of the lateral uncovered femoral head by the total width of the femoral head and multiplying that by 100 to express it as a percentage (Fig. 2G). Acetabular dysplasia is defined by an EI ≥ 25% [93].

Alpha Angle

The AA is the most commonly used measurement to define cam morphology and quantify the sphericity of the femoral head-neck junction. The AA is constructed by two lines, one from the femoral head center through the middle of the femoral neck, the femoral head-neck axis, and a second line from the center of the femoral head through the point where the contour of the femoral head-neck junction extends from the best fitting circle around the femoral head (Fig. 2D)[95]. An AA ≥ 60° threshold is commonly used in literature to define cam morphology[88].

Triangular Index Ratio

The TIR measures femoral asphericity and defines cam morphology. Compared to the AA, the TIR is measured at a specific point on the femoral head-neck junction. It is the ratio between the radius of the best-fitting circle around the femoral head and the distance between the femoral head center and the femoral head-neck junction at 0.5r along the head-neck axis (Fig. 2I). When, for instance, the resultant distance at 0.5r along the axis of the femoral neck at the head-neck junction exceeds the radius of the femoral head, this indicates that, the femoral head is aspherical, possibly indicating the presence of cam morphology[89].

Neck-shaft Angle

The NSA is the angle between the longitudinal axis of the femoral shaft and the femoral head-neck axis (Fig. 2H). It has been hypothesized that hips with a more varus neck orientation experience increased subchondral bone stress and, therefore, increased risk of degeneration in individuals with cam morphology[96]. Conversely, a relative increase in femoral neck shaft angle combined with acetabular undercoverage also leads to RHOA[96]. Coxa valga is generally defined by NSA > 140°, and coxa vara by NSA < 120° [97].

Automated morphological measurements

The bony outline of the proximal femur and acetabulum were annotated automatically on all AP pelvic radiographs with a landmarks (Fig. 2A) (BoneFinder® software (www.bone-finder.com; The University of Manchester, UK)[64]. The protocol for the 80 landmarks used in this automated hip shape annotation can be found in supplementary material 1. The landmarks were used to automatically derive the hip morphology measurements using in-house-built Python-based software[91]. This software is a pipeline to automatically determine radiographic measurements based on radiographic landmarks. The radiographic measurements are performed in accordance to the definitions provided in this manuscript[91]. To assess the impact of automated landmark placement on the morphological measurements, a second set of landmarks was created on the same set of radiographs where all landmarks were manually assessed and adjusted, if necessary, after which the morphological measurements were derived again.

Manual morphological measurements

Two researchers (JT and NSR) were trained in performing manual assessment of all previously described morphological measurements. A random set of 50 radiographs from the World COACH consortium was used to train the researchers. Radiographs were selected at random from the consortium such that the number of radiographs chosen from each cohort was proportional to the total number of radiographs available in that cohort. After all measurements were performed on all 50 radiographs by both researchers, measurements were compared under supervision of an experienced orthopedic surgeon (RA), and inconsistencies were discussed. This was repeated 3 times with the same radiographs until both researchers were proficient in performing measurements. Next, the two trained researchers (JT and NSR) performed on the 30 randomly selected radiographs from the World COACH consortium, with the same proportionality as previously mentioned. Information on whether the hips had morphological variations, hip OA, or clinical symptoms was blinded to all researchers.

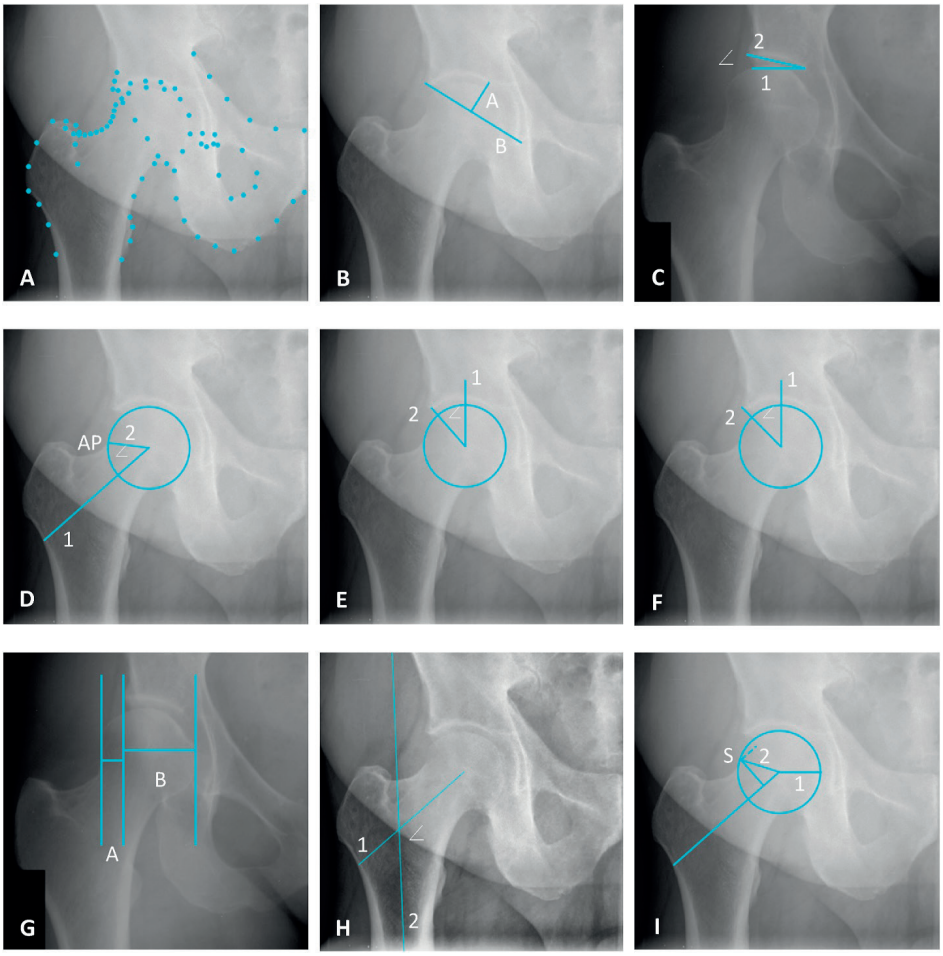


Figure 2. Definition of morphological measurements. **A:** Overview of the landmarks. **B: The ADR** – the ratio between the acetabular depth (line A) measured from the most medial point of the acetabular sourcil to line B, and the acetabular width (line B) measured from the most lateral bony edge of the acetabulum to the most caudal point of the teardrop, ADR = A/B*1000**. C: The mAI** – The angle between the horizontal reference line of the pelvis (line 1) and the line between the most lateral bony edge of the acetabulum and the most medial point of the acetabular sourcil (line 2). **D: The AA** – the angle between the femoral head-neck axis (line 1) and line 2 connecting the femoral head center and alpha point (AP), where the contour of the femoral head-neck junction leaves the best-fitting circle around the femoral head**. E: The WCEA** – The angle between line 1, a vertical line through the femoral head center perpendicular to the HRLP, and line 2 connecting the most lateral point of the acetabular sourcil and the femoral head center. **F: The LCEA** – The angle between line 1, a vertical line through the femoral head center perpendicular to the HRLP, and line 2 connecting the most lateral bony edge of the acetabulum and the femoral head center. **G: the EI** – EI = A /(A + B)*100%, where A is the distance between the most lateral point of the femoral head and the most lateral bony edge of the acetabulum, and B is the distance between the most lateral bony point of the acetabulum and the most medial point of the femoral head. **H: The NSA** – the angle between the femoral head-neck axis (line 1) and the longitudinal axis of the femoral shaft (line 2). **I: The TIR** – The ratio between the radius of the best-fitting circle around the femoral head (line 1) and the distance between the femoral head center and point S on the femoral head-neck junction at 0.5r along the femoral head-neck axis (line 2).

The measurements were repeated on the same radiographs approximately four weeks later. The radiographs were presented to the readers in a different random order each time. Measurements were performed using the DICOM viewer (Synedra View, Version 21.0.0, Synedra Information Technologies). All radiographs were presented in a blinded fashion and random order to the observers. The mean of the individual observers' first and second round of measurements was used for interobserver analyses. The mean of all four manual measurements was used as the reference standard to which the automated method was compared.

Agreement

The agreement within the two rounds of manual measurements for each observer and between observers, and between methods with regard to radiographic diagnoses solely based on morphological measurements of acetabular dysplasia, pincer and cam morphology, and coxa vara and valga was tested.

Qualitative assessment of morphological measurements

A musculoskeletal radiologist (DFH) visually inspected the second round of manual morphological measurements and the automated measurements based on the unadjusted landmarks and qualitatively rated the measurements as acceptable or unacceptable. "Acceptable" is if the radiologist would measure the same morphological measurements based on the landmark points. "Unacceptable" is if the radiologist would perform the measurements differently. This was done in order to ensure the automated measurements were correct from a clinical perspective of an MSK radiologist. In order to blind the radiologist to which method was used, Printscreens of the manual and automated measurements were visually presented in a way which made it impossible to distinguish between methods and in a random order. Printscreens were used because automated measurements were obtained in Python and manual measurements in Synedra Viewer, which would distinguish between methods. Additionally, this ensured that our reference standard of manual measurements were also approved by the MSK radiologist. An example of the ADR is shown in supplementary material 2. No additional information was disclosed about whether the measurements were performed manually or obtained by the automated method.

Statistical analysis

The agreement between the manual observers and the agreement between the automated and manual methods was visualized using Bland-Altman plots for each morphological measurement. In this study, in order to distinguish between random and systematic error, a mean difference larger than 2.5° was defined as a systematic error for mAI, AA, WCEA, LCEA and NSA. A mean difference larger than 1% of the measurement

was defined as a systematic error for ADR, EI and TIR. These thresholds are based on expert agreement. Outliers identified by the Bland-Altman plots were visually inspected to analyze whether consistencies in measurement error occurred.

Intraclass correlation coefficients (ICCs) were used to test reliability and were reported with 95% confidence intervals (CI). Intraobserver reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. Interobserver reliability between manual observers and between the automated determination of the measurements on the manually adjusted and unadjusted landmarks was tested with a 2-way random-effects model, single rater, absolute agreement ICC. Lastly, intermethod reliability between the mean of all manual and automated measurements on manually adjusted and unadjusted landmarks was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. ICCs were rated as poor (<0.50), moderate (0.50–0.75), good (0.76–0.90), or excellent (>0.90)[98].

The agreement within and between observers, and between methods with regard to radiographic diagnoses was tested using percentage agreement. Based on the qualitative rating of the measurements by the musculoskeletal radiologist, the percentage of acceptable measurements was determined for each morphological measurement by the two manual observers and the automated method, respectively. The percentage of acceptable measurements was rated as poor (<50%), moderate (50–70%), good (71–90%), or excellent (>90%).

Statistical analyses were performed using R statistical software (v4.1.0; R Core Team 2021). The ggplot2-package in R was used to create Bland-Altman plots. The irr-package in R was used to calculate the ICCs and the percentage agreement[73].

RESULTS

All morphological measurements could automatically be performed in all 30 hips, except for NSA, which could not be performed on two images as too little of the femoral shaft was depicted on the radiograph.

Agreement

The Bland-Altman plots for agreement between the two observers and the agreement between the manual and automated measurements based on unadjusted landmarks are presented in Figure 3, and the corresponding mean difference and limits of agreement are summarized in Table 1. The AA, WCEA, LCEA, mAI, and EI showed no to small

mean differences between automated and manual measurements. However, both the interobserver and intermethod agreement of ADR and the interobserver NSA and TIR showed a bias. Observer 1 consistently measured ADR and TIR higher than observer 2, while the opposite was observed for ADR. When comparing the manual and automated ADR, the mean of the manual measurements was consistently higher than the automated measurement. The intermethod limits of agreement were mainly smaller or similar to the interobserver limits of agreement for all morphological measurements except for WCEA and LCEA.

Reliability

The intra- and interobserver and intermethod reliability defined by ICCs for all measurements are shown in Table 2. The intermethod reliability between the manual and automated measurements based on both the manually adjusted and unadjusted landmarks was comparable to or better than the interobserver reliability, except for WCEA in which case the manual measurements were more reliable. Additionally, we found that manually adjusted landmarks impacted the ADR and mAI most. This led to lower reliability between manually adjusted compared to unadjusted automated ADR and mAI measurements. These measurements are calculated based on only on few specific landmarks. Conversely, measurements that do not rely on few specific landmarks from the point set like AA, NSA and TIR, showed excellent reliability between the automated measurements performed using the adjusted vs unadjusted landmarks.

Table 1: Summary of mean interobserver and intermethod bias and limits of agreement of manual morphological measurements and manual vs automated morphological measurements based on the unadjusted landmarks.

	Manual		**Manual vs Automated**	
Measurement	**Interobserver bias (mean)**	**Interobserver limits of agreement**	**Intermethod bias (mean)**	**Intermethod limits of agreement**
Acetabular depth-width ratio	13	-27 to 53	-15	-52 to 13
Modified acetabular index [°]	-1.8	-7.6 to 4.1	2.0	-3.1 to 7.0
Alpha angle [°]	-2	-22 to 18	-1	-23 to 20
Wiberg center edge angle [°]	1	-3 to 6	-2	-9 to 5
Lateral center edge angle [°]	0	-4 to 4	0	-6 to 6
Extrusion index [%]	1	-8 to 9	-1	-8 to 5
Neck-shaft angle [°]	-5	-9 to 0	-2*	-6 to 2*
Triangular index ratio	0.028	-0.058 to 0.115	-0.009	-0.078 to 0.061

Bland-Altman interobserver and intermethod bias (mean difference) and limits of agreement, n=30. *Based on 28 hips.

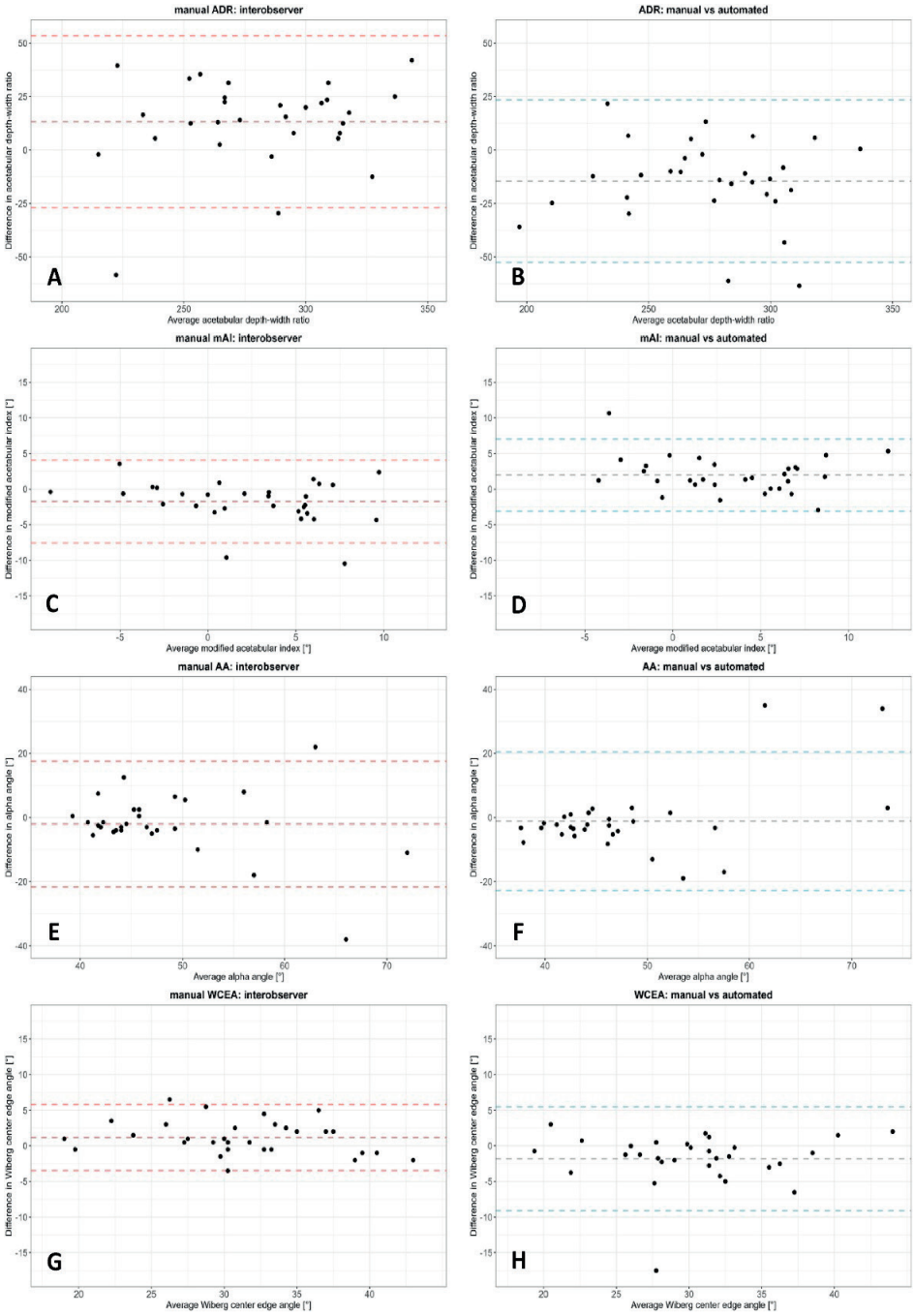


Figure 3. Bland-Altman plots of the morphological measurements. **A: The ADR** – observer 1 vs observer 2. **B: ADR** – manual vs automated measurements based on unadjusted landmarks. **C: The mAI** – observer 1 vs observer 2. **D: mAI** – manual vs automated measurements based on unadjusted landmarks. **E: The AA** – observer 1 vs observer 2. **F: AA** – manual vs automated measurements based on unadjusted landmarks. **G: The WCEA** – observer 1 vs observer 2. **H: WCEA** – manual vs automated measurements based on unadjusted landmarks.

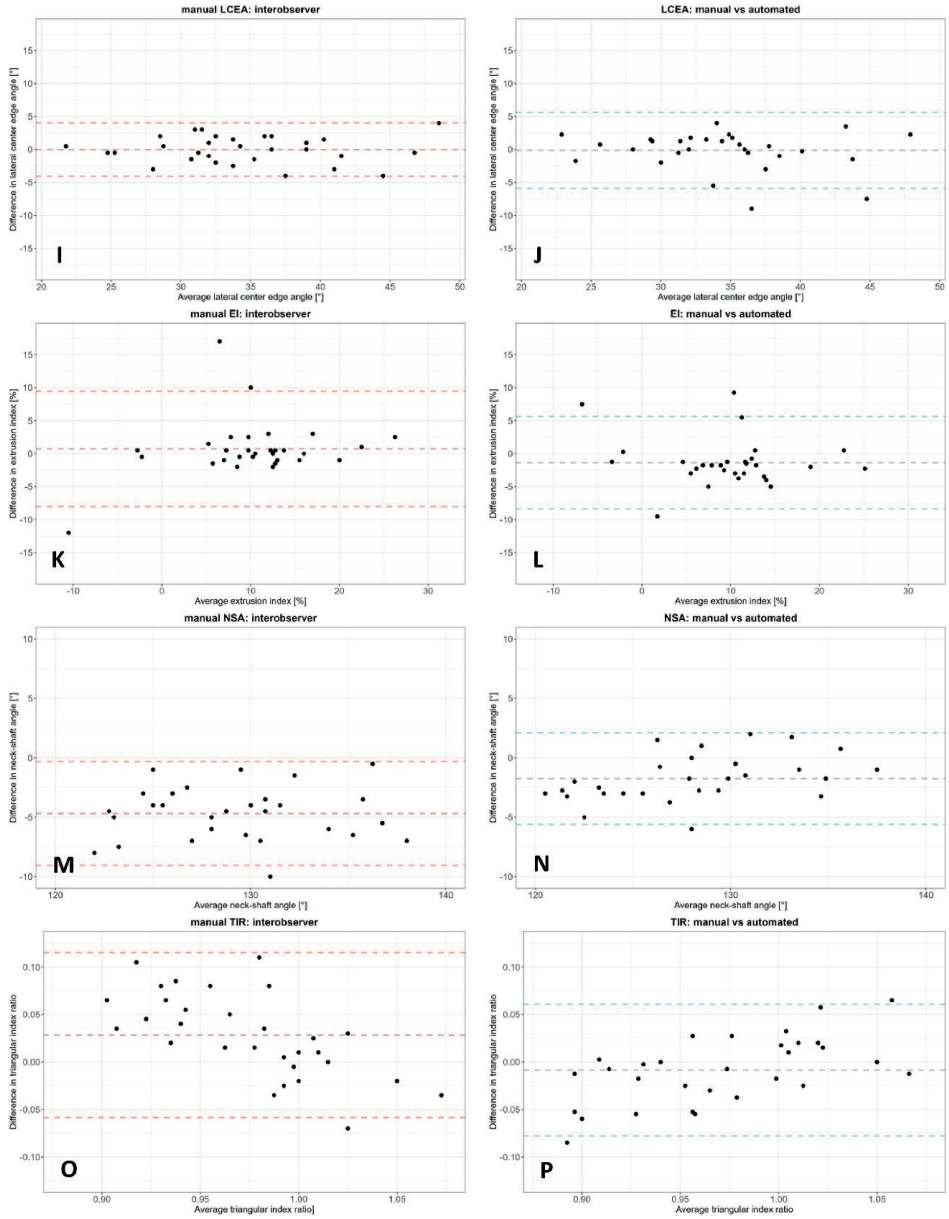


Figure 3 – continued. **I: The LCEA** – observer 1 vs observer 2. **J: LCEA** – manual vs automated measurements based on unadjusted landmarks. **K: The EI** – observer 1 vs observer 2. **L: EI** – manual vs automated measurements based on unadjusted landmarks. **M: The NSA** – observer 1 vs observer 2. **N: NSA** – manual vs automated measurements based on unadjusted landmarks. **O: The TIR** – observer 1 vs observer 2. **P: TIR** – manual vs automated measurements based on unadjusted landmarks

Table 2: Intra- and interobserver reliability between manual measurements by observer 1 and observer 2, interobserver reliability between adjusted and unadjusted landmarks and intermethod reliability between manual and automated morphological measurements.

	Manual			Automated	Manual vs automated	
	Observer 1	Observer 2	Observer 1 vs observer 2	Adjusted vs unadjusted landmarks	Unadjusted landmarks	Adjusted landmarks
Measurement	**Intraobserver ICC (95% CI)**	**Intraobserver ICC (95% CI)**	**Interobserver ICC (95% CI)**	**Interobserver ICC (95% CI)**	**Intermethod ICC (95% CI)**	**Intermethod ICC (95% CI)**
Acetabular depth-width ratio	0.67 (0.41 – 0.82)	0.89 (0.77 – 0.94)	0.79 (0.49 – 0.91)	0.70 (0.47 – 0.85)	0.78 (0.39 – 0.91)	0.80 (0.60 – 0.90)
Modified Acetabular Index	0.65 (0.36 – 0.82)	0.82 (0.61 – 0.91)	0.77 (0.48 – 0.89)	0.83 (0.67 – 0.91)	0.75 (0.34 – 0.90)	0.82 (0.30 – 0.94)
Alpha Angle	0.36 (0.01 – 0.63)	0.67 (0.42 – 0.83)	0.43 (0.10 – 0.68)	0.98 (0.97 – 0.99)	0.46 (0.12 – 0.70)	0.5 (0.17 – 0.73)
Wiberg center edge angle	0.87 (0.74 – 0.93)	0.94 (0.88 – 0.97)	0.91 (0.78 – 0.96)	0.94 (0.88 – 0.97)	0.77 (0.54 – 0.89)	0.88 (0.70 – 0.95)
Lateral center edge angle	0.89 (0.78 – 0.95)	0.95 (0.87 – 0.98)	0.95 (0.86 – 0.98)	0.91 (0.8 – 0.96)	0.89 (0.78 – 0.94)	0.95 (0.88 – 0.98)
Extrusion Index	0.74 (0.51 – 0.87)	0.80 (0.51 – 0.91)	0.83 (0.67 – 0.91)	0.94 (0.87 – 0.97)	0.86 (0.71 – 0.93)	0.88 (0.65 – 0.95)
Neck Shaft Angle	0.89 (0.78 – 0.94)	0.86 (0.73 – 0.93)	0.58 (0 – 0.87)	0.995 (0.989 – 0.998) *	0.86 (0.44 – 0.95) *	0.88 (0.51 – 0.96) *
Triangular Index Ratio	0.26 (0 – 0.57)	0.88 (0.76 – 0.94)	0.49 (0.12 – 0.73)	0.99 (0.98 – 0.996)	0.78 (0.59 – 0.89)	0.79 (0.61 – 0.89)

Intraclass correlation coefficients (ICC) of intra- and interobserver, and intermethod reliability of the morphological measurements. ICCs are presented with 95% confidence interval (CI). The mean of all four manual measurements was used as the reference standard for the intermethod measurements. Intraobserver reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC. Interobserver reliability between both manual observers, as well as between the automated determination on adjusted and unadjusted landmarks, was tested with a 2-way random-effects model, single rater, absolute agreement ICC. Intermethod reliability was tested with a 2-ways mixed-effects model, single rater, absolute agreement ICC. All ICCs were measured using 30 hips. *ICCs measured using 28 hips. Interpretation: poor (<0.50), moderate (0.50–0.75), good (0.76–0.90), or excellent (>0.90).

Table 3: Prevalence and diagnostic intraobserver and interobserver agreement between observer 1 and observer 2, interobserver agreement between adjusted and unadjusted automated measurements, and intermethod agreement.

	Manual			Automated		Manual vs automated		
	Observer 1	**Observer 2**	**Observer 1 vs observer 2**	**Adjusted vs unadjusted landmarks**	**Reference standard**	**Unadjusted landmarks**	**Unadjusted landmarks**	**Adjusted landmarks**
Measurement	**Intraobserver percent agreement**	**Intraobserver percent agreement**	**Interobserver percent agreement**	**Interobserver percent agreement**	**Prevalence**	**Prevalence**	**Intermethod percent agreement**	**Intermethod percent agreement**
Acetabular depth-width ratio ≤ 250	90.0	83.3	86.7	93.3	16.7%	26.7%	90.0	90.0
Modified acetabular Index ≥ 13°	96.7	96.7	96.7	90.0	0%	3.3%	96.7	86.7
Alpha Angle ≥ 60°	90.0	90.0	86.7	96.7	10.0%	10.0%	86.7	83.3
Wiberg center edge angle ≤ 25°	100.0	100.0	93.3	90.0	13.3%	23.3%	96.7	93.3
Lateral center edge angle ≥ 40°	90	93.3	93.3	96.7	20.0%	16.7%	96.7	100
Extrusion Index ≥ 25%	96.7	96.7	100	100	0%	0%	100	100
Neck Shaft Angle <120° & >140°	90	96.7	90	96.7	<120°: 0% >140°: 0%	<120°: 3.3% >140°: 0%	96.7	100

The reference standard consists of the mean of all manual measurements. Intermethod percent agreement was determined using the reference standard. n = 30.

Radiographic diagnostic agreement

Percentage agreement in radiographic diagnosis based on morphological measurements is summarized in Table 3.The intermethod radiographic diagnostic agreement was better than or similar to the interobserver radiographic diagnostic agreement. Except for the radiographic diagnostic agreement of dysplasia based on mAI of the manual versus automated measurements based on the manually adjusted landmarks.

Table 4: The qualitative assessment of the morphological measurements.

	Manual		Automated
Measurement	**Observer 1**	**Observer 2**	**Unadjusted landmarks**
Acetabular depth-width ratio	77	80	73
Modified Acetabular Index	70	53	70
Alpha Angle	93	90	77
Wiberg center edge angle	73	80	63
Lateral center edge angle	70	90	80
Extrusion Index	53	47	63
Neck Shaft Angle	93	100	96*
Triangular Index Ratio	63	100	73

Percentage of acceptable measurements. Qualitative assessment was performed on 30 hips. *Based on only 28 hips. Interpretation: poor (<50%), moderate (50–70%), good (71–90%), or excellent (>90%).

Qualitative assessment

The results of the qualitative assessment as performed by the MSK radiologist are presented in Table 4. The majority of automated measurements were deemed acceptable by the musculoskeletal radiologist. The percentage of acceptable measurements was moderate to excellent for all measurements, except for the EI measurements by observer 2.

DISCUSSION

This study investigated the agreement and reliability of manual and automated morphological measurements including ADR, mAI, AA, WCEA, LCEA, EI, NSA, and TIR on AP pelvic radiographs. The presented algorithm performed equally well compared to current best practice of manual measurement by trained readers, attesting to its reliability and efficiency in rapidly computing radiological measurements on an AP pelvic radiograph.

The reported intra- and interobserver reliability of morphological measurements varies in literature. The reported ICCs in the present study were compared to the reliability of various morphological measurements in literature. The ICCs reported in literature for the Wiberg and lateral CEA (ICC = 0.7 (95% CI 0.58–0.86) to 0.98 (CI 0.97–0.99)[75,78,98-100], the NSA (ICC = 0.58 (0.31–0.76) to 0.98 (0.95–0.99))[78], the mAI (or Tönnis angle) (ICC = 0.71 (95% CI 0.45–0.83) to 0.92 (95% CI 0.85–0.95))[75,98,101], the EI (ICC = 0.68 (0.57–0.79) to 0.98 (no CI reported))[75,98-100] and the ADR (ICC = 0.62 to 0.84)[86,100,101] are similar to the ICCs found in our study. The reported reliability in literature for the AA (ICC = 0.78 (95% CI 0.61–0.87) to 0.99 (no CI reported))[102-104] is higher than observed in the present study. No reliability has been reported for the TIR, although one study did report on the triangular index height in 10 individuals (κ = 0.74–0.78)[98].

In terms of reliability and agreement in the current study, the AA showed the worst reliability in the manual method between and within observers, as well as in terms of intermethod reliability. The AA also showed large limits of agreement in the Bland-Altman plots and erratic behavior in the higher AA values (representing cam hips). These results are likely caused by small differences in femoral head circle fit, which may cause large measurement variation due to movement of the alpha point. Faber et al. showed similar outliers and erratic behavior within the Bland-Altman analysis when comparing manual and automated AA measurements[76]. Similar results, although less extreme, were found for TIR, as expected since this measurement is also largely dependent on the circle fit. However, the erratic behavior observed in the AA Bland-Altman plots in hips with cam morphology is absent in the TIR Bland-Altman plots. This may be caused by the fact that compared to the location of the alpha point, the location of point S (Fig. 2I) is less influenced by the best-fitting circle around the femoral head.

ADR and mAI are two measurements which are calculated based on only two to three landmarks and, therefore highly dependent on correct landmarks recognition and placement. This is reflected in similar reliability and limits of agreement for the intra- and interobserver, and intermethod comparisons. The outliers in these measurements were all caused by different landmarks recognition and placement of both the most lateral bony edge of the acetabulum and the most medial point of the weight-bearing sourcil. Additionally, we found that the mean of the manual measurements by the trained researchers was consistently higher than the automated measurement, implying that we may under diagnose acetabular dysplasia based on manual ADR measurements. Alternatively, it may also be the case that the medial point of the ADR on the sourcil is difficult to identify for the automated measurement. This may also influence the automated ADR.

The correct identification of the most lateral bony edge of the acetabulum also influenced the LCEA and EI measurements. The reliability was good to excellent for all analyses, and the limits of agreement were similar between the interobserver and intermethod analyses.

The WCEA, as determined using the automated method, was slightly worse than the LCEA when comparing the automated method to manual measurements. This is likely due to more difficult assessment of the sourcil, than the more distinct lateral bony acetabular rim. This is also observed in literature with higher reliability for LCEA reported compared to WCEA[75,78,98-100]. Overall, this landmark needed more adjustment than the most lateral bony part of the acetabulum during the manual assessment of landmarks placement. This was reflected in the higher reliability of the manual versus automated measurement when the WCEA was performed based on the manually adjusted landmarks.

The majority of manual measurements were deemed acceptable by the musculoskeletal radiologist. This implies that the reported manual measurement ICCs represent clinically acceptable reliability. In terms of automated measurements, we can conclude that the automated ADR, mAI, AA, LCEA, NSA and TIR measurements are valid in a clinical setting and can be applied to establish radiographic morphological hip diagnoses. According to our study, performance of manual as well as automated EI measurements does not reach the threshold for good agreement. We hypothesize that in case of less sphericity of the femoral head, the identification of the most lateral point of the femoral head becomes difficult leading to unreliability in the measurement. As there are other measurements that quantify acetabular coverage, these may be more appropriate in a clinical setting to study hip morphology.

Using automated morphological measurements may advance research and have important clinical implications. First, automated measurements may improve accuracy and consistency in morphological measurements reported in literature. Measurement variability and bias could be reduced dramatically if all measurements are performed uniformly, allowing for comparison of results across studies. This holds especially true in terms of the femoral head circle fit, which is essential in many morphological measurements. The present automated method is published open-access [23], which promotes collaboration in future hip (OA) studies. While the method is still reliant on correct landmark identification, this was also automated to achieve more consistency and speed. This method can be applied in future studies to study whether these measurements are associated with clinical outcomes such as symptomatic hip OA. The automated method was tested on supine and standing pelvic radiographs from various cohorts in the World COACH consortium, potentially making the results more generalizable to a larger popu-

lation. Furthermore, the automated method can improve efficiency by accommodating the collection of large amounts of morphological data. This will allow researchers to carry out studies with increased statistical power, advancing our understanding of hip morphology as a risk factor for hip OA.

No gold standard is available for these morphological measurements, so we extensively trained researchers to obtain measurements which could be used as a reference standard. We found order to ensure that these measurements resemble clinical practice, an MSK radiologist visually inspected all manual and automated measurements. Secondly, it should be kept in mind that this study includes a rather small set of 30 hips. A larger dataset would likely show increased variation in hip morphology and therefore provide a more robust assessment of the described methods. Furthermore, as the participants from the World COACH consortium are either from the general population or from a population selected based on having symptoms or risk factors for hip OA, the hips are a representation of the normal population. Therefore, gross bony deformations as seen in hospital populations are underrepresented in the world COACH consortium and results from the automated measures should be validated in this population first. All thresholds used to define radiographic morphological diagnoses are based on literature, but what the "right" threshold is remains unknown[93].

With regards to the qualitative assessment, the radiologist evaluated printscreens of measurements, which made it impossible to adjust contrast setting on the images as preferred by the radiologist. As a result of this, the measurements that were impossible to visually inspect were labeled as unacceptable, although in reality they may have been correct. This issue may be avoided in the future by using DICOM images on PACS viewer rather than printscreens of radiographs. Another limitation of this study is that all morphological measurements were performed on AP pelvic radiographs although it is known that some morphological diagnoses require additional radiographic views to assess hip morphology[8,93,98,100]. Furthermore, acetabular morphology is influenced by pelvic orientation, which can vary significantly in terms of tilt[105]. This provides a future opportunity to also develop automated measurements in various radiographic views.

In conclusion, automated morphological measurements are a reliable and reproducible method to quantify the ADR, WCEA, LCEA mAI, TIR, EI and NSA. This method makes morphological hip measurements viable in large population studies, as it enables reliable analysis of large amounts of data. Additionally, it may be a useful tool in clinical practice, as it reduces reader bias and the landmarks allow for insightful measurements. Access to fast, externally validated, reliable methods to quantify hip morphology may aid in the quest for modifiable risk factors for hip OA in future studies.

Funding

CL is funded by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (223267/Z/21/Z). For the purposes of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

The World COACH consortium has been funded through research grants by the Dutch Arthritis Society (grant no. 18-2-203 and 21-1-205), the Dutch Research Council (NWO Veni grant scheme no. 09150161910071), and Erasmus MC, University Medical Center Rotterdam (Erasmus MC Fellowship).

Conflict of Interest

GJ reports personal fees from Novartis outside the submitted work. SBZ reports consulting fees from Pfizer Infirst Healthcare and personal fees for being a Deputy Editor for Osteoarthritis and Cartilage outside the submitted work. CL and TC report a patent for an image processing apparatus and method for fitting a deformable shape model to an image using random forest regression voting. CL reports licensing royalties for this patent from Optasia Medical outside the submitted work. AN is an associate editor for Osteoarthritis and Cartilage and is on the OARSI Board of Directors outside the submitted work. AM is on the Editorial Board for the British Journal of Sports Medicine and the Journal of Science and Medicine in Sport outside the submitted work. HW reports being a minority shareholder of Uplanner BV and Replasia BV outside the submitted work.

Acknowledgments

We would like to thank all participants of the cohort studies that are involved in the World COACH consortium. We gratefully acknowledge all international organisations that collaborated with the cohort studies in World COACH, as well as the Osteoarthritis Research Society International (OARSI) for endorsing the World COACH consortium. We thank the (non-profit) funding bodies who financially support the World COACH consortium: the Dutch Arthritis Society (grant no. 18-2-203 and 21-1-205), the Dutch Research Council (NWO Veni grant scheme no. 09150161910071) and the Erasmus MC, University Medical Center, Rotterdam (Erasmus MC Fellowship).

CHECK

The CHECK study was initiated by the Dutch Arthritis Society 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.Maartenskliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

Chingford

We would like to thank all the participants of the Chingford Women Study, Professor Nigel Arden, Professor Tim Spector, Dr Deborah Hart, Mr Gem Lawson, Maxine Daniels and Alison Turner for their time and dedication and Arthritis Research UK for their funding support to the study and the Oxford NIHR Musculoskeletal Biomedical Research Unit for funding contributions.

JoCoOA

Support for data from the Johnston County Osteoarthritis Project was provided in part by: the Center for Disease Control and Prevention (CDC) U01DP006266 and U01DP003206; Association of Schools of Public Health/CDC S043, S1734, S3486; and National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases P60AR30701, P60AR049465, P60AR064166, and P30AR072580.

MOST

The MOST study was funded by the National Institutes of Health – National Institute on Aging grants AG19069(Michael Nevitt, University of California, San Francisco) AG18820 (David Felson, Boston University) AG18947 (Cora Lewis, University of Alabama at Birmingham) and AG18832 (James Torner, University of Iowa).

OAI

The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc.

Rotterdam Study

The Rotterdam Study is funded by Erasmus University Medical Center and Erasmus University, Rotterdam, The Netherlands Organisation for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

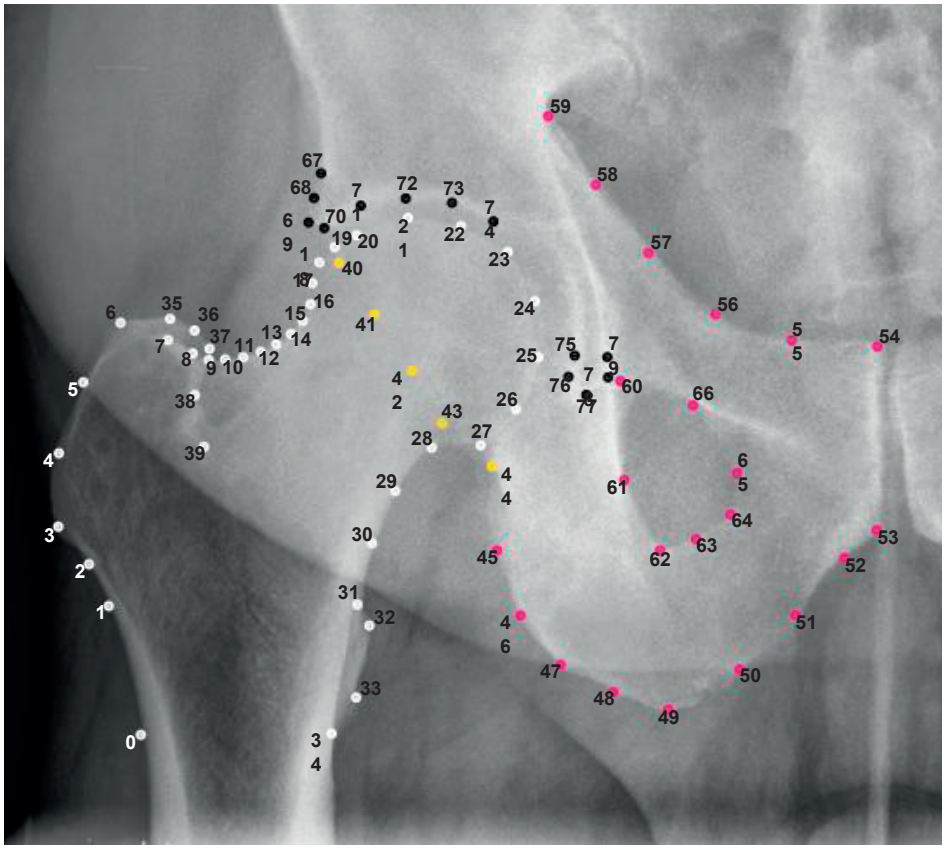
SOF

The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576.

TASOAC

The TASOAC study was supported by the National Health and Medical Research Council of Australia, Tasmanian Community Fund, Masonic Centenary Medical Research Foundation, Royal Hobart Hospital Research Foundation and Arthritis Foundation of Australia.

SUPPLEMENTARY MATERIAL 1: PROTOCOL FOR LANDMARK ANNOTATION



Proximal femur (white points)

Lesser trochanter

Point (**34**): Where the lesser trochanter starts bending off the shaft distally. If the lesser trochanter is seen behind the shaft, place this point on the cortex of the shaft at this level. If the lesser trochanter is not visible at all: missing points.

Point (**31**): Where the lesser trochanter joins the shaft proximally. If the lesser trochanter is seen behind the shaft, place this point on the cortex of the shaft at this level. If the lesser trochanter isn't visible at all: missing points.

Point (**32**)+(**33**): Respectively on the lower and upper corners of the lesser trochanter. If there are no clear corners: space them equally between (31) and (34) along the bony contour of the lesser trochanter.

Rest of proximal femur

Point (**0**) + (**1**): Respectively across (34) and (31) on the lateral femoral shaft. If point (1) would be above point (3) based on the position of point (34), place point (1) just under point (3).
Point (**3**): On the lower lateral corner of the greater trochanter.
Point (**2**): Equally spaced between (1) and (3).
Point (**6**): On the upper lateral corner of the (anterior) greater trochanter.
Point (**4**)+(**5**): Equally spaced between (3) and (6).
Point (**7**): On the medial upper corner of the anterior greater trochanter. If not visible, place this point equally spaced between (6) and (8) on the contour of the anterior greater trochanter.
Point (**8**): Where the anterior greater trochanter intersects the femoral.
Point (**18**): On the superolateral side of the femoral head, where the "best fitting circle" around the convexity of the femoral head seems to start. In case of a cam bump, osteophyte, or other irregularity: place (18) right after this bump ends, and the circle begins.
Point (**27**): On the inferomedial side of the femoral head, where the convexity of the femoral head seems to end. (The neck bends off after this point).
Point (**20-26**): Place these points equally spaced between (18) and (27) following the femoral head contour, unless there is a clear fovea dip, in which case the adjacent points, usually (24) and (25), are placed just outside of the fovea. Point (23) will be approximately placed halfway across the 'semi'-circle between (18) and (27).
Point (**9-17**): Place these points equally spaced between (8) and (18) following the lateral femoral neck contour. In case of irregularities like a cam bump or osteophyte, follow the outlining contour as closely as possible.
Point (**19**): Place this point equally spaced between (18) and (20) on the femoral head contour.
Point (**28**): At the deepest point of the inferomedial concavity of the femoral neck, so that (27-31) will follow the medial cortex of the femoral neck as closely as possible.
Point (**29**)+(**30**): Place these points equally spaced between (28) and (31), following the medial cortex of the femoral neck.

Greater trochanter, posterior part

** If the posterior greater trochanter is not visible: (35-39) missing points.

Point (**36**): On the upper medial corner of the posterior greater trochanter.
Point (**35**): Between (6) and (36), following the contour. If there is a clear corner, put it there.
Point (**37**): On the medial corner of the posterior greater trochanter, where it starts to drop downwards (caudal). This is independent of the femoral neck, so it can be before or after it dips behind the femoral neck, depending on the rotation of the proximal femur.
Point (**38**): Where the posterior greater trochanter is dropping straight down, right before it bends medially.
Point (**39**): On the end of the sclerotic line right *after* the medial bend, following the contour of the posterior greater trochanter.

Posterior wall of acetabulum (yellow points)

Point (**40**): On the uppermost visible part of the posterior wall of the acetabulum (usually right below the lateral edge of the weight-bearing surface or lateral osteophyte/pincer).
Point (**44**): Where the posterior wall joins the ischium (where the ischium usually proceeds vertically down).
Point (**41**-**43**): Place these points equally spaced between (40) and (44), following the contour of the posterior wall of the acetabulum.

Ischium & Pubis (pink points)

Point (**49**): On the most caudal point of the ischium (ischial tuberosity). If the ischial tuberosity appears as a straight line, put it in the middle of the ischial tuberosity.
Point (**45-48**): Place these points equally spaced between (44) and (49) along the contour of the ischial tuberosity.
Point (**52**): In the concavity before the symphysis.
Point (**50**)+(**51**): Place these points equally spaced between (49) and (52), following the caudal contour of the inferior pubic ramus.
Point (**53**): On the most caudal point of the pubic symphysis.
Point (**54**): On the most cranial point of the pubic symphysis.
Point (**59**): On the iliopectineal line of the pelvis, at the height where the ilioischial line splits off.
Point (**55-58**): Place these points equally spaced between (54) and (59). Follow the iliopectineal line, ignoring the ischial spine.
Point (**60**): In the superolateral corner of the obturator foramen.
Point (**62**): In the inferolateral corner of the obturator foramen.

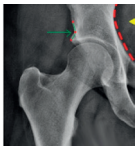
Point (**61**): Equally spaced between (60) and (62), following the contour of the lateral rim of the obturator foramen.
Point (**64**): In the inferomedial corner of the obturator foramen.
Point (**63**): Place this point equally spaced between (62) and (64), following the contour/angle of the inferior rim of the obturator foramen.
Point (**65**): In the superomedial corner of the obturator foramen.
Point (**66**): Place this point equally spaced between (65) and (60), following the contour/angle of the superior rim of the obturator foramen.

Acetabulum (black points)

Acetabular roof

*** Points (70-74) along the weight-bearing zone (sourcil) are placed on the inferior rim of the sclerotic line.*

Point (**69**): On the most lateral point of the acetabulum, this can also be a lip/osteophyte.
Point (**70**): On the most lateral point of the **weight-bearing zone** (sourcil) of the acetabulum (most lateral point of sclerotic line).
Point (**74**): On the most medial point of the weight-bearing zone (sourcil) of the acetabulum, this is also the most superolateral point of the acetabular fossa. Usually there is a clear angle in the (sclerotic) line at the transition of weight-bearing zone to fossa. If the acetabular fossa is not visible at all, just place it on the most medial point of the sclerotic line.
Point (**71-73**): Along the underside of the sourcil, place these points equally spaced between (70) and (74), following the contour of the weight-bearing zone
Point (**68**): On the 'dimple' above (70), where the acetabular lip contour has a bend. When the acetabular lip forms a straight line, equally space point (68) and (67) above point (69), with the same distance as points (71-72).
Point (**67**): Above (68), following the most lateral sclerotic line, with a similar distance between points (67-68) as points (71-72).



Pelvic teardrop

Point (**75**): On the superolateral corner of the visible teardrop (on the wall of the acetabular fossa)
Point (**77**): On the most caudal point of the teardrop.
Point (**79**): Across (75) on the other side of the teardrop.
Point (**76**)+(**78**): Across each other between (75-77-79), at the corners of the teardrop, where the more vertical (diverging) lines change direction to more oblique (converging) lines. This can be a very acute angle or more gradual.

Curve model

Proximal femur curve: 0-1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-21-22-23-24-25-26-27-28-29-30-31-32-33-34

Greater trochanter curve: 6-35-36-37-38-39

Posterior wall curve: 40-41-42-43-44

Ischium & pubis curve: 44-45-46-47-48-49-50-51-52-53-54-55-56-57-58-59

Foramen curve: 60-61-62-63-64-65-66

Acetabular roof curve: 67-68-69-70-71-72-73-74

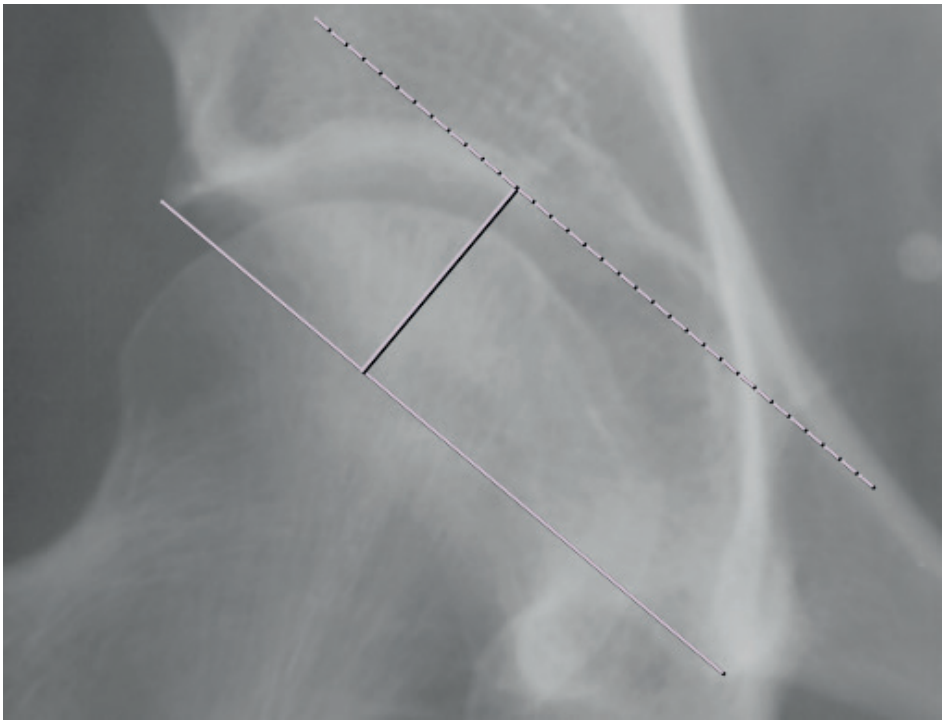
Pelvic teardrop curve: 75-76-77-78-79

General rules

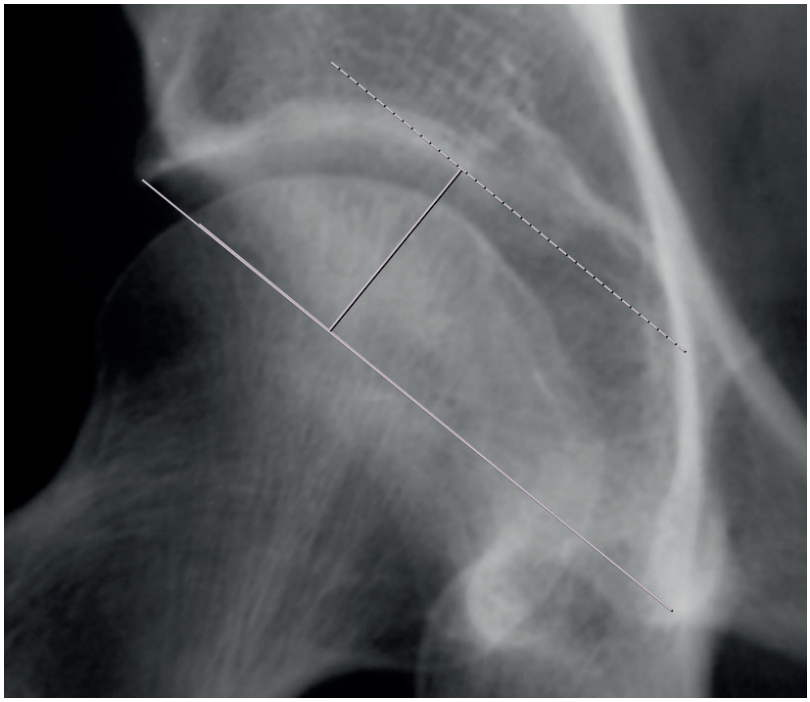
- Osteophytes of the femoral head are *included* in the model. Follow the *outermost* contour. We can later correct for these with the radiological assessment data.
- Non-identifiable landmarks: missing points (write in separate log file)
- Only follow clear bony structures, not projecting shadows.
- Every hip is different, so not all anatomical landmarks might be clearly visible in each radiograph. In case of systematic doubt or error: discuss!

SUPPLEMENTARY MATERIAL 2: EXAMPLES OF THE IMAGES FOR QUALITATIVE ASSESSMENT

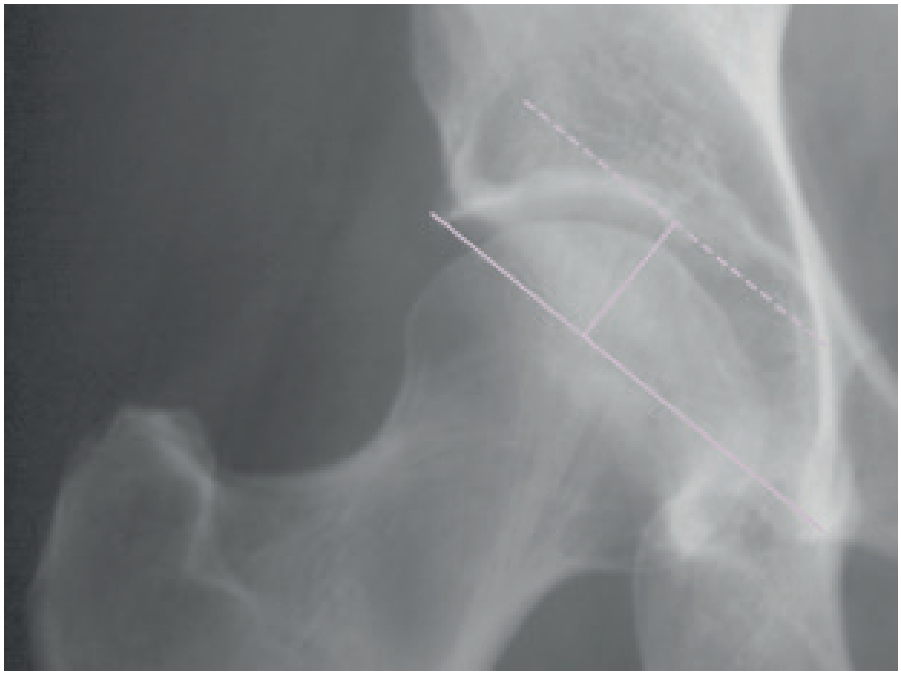
Below are depicted the visualizations of the acetabular depth-width ratio measurements as performed by observer 1, observer 2 and the automated method which were presented to the musculoskeletal radiologist for qualitative assessment of the measurement.



Visualization of the acetabular depth-width ratio measurement as performed by observer 1.



Visualization of the acetabular depth-width ratio measurement as performed by observer 2.



Visualization of the automated acetabular depth-width measurement on unadjusted landmark points.

Chapter 4

DXA images vs. Pelvic Radiographs: Reliability of Hip Morphology Measurements

F. Boel, J. Wortel, M.M.A. van Buuren, F. Rivadeneira, J.B.J. van Meurs, J. Runhaar, S.M.A. Bierma-Zeinstra, R. Agricola.

Osteoarthritis and Cartilage. 2024 Oct 24;33(2):283-292.

ABSTRACT

Objective: Dual-energy x-ray absorptiometry (DXA) images are increasingly used to study hip morphology. Whether hip morphology measurements are consistent between DXA images and radiographs is unknown. Therefore, we investigated the agreement and reliability of the measurements performed on DXA and radiographs.

Design: We included participants from the Rotterdam study, a population-based cohort study, who received a hip DXA and pelvic radiograph on the same day. The acetabular depth-width ratio (ADR), modified acetabular index (mAI), alpha angle (AA), Wiberg and lateral center edge angle (WCEA, LCEA), extrusion index (EI) and triangular index ratio (TIR) were automatically determined on both imaging modalities. The intraobserver and intermethod agreement were studied using Bland-Altman methods, and the reliability was assessed using intraclass correlation coefficients (ICC) or concordance correlation coefficients (CCC) for non-normal distributed variables. Secondly, the diagnostic agreement regarding dysplasia, cam, and pincer morphology was assessed using percent agreement.

Results: 750 hips (411 individuals, median age 67.3 years (range 52.2 – 90.6), 45.5% male) were included. The following intermethod ICCs (95% CI) were obtained: ADR 0.85 (0.74-0.91), mAI 0.75 (0.52-0.85), AA 0.72 (0.68-0.75), WCEA 0.81 (0.74-0.85), LCEA 0.93 (0.91-0.94), EI 0.88 (0.84-0.91), and TIR 0.81 (0.79-0.84). Additionally, a CCC of 0.58 (95% CI 0.53-0.62) was obtained for the AA. We found comparable intraobserver ICCs or CCCs for each morphological measurement.

Conclusion: DXA images and pelvic radiographs can both reliably be used to study hip morphology. Due to the lower radiation burden, DXA images can be an excellent alternative to pelvic radiographs for research purposes.

INTRODUCTION

Hip morphology is strongly associated with hip osteoarthritis (OA)[8-11,34,49,50]. Cam morphology and acetabular dysplasia have consistently been associated with hip OA development. These and other types of hip morphology are usually quantified by radiographic morphological measurements, such as the alpha angle (AA) and the center edge angle of Wiberg (WCEA).

Large cohort studies on hip OA usually obtain anteroposterior (AP) pelvic radiographs. Nevertheless, the image quality of a hip dual-energy x-ray absorptiometry (DXA) has increased significantly with new-generation scanners using narrow-angle fan-beam technology. It has been shown that hip OA grading on DXA images was similar to pelvic radiographs[61]. Additionally, DXA images are increasingly used to assess hip morphology, especially in large population studies[106-108]. One of the main advantages of hip DXA images is the lower radiation burden of 0.36-70 µSv compared to hip or pelvic radiographs with an effective dose of 600-700 µSv[62,109-111].

However, the image acquisition method is different between radiographs and DXA images. In radiographs, the x-ray source is kept in one position, the x-ray beam is cone-shaped, and the whole image fits in the field of view[112]. In contrast, for DXA images the x-ray beam is fan-shaped and the x-ray source moves to obtain the image, so the direction of the x-ray beam through the subject is different[113]. Notably, while AP pelvic radiographs have the x-ray source mostly centered around the pubic symphysis, DXA scans are almost always of a single hip, where the x-ray source is mainly centered on the hip joint[114]. Additionally, radiographs can be obtained while the participant is either in a weight-bearing or supine position, while the participant is always in the supine position for a DXA image[113,114].

It is unknown if morphological measurements of the hip can be assessed equally reliably on DXA images as on pelvic radiographs. Given the increasing use of hip DXA images for morphological assessment and the differences in how images are acquired between radiographs and DXA scans, we aimed to investigate the agreement and reliability of morphological hip measurements performed on DXA images and pelvic radiographs. Secondly, we investigated the diagnostic agreement between DXA images and radiographs for acetabular dysplasia, cam, and pincer morphology.

METHODS

Participants

This comparative study used data from the Rotterdam Study (RS), a prospective population-based cohort study studying determinants of disease and disability in adult individuals older than 40 years[115]. Since 1990, the RS recruited participants from the Ommoord district in Rotterdam, the Netherlands. All participants provided informed consent before participation, and the Medical Ethics Committee of the Erasmus MC approved the RS. From this cohort study, a subset of participants had a hip DXA obtained between March 2009 and June 2014. From this subset, 500 participants were selected at random. For these participants it was assessed whether they also had a pelvic radiograph performed on the same day. Hip DXAs were excluded from the current study if there was no pelvic radiograph performed on the same day, if there was no total body DXA image available (which was used to create the horizontal reference line to adjust for pelvic obliquity), if the image quality of either the DXA or radiograph was too low for identification of radiographic landmarks, if any of the radiographic landmarks were not depicted on the image, or if there was an artifact in the region of interest.

Image acquisition

The AP pelvic radiographs were weight-bearing, with the participants positioned with both feet in 10° internal rotation. The radiographs were obtained using a Solarize FV (General Electric CGR, Utrecht, the Netherlands). The right hip, left hip, and whole body DXA images were obtained using a GE-Lunar iDXA densitometer (GE Healthcare Lunar, Madison, WI, USA) and enCORE software (enCORE 2010; GE Healthcare) with participants positioned supine with legs slightly apart and big toes touching, the participant's feet were secured in this position using a Velcro band.

Radiographic hip osteoarthritis

Radiographic hip OA (RHOA) was determined on all pelvic radiographs using the Kellgren-Lawrence (KL) atlas based grading system in 5 grades (from 0 to 4) by one independent trained reader[39,116]. KL grade 0 was considered no RHOA, KL grade 1 was considered doubtful RHOA and KL grade ≥ 2 was considered definite RHOA.

Definition and calculation of radiographic measurements of hip morphology

The methods used to calculate radiographic measurements of hip morphology were developed in-house, are open-access, and were automated for consistency and accuracy[91]. The automated method was previously validated in the adult population[117]. The acetabulum and proximal femur were automatically outlined using 38 radiographic landmarks using the BoneFinder® software (www.bone-finder.com; The University of Manchester, UK)[64,117], see Figure 1A. The protocol for landmark definition can be found in Supplementary Material 1. The most cranial point of the obturator foramen and the most caudal point of the ischial tuberosity were used to calculate the horizontal reference line of the pelvis (HRLP) to correct the mAI, WCEA and LCEA measurements for any pelvic obliquity. Where the HRLP can naturally be determined on pelvic radiographs, the whole body DXA images were needed to determine the HRLP to correct for pelvic obliquity on the DXA images.

The radiographic landmark placement was visually assessed on all DXA images and radiographs independently, blinded, at least one month apart, and manual corrections were performed when necessary. The visual assessment was done to minimize the influence of incorrect landmark placement on the performance of automated measurements.

The intraobserver variability for both the DXA images and radiographs was assessed to further assess the influence of point placement on the measurements. To this end, the observer performed a second round of visual checks, where the observer visually assessed a second set of automatically placed landmarks in a random subset of 30 hips and performed manual corrections when necessary. The second round was performed independently, blinded, approximately one month after the first round of visual checks.

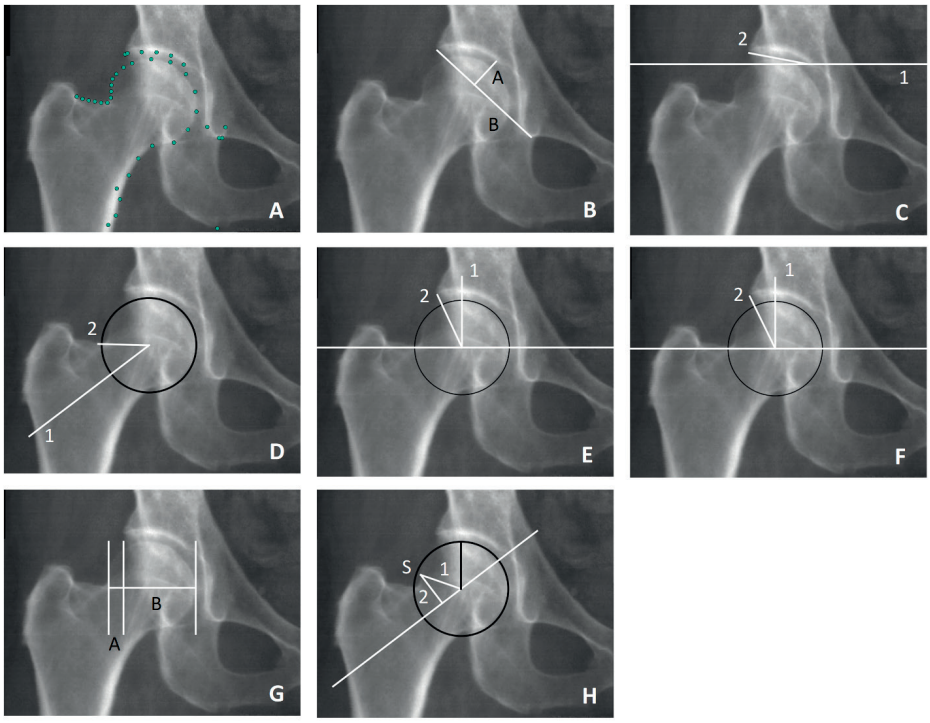


Figure 1. Example of landmark points and hip morphology measurements on a dual-energy x-ray absorptiometry (DXA) image. A: Overview of landmark points, 38 landmark points used to determine hip morphology measurements automatically. B: The acetabular depth-width ratio (ADR), ADR = (A/B)*1000, is the ratio between the acetabular depth (A), measured from the most medial point of the sourcil to line B, and the acetabular width (B), measured from the most lateral bony point of the acetabulum to the most caudal point of the teardrop. C: The modified acetabular index (mAI) is the angle between the horizontal reference line of the pelvis (HRLP) (line 1) and the line connecting the most lateral bony point of the acetabulum and the most medial point of the sourcil (line 2). D: The alpha angle (AA) is the angle between the femoral neck axis (line 1) and a line through the femoral head center and the alpha point (line 2). E: The Wiberg center edge angle (WCEA) is the angle between line 1, perpendicular to the HRLP, and line 2 through the femoral head center and the most lateral part of the acetabular sourcil. F: The lateral center edge angle (LCEA) is the angle between line 1, perpendicular to the HRLP, and line 2 through the femoral head center and the most lateral bony part of the acetabulum. G: The extrusion index (EI), EI = A/(A+B)*100%, is the ratio between the uncovered part of the femoral head (A) and the width of the femoral head (A+B). H: The triangular index ratio (TIR) is the ratio between the radius of the femoral head and the length of line 1, the line connecting the intersection point (point S) of the cortex of the femoral head and line 2 and the femoral head center.

We performed the following radiographic measurements on all DXA images and radiographs: the acetabular depth-width ratio (ADR), the modified acetabular index (mAI), the alpha angle (AA), the center edge angle of Wiberg (WCEA), the lateral center edge angle (LCEA), the extrusion index (EI), and the triangular index ratio (TIR), see Figure 1. One or more out of ADR ≤ 250, mAI ≥ 15°, EI ≥ 25% and WCEA ≤ 25° were used to define acetabular dysplasia[92,118]. An AA ≥ 60° or TIR ≥ 1.05 was defined as cam morphology[75,88,119] and an LCEA ≥ 40° was defined as pincer morphology[105].

Statistical analyses

We report the mean values for all measurements and standard deviations for each imaging modality. The agreement between the imaging modalities, as well as the intraobserver agreement, were visualized using Bland-Altman plots with limits of agreement. Additionally, the 95% confidence intervals (CI) of the mean differences and limits of agreements were determined. Histograms of the differences were created to assess whether they approximated a normal distribution. A mean difference larger than 2.5° for the mAI, AA, WCEA and LCEA and a mean difference larger than 1% of the measurement for the ADR, EI and TIR was defined as a systematic error. Any large differences in the measurements, as identified using the Bland-Altman plots, were visually assessed to see if any specific reasons for these differences could be found.

The reliability was tested through intraclass correlation coefficients (ICCs) and reported with 95% CI for all measurements. The AA measurement was transformed using Turkey's Ladder of Powers transformation to approximate a normal distribution, and ICCs were determined on the transformed data. The intermethod reliability was tested with a 2-way mixed-effects model, single rater, absolute agreement ICC, and the intraobserver reliability was tested with a 2-way random-effects model, single rater, absolute agreement ICC. ICCs were rated as poor (<0.50), moderate (0.50-0.75), good (0.76-0.90), or excellent (>0.90).

We conducted a sensitivity analysis using only one randomly selected hip from each participant to assess the impact of including two hips from one participant on the found results. The results of the sensitivity analysis are presented in Supplementary material 2.

The prevalence of diagnoses of acetabular dysplasia, cam morphology, and pincer morphology was determined and compared between the imaging modalities using percent agreement. The percent agreement was expressed as the number of corresponding diagnoses between DXA images and radiographs divided by the total number of hips assessed. Additionally, two-way frequency tables were created for each diagnosis and presented in Supplementary material 3. All statistical analyses were performed using R Statistical Software (v4.1.0; R Core Team 2021). The ICCs were calculated using the irr-package[73], the Turkey's Ladder of Powers transformation was performed using the rcompanion-package[120], the Bland-Altman analyses were performed using the blandr-package[121] and the Bland-Altman plots were created using the ggplot2-package[74].

RESULTS

411 individuals with 750 hips were included in the current study; the inclusion and exclusion can be found in Figure 2. The participants had a median age of 67.3 (range 52.2 – 90.6) years, and 45.5% were male, and 4.4% had definite RHOA see Table 1.

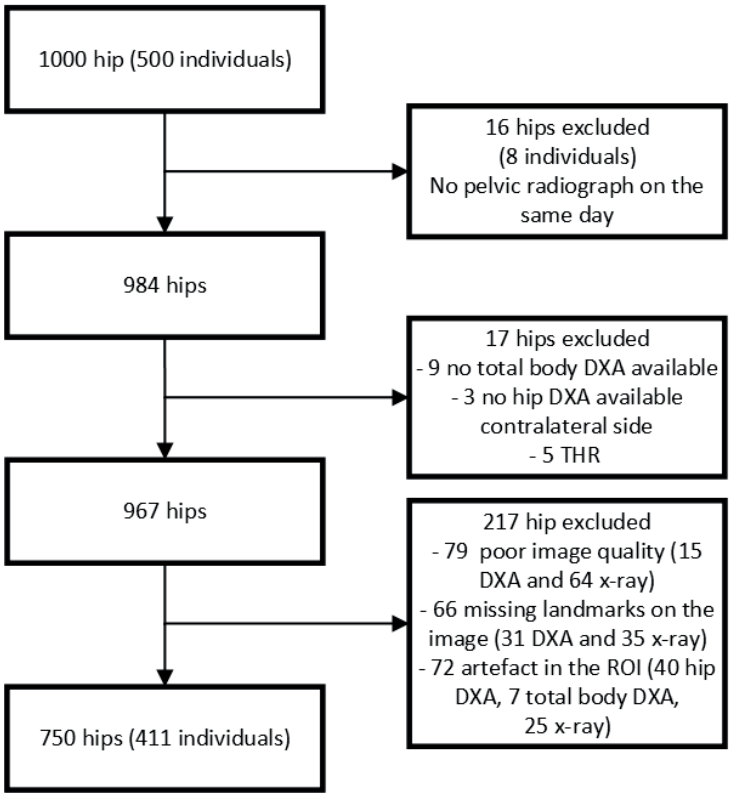


Figure 2. Flowchart of inclusion. DXA: dual-energy x-ray absorptiometry. THR: Total hip replacement. ROI: region of interest.

Table 1. Participant characteristics

	N (%)
Sample size	
Participants	411
Hips	750
Age, years (median (range))	67.3 (52.2 – 90.6)
Male	187 (45.5)
BMI, kg/m^2 (median (range))	26.2 (16.9 – 39.5)

Table 1. Participant characteristics (*continued*)

	N (%)
Hip right	381 (50.8)
RHOA*	
No	611 (81.5)
Doubtful	106 (14.1)
Definite	33 (4.4)

BMI: body mass index. RHOA: radiographic hip osteoarthritis. *On hip level.

Agreement

The Bland-Altman plots showing the agreement between the morphological measurements derived from the DXA images and radiographs and the intraobserver agreement within each imaging modality are shown in Figure 3. Additionally, the mean values of each measurement on both DXA and radiograph, as well as the intraobserver and intermethod mean difference with limits of agreement and 95% CIs, are summarized in Table 2. No systematic error was found for the measurements except for the EI between DXA images and radiographs. The EI was biased towards a lower value measured on DXA images than radiographs. Moreover, there was no agreement in the AA for angles between 60° and 80°.

The limits of agreement for the intraobserver agreement within each imaging modality consistently demonstrated equal or narrower limits of agreement compared to the intermethod agreement between the DXA images and radiographs. Similar results were found in the sensitivity analysis, see Supplementary material 2.

In the analysis of values exceeding the limits of agreement, two primary sources of measurement disparity were defined for every measurement. First, there was a (slight) difference in landmark point identification between the DXA images and radiographs present. This was particularly impactful on measurements reliant on individual landmark points for determination of the measurement, such as the ADR. Secondly, variation in the best-fitting circle around the femoral head resulted in measurement differences, where a slight difference in circle fit could result in different values, especially for the AA (Figure 4).

Reliability

The intraobserver and intermethod reliability for all measurements are shown in Table 3. The intraobserver reliability was comparable between DXA images and radiographs, indicating a comparable level of precision within each modality. The intraobserver reliability was better than the intermethod reliability of the DXA images compared to the

radiographs. However, the ICCs for the intermethod reliability were overall good, implying that measurements performed on both imaging modalities were similar. Similar reliability was found in the sensitivity analysis, except for the ADR, where the ICC was higher in the sensitivity analysis (0.85, 95% CI 0.74-0.91), see Supplementary material 2.

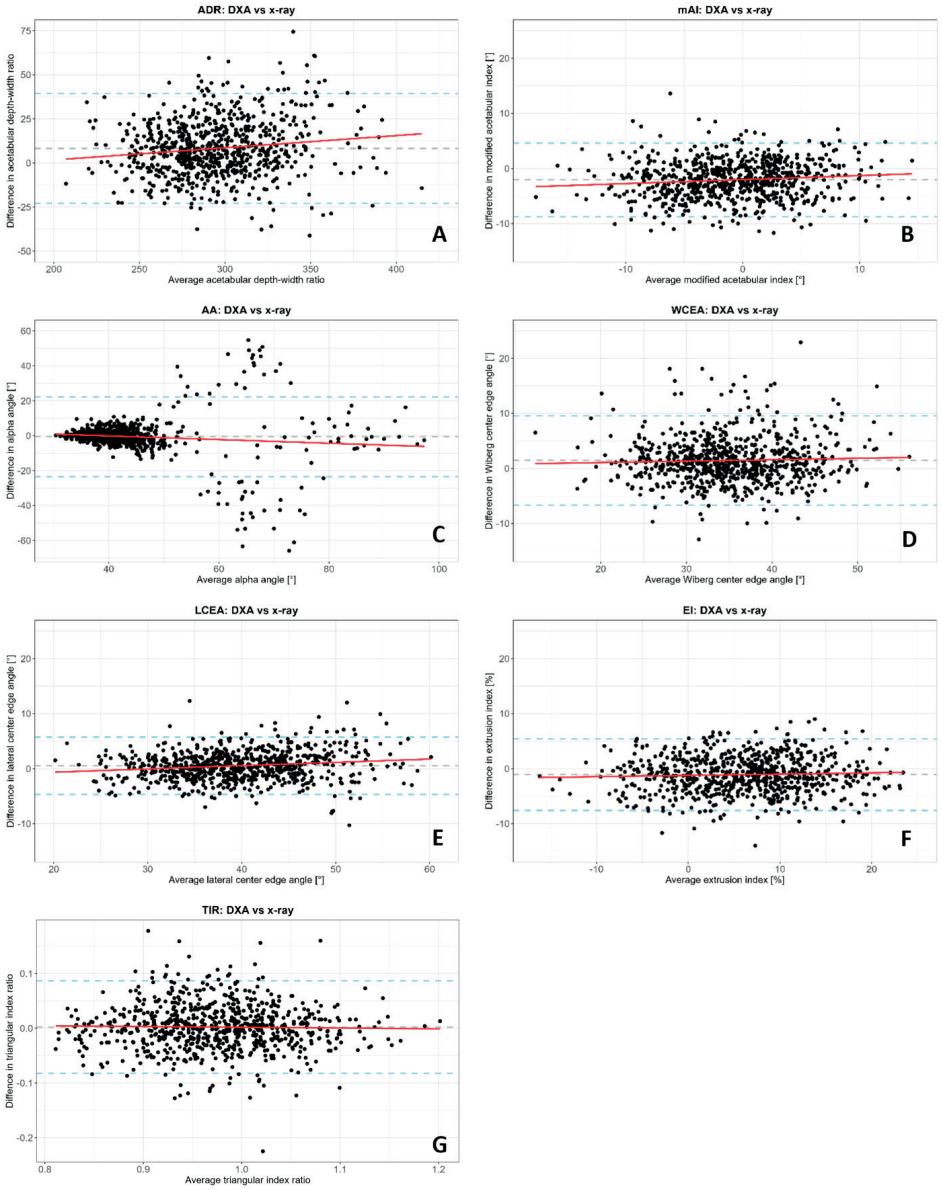


Figure 3. Bland-Altman plots of the intermethod agreement between the DXA images and radiographs for all morphological hip measurements. N=750. A: The acetabular depth-width ratio (ADR). B: The modified acetabular index (mAI). C: The alpha angle (AA). D: The Wiberg center edge angle (WCEA). E: The lateral center edge angle (LCEA). F: The extrusion index (EI). G: The triangular index ratio (TIR).

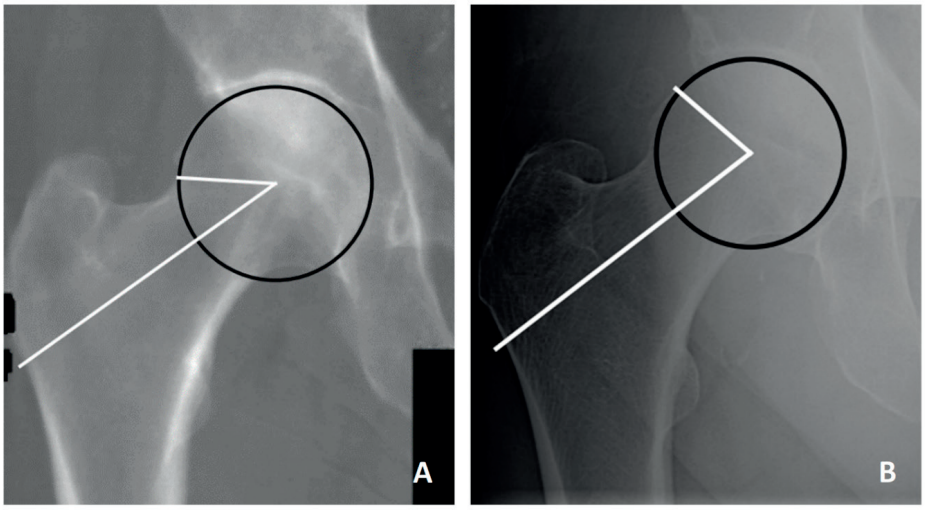


Figure 4. This is an example of how a slight difference in circle fit can result in a large difference in the measured alpha angle (AA). A: AA as determined on the right hip DXA image. B: AA as determined on the pelvic radiograph.

Table 2. Intraobserver and intermethod agreement.

	DXA			Pelvic radiograph			DXA vs pelvic radiograph	
Measurement	Mean (± SD)	Intraobserver mean difference (95% CI)	Intraobserver limits of agreement (95% CI)	Mean (± SD)	Intraobserver mean difference (95% CI)	Intraobserver limits of agreement (95% CI)	Intermethod mean difference (95% CI)	Intermethod limits of agreement (95% CI)
Acetabular depth-width ratio	298.8 (33.5)	-4.5 (-8.7 to -0.2)	-26.4 (-33.8 to -19.1) to 17.5 (10.1 to 24.9)	290.6 (31.4)	-3.2 (-7.5 to 1.2)	-25.6 (-33.1 to -18.1) to 19.3 (11.7 to 26.8)	8.2 (7.1 to 9.4)	-23.0 (-25.0 to -21.1) to 39.4 (37.5 to 41.4)
Modified acetabular index [°]	-1.8 (5.6)	2.5 (1.7 to 3.2)	-1.5 (-2.8 to -0.2) to 6.4 (5.1 to 7.7)	0.3 (5.2)	-0.1 (-0.7 to 0.6)	-3.5 (-4.7 to -2.4) to 3.4 (2.2 to 4.5)	-2.1 (-2.3 to -1.8)	-8.8 (-9.2 to -8.3) to 4.6 (4.2 to 5.0)
Alpha angle [°]	44.3 (12.2)	-1.6 (-5.1 to 1.8)	-19.1 (-25.0 to -13.2) to 15.9 (10.0 to 21.7)	44.9 (13.2)	0.8 (-2.6 to 4.1)	-16.5 (-22.4 to -10.7) to 18.1 (12.3 to 23.9)	-0.6 (-1.5 to 0.2)	-23.5 (-24.9 to -22.1) to 22.2 (20.8 to 23.7)
Wiberg center edge angle [°]	35.8 (7.1)	-2.0 (-3.2 to -0.7)	-8.6 (-10.8 to -6.4) to 4.7 (2.4 to 6.9)	34.4 (6.9)	-0.4 (-1.7 to 1.0)	-7.4 (-9.7 to -5.0) to 6.7 (4.3 to 9.0)	1.5 (1.2 to 1.7)	-6.7 (-7.2 to -6.2) to 9.6 (9.1 to 10.1)
Lateral center edge angle [°]	39.9 (7.2)	0.4 (-0.05 to 0.8)	-1.8° (-2.6 to -1.1) to 2.6° (1.9 to 3.4)	39.4 (6.8)	-0.1 (-0.5 to 0.3)	-2.1 (-2.8 to -1.4) to 1.8 (1.2 to 2.5)	0.5 (0.3 to 0.7)	-4.7 (-5.0 to -4.4) to 5.7 (5.4 to 6.0)
Extrusion index [%]	5.5 (7.1)	-0.9 (-1.5 to -0.3)	-4.1 (-5.2 to -3.0) to 2.3 (1.3 to 3.4)	6.6 (7.0)	0.1 (-0.3 to 0.5)	-1.9 (-2.6 to -1.2) to 2.1 (1.5 to 2.8)	-1.1 (-1.3 to -0.9)	-7.6 (-8.0 to -7.2) to 5.4 (5.0 to 5.8)
Triangular index ratio	0.976 (0.070)	0.006 (-0.005 to 0.017)	-0.053 (-0.072 to 0.033) to 0.065 (0.045 to 0.084)	0.974 (0.071)	0.0007 (-0.010 to 0.012)	-0.055 (-0.074 to -0.037) to 0.057 (0.038 to 0.076)	0.002 (-0.001 to 0.005)	-0.082 (-0.088 to -0.077) to 0.087 (0.081 to 0.092)

Bland-Altman interobserver and intermethod agreement between the dual-energy x-ray absorptiometry (DXA) images and pelvic radiographs of the hip morphology measurements. The agreement is presented as the mean difference with 95% confidence intervals (CI) and corresponding limits of agreement with 95% CI. For intraobserver agreement n=30, and for intermethod agreement n=750. SD: standard deviation.

Table 3. Intraclass correlation coefficients

	Intraobserver DXA		Intraobserver pelvic radiograph		DXA vs pelvic radiograph	
Measurement	ICC	95% CI	ICC	95% CI	ICC	95% CI
Acetabular depth-width ratio	0.92	0.83-0.96	0.93	0.86-0.97	0.85	0.74-0.91
Modified acetabular index [°]	0.87	0.18-0.96	0.94	0.87-0.97	0.75	0.52-0.85
Alpha angle *	0.74	0.52-0.87	0.85	0.70-0.93	0.72	0.68-0.75
Wiberg center edge angle [°]	0.88	0.68-0.95	0.89	0.78-0.95	0.81	0.74-0.85
Lateral center edge angle [°]	0.99	0.97-0.993	0.99	0.97-0.99	0.93	0.91-0.94
Extrusion index [%]	0.96	0.89-0.98	0.99	0.98-0.995	0.88	0.84-0.91
Triangular index ratio	0.85	0.70-0.93	0.88	0.75-0.94	0.81	0.79-0.84

Intraclass correlation coefficients (ICC) of intraobserver and intermethod reliability between the dual-energy x-ray absorptiometry (DXA) images and pelvic radiographs of the measurements of hip morphology. *Transformed using Tukey's Ladder Power transformation with lambda –3.1. For intraobserver reliability n=30, and for intermethod reliability n=750. ICCs are presented with 95% confidence intervals (CI). Intraobserver reliability was tested with a 2-way random-effects model, single rater, absolute agreement ICC. Intermethod reliability was tested with a 2-ways mixed-effects model, single rater, absolute agreement ICC. Interpretation: poor (<0.50), moderate (0.50-0.75), good (0.76-0.90), or excellent (>0.90).

Diagnostic agreement

Table 4 illustrates the prevalence and the intermethod diagnostic agreement for all measurements. For all measurements, except for the LCEA, the prevalence determined on DXA images was lower than that determined on pelvic radiographs.

Table 4. Prevalence and diagnostic agreement

Diagnosis	Prevalence DXA	Prevalence pelvic radiograph	Percent agreement (%); DXA vs pelvic radiograph
Acetabular dysplasia			
ADR ≤ 250	5.7% (43 out of 750)	8.8% (66 out of 750)	94
mAI ≥ 13°	0.3% (2 out of 750)	0.8% (6 out of 750)	99.2
EI ≥ 25%	0% (0 out of 750)	0% (0 out of 750)	100
WCEA ≤ 25°	5.6% (42 out of 750)	8.6% (65 out of 750)	92.9
Cam morphology			
AA ≥ 60°	8.4% (63 out of 750)	8.9% (67 out of 750)	92.3
TIR ≥ 1.05	13.2% (99 out of 750)	15.3% (115 out of 750)	90.9
Pincer morphology			
LCEA ≥ 40°	47% (353 out of 750)	45.2% (339 out of 750)	90.4

Intermethod percentage agreement of diagnosis between the two imaging modalities. ADR: acetabular-depth width ratio. mAI: modified acetabular index. EI: extrusion index. WCEA: center edge angle of Wiberg. AA: alpha angle. TIR: triangular index ratio. LCEA: lateral center edge angle.

DISCUSSION

We compared the agreement and reliability of automated hip morphology measurements between DXA images and pelvic radiographs on 750 hips from 411 individuals which were performed on the same day. Generally, the morphological measurements performed on DXA images and pelvic radiographs are mainly similar, although some large differences were present. The EI showed a slight systematic error between the DXA images and pelvic radiographs. Additionally, the prevalence of acetabular dysplasia and cam morphology determined on DXA images was lower than the prevalence based on pelvic radiographs.

To our knowledge, no studies have previously compared hip morphology measurements between DXA images and radiographs. Powell et al.[101] compared the ADR, mAI, LCEA, and EI between EOS imaging and radiographs. They found small biases and similar limits of agreement to those we have found while comparing DXA images and radiographs.

All morphological hip measurements have been studied previously regarding intra- and interobserver agreement and reliability[86,101,102,122-124]. Additionally, some studies reported intermethod agreement and reliability where manual measurements were compared to automated measurements[76-78,80,99]. The mean difference and limits of agreement for each measurement, as defined by the Bland-Altman plots observed in this study, were similar to those found in the literature[76,77,80,86,99,101,123]. The reported intraobserver and

intermethod reliability were comparable to or better than previously reported in the literature, although these were most often manual measurements[77,78,86,99,101,102,122-124]. The exceptions were the intraobserver and intermethod reliability of the alpha angle[76]. This is likely due to the susceptibility to differences in the best-fitting circle around the femoral head.

The observed differences between the measurements determined on DXA images compared to pelvic radiographs can partly be explained by landmark placement differences and the best-fitting circle around the femoral head. However, other possible reasons could explain the observed differences between the measurements. Namely, there are inherent differences between the two imaging modalities despite the DXA images and pelvic radiographs being obtained on the same day. First, the pelvic radiographs are weight-bearing, while the DXA images are non-weight-bearing. Secondly, pelvic positioning may contribute to differences in the projected hip joint. While we corrected for pelvic obliquity using the horizontal reference line, no correction was made for pelvic tilt or rotation. Pelvic tilt has been shown to affect the LCEA and mAI[120], while pelvic rotation has been shown to affect the WCEA[121]. Thirdly, the dissimilarities in beam positioning and shape between the two imaging modalities may further contribute to the observed differences. In pelvic radiographs, the x-ray beam is focused on the center of the pelvis, and the x-ray beam is cone-shaped, capturing the whole image in a single exposure. Conversely, a DXA image is made for each hip separately, where the x-ray beam is focused on the center of the hip and the image is acquired in rows using a fan beam. While iDXA densitometers employ narrow-angle fan beams that reduce distortion, differences in the x-ray beam can still contribute to the observed differences between the hip morphology measurements performed on DXA images compared to pelvic radiographs. However, despite these differences, which might affect the projected hip shape on an image, we still found high consistency across the morphological measurements between DXA and pelvic radiographs.

The current study has some limitations. First, the automatic search model for landmark point placement was trained using pelvic radiographs. Due to this, more manual adjustments were needed on the DXA images, potentially introducing more inaccuracy and error to the measurements. However, the intraobserver agreement and reliability were similar between the DXA images and pelvic radiographs, indicating that this probably did not influence the results. However, it might be important to create an open-access automatic search model on DXA images specifically to eliminate this bias in future studies. Secondly, all landmark placements were assessed manually and adjusted when necessary, which can introduce observer bias. This is the case based on the intraobserver analyses, where slight differences in landmark placement are present for both the DXA

images and the pelvic radiographs. The ADR, mAI, AA and EI can specifically be very sensitive to landmark placement. Next, no pelvic tilt and rotation correction was performed for the measurements. Adding a correction for pelvic tilt and rotation might improve the reliability and agreement of the measurements on both imaging modalities. Lastly, reproducibility of the findings can be influenced by varying image resolution and quality between different DXA and radiograph machines and software versions. Future research should therefore be performed to validate the findings in different populations, using various imaging acquisition protocols. Additionally, continuous improvement of imaging modalities can enhance measurement accuracy and reliability, while minimizing the costs and radiation burden.

As far as we know, this study is the first to compare hip morphology measurements between DXA images and pelvic radiographs. Additional strengths of the current study are the large number of included hips, image acquisition for both modalities on the same day, and automatic determination of the hip morphology measurements.

Based on the findings presented in this study, we propose that DXA images are a good alternative to radiographs for studying hip morphology despite the intrinsic differences in weight-bearing acquisition of the images. However, it's important to note that the agreement between all measurements extends beyond one standard deviation of the measured values. While this level of variability is consistent with existing research, the found limits of agreement can be considered large. This observation underscores the need for determining relevant differences in hip morphology measurements. Defining relevant differences would provide a clearer context for interpreting measurement variability and its clinical significance.

Both DXA images and radiographs have been used to evaluate hip morphology as a potential risk factor for RHOA. Specifically, cam morphology, defined by an alpha angle ≥ 60°, has been consistently associated with RHOA across both imaging modalities[8,11,58]. Using DXA images over radiographs will reduce radiation exposure, making it an advantageous choice for exploring hip morphology in pediatric populations and large-scale cohort studies. Studying the hip joint within the pediatric population can provide insight into the development of the hip. Additionally, it might provide more insight into different types of hip morphology, like acetabular dysplasia, cam and pincer morphology, and their development. An additional benefit of DXA lies in the simultaneous utility of the images for assessing bone mineral density, enhancing the comprehensiveness of musculoskeletal research.

In conclusion, we suggest that both DXA images and pelvic radiographs can be used to study hip morphology. The difference between the imaging modalities seems to be mainly due to the reliability of the measurements due to landmark point placement error, differences in pelvic positioning, and differences in the best-fitting circle around the femoral head. While pelvic radiographs might be the preference in clinical practice, DXA images can be a good alternative for research purposes due to the lower radiation burden. Further, our study opens the possibility for opportunistic screening of OA parameters in subjects undergoing DXA scans. This makes it feasible to study hip morphology within larger populations, aiding in our insight into the morphology of the hip and the risk of developing diseases such as hip OA.

Acknowledgements

The authors are grateful to the study participants, the Rotterdam Study staff, and the participating general practitioners and pharmacists.

Author contributions

RA, SBZ and FB conceptualized and designed the study. FR and JBJM facilitated the collection and assembly of the data. JW and FB assembled the data. FB analyzed and interpreted the data and drafted the manuscript. All the authors revised and approved the manuscript.

Role of funding source

The Rotterdam Study is funded by Erasmus University Medical Center and Erasmus University, Rotterdam, The Netherlands Organisation for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

Competing interest statement

The authors declare no conflict of interests.

SUPPLEMENTARY MATERIAL 1: LANDMARK PROTOCOL DXA VS X-RAY STUDY

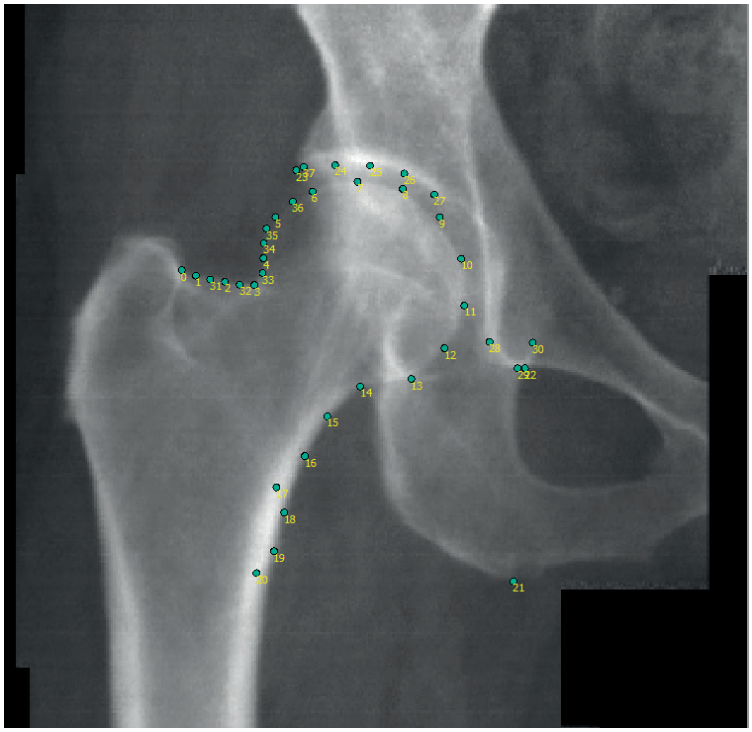


Figure 1. Example of the landmark points with point indices on a right hip DXA image.

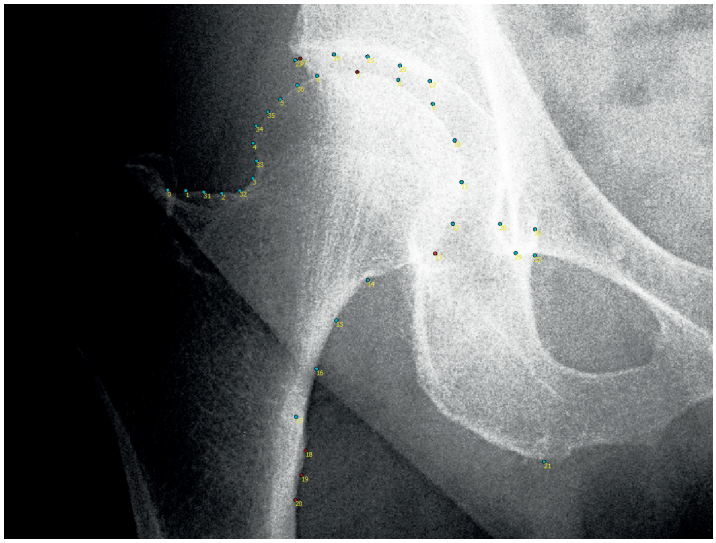


Figure 2. Example of the landmark points with point indices on a right hip of a pelvic radiograph.

Lesser trochanter

Point (**20**): Where the lesser trochanter starts bending off the shaft distally. If the lesser trochanter is seen behind the shaft, place this point on the cortex of the shaft at the level of this bend. If the lesser trochanter is not visible at all: missing points.
Point (**17**): Where the lesser trochanter joins the shaft proximally. If the lesser trochanter is seen behind the shaft, place this point on the cortex of the shaft at this level. If the lesser trochanter isn't visible at all: missing points.
Point (**18**)+(**19**): Respectively on the lower and upper corners of the lesser trochanter. If there are no clear corners: space them equally between (20) and (17).

Rest of proximal femur

Point (**0**): Where the anterior greater trochanter joins the femoral neck (usually at an angle and at a sclerotic corner).
Point (**5**): On the superolateral side of the femoral head, where the "best fitting circle" around the convexity of the femoral head seems to start. In case of a cam bump, osteophyte, or other irregularity: place (5) right after this bump ends, and the circle begins.
Point (**13**): On the inferomedial side of the femoral head, where the convexity of the femoral head seems to end. (The neck bends off after this point).
Point (**6-12**): Equally spaced between (5) and (13), unless there is a clear fovea dip, in which case the adjacent points are placed just outside of the fovea. Point (**9**) will be placed halfway across the 'semi'-circle between (5) and (13).
Point (**1**)+(**31**)+(**2**)+(**32**)+(**3**)+(**33**)+(**4**)+(**34**)+(**35**): Equally spaced between (0) and (5). In case of irregularities like a cam bump or osteophyte, follow the outlining contour as closely as possible.
Point (**36**): Equally spaced between (5) and (6).
Point (**14**): At the deepest point of the inferomedial concavity of the femoral neck, so that (13-17) will follow the medial cortex of the femoral neck as closely as possible.
Point (**15**)+(**16**): Equally spaced between (14) and (17), following the medial cortex of the femoral neck.

Ischium & Pubis

Point (**21**): On the most caudal point of the ischium (ischial tuberosity). If this is a straight line, put it in the middle.
Point (**22**): In the superolateral corner of the obturator foramen.

Acetabular roof

*** Points (37-27) along the weight-bearing zone (sourcil) are placed on the inferior rim of the sclerotic line.*

Point (**23**): On the most lateral point of the acetabulum, this can also be a lip/osteophyte.
Point (**37**): On the most lateral point of the **weight-bearing zone** (sourcil) of the acetabulum (most lateral point of sclerotic line).
Point (**27**): On the most medial point of the weight-bearing zone (sourcil) of the acetabulum, this is also the most superolateral point of the acetabular fossa. Usually there is a clear angle in the (sclerotic) line at the transition of weight-bearing zone to fossa. If the acetabular fossa is not visible at all, just place it on the most medial point of the sclerotic line.
Point (**24-26**): Along the underside of the sourcil, equally spaced between (37) and (27), following the contour of the weight-bearing zone

Pelvic teardrop

Point (**28**): On the superolateral corner of the visible teardrop (on the wall of the acetabular fossa)
Point (**29**): On the most caudal point of the teardrop.
Point (**30**): Across (28) on the other side of the teardrop.

General rules

- Osteophytes of the femoral head are *included* in the model. Follow the *outermost* contour. We can later correct for these with the radiological assessment data.
- Non-identifiable landmarks: missing points (write in separate log file)
- Only follow clear bony structures, not projecting shadows.

SUPPLEMENTARY MATERIAL 2: SENSITIVITY ANALYSIS

In this supplement the results of the sensitivity analysis are presented. In these analyses only one hip per participant was included. This resulted in the analysis of 411 hips from 411 individuals, see the flow diagram in Figure 1. The participants had a median age of 67.3 (range 52.2 – 90.6) years, and 45.5% were male, and 4.4% had definite RHOA see Table 1. The mean values of each measurement on both DXA and radiograph, as well as the Bland-Altman intermethod mean difference with limits of agreement and 95% CIs, are summarized in Table 2. The intermethod reliability for all measurements is presented in Table 3.

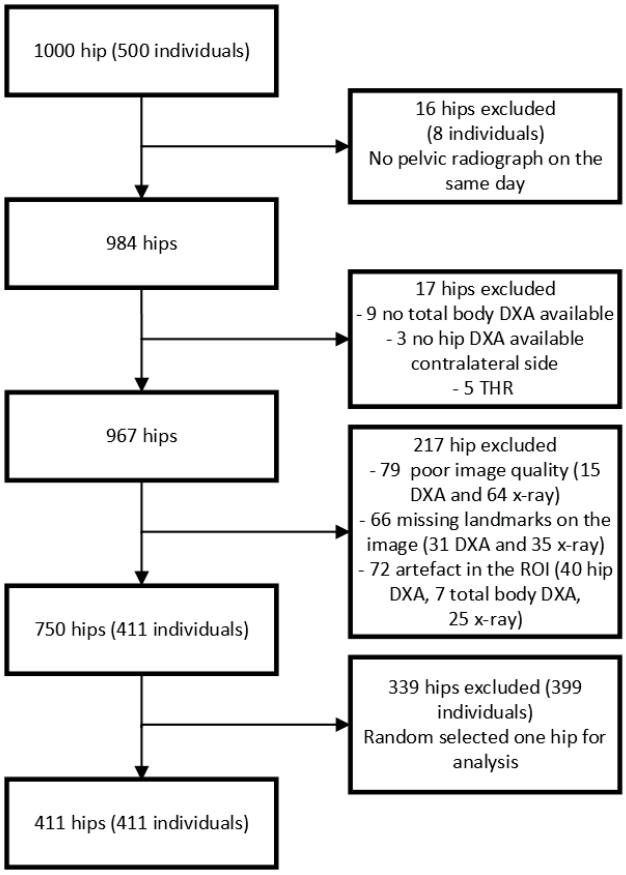


Figure 1. Flowchart of inclusion for the sensitivity analysis. DXA: dual-energy x-ray absorptiometry. THR: Total hip replacement. ROI: region of interest.

Table 1. Participant characteristics

	N (%)
Sample size	
Participants	411
Hips	411
Age, years (median (range))	67.3 (52.2 – 90.6)
Male	187 (45.5)
BMI, kg/m^2 (median (range))	26.2 (16.9 – 39.5)
Hip right	226 (55.0)
RHOA	
No	330 (80.3)
Doubtful	60 (14.6)
Definite	21 (5.1)

BMI: body mass index. RHOA: radiographic hip osteoarthritis.

Table 2. Intermethod agreement.

	DXA	**Pelvic radiograph**	**DXA vs pelvic radiograph**	
Measurement	Mean (± SD)	Mean (± SD)	Intermethod mean difference (95% CI)	Intermethod limits of agreement (95% CI)
Acetabular depth-width ratio	298 (33.7)	290 (31.2)	7.9 (6.4 to 9.5)	-23.3 (-26.0 to -20.7) to 39.1 (36.5 to 41.8)
Modified acetabular index [°]	-1.78 (5.5)	0.21 (5.1)	-2.0 (-2.3 to -1.7)	-8.5 (-9.0 to -7.9) to 4.5 (4.0 to 5.1)
Alpha angle [°]	44.5 (12.2)	45.0 (13.3)	0.02 (-1.1 to 1.1)	-22.3 (-24.2 to -20.4) to 22.3 (20.5 to 24.2)
Wiberg center edge angle [°]	35.8 (6.9)	34.5 (6.8)	1.5 (1.1 to 1.9)	-6.7 (-7.4 to -6.0) to 9.8 (9.1 to 10.5)
Lateral center edge angle [°]	39.9 (7.0)	39.4 (6.7)	0.70 (0.45 to 0.96)	-4.5 (-4.9 to -4.1) to 5.9 (5.5 to 6.4)
Extrusion index [%]	5.9 (6.9)	6.8 (6.8)	-1.3 (-1.6 to -1.0)	-7.8 (-8.3 to -7.2) to 5.2 (4.6 to 5.7)
Triangular index ratio	0.980 (0.069)	0.974 (0.071)	0.007 (0.003 to 0.011)	-0.076 (-0.083 to -0.069) to 0.090 (0.083 to 0.097)

Bland-Altman intermethod agreement between the dual-energy x-ray absorptiometry (DXA) images and pelvic radiographs of the hip morphology measurements. The agreement is presented as the mean difference with 95% confidence intervals (CI) and corresponding limits of agreement with 95% CI. N=411. SD: standard deviation.

Table 3. Intraclass correlation coefficients

	DXA vs pelvic radiograph	
Measurement	ICC	95% CI
Acetabular depth-width ratio	0.85	0.74-0.91
Modified acetabular index [°]	0.75	0.53-0.85
Alpha angle *	0.71	0.66-0.75
Wiberg center edge angle [°]	0.80	0.73-0.85
Lateral center edge angle [°]	0.92	0.90-0.94
Extrusion index [%]	0.87	0.81-0.91
Triangular index ratio	0.82	0.79-0.85

Intraclass correlation coefficients (ICC) of intermethod reliability between the dual-energy x-ray absorptiometry (DXA) images and pelvic radiographs of the measurements of hip morphology. *Transformed using Tukey's Ladder Power transformation with lambda –3.1. N=411. ICCs are presented with 95% confidence intervals (CI). Intraobserver reliability was tested with a 2-way random-effects model, single rater, absolute agreement ICC. Intermethod reliability was tested with a 2-ways mixed-effects model, single rater, absolute agreement ICC. Interpretation: poor (<0.50), moderate (0.50-0.75), good (0.76-0.90), or excellent (>0.90).

SUPPLEMENTARY MATERIAL 3: FREQUENCY TABLES

In this supplement the frequency tables are presented for acetabular dysplasia, cam and pincer morphology as determined on DXA images and radiographs.

Acetabular dysplasia

ADR ≤ 250		Radiograph	
		No	Yes
DXA image	No	673	34
	Yes	11	32

mAI ≥ 15°		Radiograph	
		No	Yes
DXA image	No	743	5
	Yes	1	1

EI ≥ 25%		Radiograph	
		No	Yes
DXA image	No	750	0
	Yes	0	0

WCEA ≤ 25°		Radiograph	
		No	Yes
DXA image	No	670	38
	Yes	15	27

Cam morphology

AA ≥ 60°		Radiograph	
		No	Yes
DXA image	No	656	31
	Yes	27	36

TIR ≥ 1.05		Radiograph	
		No	Yes
DXA image	No	609	42
	Yes	27	73

Pincer morphology

LCEA ≥ 40°		Radiograph	
		No	Yes
DXA image	No	368	29
	Yes	43	310

Part 2

Hip Morphology in Children

Chapter 5

Prevalence and radiological definitions of acetabular dysplasia after the age of 2 years: a systematic review

S. de Vos-Jakobs, F. Boel, W.M. Bramer, S.M.A. Bierma-Zeinstra, R. Agricola.

Journal of Pediatric Orthopaedics. 2024 Jul 1;33(4):334-339.

ABSTRACT

Acetabular dysplasia is one of the most common causes of early hip osteoarthritis and hip replacement surgery. Recent literature suggests that acetabular dysplasia does not always originate at infancy, but can also develop later during childhood. This systematic review aims to appraise the literature on prevalence numbers of acetabular dysplasia in children after the age of 2 years. A systematic search was performed in several scientific databases. Publications were considered eligible for inclusion if they presented prevalence numbers on acetabular dysplasia in a general population of healthy children aged 2 18 years with description of the radiological examination. Quality assessment was done using the Newcastle-Ottawa score. Acetabular dysplasia was defined mild when: the center edge angle of Wiberg (CEA-W) measured 15-20°, the CEA-W ranged between -1 to -2SD for age, or based on the acetabular index using thresholds from the Tönnis table. Severe dysplasia was defined by a CEA-W < 15°, < -2SD for age, or acetabular index according to Tönnis. Of the 1837 screened articles, four were included for review. Depending on radiological measurement, age and reference values used, prevalence numbers for mild acetabular dysplasia vary from 13.4 to 25.6% and for severe acetabular dysplasia from 2.2 to 10.9%. Limited literature is available on prevalence of acetabular dysplasia in children after the age of 2 years. Prevalence numbers suggest that acetabular dysplasia is not only a condition in infants but also highly prevalent later in childhood.

INTRODUCTION

Developmental dysplasia of the hip (DDH) is the single most common musculoskeletal disorder in infants and young children. It occurs in 5–10% of live births throughout Western countries. DDH includes a broad spectrum of hip pathology from hip dislocation up to stable hips with acetabular dysplasia[17].

Treatment of DDH is based on preservation of the native hip joint and resolving acetabular dysplasia. Treating acetabular dysplasia is important in order to establish a wide load-bearing acetabular surface for evenly distributed weightbearing and therefore diminishing risk for osteoarthritis and total hip replacement in the long term[52].

However, despite all efforts in treating DDH during childhood, prevalence rates for acetabular dysplasia still remain high in adults (5–21%)[13,52,125,126]. This suggests that either current treatment is insufficient or that a large number of children who eventually develop acetabular dysplasia remain out of scope.

Interestingly, DDH is currently considered to originate in infants and babies. Screening programs, therefore, focus on diagnosis and treatment in the first months of life but do not take late-onset or developmental factors into account. However, other studies suggest that acetabular dysplasia can also develop later during growth[126,127] and might be influenced by environmental factors during childhood[106]. This might be one of the reasons that acetabular dysplasia often remains undiagnosed and untreated and therefore might (partially) explain the high prevalence of acetabular dysplasia in adults[10].

Another reason for the high prevalence numbers in adults can be due to radiological measurements used to quantify acetabular dysplasia. Most radiological measurements used for (residual) acetabular dysplasia are based on reference values for adults. Only Tönnis table for acetabular index provides data specific for age, gender and laterality during childhood up to 7 years of age[128]. If center-edge angle of Wiberg (CEA-W) or the lateral center-edge-angle (LCEA) are used in children, reference values are based on adult values[127,129].

Therefore, this systematic review aims to appraise the literature on prevalence of acetabular dysplasia in the general population of children after the age of 2 years. Second, we aim to describe the radiological measurements used to diagnose acetabular dysplasia during childhood.

METHODS

The protocol for this systematic research was published in the PROSPERO database, reference number CRD42021282217.

Data sources and study selection

The methods are described based on the Preferred reporting items for systematic reviews and meta-analyses (PRISMA) checklist[130] and the PRISMA-S extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews[131]. An exhaustive search strategy was developed by an experienced information specialist (W.M.B.). The original search was developed in October 2021 in Embase.com, optimized for sensitivity, then translated to other databases and later updated in May 2022 following the method as described by Bramer et al.[132,133]. The search was carried out in the databases Embase.com (date of inception 1971), Medline ALL via Ovid (1946 to Daily Update), Web of Science Core Collection and the Cochrane Central Register of Controlled Trials via Wiley (date of inception 1992).

The search strategies for Embase and Medline used relevant thesaurus terms from Emtree and Medical Subject Headings (MeSH), respectively. In all databases, terms were searched in titles and abstracts of references. The search contained terms for (1) hip dysplasia or congenital hip dislocation and (2) either a combination of incidence or epidemiology in children or terms related to diagnostic delay or late presentation. Terms were combined with Boolean operators AND and OR and proximity operators were used to combined terms into phrases. The full search strategies of all databases are available in the supplementary materials (Appendix A, Supplemental digital content 1, http://links.lww.com/JPOB/A83). The searches in Embase and Web of Science were limited to exclude conference papers. In all databases, non-English articles, and animal-only articles were excluded from the search results. No study registries were searched, but Cochrane CENTRAL retrieves the contents of ClinicalTrials.gov and WHO's International Clinical Trials Registry Platform. According to the methodology proposed by Bramer et al.[134-136] the following steps were taken: (1) the reference lists of retrieved non-included relevant review articles and of the included references, as well as articles citing these papers have been scanned for relevant references missed by the search; (2) the references were imported into EndNote and duplicates were removed; (3) two reviewers (S.d.V. and F.B.) independently screened titles and abstracts in EndNote. Any discrepancies in the verdict were resolved by discussion with a third reviewer (RA). Next, full texts were retrieved for (preliminary) included articles. Definite inclusion was done by reading full text of the remaining articles by two independent reviewers (S.d.V. and F.B.) and discrepancies were again resolved by a third reviewer (R.A.).

The inclusion and exclusion criteria are listed in Table 1.

Table 1. In- and exclusion criteria.

Inclusion	Exclusion
Reported prevalence of acetabular dysplasia in children aged 2-18 years (or subgroup analysis within the age range)	Comorbidities compromising hip development (such as neuromuscular disorders or syndromal diseases)
Data from which prevalence numbers can be calculated (e.g. incidence and/or reference values)	Non-ambulatory children
General population (including children with and without acetabular dysplasia)	Solely hip dislocation
Diagnosis acetabular dysplasia confirmed by any imaging modality (radiograph, DXA, ultrasound, CT, MRI)	

Quality assessment

Quality assessment of the included articles was done, using the Newcastle-Ottawa scale (NOS) for cross-sectional studies[137]. This questionnaire was specified for the topic (Appendix B, Supplemental digital content 2, http://links.lww.com/JPOB/A84). Both reviewers (S.d.V. and F.B.) independently calculated a NOS score and discrepancies were again solved by the third reviewer (R.A.). The scores for 3 aspects of quality (selection, comparability and outcome) were separately used for the definite estimation of quality.

Studies that scored a total of 7 or 8 points were considered to have a low risk of bias; 6 points were considered to have a medium risk of bias; 5 points or less were considered to have a high risk of bias[137].

Data extraction

Before reading the articles, a data extraction form was composed by the authors (Appendix C, Supplemental digital content 3, http://links.lww.com/JPOB/A85). This information was extracted by 2 reviewers (S.d.V. and F.B.) independently and discussed in order to achieve agreement. When provided, prevalence numbers for mild acetabular dysplasia (CEA-W or LCEA 15–20° or −1 to 2 SD; acetabular index determined by Tönnis) and severe acetabular dysplasia (CEA-W or LCEA < 15° or <2SD; acetabular index determined by Tönnis) will be reported separately for each study. If appropriate, prevalence data of studies will be pooled.

RESULTS

The systematic search resulted in a total of 1837 articles. Figure 1 shows the flow from the initial searches to the final inclusion of four articles.

Quality assessment

The quality assessment for the four included articles is summarized in Table 2. Only the study by Chung et al.[106] had an overall low risk of bias.

Prevalence

Pooling was not appropriate due to the heterogeneous character of the data. For one of the studies[141] the prevalence was calculated by ourselves, using the data and numbers from the reference values reported in the article. In Table 3, the study characteristics and outcomes are summarized.

Chung et al.[106] described the prevalence of acetabular dysplasia in a randomly selected cross-sectional subgroup of 9-year-olds from an ongoing prospective population-based cohort (Generation R). Besides CEA-W, acetabular depth-width ratio (ADR) was used to determine acetabular development. Since the cutoff values for ADR were chosen so, that the prevalence of acetabular dysplasia was similar to the prevalence measured with CEA-W, these measurements are not included in our systematic review.

Akel et al.[138] derived their population from a database of lower abdomen and pelvis radiographs for non-dysplasia-related causes. Age groups were defined per year. Cut-off values and prevalence vary by age group. More detailed information for the various age groups is attached in Appendix D, Supplemental digital content 4, http://links.lww.com/JPOB/A86.

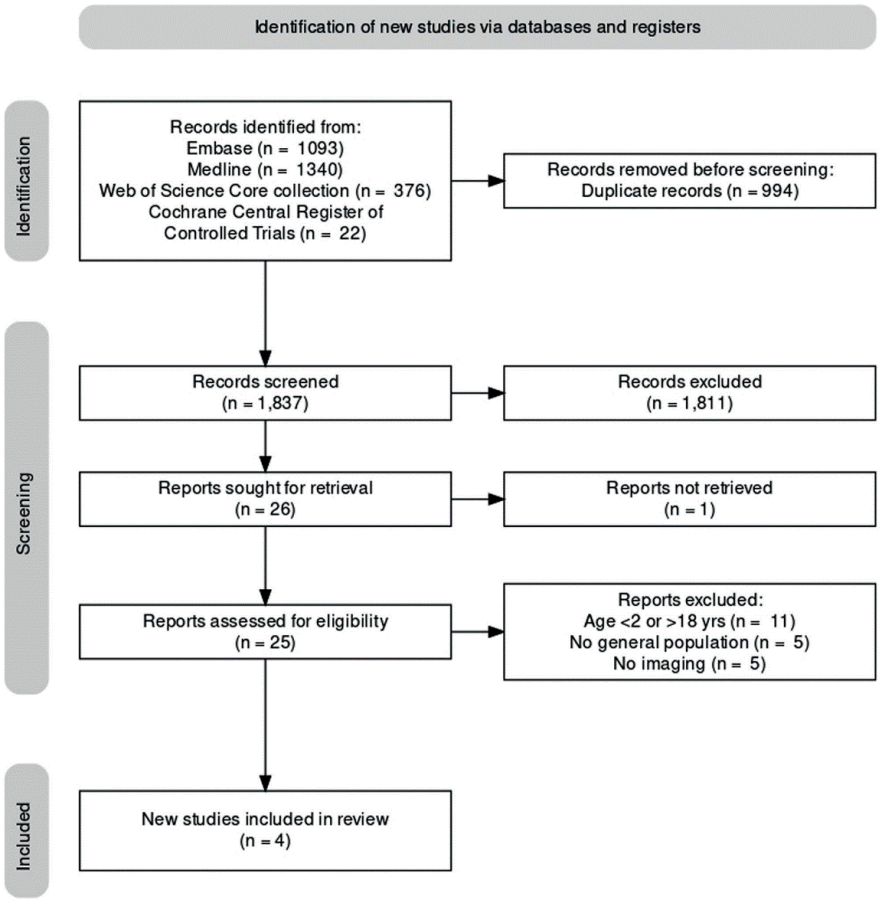


Figure 1. PRISMA flow diagram for article inclusion.

Table 2. Outcomes of NOS for quality assessment of cross-sectional studies

The Newcastle-Ottawa Scale, scores per study				
			NOS score	
Authors, year	**Study design**	**Selection**	**Comparability**	**Outcome**
Chung[106], 2021	Cross-sectional cohort	☆☆	☆☆	☆☆☆
Akel[138], 2013	Cross-sectional cohort	☆	☆☆	☆☆
Tugrul[139], 2020	Cross-sectional cohort	☆	☆☆	☆
Shi[140], 2010	Cross-sectional cohort	☆	☆☆	-

Tugrul et al.[139] used the same database as Akel et al. and defined similar age groups (ages 5–14 years, groups of 1 year). cutoff values for dysplasia were estimated by their own measurements (CEA-W −1 to −2 SD for mild dysplasia, CEA-W < −2SD for severe dys-

plasia). This resulted in reference values varying by age, gender and laterality (Appendix E, Supplemental digital content 5, http://links.lww.com/JPOB/A87).

Shi et al.[140] designed their study to establish reference values for CEA-W for the Chinese population per age group. In order to do so, they used a database of radiographs 'for routine examination or exclusion of pelvic trauma'. With the use of the 95% confidential interval, prevalence was calculated.

DISCUSSION

The overall prevalence of mild acetabular dysplasia in children aged 2 years and older is estimated between 13.4 and 25.6% and for severe acetabular dysplasia between 2.2 and 10.9% when measured by CEA-W or acetabular index. While acetabular dysplasia in otherwise healthy children is currently thought to develop during infancy and improve over time, these results show that prevalence remains high in later childhood. In the reviewed literature CEA-W and acetabular index are most widely used, but also ADR can be measured as indicator for acetabular dysplasia.

Limited number of studies available on acetabular dysplasia during childhood

In the systematic search, only four articles met our inclusion criteria indicating that the number of studies on prevalence of acetabular dysplasia after the age of 2 years is limited. In contrast, large numbers of studies were published on prevalence of DDH during infancy[141,142] or prevalence of late-diagnosed hip dislocation[143,144].

Also, when numbers on DDH after the age of 2 years were presented, this was not in the general population, but in more biased populations such as hospitalized patients. For that reason further (longitudinal) research in the general population is essential to acquire more information on the development of acetabular dysplasia during growth.

Patient selection and representation

Of the four included studies, only one study (Chung et al.[106]) was an unadulterated general (multi-ethnic) population study where high-resolution DXA's were derived for research purpose only and not for treatment or diagnostic purposes. All other studies used radiographs that were made for other purposes but considered it a sample of the general population as no hip complaints were reported. Akel et al.[138] and Tugrul et al.[139] both used radiographs from the same database derived for 'non-dysplasia related causes', but no information was provided on the indication for radiographs.

Table 3. Summary of study data

Summary of studies										
	Cohort characteristics				**Imaging characteristics**		**Outcomes**			
Author, year	Sample size (no. of hips)	Age (years)	Gender	Ethnicity	Imaging modality	Measurement	Cutoff value severe dysplasia	Prevalence severe dysplasia (%)	Cutoff value mild dysplasia	Prevalence mild dysplasia (%)
Chung[106], 2021	1026	9.86	♀ 49.8% ♂ 50.2%	European 79.5% Asian 7.7% African 11.2% Other 1.6%	HR-DXA	CEA-W	< 15°	4.8	15-20°	25
Akel[138], 2013	4622	2-8	♀ 52.9% ♂ 47.1%	Turkish	Radiograph	Acetabular index	Defined by Tönnis > + 2 SD	2.6 – 10.9 2.9 – 4.8	Defined by Tönnis + 1 to 2 SD	14.7 – 29.9 14.9 – 21.6
Tugrul[139], 2020	3192	5-14	♀ 46.1% ♂ 53.9%	Turkish	Radiograph	CEA-W	< - 2 SD	2.21	- 1 to 2 SD	14.06
Shi[140], 2010	1892	4-17	♀ 38.5% ♂ 61.5%	Multi-ethnical Chinese	Radiograph	CEA-W	< - 2 SD	2.2	- 1 to 2 SD	13.4

CEA-W, center-edge angle of Wiberg; HR-DXA, high-resolution dual-energy X-ray absorptiometry.

Healthy, non-complaining children will probably not routinely have this radiograph obtained and therefore a potential selection bias cannot be ruled out. Similarly, Shi et al. used radiographs taken for 'routine examination or exclusion of pelvic trauma', probably in an emergency setting, but this information is not provided.

Outcomes of the study of Shi et al.[140] might be less representative for populations outside China. Far more male participants than female participants were included, and it is known that acetabular dysplasia is more common in females than in males[17]. Therefore, prevalence numbers for the general population might be underestimated. On the other hand, prevalence of acetabular dysplasia is known to be higher in Asian populations compared to Caucasian populations[145]. Altogether, we conclude that the prevalence numbers from this study are less representative of non-Chinese population than the prevalence numbers of the other reviewed studies.

Prevalence numbers and reference values

Both the study of Tugrul et al.[139] and Shi et al.[140] used their own calculated cutoff values to estimate the prevalence numbers of acetabular dysplasia based on the 95% confidential interval. As a result of this method, one can anticipate that prevalence numbers will be close to 13.6% for mild acetabular dysplasia (<−1 SD) and 2.2% for severe acetabular dysplasia (<−2 SD). In normally distributed data these percentages represent ±1 SD and ±2 SD, respectively. For that reason, the prevalence number derived from these studies are less informative than from the study of Chung et al.[106] and Akel et al.[138].

Also, for the estimation of cutoff values, this might not be the optimal approach. With this method, the assumption is made that 13.6% of the population has mild acetabular dysplasia and 2.2% has severe dysplasia, but this might be incorrect, given prevalence numbers of acetabular dysplasia in adults[10,125,126]. As the spectrum of DDH is more common in females than in males[17], this calculation method results in gender-specific normal values (−1 SD and −2 SD are at different values), ultimately leading to either overestimation of acetabular dysplasia in males or underestimation of acetabular dysplasia in females. A more reliable method to establish normal values would be using a certain outcome in time such as developing hip symptoms in young adulthood or osteoarthritis later in life. At skeletal maturity hip joints of both males and females equally require a wide load-bearing acetabular surface for diminishing risk of hip complaints in the long term[52]. Preferably longitudinal studies should be performed, where these pathological outcomes can be correlated to acetabular development and threshold values in children.

Only Akel et al.[138] present prevalence numbers based on previously verified cutoff values for the specific age groups (Tönnis' table for acetabular index). Still, the prevalence of acetabular dysplasia in this study is equally as high as in the other reviewed studies.

Definition of radiological measurements

The studies of Chung et al.[106], Tugrul et al.[139] and Shi et al.[140] refer to their measurements as 'center-edge angle of Wiberg'. Chung et al.[106] and Tugrul et al.[139] specify in their text how the measurement is performed, while Shi et al.[140] don't provide details on the measurement. Based on the recent consensus on measurement of the center-edge angle[53], we conclude that both Chung et al.[106] and Tugrul et al.[139] have actually reported the lateral center-edge angle (LCEA) instead of the CEA-W. LCEA refers to the most lateral point of the acetabulum and CEA-W refers to the most lateral point of the acetabular source, so these values aren't always equal. Especially in dysplastic hips, CEA-W is often lower than LCEA. These studies use reference values for (the lower) CEA-W and compare them with a measured (mostly higher) LCEA, this may lead to an underestimation of the prevalence of acetabular dysplasia in both studies.

CONCLUSION

Prevalence of mild acetabular dysplasia in children over 2 years of age is 13.4–25.6%, and for severe dysplasia prevalence is 2.2–10.9%. Very limited data is available, but based on the reviewed data, prevalence of acetabular dysplasia varies strongly by age, method of measuring, and estimated cutoff values.

Either way, acetabular dysplasia not only seems to be a condition in infants but is also of great importance later in childhood. For this reason, health care practitioners should be more suspicious for acetabular dysplasia, also when DDH is ruled out during infancy. Future longitudinal studies in general, multi-ethnic populations are essential for evaluation of acetabular development, its determinants and prognostic implications.

Conflicts of interest

There are no conflicts of interest.

APPENDIX A – COMPLETE SEARCH STRATEGY

Embase.com

('hip dysplasia'/exp OR 'congenital hip dislocation'/de OR ('congenital joint dislocation'/de AND (hip/de OR 'hip dislocation')) OR 'hip dislocation'/exp/dm_cn OR 'acetabular dysplasia'/de OR (((hip OR hips OR Acetabul*) NEAR/6 (dysplas*)) OR ((hip OR hips OR Acetabul*) NEAR/6 (development* OR congenital*) NEAR/6 (dislocat*))):ab,ti) AND (((incidence/exp OR epidemiology/de OR 'hip dysplasia'/exp/dm_ep OR 'hip dislocation'/exp/dm_ep OR prevalence/de OR screening/exp OR morbidity/de OR 'geographic distribution'/de OR (inciden* OR epidemiolog* OR prevalen* OR screening OR frequen* OR morbidit* OR (geograph* NEAR/3 distribut*)):Ab,ti) AND ('preschool child'/exp OR child/de OR adolescent/de OR childhood/de OR adolescence/de OR (preschool* OR pre-school* OR (child* NEAR/6 (2 OR 3 OR 4 OR 5) NEXT/1 (year*)) OR (child* NEAR/6 age* NEXT/1 (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16)) OR (age* NEXT/1 (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16) NEXT/1 year*) OR (child* NEAR/6 (24 OR 36 OR 48 OR 60 OR 72 OR 84 OR 96 OR 108 OR 120 OR 132 OR 144 OR 156 OR 168 OR 180 OR 192) NEXT/1 month*)):Ab,ti)) OR ('missed diagnosis'/exp OR 'delayed diagnosis'/de OR 'diagnostic delay'/de OR 'onset age'/de OR 'diagnostic error'/exp OR (((late* OR delay* OR missed OR Enhanc* OR error*) NEAR/3 (presentation* OR diagnos* OR detect* OR onset*)) OR (false NEAR/3 negative*)):Ab,ti)) NOT [conference abstract]/lim NOT ([animals]/lim NOT [humans]/lim) AND [english]/lim

Medline ALL Ovid

(Hip Dislocation/ OR Hip Dislocation, Congenital / OR Developmental Dysplasia of the Hip/ OR (Joint Dislocations AND (Hip / OR Hip Dislocation/)) OR (((hip OR hips OR Acetabul*) ADJ6 (dysplas*)) OR ((hip OR hips OR Acetabul*) ADJ6 (development* OR congenital*) ADJ6 (dislocat*))).ab,ti.) AND (((Incidence/ or Epidemiology/ or Prevalence/ or Mass Screening/ or Morbidity/ OR Hip Dislocation/ep OR (inciden* OR epidemiolog* OR prevalen* OR screening OR frequen* OR morbidit* OR (geograph* ADJ3 distribut*)).ab,ti.) AND (Child, Preschool/ OR Child/ OR Adolescent/ OR (preschool* OR pre-school* OR (child* ADJ6 (2 OR 3 OR 4 OR 5) ADJ (year*)) OR (child* ADJ6 age* ADJ (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16)) OR (age* ADJ (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16) ADJ year*) OR (child* ADJ6 (24 OR 36 OR 48 OR 60 OR 72 OR 84 OR 96 OR 108 OR 120 OR 132 OR 144 OR 156 OR 168 OR 180 OR 192) ADJ month*)).ab,ti.)) OR (Missed Diagnosis/ or Delayed Diagnosis/ OR Age of Onset/ OR exp Diagnostic Errors/ OR (((late* OR delay* OR missed OR Enhanc* OR error*) ADJ3 (presentation* OR diagnos* OR detect* OR onset*))

OR (false ADJ3 negative*)).ab,ti.)) NOT (conference abstract) NOT (exp animals/ NOT humans/) AND english.la.

Web of Science Core Collection*

*Science Citation Index Expanded (1975-present) ; Social Sciences Citation Index (1975-present) ; Arts & Humanities Citation Index (1975- present) ; Conference Proceedings Citation Index- Science (1990-present) ; Conference Proceedings Citation Index- Social Science & Humanities (1990-present) ; Emerging Sources Citation Index (2015-present)

TS=(((((hip OR hips OR Acetabul*) NEAR/5 (dysplas*)) OR ((hip OR hips OR Acetabul*) NEAR/5 (development* OR congenital*) NEAR/5 (dislocat*)))) AND ((((inciden* OR epidemiolog* OR prevalen* OR screening OR frequen* OR morbidit* OR (geograph* NEAR/2 distribut*))) AND ((preschool* OR pre-school* OR (child* NEAR/5 (2 OR 3 OR 4 OR 5) NEAR/1 (year*)) OR (child* NEAR/5 age* NEAR/1 (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16)) OR (age* NEAR/1 (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16) NEAR/1 year*) OR (child* NEAR/5 (24 OR 36 OR 48 OR 60 OR 72 OR 84 OR 96 OR 108 OR 120 OR 132 OR 144 OR 156 OR 168 OR 180 OR 192) NEAR/1 month*)))) OR ((((late* OR delay* OR missed OR Enhanc* OR error*) NEAR/2 (presentation* OR diagnos* OR detect* OR onset*)) OR (false NEAR/2 negative*)))))

Cochrane Central Register of Controlled Trials

((((hip OR hips OR Acetabul*) NEAR/6 (dysplas*)) OR ((hip OR hips OR Acetabul*) NEAR/6 (development* OR congenital*) NEAR/6 (dislocat*))):ab,ti) AND ((((inciden* OR epidemiolog* OR prevalen* OR screening OR frequen* OR morbidit* OR (geograph* NEAR/3 distribut*)):Ab,ti) AND ((preschool* OR pre-school* OR (child* NEAR/6 (2 OR 3 OR 4 OR 5) NEXT/1 (year*)) OR (child* NEAR/6 age* NEXT/1 (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16)) OR (age* NEXT/1 (2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16) NEXT/1 year*) OR (child* NEAR/6 (24 OR 36 OR 48 OR 60 OR 72 OR 84 OR 96 OR 108 OR 120 OR 132 OR 144 OR 156 OR 168 OR 180 OR 192) NEXT/1 month*)):Ab,ti)) OR ((((late* OR delay* OR missed OR Enhanc* OR error*) NEAR/3 (presentation* OR diagnos* OR detect* OR onset*)) OR (false NEAR/3 negative*)):Ab,ti))

APPENDIX B – NEWCASTLE-OTTAWA SCALE FOR QUALITY ASSESSMENT OF CROSS SECTIONAL STUDIES

Selection (maximum 3)

1. Representativeness of the sample
   a. Truly representative of the average in the target population (random sample orwhole population) *
   b. Somewhat representative of the average in the target population *
   c. Selected group
   d. No description of the sampling strategy
2. Sample size
   e. Justified and satisfactory (>1000 total)*
   f. Not justified
3. Non-included subjects
   g. Comparability between subjects and non-subjects characteristics is established*
   h. The response rate is unsatisfactory
   i. No description of the response rate

Comparability (maximum 2)

1. The subjects in different outcome groups are comparable, based on the study design or analysis. Confounding factors are controlled
   a. Study controls for age and gender (or analysis separated by gender)*
   b. Study controls for any additional factor *

Outcome (maximum 3)

1. Assessment of the outcome (hip morphology measurement)
   a. Blind*
   b. Description measurement*
   c. No description
2. Statistical test
   d. The statistical test used to analyze the data was clearly described and appropriate, and the measurement of the association was presented, including confidence intervals and the probability level (p value)*
   e. The statistical test is not appropriate, not described or incomplete

APPENDIX C – DATA EXTRACTION FORM

	Population characteristics					Imaging		
	Study design	Number	Age group	Cohort/ sample	Ethnicity	Mo-dality	Measurement(s)	Observer
		Hips Patients						No. and type
Akel								
Shi								
Tugrul								
Chung								

	Prevalence											
	Reported						Calculated					
	Mild (n)	Mild (%)	Severe (n)	Severe (%)	Mean	SD	Cut-off Mild	Cut-off severe	Mild (n)	Mild (%)	Severe (n)	Severe (%)
Akel												
Shi												
Tugrul												
Chung												

APPENDIX D – SPECIFICATIONS FROM AKEL ET AL.

Tables copied from original article[138].

	Male						Female					
	Right			Left			Right			Left		
Age (yrs.)	N	Mild dys.	Severe dys.	N	Mild dys.	Severe dys.	N	Mild dys.	Severe dys.	N	Mild dys.	Severe dys.
6 mo.–1 yr	<24	24-28	>28	<26	26-30	>30	<28	28-32	>32	<29	29-34	>34
2	<23	23-27	>27	<23	23-27	>27	<25	25-29	>29	<26	26-31	>31
3	<21	21-24	>24	<21	21-25	>25	<23	23-27	>27	<23	23-26	>26
4	<20	20-24	>24	<20	20-24	>24	<21	21-25	>25	<21	21-25	>25
5	<19	19-22	>22	<19	19-23	>23	<20	20-23	>23	<20	20-24	>24
6	<18	18-21	>21	<18	18-22	>22	<20	20-24	>24	<20	20-23	>23
7	<18	18-21	>21	<18	18-21	>21	<18	18-22	>22	<18	18-22	>22
8	<17	17-21	>21	<17	17-21	>21	<18	18-23	>23	<19	19-23	>23

N: normal, dys.: dysplasia

Cut-off vales from own calculation.

	Study values				Tönnis				p	p (Mild)	p (Severe)
	Mild dysplasia rate		Severe dysplasia rate		Mild dysplasia rate		Severe dysplasia rate				
Age (yrs.)	N	Rate (%)	N	Rate (%)	N	Rate (%)	N	Rate (%)			
1	89	19.6	18	3.9	125	27.4	43	9.4	<0.001	<0.001	<0.001
2	63	15.3	20	4.9	107	26.1	45	10.9	<0.001	<0.001	<0.001
3	65	17	18	4.7	56	14.7	20	5.2	0.336	0.108	0.687
4	46	14.9	16	5.2	92	29.9	21	6.8	<0.001	<0.001	0.125
5	60	19.1	13	4.2	77	24.6	8	2.6	0.250	0.035	0.125
6	66	21.6	12	3.9	81	26.4	14	4.6	<0.001	<0.001	0.754
7	52	16.8	13	4.2	66	21.3	8	2.6	0.414	0.007	0.180
8	52	18.5	8	2.8							
Total	493	17.8	118	4.3	604	24.3	159	6.4	<0.001	<0.001	<0.001

N: number

Acetabular dysplasia rates.

APPENDIX E – SPECIFICATIONS TUGRUL ET AL.

Table copied from original article[139].

	Male						Female					
	Right			Left			Right			Left		
Age	N	Mild	Severe	N	Mild	Severe	N	Mild	Severe	N	Mild	Severe
5	>20	15-20	<15	>20	15-20	<15	>19	14-19	<14	>19	14-19	<14
6	>20	15-20	<15	>20	15-20	<15	>19	14-19	<14	>20	15-20	<15
7	>20	15-20	<15	>21	16-21	<16	>20	15-20	<15	>21	16-21	<16
8	>20	15-20	<15	>21	16-21	<16	>21	16-21	<16	>21	16-21	<16
9	>21	16-21	<16	>22	17-22	<17	>21	16-21	<16	>22	17-22	<17
10	>21	16-21	<16	>22	17-22	<17	>21	16-21	<16	>22	17-22	<17
11	>22	17-22	<17	>22	17-22	<17	>21	16-21	<16	>23	18-23	<18
12	>23	18-23	<18	>23	18-23	<18	>22	17-22	<17	>23	18-23	<18
13	>23	18-23	<18	>23	18-23	<18	>22	17-22	<17	>23	18-23	<18
14	>23	18-23	<18	>23	18-23	<18	>22	17-22	<17	>23	18-23	<18

N: normal

Cut-off values for acetabular dysplasia from Tugrul et al.[139]

Chapter 6

Prevalence of acetabular dysplasia in early adolescents in a general population: a Generation R study

S. de Vos-Jakobs, F. Boel, D. Chen, J.C.M. van Haasteren, Jaap J. Tolk, S.M.A. Bierma-Zeinstra, R. Agricola.

Submitted.

ABSTRACT

Background: Acetabular dysplasia is an important risk factor for hip complaints and hip deterioration in adults. As early treatment of this condition, before onset of degeneration, is the key to hip preservation, knowledge of the natural course of acetabular dysplasia is essential.

Purpose: The purpose of this study was to estimate prevalence numbers of acetabular dysplasia in early adolescents of the general population and in subgroups (sex assigned at birth, ethnicity, and skeletal maturity).

Methods: This study was part of the Generation R project, a population based cohort study in Rotterdam (NL). In early adolescence high-resolution dual-energy x-ray absorptiometry (DXA) of the right hip was done, providing images highly comparable to radiographs. All participants with sufficient DXA images were included for measurement of the lateral centre-edge angle (LCEA) as an indicator for acetabular roof coverage. Acetabular dysplasia was defined as LCEA <20°. Multivariate analysis was done for the subgroups.

Results: A total of 3896 participants were included with a mean age of 13.6 years (SD 0.3 years) and 46.8% males. Prevalence of acetabular dysplasia was 6.4%. This was higher in skeletally immature compared to skeletally mature participants (OR = 2.94). Sex assigned at birth and ethnicity had no statistical significant association with acetabular dysplasia.

Conclusion: As prevalence is higher in early adolescents compared to infants, we conclude that acetabular dysplasia may develop during childhood growth, rather than solely manifesting in infancy. Therefore, a low threshold for radiographic pelvic examination in adolescents should be adhered to, as treatment for hip preservation is still an option at this age.

INTRODUCTION

Background

Acetabular dysplasia is a disorder in the spectrum of developmental dysplasia of the hip (DDH). The spectrum ranges from mild acetabular dysplasia to complete dislocation of the hip joint. Acetabular dysplasia is defined as insufficient coverage of the acetabular roof over the femoral head[92]. Peak loading of this insufficient acetabular roof may lead to premature failure of the articular cartilage and osteoarthritis, resulting in pain, limping, limited activity, and eventually hip replacement surgery at a young age[10,12]. Acetabular dysplasia is one of the most common causes of disability due to osteoarthritis at young age (<50 years)[10,12,13].

Identification of acetabular dysplasia before onset of symptoms is essential during childhood and adolescence, because then treatment options are available for hip preservation and prevention of osteoarthritis, before degeneration begins. However, early identification of acetabular dysplasia is challenging, because at that stage there often are no clinical complaints or abnormal findings[146].

Rationale

For many years, DDH was considered a condition developing in infancy and naturally improving throughout growth. Therefore, screening programs, for early detection and treatment of DDH, are widely implemented in infants[18,19]. However, some studies suggest that acetabular dysplasia not only commences in infancy, but can also develop later during childhood[127,146]. If this concept holds true, these late cases of acetabular dysplasia are missed in the current screening programs, but do potentially also lead to pain, limping, and a higher risk for early-onset osteoarthritis. The extent of this problem of late-developed acetabular dysplasia is yet unknown.

Currently, prevalence of acetabular dysplasia has been reported in infants up to 2 years of age and in adults[12,13,23,126], but prevalence numbers for different ages during childhood remain unclear[147]. To oversee the complete extent of the problem of acetabular dysplasia, prevalence numbers from the general population at multiple ages are needed.

The primary aim of this study is to present prevalence of acetabular dysplasia amongst a large sample of Dutch early adolescents from the general, multi-ethnic population of Rotterdam, the Netherlands. The secondary aims are (1) to calculate prevalence of acetabular dysplasia in subgroups of sex assigned at birth, ethnicity, and skeletal maturity and (2) to report and critically appraise potential differences within each subgroup.

PATIENTS AND METHODS

Study population, inclusion, and exclusion

For this research the population-based cohort "Generation R" was used. This is a unique prospective birth cohort of 9,778 children from Rotterdam, The Netherlands. Foetuses of pregnant women were included before birth between April 2002 and August 2006. The children were prospectively followed after birth during childhood at predetermined ages (Focus points). More detailed information on the Generation R cohort is described in the design papers[47,148]. The study protocol was approved by the Medical Ethical Committee from the Erasmus Medical Centre, Rotterdam, The Netherlands (registered by number MEC 2015-749 NL55105.078.15). Written informed consent was obtained from all parents and participants.

This current study was done in the Focus-13 cohort consisting of early adolescents (approximately 13 years old). All 4929 Focus-13 participants who visited the research centre were considered for this study. Demographic data was obtained at the first visits (sex assigned at birth, based on birth record; and ethnicity, based on a questionnaire filled out by the parents) and at the Focus-13 visit (age). Ethnicity was registered and grouped, based on demographic criteria of the Dutch Central Bureau for Statistics[149]. Also, at the Focus-13 visit, all participants were invited for a high-resolution dual-energy x-ray absorptiometry (DXA) of the total body and the right hip with use of iDXA scanner (GE Healthcare, Madison, WI, USA). All participants with DXA images were eligible for inclusion. Exclusion criteria were: (1) absence of one or both of the DXA images, (2) incomplete DXA image of the acetabulum, and (3) artefact in DXA image interfering with measurements.

Radiological measurements

Measurements for acetabular dysplasia were done on the high resolution DXA images of the right hip. These images have proved to be highly comparable to radiographs of the hip when assessing hip morphology[150]. The total body DXA was only used for determination of a horizontal pelvic reference line.

Acetabular coverage was measured by the lateral centre-edge angle (LCEA), indicating the coverage of the femoral head by the bony acetabular roof. Correction for pelvic obliquity was done based on the horizontal reference line of the pelvis. By definition, LCEA is the angle measured between (1) the line through the centre of the femoral head, perpendicular to the horizontal reference line of the pelvis and (2) the line from the centre of the femoral head to the most lateral point of the bony acetabular roof (Fig 1)[53]. As LCEA <20.0° is associated with increased risk for osteoarthritis in skeletally mature

individuals, acetabular dysplasia was defined accordingly with a dichotomous outcome (dysplasia or no dysplasia)[53,151].

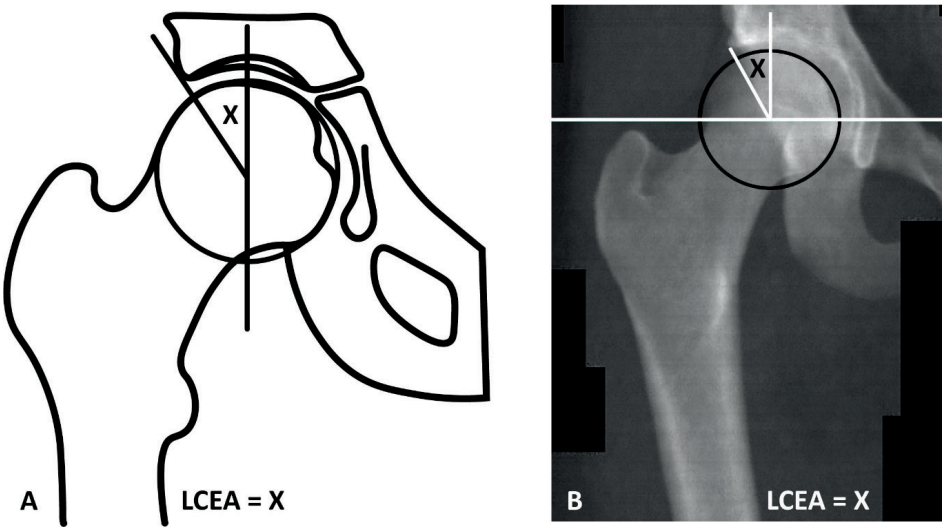


Figure 1. A. Illustration of the lateral centre-edge angle (LCEA). **B.** LCEA as performed on a dual-energy x-ray absorptiometry (DXA) image of the right hip with horizontal reference line based on total body scan.

Skeletal maturity of the hip was defined by the status of the triradiate cartilage. Once the triradiate cartilage was fused, the participant's hip was considered skeletally mature. The triradiate cartilage status was scored by two paediatric orthopaedic surgeons at two different time points, six months apart, as a dichotomous outcome (open or fused). When the status of the triradiate cartilage could not be identified, e.g. due to technical reasons, it was registered as unclear. These cases were excluded from the subgroup analysis.

Automated measurements

The measurements of LCEA were performed automatically using the method developed by Boel et al.[91]. This open-access method allows for fast and reproducible calculation of the LCEA with an ICC of 0.95 (95% CI 0.87 – 0.98)[91]. In short, the LCEA was calculated, based on a set of radiographic landmarks outlining the contour of the hip, which were automatically placed using the BoneFinder® software (www.bone-finder.com; The University of Manchester, UK)[64]. The landmark point placement was manually checked by the researchers for all hips and corrections were made if necessary.

Statistical Analysis

After LCEA was calculated for all participants, the prevalence of acetabular dysplasia was calculated for the whole group as a percentage with a 95% confidence interval. The reliability of both interobserver and intraobserver assessments for the triradiate cartilage status was determined by calculating intraclass correlation coefficients (ICC). Intraobserver reliability was assessed for a single rater, while interobserver reliability involved multiple raters. Our analysis focused on absolute agreement and utilized a mixed model with average measures. Subgroup analyses were done for sex, grouped ethnicity, and skeletal maturity of the hip. To test potential differences in prevalence of acetabular dysplasia within the subgroups, univariate binary logistic regression was done. Suspecting potential confounding associations between sex, ethnicity, and skeletal maturity of the hip, we incorporated these variables into a multivariable model through binary logistic regression analysis. The dependent variable was the presence of dysplasia, represented as a dichotomous outcome (yes/no). The covariates included sex, grouped ethnicity, and skeletal maturity of the hip. Male sex and skeletally mature hips were set as the reference groups. Ethnicity was marked as polynomial covariate. Results are presented as odds ratios (ORs) with 95% confidence intervals (CIs). All analyses were conducted in SPSS (IBM SPSS Statistics, v28.0.1.0). The significance level was set at $p < 0.05$.

RESULTS

Based on the in- and exclusion criteria, 3986 hips were included. Figure 2 shows the flowchart for exclusion.

In Table 1, the demographic data of the study group is presented together with data on skeletal maturity and LCEA.

Table 1. Participants' characteristics

Characteristic		**Mean (SD) or n (%)**
Age (years)	Mean (SD)	13.6 (0.3)
	Range	12.6 – 17.0
Gender (n, %)	Male	1864 (46.8)
	Female	2122 (53.2)
Ethnicity (n, %)	Western	2872 (72.1)
	African	644 (16.2)
	Asian	237 (5.9)
	Other	233 (5.8)

Table 1. Participants' characteristics (*continued*)

Characteristic		Mean (SD) or n (%)
Triradiate cartilage status (n, %)	Open	995 (25.0)
	Fused	2977 (74.7)
	Unclear	14 (0.4)
LCEA (degrees)	Mean (SD)	28.7 (5.9)
	Range	4.4 – 57.1

SD: standard deviation, LCEA: lateral centre-edge angle. Note that the sum of percentages may not add up to 100% due to rounding of the numbers.

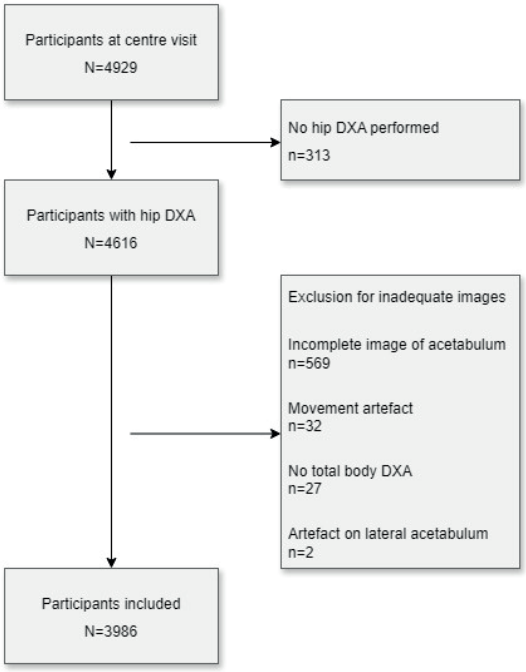


Figure 2. Flowchart of inclusion. DXA: dual-energy x-ray absorptiometry.

There was excellent intra- and interobserver reliability for scoring of triradiate cartilage (open or fused) with an interobserver ICC of 0.95 (95% CI 0.91-0.97) and intraobserver ICC of 0.97 (95% CI 0.95-0.98).

Based on LCEA <20.0°, the prevalence of acetabular dysplasia in this complete cohort is 6.4% (95% CI 5.6-7.1%). Distribution of LCEA is presented in Figure 3 for the total group stratified for skeletally mature and immature hips. This figure shows that the distribution shifts to the left in skeletally immature participants, indicating lower LCEA and more acetabular dysplasia.

Univariate subgroup analysis showed that the prevalence of acetabular dysplasia is twice as high in males compared to females. There were no differences observed in prevalence of acetabular dysplasia between ethnicity groups. Skeletally immature hips showed a threefold higher prevalence of acetabular dysplasia than mature hips.

In multivariate analysis (Table 2), skeletally immature hips remained significantly associated with acetabular dysplasia (aOR=2.94, 95% CI 2.14-4.03), but sex no longer showed statistical significant association (aOR=0.74, 95% CI 0.53-1.02). Ethnicity remained an insignificant factor for prevalence of acetabular dysplasia.

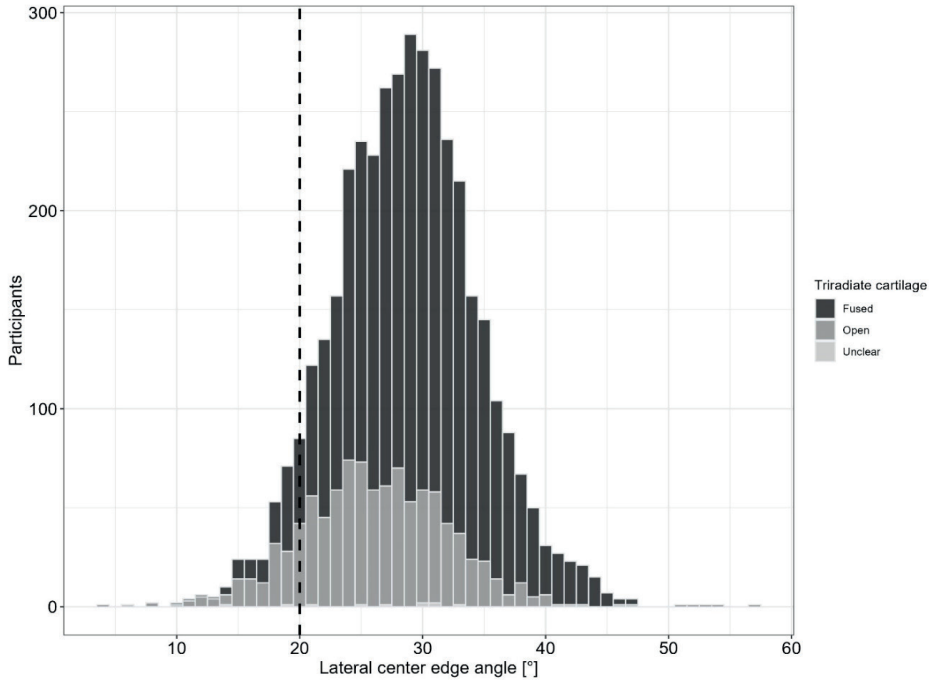


Figure 3. Distribution of lateral centre-edge angle within the study population stratified for open and fused triradiate cartilage. The dashed line indicates the cut-off value of <20.0°.

Table 2. Subgroup analyses

	n	Prevalence AD (%)	OR (95% CI)	p-value
Male	1864	9.1		
Female	2122	4.0	0.74 (0.53 – 1.02)	p = 0.07
European	2863	6.7		p = 0.56
African	642	6.2	1.54 (0.80 – 2.96)	p = 0.20
Asian	236	4.6	1.55 (0.76 – 3.18)	p = 0.23
Other	231	4.3	1.24 (0.5 – 3.00)	p = 0.63
Skeletally immature	995	13.1		
Skeletally mature	2977	4.1	2.94 (2.14 – 4.03)	p < 0.001

Multivariate analysis for prevalence of acetabular dysplasia in various subgroups. AD: acetabular dysplasia, OR: odds ratio, CI: confidence interval.

DISCUSSION

This study is the first to assess the prevalence of acetabular dysplasia in a large cross-sectional cohort of early adolescents from the general population. The results of this unique and highly diverse cohort offer unique insights, revealing that acetabular dysplasia is a common condition in the general early adolescent population with a prevalence of 6.4%. Analyses in subgroups show significantly higher prevalence in skeletally immature hips and hips belonging to male participants. Ethnicity appears of no influence on the results. The multivariate analysis reveals that the high numbers in males can be attributed to a higher proportion of skeletally immature participants within the male subgroup, as males tend to reach skeletal maturity at a higher age than females[152].

This high prevalence in early adolescence is an important finding as acetabular dysplasia is a known major contributor to later hip deterioration[10,12]. Hip joint deterioration has substantial societal costs, increases the risk of social isolation, and profoundly impacts daily life through factors such as pain, limping, and restricted activity[13,153]. When discovered early, before onset of radiographic degeneration, treatment of acetabular dysplasia might prevent these long term effects. Until now, the prevalence of acetabular dysplasia and its extent during childhood have remained unclear[147].

Literature on acetabular dysplasia in children and adolescents is scarce, so the exact extent and impact of acetabular dysplasia in early adolescence remained unclear. A recent systematic review on prevalence of acetabular dysplasia in children aged 2 years and above reported large variance in inclusion criteria and radiographic measurements and high risk of bias in the studies previously available[147]. Selection bias was a high risk, as study populations were mainly derived from hospital databases and not from the general population. Either way, the prevalence numbers vary strongly (2.2% - 25.6%)

within the age range of 2-18 years. Merely one of the included studies reported on the general population, but participants were reasonably younger (mean age of 9.86 years) in that population[106]. Also, the aim of that study was not to evaluate prevalence numbers, but to study associations[106].

In comparison to this recent literature, the current study reveals that prevalence numbers in early adolescents are substantially higher than in infants (1.45% for acetabular dysplasia and 3.90% for the complete spectrum of DDH)[16] and comparable to those in adults (6.7%)[154]. Treatment of infants with DDH would be expected to result in decreasing prevalence of acetabular dysplasia with age. However, this study in early adolescents and other studies in adults show increased prevalence of acetabular dysplasia compared to infants. This raises the subsequential hypothesis that late cases of acetabular dysplasia develop throughout childhood growth, implying that adolescent and adult dysplasia might not fully be considered the same disease as infant dysplasia.

In the current study, we find a strong association between skeletal immaturity of the hip and the prevalence of acetabular dysplasia (aOR 2.94). As the LCEA is generally higher in skeletally mature hips compared to skeletally immature hips, this suggests an increase of LCEA (and thus decrease in acetabular dysplasia) in the final phase of hip development. This final phase involves the ossification of the acetabular lip, situated at the most lateral part of the acetabulum[152]. The acetabular lip consistently leads to improvement in LCEA as LCEA is measured using the most lateral point of the lip (Fig. 1). Although the extent of improvement in LCEA by the acetabular lip varies, this can be interpreted as acetabular dysplasia gradually resolving with growth; a dysplastic immature hip (LCEA < 20.0°) might have the potential to develop into a normal mature hip (LCEA ≥ 20.0°). In our opinion, however, the labelling of hips as dysplastic should be reserved for those hips that are at risk for future deterioration, and would ideally not include hips that normalize spontaneously towards skeletal maturity. We emphasize that the definition of acetabular dysplasia should be adjusted to age or skeletal maturation stage instead of a single definition which is applied to all hips. Age-specific cut-off values for LCEA, ideally combined with identification of other risk factors, should aid in identifying hips at risk for future deterioration (alike age-specific cut-off values for acetabular index in early hip development[128]).

Future research should concentrate on more detailed information about development throughout growth. This is essential to improve our knowledge of late-developed acetabular dysplasia. Longitudinal studies where the general population is followed over time, like the Generation R cohort, are essential to unravel the full development of the hip joint from infancy to adulthood. Also, with this comprehensive view of the child's hip

development, one can initiate or optimize screening programs to diagnose acetabular dysplasia at an early stage. At that point treatment for prevention of later problems (e.g. pain, limping and osteoarthritis) is still available, as failure of articular cartilage is not yet present. Also, early recognition might lead to less invasive methods for treatment in the future.

The major strength of our study is the unique design and sample size enabling accurate prevalence estimation in the general population. Secondly, the automated, open access method for measurement of LCEA enhances reproducibility and efficiency in large cohorts.

A limitation of the current study is that pelvic radiographs were not available, but instead DXA images were used. Previous research from our study group indicates that DXA images are a reliable alternative to radiographs for performing measurements of hip morphology[150]. The combination of low radiation dose and reliable measurements, makes DXA a very good imaging modality for population studies. A second limitation is that only images of the right hip were obtained. Analysing the prevalence in solely right hips underestimates the overall prevalence, as acetabular dysplasia of the left hip might still be present in the images of non-dysplastic right hips, especially given the fact that in infants, DDH is more commonly observed in the left hip. This might explain the discrepancy between acetabular dysplasia in skeletally mature early adolescents in this cohort with only right hip images (4.1%) and adults with pelvis images (6.7%). Also, only an anteroposterior view of the hip was available. As stated by Herfkens et al. this might lead to an underestimation of prevalence of acetabular dysplasia, because only dysplasia of the lateral acetabular rim is visible and not dysplasia of the anterior acetabular rim[126]. As both these limitations potentially lead to underestimation, our hypothesis of late-developing acetabular dysplasia remains credible.

Based on the high prevalence numbers in this study, we suggest that all adolescents presenting with pain, limping, or limitations of activity whatsoever, should have a radiograph taken for evaluation of their hips. At an early stage, before deterioration and radiographic degeneration commences, treatment for hip preservation is available and can strongly decrease individual and societal consequences of hip degeneration[155,156]. Acetabular dysplasia should be considered a common condition in early adolescents, even when screening programs earlier in life were successful.

Conclusions

Prevalence of acetabular dysplasia in a cross-sectional cohort of 3986 early adolescents with a mean age of 13.6 years is 6.4%. Skeletal maturity was strongly associated with acetabular dysplasia (aOR = 2.94), where skeletally immature participants have threefold higher prevalence of acetabular dysplasia (13.1%) than skeletally mature participants (4.1%).

As this prevalence is higher than in infants and comparable to adults, it suggests that acetabular dysplasia develops later during childhood growth. Therewith, acetabular dysplasia should always be considered in children with hip complaints, limping, or limited activities, even after sufficient screening in infancy.

Conflict of Interest Statement

Two of the authors (FB & SBZ) have received funding from independent research grants paid to the Institution from Dutch Arthritis Association, ZonMw and EU. Also, personal fees were received by SBZ as consultant for Pfizer Infirst Healthcare and as deputy editor for Osteoarthritis and Cartilage. All other authors certify that there are no funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article related to the author or any immediate family members.

Ethical Review Committee Statement

The research protocol was approved by the medical ethical committee and registered by number MEC 2015-749 NL55105.078.15.

Acknowledgements

The Generation R Study is conducted by the Erasmus MC, University Medical Center Rotterdam in close collaboration with the Erasmus University Rotterdam and the city of Rotterdam. We gratefully acknowledge the contribution of children and parents. The general design of Generation R Study is made possible by long-term financial support from Erasmus MC, University Medical Center Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw) and the Ministry of Health, Welfare and Sport.

Chapter 7

The prevalence of pincer morphology in early adolescents from the general population: a population-based study (Generation R)

D. Chen, F. Boel, S. de Vos-Jakobs, P. van Klij, M.M.A. van Buuren, S.M.A. Bierma-Zeinstra, R. Agricola.

Arthritis care & research. 2025 Jul.

ABSTRACT

Objective: Pincer morphology can lead to femoroacetabular impingement syndrome (FAIS) and may be a modifiable risk factor for hip osteoarthritis (OA). Currently, no studies investigate the prevalence of pincer morphology in early adolescence – the period when this bony shape likely develops. The purpose of this study was to estimate the prevalence and birth-assigned sex distribution of pincer morphology in early adolescents from the general population in the Netherlands.

Methods: This study was embedded in the Generation R study, a population-based prospective cohort in Rotterdam, the Netherlands. Around the age of 13 years, participants underwent high-resolution dual-energy x-ray absorptiometry (DXA) of their full-body and right hip. The lateral center edge angle (LCEA) was automatically determined based on landmarks outlining the hip contour, and pincer morphology was defined as a LCEA ≥ 40°. The overall and birth-assigned sex-specific prevalence was presented as a percentage with 95% confidence interval (CI).

Results: A total of 3,986 adolescents (median age 13.5 years [2.5 - 97.5% percentiles, 13.2 - 14.6]; 46.8% males) were included. The overall prevalence of pincer morphology was 3.1% (95% CI 2.6% - 3.6%). The prevalence in male and female adolescents was 3.0% (95% CI 2.2% - 3.7%) and 3.3% (95% CI 2.5% - 4.0%), respectively.

Conclusion: Among early adolescents from the general population in the Netherlands, the estimated prevalence of pincer morphology was 3.1%. Male and female adolescents had a similar prevalence of pincer morphology. These findings could inform the timing of prevention strategies for pincer morphology, potentially reducing the risk of FAIS and hip OA.

SIGNIFICANCE AND INNOVATIONS

- Based on a cross-sectional analysis of 3,986 adolescents aged 13 years in a population-based cohort (Generation R), we found that the overall prevalence of pincer morphology was 3.1% (95% CI 2.6% - 3.6%), with similar prevalence in male (3.0% [95% CI 2.2% - 3.7%]) and female adolescents (3.3% [95% CI 2.5% - 4.0%]).
- Given the relatively low prevalence in this population of early adolescents, pincer morphology might also develop further during skeletal maturation of the hip like cam morphology.
- Our findings could inform the timing of prevention strategies for pincer morphology, potentially reducing the risk of femoroacetabular impingement syndrome and hip osteoarthritis.

INTRODUCTION

Osteoarthritis (OA) is a major cause of disability worldwide[157,158] and femoroacetabular impingement syndrome (FAIS) is an important risk factor of hip OA[159]. FAIS is a motion-related disorder that affects the hip joint and results from abnormal contact between the acetabulum and the femoral head[84]. Depending on the anatomical morphology, FAIS can be distinguished based on the presence of cam and/or pincer morphology. Cam morphology involves an extra bone formation at the anterolateral femoral head-neck junction, while pincer morphology exhibits either focal or global overcoverage of the femoral head by the acetabulum. Pincer morphology is not only associated with FAIS but also may lead to labral tears and circumferential acetabular cartilage damage[44,160]. Although the relationship between pincer morphology and hip OA is controversial[8,10,75,161], a recent study from the Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) consortium reported that hips with pincer morphology have 1.59 (95% confidence interval (CI) 1.16 - 2.20) times higher odds of developing radiographic hip OA within eight years[162].

Until now, the exact age at which pincer morphology starts to develop has not yet been determined. A retrospective study based on abdominal computed tomography scans from 225 patients aged 2 to 19 without hip complaints reported that pincer morphology first developed at the age of 12 years[33]. To the best of our knowledge, no large population-based epidemiological studies have examined the prevalence of pincer morphology in the early adolescent population. Interestingly, there are more studies available on the development of cam morphology, which show that cam morphology is an adaptive response to vigorous hip loading and develops gradually over time during adolescence[163,164]. However, it is unknown whether this concept holds for pincer morphology as well. Understanding the prevalence of pincer morphology in early adolescence could be critical for informing the timing of prevention strategies that may help slow its progression to hip OA.

We therefore aimed to estimate the overall and birth-assigned sex-specific prevalence of pincer morphology in a general population of 3,986 early adolescents.

METHODS

Study design and participants

The data used in this study were derived from the Generation R study, a population-based prospective cohort aiming to investigate the growth, development and health of children from fetal life onwards in Rotterdam, the Netherlands. A total of 9,778 pregnant women whose delivery was expected between April 2002 and January 2006 were enrolled in the study, and their children became part of the Generation R cohort. Detailed designs and methods on the Generation R cohort were previously described elsewhere[47]. Around the age of 5 (Focus 5), 9 (Focus 9), and 13 years (Focus 13), all participating children and their parents were invited to visit the Generation R research center in the Erasmus MC – Sophia Children's Hospital. For the current study, we included all participants who had both full-body and right hip high-resolution dual-energy x-ray absorptiometry (DXA) scans available at the Focus 13 visit. Exclusion criteria were: 1) participants with an incomplete acetabulum depicted on the DXA image; 2) the presence of movement artifacts; 3) the presence of an artifact in the region of interest. This study was approved by the Medical Ethical Committee of Erasmus Medical Center (MEC-2015-749) and written informed consents were obtained from the participants and their parents. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines[165].

DXA scan

All participants underwent DXA scans of their full-body and right hip by well-trained investigators using the General Electric (GE)-Lunar iDXA densitometer (GE Healthcare, Madison, WI, USA). The full-body and right hip DXA scans were done sequentially with the participants in the same position. Before scanning, all participants were required to remove heavy clothing, shoes and any metal accessories. Afterwards, they were placed in a supine position with the hands flat at their side. The legs were slightly separated and rotated internally, with the big toes touching. The feet were fastened in this position with a Velcro strip to avoid movement.

Pincer morphology

Pincer morphology was determined by the lateral center edge angle (LCEA) on the DXA images. The presence of pincer morphology was defined as a LCEA ≥ 40° [10,33,46]. The LCEA was automatically calculated based on landmarks outlining the hip contour with in-house developed software (Fig. 1)[91]. All landmarks were placed automatically using the BoneFinder® software (www.bone-finder.com; The University of Manchester, UK)[64]. A visual inspection was conducted to ensure correct placement of landmarks and manual adjustments (DC and FB) were made when necessary. The LCEA was determined using

the following steps: first, the center of the femoral head was automatically computed based on the best-fitting circle around the femoral head. Next, the LCEA was formed by the line from the center of the femoral head to the most lateral bony edge of the acetabulum, and the line from the center of the femoral head perpendicular to the horizontal reference line of the pelvis as determined on the full-body DXA image (Fig. 1)[91]. Details of this method have been published previously and showed an intermethod reliability between automatic measurement and manual measurement of 0.95 (95% CI 0.87 - 0.98)[91].

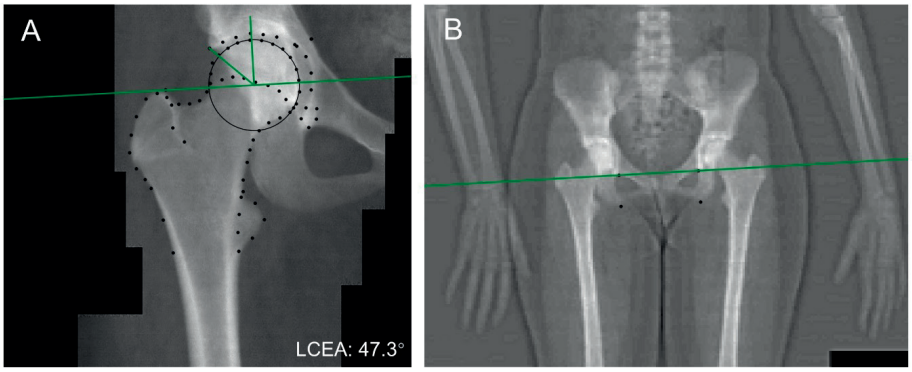


Figure 1. The measurement of lateral center edge angle (LCEA). On the hip dual-energy x-ray absorptiometry (DXA) image, we outlined the shape of the proximal femur and acetabulum with 80 landmarks. The center of the femoral head was automatically determined by the best-fitting circle, which was based on the landmarks outlining the femoral head. The LCEA of 47.3 degrees in this sample (A) is formed by the intersection of two lines: one from the center of the femoral head to the most lateral bony edge of the acetabulum, and the other from the center of the femoral head perpendicular to the horizontal reference line of the pelvis as determined on the full-body DXA image (B).

Data availability statement

The datasets used in the current study are available to researchers upon reasonable request to the management team of the Generation R Study. More detailed information is available on the following website (https://generationr.nl/researchers/collaboration/).

Statistical analyses

We compared the characteristics of included and excluded adolescents using Mann-Whitney U tests for non-normally distributed continuous data, Student's t-tests for normally distributed continuous data, and chi-square tests for categorical data. We estimated the prevalence of pincer morphology and presented it as a percentage with 95% CI stratified by sex. The 95% CI was computed assuming a binomial distribution. The chi-square test was used to examine the difference in birth-assigned sex-specific prevalence. The overall and birth-assigned sex-specific distribution of LCEA was assessed using histograms. To address potential selection bias on our findings, we calculated the weighted prevalence

using inverse probability weighting (IPW) adjusting for birth-assigned sex, body mass index (BMI) and height[263]. We performed statistical analyses using R Statistical software (v4.2.1; R Core Team2022). The 95% CIs were computed using the binom-package[166], histograms were generated using the ggplot2-package[74], and inverse probability of weights was computed using the ipw-package[264]. A p-value < 0.05 was considered as statistically significant.

RESULTS

Characteristics of study participants

A total of 3,986 participants had sufficient quality right hip and full-body DXA images available and were analyzed in this study (Fig. 2). Table 1 shows the characteristics of the included early adolescents.

Table 1. Demographic characteristics of included adolescents.

Characteristic	Included adolescents (n=3,986)	Excluded adolescents (n=943)	P value
Age, years	13.5 (13.2 – 14.6)	13.6 (13.2-14.7)	0.567
Male, %	1,864 (46.8%)	556 (59.0%)	<0.001
Female, %	2,122 (53.2%)	387 (41.0%)	
BMI, kg/m^2	19.2 (15.2 – 29.3)	18.9 (15.1-28.3)	0.003
Height, cm	164.0 (7.9)	165.9 (8.2)	<0.001

Data are presented as frequencies (percentages) for categorical variables, means (SD) for continuous normally distributed variables and medians (2.5 - 97.5 percentiles) for continuous, non-normally distributed variables. BMI: body mass index.

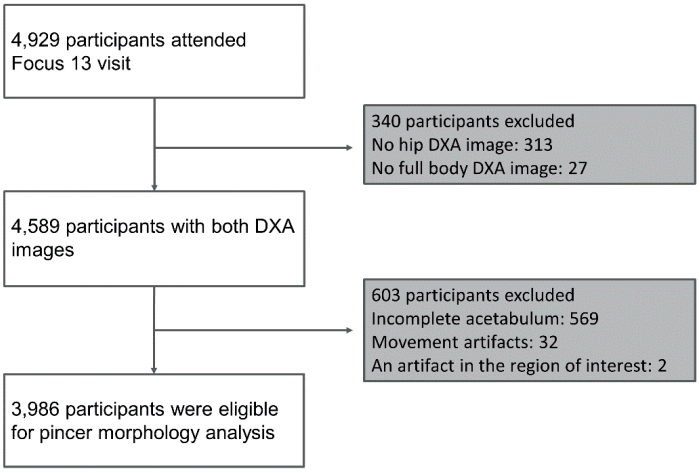


Figure 2. Flowchart of participants. DXA: dual-energy x-ray absorptiometry.

A total of 4,929 participants attended the Focus 13 visit, of whom 3,986 participants (81%) had sufficient quality right hip and full-body DXA images available and were analyzed in this study (Fig. 2). Table 1 shows the characteristics of the included and excluded early adolescents. The median age of included adolescents was 13.5 years (2.5th - 97.5th percentile, 13.2 - 14.6), and 1,864 (46.8 %) were male adolescents. The median BMI of included adolescents was 19.2 kg/m^2 (2.5th - 97.5th percentile, 15.2 - 29.3), and the mean height was 164.0 cm (SD, 7.9). The excluded adolescents were more frequently males, slightly taller and had slightly lower BMI than the included adolescents (Table 1).

Prevalence of pincer morphology

The overall and birth-assigned sex-specific LCEA distributions are presented in Figure 3. The overall prevalence of pincer morphology was 3.1% (95% CI, 2.6% - 3.6%) in early adolescents. In the sex subgroup analysis, the prevalence of pincer morphology was similar in male (3.0% [95%CI 2.2% - 3.7%]) and female participants (3.3% [95%CI 2.5% - 4.0%]) (Table 2). The weighted prevalence was similar to the unweighted prevalence in the total group and birth-assigned sex subgroup, as shown in Table 2.

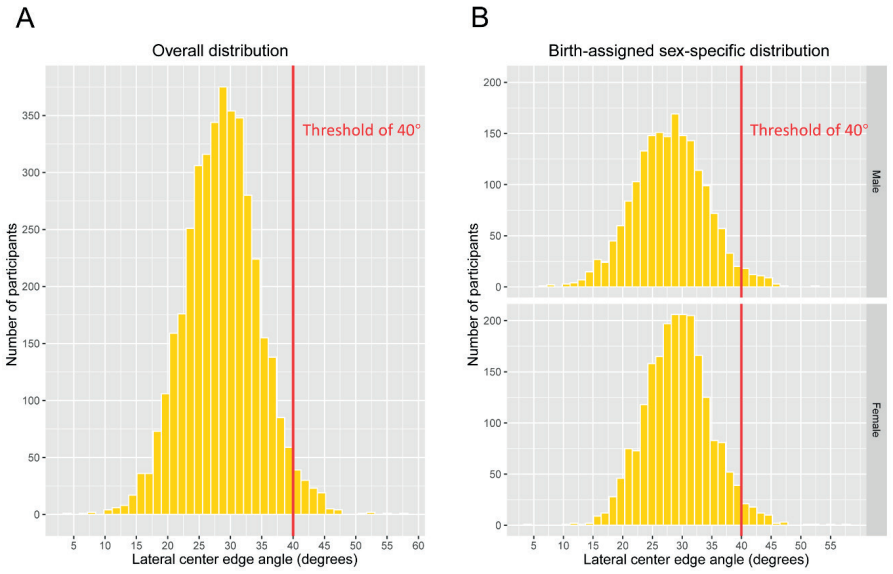


Figure 3. The overall (A) and birth-assigned sex-specific distribution (B) of the lateral center edge angle among the included early adolescents. The red solid line indicates the 40° threshold of the lateral center edge angle.

Table 2. The prevalence of pincer morphology among early adolescents stratified by birth-assigned sex.

Characteristic		LCEA ≥ 40°		
	No. of cases / total no.	Prevalence, % (95% CI)	X^2	P value
Unweighted prevalence				
Overall	124/3,986	3.1% (2.6% – 3.6%)		
Sex				
Male	55/1,864	3.0% (2.2% – 3.7%)	0.298	0.585
Female	69/2,122	3.3% (2.5% – 4.0%)		
Weighted prevalence				
Overall	-	3.1% (2.6% – 3.7%)		
Sex				
Male	-	3.0% (2.2% – 3.8%)	0.203	0.653
Female	-	3.3% (2.5% – 4.0%)		

LCEA: lateral center edge angle; CI: confidence interval.

DISCUSSION

This study included 3,986 early adolescents from the general population of Rotterdam, the Netherlands, and used high-resolution DXA scans to determine the prevalence of pincer morphology. We found that the overall prevalence of pincer morphology was 3.1% (95% CI 2.6% - 3.6%) and that it was similar in male (3.0% [95% CI 2.2 - 3.7%]) and female participants (3.3% [95% CI 2.5 - 4.0%]).

The previously reported prevalence of pincer morphology varies widely in literature due to different study populations and inconsistent definitions. To date, there is a paucity of literature on the prevalence of pincer morphology in early adolescents. A population-based study of 2,081 participants (mean age, 18.6 years) in Norway found that 24% of them had a pincer morphology as defined by the presence of a posterior wall sign, crossover sign or excessive acetabular coverage on the anteroposterior (AP) pelvic radiograph[167]. In a cross-sectional study of 6,807 individuals with a mean age of 62.7 years from the UK Biobank, pincer morphology was defined as a LCEA ≥ 45° and found in 8.5% of participants using DXA scans of the left hip[58]. Our results showed the prevalence of pincer morphology was 3.1% in early adolescents, lower than most of the previously reported prevalence in adults. Several factors could explain this result. First, the coverage of the acetabulum may still increase during skeletal maturation, which could lead to the development of pincer morphology. Some prospective studies have investigated the development of cam morphology, and revealed it gradually increases in size during adolescence[163,164]. Interestingly, the low prevalence of pincer morphology in our study suggests that it might also gradually develop during skeletal growth similar to cam

morphology. Prospective studies are needed to confirm this observation. Secondly, we only utilized the LCEA, one of the most commonly used objective measures, to quantify pincer morphology. Other definitions, such as the crossover sign, posterior wall sign and coxa profunda, have also been reported to determine pincer morphology. However, previous studies on these measures reported a poor reliability and specificity[31,168,169]. The definition of pincer morphology lacks a validated measure and accompanying thresholds, such as for the LCEA. In adults, the LCEA threshold to define pincer morphology varies across studies. Some large prospective studies used a threshold of 33.7° or 40° [10,52] to investigate the association of pincer morphology and hip OA, while other studies used a threshold of 45° [58,162]. The recent Lisbon Agreement provided guidance and criteria for defining pincer morphology, including global overcoverage identified by the presence of protrusio acetabuli or a Wiberg center edge angle (WCEA) ≥ 40° (or ≥ 35° with an acetabular index < 0°)[53]. We used the same threshold value but chose to use the LCEA instead of the WCEA in this study since pincer impingement usually occurs between the femoral head-neck junction and the most lateral bony edge of the acetabulum.

Research on sex differences in the prevalence of pincer morphology has yielded inconclusive results. A population-based study of 2,596 participants (mean age, 63 years) in the U.S. found that 7% of male participants and 10% of female participants had pincer morphology as defined by a LCEA ≥ 40° or acetabular protrusio in the AP pelvic radiograph[25]. Faber et al. included 6,807 participants (mean age, 62.7 years) from UK biobank in a cross-sectional study, where pincer morphology was defined as LCEA ≥ 45° and present in 8.9% of male participants and 8.1% of female participants on the DXA images of left hips[58]. Our results are in line with these findings and also indicate the similar prevalence of pincer morphology between males and females at a mean age of 13 years. However, results from other studies are contrary to our findings. In a Norwegian population-based cohort study, Laborie et al. studied the AP pelvic radiograph of 2,081 subjects aged 17.2 to 20 years. They found that 34.3% of male participants had at least one sign of pincer morphology (posterior wall sign, crossover sign or excessive acetabular coverage) compared to 16.6% of female participants[167]. The reason for these inconsistent results, which complicate direct comparison, are the heterogeneous populations, diverse imaging modalities (radiographs vs. DXA scans), and variety in definitions for pincer morphology (LCEA thresholds or specific radiographic signs). Although our study found the similar prevalence of pincer morphology between sexes in early adolescence, follow-up of the current cohort can reveal if these sex-specific differences become apparent during or after skeletal maturation.

Selection bias is a critical consideration in this study, as 19% (943/4,929) of adolescents who visited the Generation R research center were excluded due to missing DXA scans or LCEA measurements. To address this selection bias, we applied IPW using a logistic regression model adjusting for sex, BMI, and height—factors that showed statistically significant differences between included and excluded participants (Table 1). The similarity between weighted and unweighted prevalence estimates suggests that the selection bias had a limited impact on our findings. However, generalizability to the broader Dutch population requires caution as our study did not weight results to the entire Dutch adolescent population.

The major strength of this study is the population-based design with a large sample size in an early adolescent population. Moreover, automated measurement is fast and highly reproducible for LCEA calculation which helps eliminate observer bias that is inherently present with manual measurement. Our study has some limitations that should be addressed. First, our study only included the right hips for analyses, so the prevalence of pincer morphology was based on hip level rather than person level. Moreover, it is important to note that analyzing only in right hips underestimates the prevalence of pincer morphology at the person level since participants may have unilateral pincer morphology in their left hips. Secondly, DXA is a less common imaging modality for determining pincer morphology. However, it provides sufficient resolution to identify hip morphologies with less radiation burden than radiographs and has been validated against AP pelvic radiographs in adults (ICC for LCEA: 0.93 [95% CI 0.91 - 0.94])[150]. Unlike three-dimensional imaging modalities or additional lateral radiographs, the anterior center edge angle could not be obtained from the AP DXA hip images, so this may lead to an underestimation of the prevalence of pincer morphology[126]. Thirdly, care should be taken when generalizing our results to late adolescents due to the potential for further development of pincer morphology when the study population becomes older.

In conclusion, our study provides a large-scale objective evaluation of pincer morphology among early adolescents from the general population in the Netherlands. Pincer morphology was present in 3.1% of early adolescents, and is similarly prevalent in male and female participants. Our study provides valuable data for this age group, serving as a reference for investigating the development of pincer morphology in males and females throughout adolescence. Given its higher prevalence in adults, pincer morphology might also develop further during skeletal maturation of the hip like cam morphology. Our findings could inform the timing of prevention strategies for pincer morphology, potentially reducing the risk of FAIS and hip OA.

Acknowledgements

The Generation R Study is conducted by the Erasmus MC, University Medical Center Rotterdam in close collaboration with the Erasmus University Rotterdam and the city of Rotterdam. We gratefully acknowledge the contribution of children and parents. The general design of Generation R Study is made possible by long-term financial support from Erasmus MC, University Medical Center Rotterdam, the Netherlands, Organization for Health Research and Development (ZonMw) and the Dutch Ministry of Health, Welfare and Sport.

CRediT authorship contribution statement

D. Chen: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing.

F. Boel: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Software, Writing – original draft, Writing – review & editing.

S. de Vos-Jakobs: Data curation, Writing – review & editing.

P. van Klij: Conceptualization, Writing – review & editing.

M.M.A. van Buuren: Conceptualization, Writing – review & editing.

S.M.A. Bierma-Zeinstra: Conceptualization, Project administration, Supervision, Writing –review & editing.

R. Agricola: Conceptualization, Methodology, Supervision, Writing – original draft, Writing –review & editing.

All authors approved the final version of the manuscript.

Funding

DC was funded by the China Scholarship Council (CSC) PhD Fellowship for his PhD study at Erasmus Medical Center, Rotterdam, the Netherlands. (Grant number: 202108430032).

Chapter 8

Known risk factors for developmental dysplasia of the hip are not associated with early adolescent acetabular dysplasia: findings from a prospective open population cohort

F. Boel, S. de Vos-Jakobs, D. Chen, J. Runhaar, S. Bierma-Zeinstra, R. Agricola.

Submitted.

ABSTRACT

Background: We wanted to investigate whether known risk factors for developmental dysplasia of the hip (DDH) are also associated with acetabular dysplasia in early adolescence, as we hypothesized that acetabular dysplasia during and after infancy are different types of acetabular dysplasia with different risk factors. We also investigated BMI and physical activity since these were shown to be associated with acetabular dysplasia in childhood.

Methods: All participants around age 13 from the prospective, population-based cohort study Generation R, with a dual-energy x-ray absorptiometry (DXA) image of the right hip and full-body were eligible for inclusion. Acetabular dysplasia was defined by a lateral centre edge angle <20°, automatically determined on the hip DXA images using validated methods. The association of known DDH risk factors (female sex assigned at birth, low birth weight, breech presentation, caesarean section and firstborn), weight status and sport participation over time with the presence of acetabular dysplasia at 13 years was investigated using logistic regression, adjusted for age at outcome, skeletal maturity and ethnicity.

Findings: 3986 early adolescents with a mean age of 13.6 ± 0.4 years, 53% female, were included in this study. The prevalence of acetabular dysplasia was 6.4%. Known risk factors for DDH were not significantly associated with acetabular dysplasia around age 13, and neither was sports participation. Being overweight between the ages of 6 and 13 years compared to participants of normal weight was protective for acetabular dysplasia, aOR 0.39 (95% CI 0.18 – 0.85).

Interpretation: Known risk factors for DDH and sport participation were not associated with acetabular dysplasia in early adolescents, while weight status over time was. This might indicate that acetabular dysplasia during and after infancy are different types of acetabular dysplasia with different risk factors.

Funding: The Dutch Arthritis Association.

INTRODUCTION

A shallow acetabulum with insufficient femoral head coverage characterizes acetabular dysplasia. Acetabular dysplasia is part of the spectrum of developmental dysplasia of the hip (DDH). DDH includes hip dislocation, hip subluxation, hip instability, and a stable hip with insufficient acetabular coverage[14,15]. Acetabular dysplasia can result in instability of the hip and increased stress on the labrum and articular cartilage. Therefore, acetabular dysplasia is associated with hip pain, decreased function and early-onset hip osteoarthritis[8-11].

DDH is generally thought to develop during the perinatal period and infancy. The prevalence of DDH in infants, according to a recent meta-analysis, is 1.4 % (95% CI 0.86 – 2.3)[16]. Risk factors commonly associated with DDH are female sex assigned at birth, breech presentation, a family history of DDH, caesarean section and firstborn status[14,17]. Low birth weight is generally considered to be protective for DDH[14]. Early detection of DDH allows for simple and effective non-surgical abduction treatment using a harness or cast to stimulate normal development of the acetabulum. To this end, various screening programs to detect and treat DDH in infancy have been implemented worldwide.

However, the prevalence of acetabular dysplasia in adults from the general population is higher than the prevalence of DDH in infants, ranging between 3.3 – 9.4% [23-25]. Possible explanations can be missed diagnoses in infancy, residual dysplasia, or that acetabular dysplasia also develops later during skeletal growth in childhood[26-28]. It remains unclear why and how often acetabular dysplasia occurs after infancy. Additionally, it is unknown whether DDH and acetabular dysplasia in adolescents and adults have the same aetiology, and thus share the same risk factors.

To this end, we aimed to investigate whether known risk factors of DDH are also associated with acetabular dysplasia in early adolescence. We also investigated BMI and physical activity since these were shown to be associated with acetabular dysplasia in childhood[106].

METHODS

Participants

The current study is embedded in the Generation R Study, a population-based prospective cohort study that follows participants from foetal life until young adulthood in the multi-ethnic population of Rotterdam, the Netherlands[47]. The study was approved by

the Medical Ethics Committee of the Erasmus MC (MEC-2007-413, MEC-2012-165, MEC-2015-749). All participants provided written informed consent.

All participants were invited to complete questionnaires and visit the Generation R research centre every 3-4 years from the age of 6. Three visits around the ages of 6, 9 and 13 years have been completed. All participants with a visit around the age of 13 years and high-resolution dual-energy x-ray absorptiometry (DXA) images of the right hip and full-body were eligible for inclusion in the current study. Participants were excluded from the current study if 1) the acetabulum was not fully depicted on the right hip DXA images, 2) the right hip or full-body DXA images contained a movement artefact, or 3) there was an artefact in the region of interest.

Acetabular dysplasia

Acetabular dysplasia is most commonly determined using the lateral centre edge angle (LCEA), which is a measure of femoral head coverage by the acetabulum. Acetabular dysplasia was defined by an LCEA < 20.0° [170,171]. The LCEA was defined as the angle between the line connecting the most lateral bony point of the acetabulum with the femoral head centre and the line through the femoral head centre perpendicular to the horizontal reference line of the pelvis (HRLP), see Figure 1. We automatically determined the LCEA on all right hip DXA images with a reliable and valid method[91]. We corrected for pelvic obliquity using the HRLP determined on the full-body DXA image.

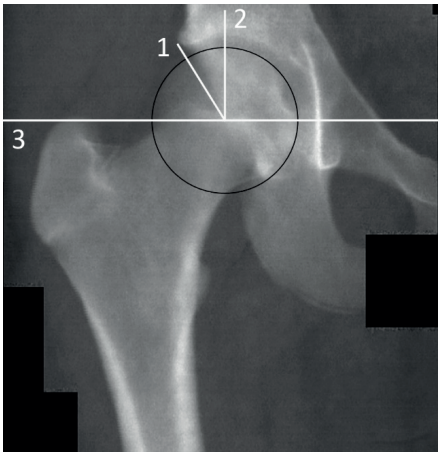


Figure 1. The lateral centre edge angle (LCEA) as determined on a right hip dual-energy x-ray absorptiometry (DXA) image. The LCEA is defined as the angle between line 1 connecting the most lateral bony point of the acetabulum with the femoral head centre and line 2 through the femoral head centre perpendicular to the horizontal reference line of the pelvis (line 3).

Image acquisition

The DXA images were acquired using a GE-Lunar iDXA densitometer (GE Healthcare, Madison, WI, USA) and enCORE software (enCORE 2010; GE Healthcare). The participants were positioned supine with legs slightly apart and internally rotated so that the big toes were touching and secured in this position using a Velcro strap around both feet. A unilateral anteroposterior (AP) DXA of the right hip and an AP full-body DXA image were acquired consecutively for each participant.

Variables of interest

The perinatal variables were obtained by Generation R through the medical records retrieved from hospitals and midwife practices. The variables included sex assigned at birth, birth weight, presentation of the foetus, delivery mode, and parity of the mother. Birth weight was standardised based on sex assigned at birth and gestational age[172]. Next, a dichotomous variable was created describing whether the child had a low (<2500 g) or normal (≥ 2500 g) birth weight. Foetal presentation at birth was categorised as breech and other presentation. The delivery mode was categorised as vaginal delivery, elective caesarean section or urgent caesarean section. Lastly, maternal parity was used to determine if the child was the firstborn.

The participants' height and weight were measured at each visit following standardised protocols and BMI was calculated. The International Obesity Task Force BMI cut-offs (IOTF grade) were used to assess weight status as normal, overweight or underweight adjusted for age and sex assigned at birth[173].

Information on sports participation was obtained using questionnaires sent at each of the three time points. The question was posed whether the participant engaged in sports. The primary caregiver filled in the questionnaire for ages 6 and 9, while the participant filled in the questionnaire for age 13. Based on the questionnaire data, a dichotomous variable was created for each age whether the participant participated in sports.

Age at outcome, skeletal maturity (triradiate cartilage status) at outcome, and ethnicity were considered confounders. The triradiate cartilage status was rated on the right hip DXA images as open or closed by a paediatric orthopaedic surgeon, with an intraobserver reliability of 0.97 (95% CI 0.95 – 0.98) and an interobserver reliability of 0.95 (95% CI 0.91 – 0.97)[171]. The triradiate cartilage status was used as a measure of the skeletal maturity of the acetabulum. Ethnicity was defined according to the classification of Statistics Netherlands. The participants' ethnicity was assessed based on the country of birth of the participants and their parents[174]. Ethnicity was categorised into four groups:

Western (Dutch, Turkish, American western, other European, Oceanic), African (Cape Verdean, Moroccan, Antillean, Surinamese-Creole, other African), Asian (Indonesian, Surinamese-Hindu, other Asian) and other (Surinamese-unspecified, American non-western)[175].

Statistical analyses

We assessed the association of perinatal variables (sex assigned at birth, low birth weight, breech presentation, delivery mode and firstborn status), IOTF grade over time and sport participation over time with the presence of acetabular dysplasia using logistic regression. Male sex assigned at birth, normal birth weight, other presentation than breech, vaginal delivery, and not firstborn were defined as the reference category for each categorical variable, respectively. To model IOTF grade and sport participation over time, exposure patterns were created, as suggested by Meinert et al.[176]. Based on the IOTF grades of the three visits, seven categories were created to describe the IOTF grade patterns over time: always normal, previously normal currently overweight, previously normal currently underweight, only previously overweight, only previously underweight, always overweight, always underweight. Always normal IOTF grade was used as the reference category. Similarly, based on the three dichotomous sports participation variables, sports participation over time was categorised into four categories: never played sports, only previously played sports, only currently plays sports, and always played sports. Never played sports was used as the reference category. Multicollinearity of all predictors was assed using variance inflation factors. All variance inflation factors ranged between one and two.

Missingness of the data ranged from 0% for participant characteristics and variables related to the visit around age 13 to 21% for questionnaire data on sports participation. Only 55% of participants (2194 out of 3986) were complete cases. An overview of all missing data can be found in supplementary material 1. Data are primarily missing due to nonresponse. Data was assumed to be missing at random. Multiple imputation by chained equations (MICE) was used to impute the missing values and 50 imputed datasets, with 10 iterations each, were created[177]. MICE was performed based on all variables of interest, confounders, and the following auxiliary variables: ethnicity of the mother, gestational age at birth, and age at recording of BMI for the visits around the ages of 6 and 9. MICE results were compared to observed results to assess if the imputed results were plausible.

The results from all imputed datasets were pooled and presented as odd ratios (ORs) with 95% CIs adjusted for age at outcome, skeletal maturity and ethnicity. Analyses were performed using R Statistical Software (v4.1.0; R Core Team 2021)[178]. MICE was

performed using the mice-package[179]. Multicollinearity assessments were performed using the VIF function from the car-package[180]. The glm-function in base R was used to perform logistic regression. The visualisation was created using the ggplot2-package[74].

RESULTS

3986 adolescents were included in the current study; both complete case and imputed participant characteristics are shown in Table 1. The prevalence of acetabular dysplasia was 6.4%. A flowchart of the included participants is shown in Figure 2. None of the known risk factors of DDH were significantly associated with acetabular dysplasia in early adolescence. Being overweight from around age 6 to age 13 was negatively associated with acetabular dysplasia, OR 0.39 (95% CI 0.18 – 0.85) adjusted for age at outcome, skeletal maturity and ethnicity. Sport participation was not significantly associated with acetabular dysplasia around age 13 years, see Figure 3.

Table 1: Participant characteristics

Characteristic	N with available data	Values, N (%)	Imputed values, N (%)
Sex assigned at birth, female	3986	2122 (53.2%)	NA
Ethnicity	3892		
Western		2895 (74.4%)	2955 (74.1%)
African		644 (16.5%)	668 (16.8%)
Asian		237 (6.1%)	243 (6.1%)
Other		116 (3.0%)	120 (3.0%)
Low birth weight	3980	230 (5.8%)	231 (5.8%)
Breech birth	3656	151 (4.1%)	174 (4.4%)
Delivery mode	3471		
Vaginal delivery		3007 (86.6%)	3450 (86.6%)
Elective caesarean section		173 (5.0%)	202 (5.0%)
Urgent caesarean section		291 (8.4%)	334 (8.4%)
Firstborn	3866	2232 (57.7%)	2301 (57.7%)
IOTF grade	3467		
Always normal		2131 (61.5%)	2421 (60.7%)
Previously normal currently overweight		272 (7.8%)	326 (8.2%)
Previously normal currently underweight		291 (8.4%)	322 (8.1%)
Only previously overweight		277 (8.0%)	321 (8.0%)
Only previously underweight		141 (4.1%)	163 (4.1%)
Always overweight		288 (8.3%)	349 (8.8%)
Always underweight		67 (1.9%)	84 (2.1%)

Table 1: Participant characteristics (*continued*)

Characteristic	N with available data	Values, N (%)	Imputed values, N (%)
Sports participation	2679		
Never played sports		100 (3.7%)	206 (5.2%)
Only previously played sports		1133 (42.3%)	1636 (41.0%)
Only currently plays sports		331 (12.4%)	586 (14.7%)
Always played sports		1115 (41.6%)	1558 (39.1%)
Age at outcome in years (mean (SD))	3986	13.6 (0.4)	NA
Triradiate cartilage status	3986		
Open		999 (25.0%)	NA
Closed		2987 (75.0%)	NA
Acetabular dysplasia (LCEA < 20.0°)	3986	254 (6.4%)	NA

The imputed values presented are the mean values of all 50 imputed datasets with N=3986. IOTF: international obesity task force. SD: standard deviation. LCEA: lateral centre edge angle. NA: not applicable, there were no missing values in the variable.

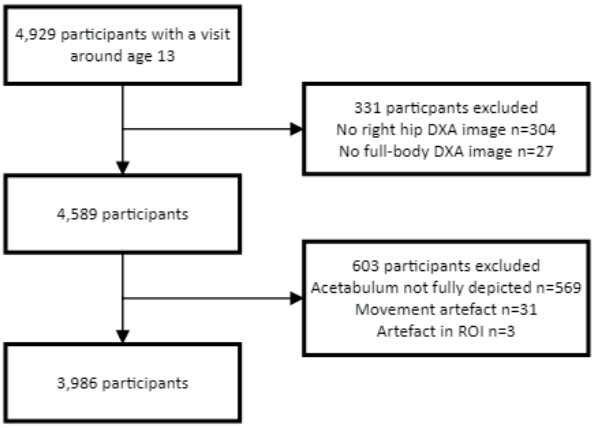


Figure 2. Flowchart of participant inclusion. DXA: dual-energy x-ray absorptiometry. ROI: region of interest.

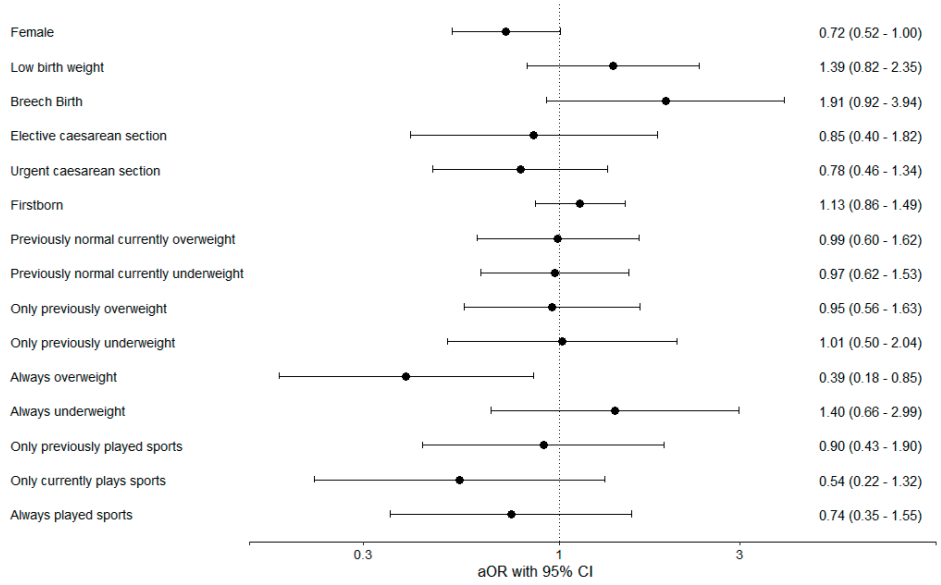


Figure 3. Plot of the resulting adjusted odds ratios (aORs) and 95% confidence intervals (95% CIs) from the multivariable logistic regression with outcome acetabular dysplasia around age 13 years. The model was adjusted for age at outcome, skeletal maturity and ethnicity. The reference categories were male sex assigned at birth, normal birth weight, other presentation than breech, vaginal delivery, not firstborn, always normal IOTF grade, and never played sports, respectively.

DISCUSSION

We found that none of the known DDH risk factors were associated with the presence of acetabular dysplasia in early adolescents in an open-population cohort from Rotterdam, the Netherlands. There was a negative association between having been overweight from age 6 years onwards and the presence of acetabular dysplasia around age 13 years compared to participants who were always of normal weight. This might indicate that being overweight for 6-7 years during childhood and early adolescence is protective for the presence of acetabular dysplasia around age 13 years. Meanwhile, no association was found between sports participation during childhood and the presence of acetabular dysplasia around age 13.

Previously, when assessing acetabular dysplasia within 1188 children around age 9 years from the Generation R Study, they also found a protective association between being overweight and the presence of acetabular dysplasia at age 9 [106]. Lee et al.[27] studied patients undergoing periacetabular osteotomy for acetabular dysplasia. They found that risk factors commonly associated with DDH, namely female sex assigned at birth and breech presentation, were less prevalent in adolescent or adult diagnosed acetabular dysplasia patients compared to the infant-diagnosed group. Humphry et al. studied

acetabular dysplasia after the age of 2 years. They found that 92% of patients (mean age 4.4 ± 0.78 years, range 2.0-6.6 years) diagnosed with mild or severe acetabular dysplasia had no prior DDH diagnosis, even though they all received ultrasound screening during infancy[26]. Contrary to our results, they did find a relationship with breech presentation, but their study population was much younger. Similarly, Laborie et al.[181] presented 18-year follow-up data in 2340 participants from a randomized controlled trial and found that female sex assigned at birth, breech presentation, left hip side, the alpha angle measured on ultrasound at birth and late abduction treatment were predictive for acetabular dysplasia in young adulthood. However, they did not correct for previous DDH diagnosis and treatment, so the observed relationship could be a result of residual DDH. The difference in the assessed hip side could also explain the variation in the results; we only assessed right hips, whereas Laborie et al. assessed both hips.

It has been hypothesized that acetabular dysplasia in infancy versus adolescence or adulthood are two distinct forms of acetabular dysplasia. The results of the current study that risk factors associated with DDH are not statistically significantly related to acetabular dysplasia around age 13 support this hypothesis. Additionally, female sex assigned at birth and a caesarean section seem to trend towards a protective effect in our current study, while they are risk factors commonly associated with DDH[14,17]. Conversely, low birth weight trends towards a risk factor in our current study but is considered to have a protective effect for DDH. However, DDH risk factors breech presentation at birth and being firstborn both trend towards a positive association with acetabular dysplasia around age 13. These findings highlight the need to better understand the development of acetabular dysplasia during childhood and its associated risk factors. Insight into adolescent-onset acetabular dysplasia could help identify patients before the onset of symptoms and maybe even prevent joint deterioration due to acetabular dysplasia.

The current study has several limitations. Firstly, only right hip DXA images were obtained for the Generation R cohort study, so only right hips were studied. This could possibly result in an underestimation of the prevalence of acetabular dysplasia and its relation to DDH risk factors, since DDH is known to have a higher occurrence in left hips[17]. However, adult acetabular dysplasia seems to have a similar prevalence in left and right hips and is more often bilateral[23,25]. Additionally, hip morphology has traditionally been assessed on radiographs. However, we previously showed good reliability and agreement of the LCEA measured on hip DXA images and pelvic radiographs, ICC 0.93 (95% CI 0.91 – 0.94)[150]. Next, there is no clear definition for acetabular dysplasia around age 13, especially with regard to the cut-off. However, we think that the LCEA is a good, reproducible measurement of femoral head coverage by the acetabulum, and a cut-off value of 20° is often used in literature[23-25,182]. Moreover, DDH diagnoses and possible treatment are unknown

for the participants, so we could not take this into account in our analyses. A part of the acetabular dysplasia found in the current study could be residual dysplasia from DDH. In addition, only sports participation was taken into account, meaning that the frequency, intensity and type of sport were not considered in the current study. Lastly, the study population consists of only people living in the Netherlands who are mostly of Western ethnicity; this means the findings might not apply to other populations.

Major strengths of the current study are its sample size, allowing to study multiple factors and enabling categorisation of variables with sufficient power, and the open-population nature of the cohort. Additionally, the LCEA measurements were performed using an automated validated method, ensuring consistent measurements and reducing reader bias. Lastly, the current study is the first study, to our knowledge, investigating the association between acetabular dysplasia in early adolescents and DDH risk factors, weight status and sports participation in an open population cohort.

Known risk factors for DDH were not significantly associated with acetabular dysplasia in early adolescence. Being overweight for an extended period during childhood and early adolescence compared to people of normal weight was negatively associated with acetabular dysplasia around age 13 years. Additionally, sports participation throughout childhood and early adolescence was not significantly associated with acetabular dysplasia. These results support the hypothesis that acetabular dysplasia during and after infancy are different types of acetabular dysplasia with different risk factors.

Acknowledgements

The Generation R Study is conducted by the Erasmus MC, University Medical Center Rotterdam in close collaboration with the Erasmus University Rotterdam and the city of Rotterdam. We gratefully acknowledge the contribution of children and parents. The general design of Generation R Study is made possible by long-term financial support from Erasmus MC, University Medical Center Rotterdam, the Netherlands, Organization for Health Research and Development (ZonMw) and the Dutch Ministry of Health, Welfare and Sport.

Data sharing

This manuscript is based on data from the Generation R Study. The Generation R Study has an open policy in regard to collaboration with other research groups. Requests for collaboration regarding the data should primarily be pointed to Generation R, https://generationr.nl/researchers/collaboration/. The automated measurements are published open access, for collaboration regarding the measurements please contact the corresponding author.

Declaration of interests

This research is supported by the Dutch Arthritis Association [grant number 19-2-202, 2019] and the European Research Council (ERC) grant “HIPSTAR” ERC-adv: 101054778.

SUPPLEMENTARY MATERIAL 1: DATA MISSINGNESS

Table 1. Number of complete cases and missing values for all variables.

Variable	Total number of complete cases	Missing values, N	Missing values, %
Sex assigned at birth	3986	0	0%
Ethnicity child	3892	94	2%
Ethnicity mother	3889	97	2%
Presentation at birth	3656	330	8%
Delivery	3471	515	13%
Firstborn	3866	120	3%
Gestational age at birth	3964	22	0.6%
Gestation weight	3979	7	0.2%
Triradiate cartilage status	3982	0	0%
LCEA at age 13	3986	0	0%
Age at visit around age 6	3697	289	7%
Age at visit around age 9	3688	298	7%
Age at visit around age 13	3986	0	0%
BMI at visit around age 6	3697	289	7%
BMI at visit around age 9	3660	326	8%
BMI at visit around age 13	3986	0	0%
IOTF grade at visit around age 6	3697	289	7%
IOTF grade at visit around age 9	3660	326	8%
IOTF grade at visit around age 13	3986	0	0%
IOTF grade pattern	3467	519	13%
Sport participation around age 6	3377	609	15%
Sport participation around age 9	3158	828	21%
Sport participation around age 13	3388	598	15%
Sport participation pattern	2679	1307	33%

LCEA: lateral centre edge angle. BMI: body mass index. IOTF: international obesity task force.

Chapter 9

Hip Shape is associated with the Change in Acetabular Coverage in Early Adolescence: A Longitudinal Study

F. Boel, S. de Vos-Jakobs, D. Chen, F. Rivadeneira, S.M.A. Bierma-Zeinstra, J. Runhaar*, R. Agricola*.

*Shared last author

Submitted.

ABSTRACT

Although cross-sectional studies suggested that acetabular coverage changes during childhood, longitudinal data and insights into underlying mechanisms are lacking. We investigated the change of acetabular coverage during childhood and its association with birth-assigned sex, weight status, triradiate cartilage orientation (TCO), and head-shaft angle (HSA). Participants from the population-based cohort Generation R with a dual-energy x-ray absorptiometry (DXA) image of the right hip and full-body available around age 9 and 13 years were included. Acetabular coverage was automatically determined on all DXA images using the lateral center edge angle (LCEA). Additionally, TCO and HSA were automatically measured around age 9. A marginal model, adjusted for skeletal maturity at age 13, was used to analyze the change in LCEA over time and its associations with sex, weight status at ages 9 and 13, and TCO and HSA around age 9. Time was modeled using the participants' age. Five hundred sixteen children were included, of whom 50% were female. On average, the LCEA increased with age, and females had higher LCEA values. Overweight and obese participants exhibited higher LCEA values compared to normal-weight participants. Each unit increase in TCO increased the LCEA by 0.10° on average. HSA was not significantly associated with LCEA change. This longitudinal study demonstrated that the acetabular coverage changes throughout childhood. Birth-assigned sex, weight status, and TCO were associated with this change. Further research is needed to explore the long-term implications of these developmental changes and establish appropriate cut-off values for acetabular dysplasia and pincer morphology in pediatric populations.

INTRODUCTION

The hip is a ball and socket joint formed by the proximal femur and the acetabulum. The development of the hip joint during childhood is a complex process, dependent on the interaction between the acetabulum and proximal femur[2]. Abnormal hip joint morphology can arise during growth, where alterations in the development of one of the hip joint components can lead to changes in the shape of the other.

Acetabular dysplasia (acetabular undercoverage of the femoral head) and pincer morphology (overcoverage) are two hip morphologies associated with pain, reduced function, and the development of hip osteoarthritis[10,11,183,184]. Developmental dysplasia of the hip (DDH) is thought to develop during fetal life and infancy and is often diagnosed in infancy due to various screening programs implemented worldwide. The prevalence of acetabular dysplasia in adolescents and adults (3.3-9.4%) is higher than that of DDH in infancy (1.4%), suggesting that dysplasia can also develop independently of DDH[16,23,25,106,171]. Additionally, DDH and acetabular dysplasia at a later age have different risk factors[27,185]. Pincer morphology, commonly diagnosed in young adults after the onset of symptoms, has a reported prevalence of 3-74% in adults, with development thought to begin around age 12 [30,33].

Several factors could influence the shape of the acetabulum. The triradiate cartilage is the growth plate of the acetabulum, and its orientation may influence the acetabular coverage of the femoral head. Similarly, the proximal femur shape could affect the acetabular coverage due to the interplay between the acetabulum and the proximal femur. While one study has linked body mass index (BMI) to acetabular overcoverage in adults[186], no such association was found in another study, which only found an association with proximal femur morphology[187]. Therefore, the association between BMI and acetabular coverage remains unclear.

The change of acetabular coverage throughout childhood and the underlying mechanisms remain unclear. Therefore, this study investigated how acetabular coverage changes over time in children from the general population. Additionally, its association with birth-assigned sex, weight status, triradiate cartilage orientation, and proximal femur shape was studied.

METHODS

Study population

The study population was drawn from the Generation R Study, a prospective, population-based cohort from Rotterdam, the Netherlands. The Generation R Study follows participants from fetal life through adulthood[47]. The research protocol was reviewed and approved by the Medical Ethics Committee of Erasmus MC (MEC-2012-165, MEC-2015-749). All study participants and their parent(s) or guardian(s) provided written informed consent.

Participants were invited to visit the Generation R research center around 9 and 13 years old. The participant's body mass index (BMI) was determined at each visit, and dual-energy x-ray absorptiometry (DXA) imaging of the full-body and right hip was performed. The International Obesity Task Force BMI cut-offs (IOTF grade) were used to assess weight status as normal, underweight, overweight or obese adjusted for age and birth-assigned sex[173]. Participants were eligible for inclusion in the current study if DXA imaging was performed at both visits . Participants were excluded from the current study if the acetabulum was not depicted on the right hip DXA image, if there was a movement artifact in the right hip or full-body DXA image, if there was an artifact in the region of interest, or if one of the hip morphology measurements could not be performed.

Image acquisition and processing

All DXA images were acquired using a GE-Lunar iDXA densitometer (GE Healthcare, Madison, WI, USA) and enCORE software (enCORE 2010; GE Healthcare). The participants were in supine position with legs internally rotated approximately 15° with their big toes touching and secured using a Velcro strap around both feet. An anteroposterior full-body and right hip DXA were acquired consecutively.

The hip morphology was automatically determined on the DXA imaging using a previously described, automated and validated method[91]. Although pelvic radiographs are traditionally used to determine hip morphology measurements, DXA images are a good alternative[150]. The lateral center edge angle (LCEA) was used to define the acetabular coverage of the femoral head at both age 9 and age 13, as shown in Figure 1A. The head-shaft angle (HSA) was used to evaluate the proximal femur orientation, Figure 1B. A higher HSA is a sign of a valgus hip, which could be associated with a more dysplastic hip[188]. The triradiate cartilage orientation (TCO) is a measure of the triradiate cartilage, where a TCO of zero degrees is descriptive of a horizontally oriented growth plate, Figure 1C. A more horizontally orientated growth plate is thought to result in less acetabular coverage in the full-grown hip. The horizontal reference line of the pelvis, as determined

on the full-body DXA images, was used to correct the LCEA and TCO for any pelvic obliquity.

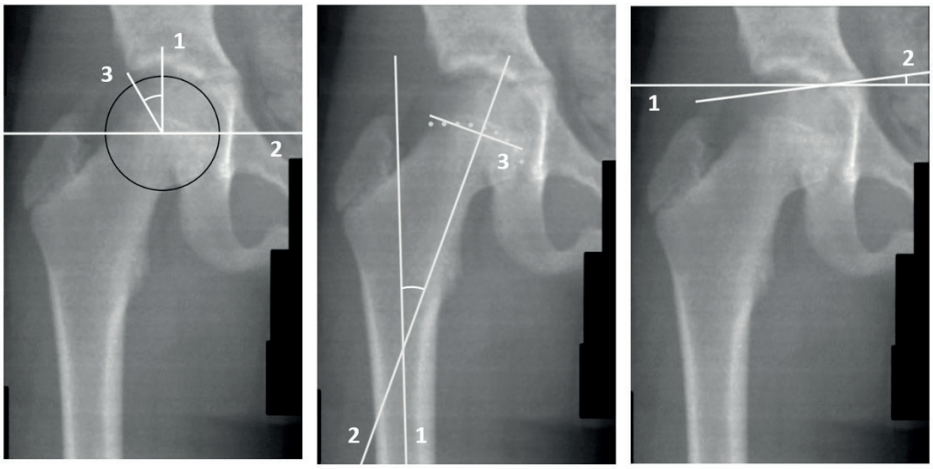


Figure 1. Hip morphology measurements as determined on the right hip DXA image of a 9-year-old. **A: The lateral center edge angle (LCEA)** - the angle between line 1, perpendicular to the horizontal reference line of the pelvic (HRLP) as determined on the full-body DXA (line 2) through the center of the femoral head, and line 3 from the center of the femoral head to the most lateral bony edge of the acetabulum. **B: The head-shaft angle (HSA)** - the angle between the shaft axis, line 1, and line 2, which is perpendicular to the linear line fitted through the femoral head growth plate landmarks. **C: The triradiate cartilage orientation** - the angle between the HRLP (line 1) and line 2 through the most laterosuperior point and the most mediosuperior point of the triradiate cartilage.

The triradiate cartilage and the femoral head growth plate were scored as open or closed by experienced readers on the DXA hips imaging around ages 9 and 13. The intra- and interobserver reliability of the triradiate cartilage status were 0.97 (95% CI 0.95 – 0.98) and 0.95 (95% CI 0.91 – 0.97), respectively[171]. The intra- and interobserver reliability of the femoral head growth plate status were 0.90 (95% CI, 0.85 - 0.94) and 0.87 (95% CI, 0.78 - 0.92), respectively[189]. Since the HSA and TCO can only be performed on an open growth plate, all participants with closed growth plates on the DXA image at age 9 were excluded from the current study.

Statistical analyses

A histogram of the change in LCEA for all participants was created, where the change in LCEA was defined as 'the LCEA at age 13 years minus the LCEA at age 9 years'. We assessed how the LCEA changed over time and how this is related to the birth-assigned sex, weight status around age 9 and age 13, TCO around age 9, and HSA around age 9 using a repeated measurement model. The participant's age was used to model time. The triradiate cartilage status (open/closed) at the visit around age 13 was considered a confounder to correct for differences related to skeletal maturity. A marginal model

based on generalized least squares and a linear mixed-effects model were built and optimized; see supplementary material 1. The marginal model with a continuous AR1 correlation matrix and age modeled nonlinearly best fit the included data. Age was modeled using natural cubic splines with three degrees of freedom, and the boundary knots were set at the 5% and 95% quantiles, respectively. The model assumptions were checked using residual and QQ plots. No model assumptions were violated. The model coefficients were presented with standard errors and p-values. Additionally, to provide insight into the relationship between age and LCEA, the effect plot showing the model results and a scatterplot with a loess curve and 95% confidence interval (CI) were visualized, see supplementary material 2.

A sensitivity analysis was performed to further examine the impact of skeletal maturity, as defined by triradiate cartilage status, by creating subgroups with an open triradiate cartilage and a closed triradiate cartilage around age 13. The same marginal model used for the primary analysis was built for each subgroup.

Analyses were performed using R Statistical Software (v4.1.0; R Core Team 2021)[178]. The gls-function of the nlme-package was used to create the marginal model[190]. The histogram and scatterplot were created using the ggplot2-package[74]. The effect plot was created using the ggeffects-package[191].

RESULTS

Five hundred sixteen participants were included in the current study. Participant characteristics are shown in Table 1. The flowchart of participant's in- and exclusion can be found in Figure 2. The change in LCEA within each participant is visualized in Figure 3.

Birth-assigned sex, overweight or obese compared to normal weight and TCO were significantly associated with the change of LCEA, while HSA was not associated, see Table 2. Females had, on average, a 1.03 degree higher LCEA than males of the same age, with the same HSA, weight status, TCO, and triradiate cartilage status. Participants with overweight or obesity had, on average, a 1.42 and 2.08 degree higher LCEA, respectively, than participants of normal weight of the same age and birth-assigned sex, with the same HSA and TCO. Each unit increase in TCO increased the average LCEA by 0.10 degrees for participants of the same birth-assigned sex and age with the same weight status, HSA and triradiate cartilage status.

Sensitivity analysis

The subgroups comprised 154 participants and 362 participants with an open or closed triradiate cartilage, respectively (Table 3). More males than females had an open triradiate cartilage around age 13. Additionally, participants with an open triradiate cartilage around age 13 seemed more normal or underweight than those with a closed triradiate cartilage around age 13. Lastly, the LCEA in participants with an open triradiate cartilage seemed lower for both ages than those with a closed triradiate cartilage.

Table 1. Participant characteristics

	Included		**Excluded**	
Characteristic	**Visit age 9**	**Visit age 13**	**Visit age 9**	**Visit age 13**
N	516		5346	4413
Birth-assigned sex, female	259 (50.2)		2688 (50.3)	2250 (51.0)
Age, years	9.8 (0.5)	13.5 (0.3)	9.8 (0.4)	13.6 (0.4)
IOTF grade*				
Normal	391 (75.8)	370 (71.7)	3891 (75.0)	3231 (73.4)
Underweight	43 (8.3)	65 (12.6)	352 (6.8)	446 (10.1)
Overweight	71 (13.8)	65 (12.6)	749 (14.4)	570 (12.9)
Obese	11 (2.1)	16 (3.1)	197 (3.8)	156 (3.5)
LCEA, degrees	21.6 (5.4)	28.3 (6.0)	-	-
TCO, degrees	19.3 (7.2)	-	-	-
HSA, degrees	152.6 (5.9)	-	-	-
Triradiate cartilage status, open	516 (100)	154 (29.8)	-	-

Dichotomous variables are presented as n (%) and continuous variables as mean (SD). IOTF grade: International Obesity Task Force BMI cut-offs, LCEA: lateral center edge angle, TCO: triradiate cartilage orientation, HSA: head-shaft angle. *IOTF grade data unavailable for 157 excluded participants at visit age 9 and 10 excluded participants at visit age 13.

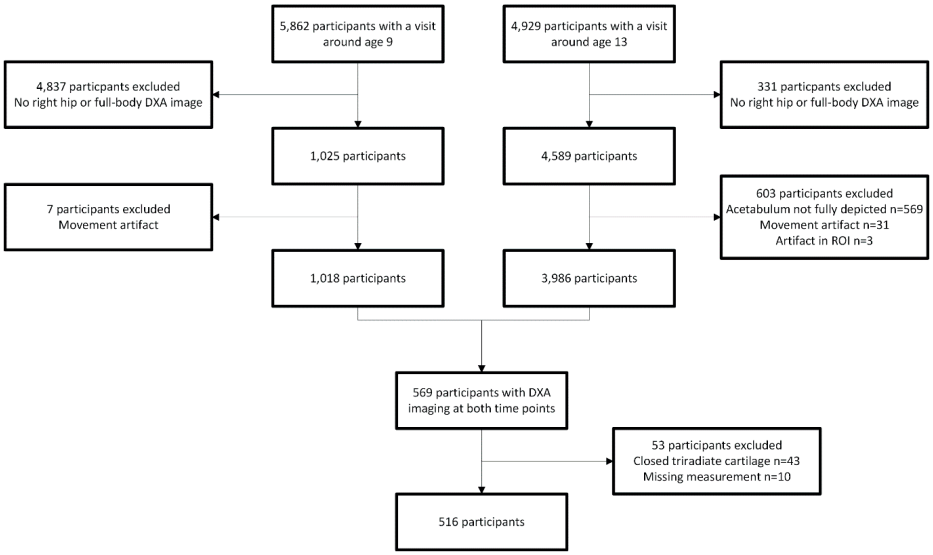


Figure 2. Flowchart of participant inclusion. DXA: dual-energy x-ray absorptiometry. ROI: region of interest.

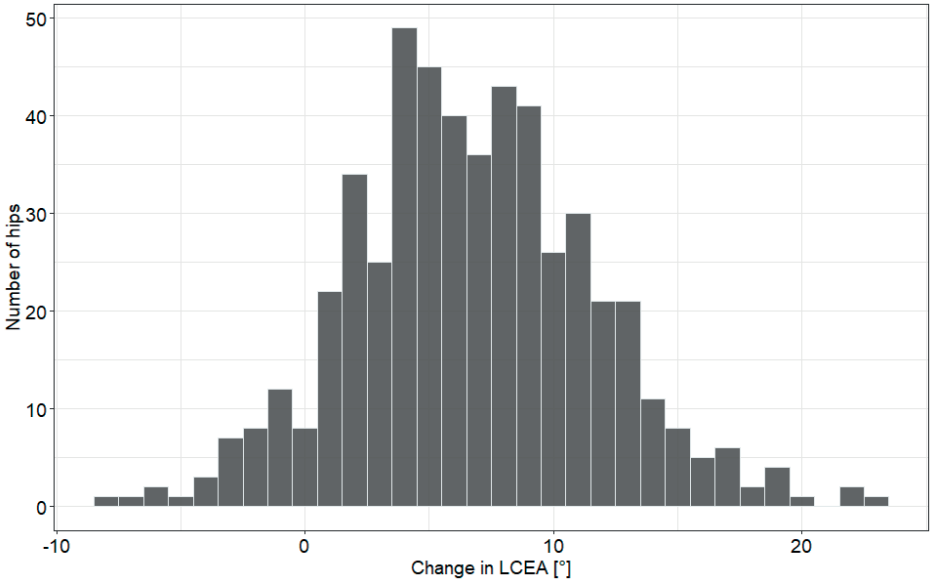


Figure 3. Histogram of the change in lateral center edge angle (LCEA) as defined by the difference between the LCEA at ages 9 and 13.

Table 2. Generalized Least Squares Regression Results for LCEA change over time.

	β	Standard error	p-value
Age, years			
spline degree 1	7.21	0.91	< 0.0001
spline degree 2	3.75	1.25	0.0026
spline degree 3	8.82	0.65	< 0.0001
Female	1.03	0.50	0.040
HSA, degrees	0.004	0.04	0.92
IOTF grade			
Underweight	-0.70	0.56	0.21
Overweight	1.42	0.50	0.0047
Obese	2.08	1.00	0.038
TCO, degrees	0.10	0.03	0.0009

HSA: head-shaft angle. IOTF grade: International Obesity Task Force BMI cut-offs. TCO: triradiate cartilage orientation. Normal weight was used as the reference category for the IOTF grade. All values are corrected for triradiate cartilage status.

Table 3. Participant characteristics of subgroups with an open or closed triradiate cartilage around age 13.

	Open		Closed	
Characteristic	**Visit age 9**	**Visit age 13**	**Visit age 9**	**Visit age 13**
N	154		362	
Birth-assigned sex, female	18 (11.7)		241 (66.6)	
Age, years	9.8 (0.4)	13.4 (0.3)	9.8 (0.5)	13.5 (0.3)
IOTF grade*				
Normal	133 (86.4)	113 (73.4)	258 (71.3)	257 (71.0)
Underweight	16 (10.4)	32 (20.8)	27 (7.5)	33 (9.1)
Overweight	4 (2.6)	8 (5.2)	67 (18.5)	57 (15.7)
Obese	1 (0.6)	1 (0.6)	10 (2.8)	15 (4.1)
LCEA, degrees	19.9 (4.9)	21.5 (5.8)	22.3 (5.4)	29.5 (5.8)
TCO, degrees	19.1 (7.2)	-	19.4 (7.3)	-
HSA, degrees	152.0 (5.7)	-	152.8 (6.0)	-
Triradiate cartilage status, open	154 (100)	154 (100)	362 (100)	0 (0.0)

Dichotomous variables are presented as n (%) and continuous variables as mean (SD). IOTF grade: International Obesity Task Force BMI cut-offs, LCEA: lateral center edge angle, TCO: triradiate cartilage orientation, HSA: head-shaft angle.

In the subgroup of participants with an open triradiate cartilage around age 13, only TCO was significantly associated with the change in LCEA, see Table 4. Each unit increase in TCO increased the average LCEA by 0.15 degrees for participants of the same age and birth-assigned sex, with the same HSA and weight status.

In the subgroup of participants with a closed triradiate cartilage around age 13, birth-assigned sex, overweight compared to normal weight, and TCO were significantly as-

sociated with the change in LCEA, see Table 4. Females have, on average, a 1.23 degree higher LCEA than males of the same age with the same HSA, weight status and TCO. Participants who were overweight had, on average, a 1.64 degree higher LCEA than participants of normal weight of the same age and birth-assigned sex, with the same HSA and TCO. Lastly, each unit increase in TCO increased the average LCEA by 0.04 degrees for participants of the same age and birth-assigned sex, with the same HSA and weight status.

Table 4. Generalized Least Squares Regression Results for LCEA change over time for subgroup analysis with an open or closed triradiate cartilage around age 13.

	Open			**Closed**		
	β	**Standard error**	**p-value**	**β**	**Standard error**	**p-value**
Age, years						
spline degree 1	4.09	2.17	0.06	8.43	1.01	< 0.0001
spline degree 2	3.35	2.39	0.16	4.16	1.49	0.0052
spline degree 3	7.95	1.12	< 0.0001	9.17	0.79	< 0.0001
Female	-0.24	1.19	0.84	1.23	0.56	0.0279
HSA, degrees	-0.05	0.07	0.46	0.03	0.04	0.56
IOTF grade						
Underweight	-0.69	0.88	0.43	-0.41	0.72	0.56
Overweight	1.27	1.55	0.41	1.64	0.54	0.0026
Obese	2.31	2.94	0.43	2.04	1.07	0.0568
TCO, degrees	0.15	0.05	0.006	0.07	0.04	0.046

HSA: head-shaft angle. IOTF grade: International Obesity Task Force BMI cut-offs. TCO: triradiate cartilage orientation. Normal weight was used as the reference category for the IOTF grade. All values are corrected for triradiate cartilage status.

DISCUSSION

This longitudinal study demonstrated that acetabular coverage, as measured by LCEA, changes throughout childhood and is associated with birth-assigned sex, weight status and triradiate cartilage orientation. Notably, HSA showed no association with LCEA change.

To our knowledge, this is the first longitudinal study investigating the change of the LCEA during childhood and its relationship to birth-assigned sex, weight status, HSA, and TCO. Previous cross-sectional studies demonstrated a similar increase in LCEA with age[33,192]. However, in cross-sectional studies investigating sex differences in LCEA, males had, on average, higher LCEA values than females[122,193]. Laborie et al.[193] studied 19-year-olds and Nishimura et al.[122] studied 12 to 18-year-olds, while our study population is younger. This apparent contradiction can likely be explained by considering the influ-

ence of skeletal maturity. Generally, females will be more skeletally mature than males of the same age. Consequently, females might exhibit higher LCEA values in younger populations due to their advanced skeletal development. While our analysis controlled for skeletal age using triradiate cartilage status at age 13, it is important to acknowledge that this measure does not fully capture the skeletal age of the hip. The os acetabuli, a secondary ossification center at the acetabular rim, also contributes to the growth of the acetabulum and, thus, acetabular coverage. This ossification center is thought to appear around age 9 in females and age 11 in males and will be entirely fused around age 11 in females and age 13 in males[4]. The ossification timeline of the os acetabuli, particularly during the age range of our study participants, could explain the observed sex differences in LCEA. As the os acetabuli is not present from birth, its absence on imaging could indicate either it has not yet appeared, or it is already fused, making it challenging to score the status of this ossification center. Consequently, the status of this ossification center was not included as a confounder in the current study.

Further supporting the role of skeletal maturity, our subgroup analysis of participants with an open triradiate cartilage found no significant association between birth-assigned sex and LCEA change. This suggests that the sex differences observed in our primary analysis could be driven by skeletal maturity, influenced by factors like the os acetabuli, rather than inherent sex differences in hip morphology.

Participants with overweight or obesity had, on average, 1.42 and 2.08 degrees higher LCEA, respectively, than normal-weight participants of the same age and sex, with the same TCO, HSA, and triradiate cartilage status. This aligns with previous research in adults demonstrating an association between obesity and acetabular overcoverage[186]. Similarly, a cross-sectional analysis of the Generation R data around age 9 found a negative association between BMI and acetabular dysplasia[106]. The influence of BMI on acetabular coverage might be due to excessive mechanical load impacting bone growth[186,187,194]. Additionally, metabolic and endocrine differences related to high BMI may influence growth plates and hip morphology. Cross-sectional studies have shown an association between accelerated skeletal maturation and overweight and obesity[195,196], further highlighting the impact of high BMI on growth plates. This was reflected in the current study, where nearly all (96%) participants with an open triradiate cartilage around age 13 were underweight or of normal weight.

Each degree increase in TCO was associated with an average increase of 0.10° of the LCEA when all other variables were the same. This indicates that a more vertical orientation of the triradiate cartilage around age 9 is associated with a higher LCEA value. This relationship was also found within the sensitivity analysis, where each degree increase

in TCO was associated with an increase of 0.15° and 0.07° in the LCEA within the subgroups with an open and closed triradiate cartilage around age 13, respectively. This suggests that triradiate cartilage orientation may be a factor in acetabular development and a potential early indicator for future acetabular dysplasia or pincer morphology.

The current study has some limitations. Firstly, only a small part of the Generation R Study population could be included in the current study resulting in possible selection bias. Due to logistical constraints, the DXA images were only performed in a subset of participants visiting the research center around age 9 years. However, the birth-assigned sex distribution, age, and BMI of the included and excluded participants were comparable and 516 participants could still be included. Secondly, only right hip DXA images were obtained in the Generation R study, so only right hips were included in the current study. A study on 19-year-olds showed that the average LCEA in right hips is lower than in left hips in both males and females[193]. However, we expect similar developmental trends in both hips since the found trajectory was similar to studies including both hips[33,192]. Thirdly, the TCO is a novel measurement devised by the authors of the current study since, to the best of our knowledge, no measurement existed to assess the triradiate cartilage orientation. This makes it impossible to compare the found results on TCO to literature. Lastly, while the automated method of the LCEA was validated against manual measurements, the HSA and TCO were not formally validated. During development, the automated HSA and TCO underwent qualitative assessments by two experienced pediatric orthopedic surgeons. This was done through a visual assessment of the automated measurements.

The current study also has some major strengths. Firstly, the longitudinal nature of the data allowed us to study the development of acetabular coverage within participants. Secondly, the automated measurements allowed for consistent measurements with reduced reader bias. Lastly, the population-based nature of the Generation R cohort enabled us to study the acetabular coverage in the general population.

The LCEA is used to quantify both acetabular dysplasia and pincer morphology on two-dimensional hip images. However, while some consensus exists on the correct cut-off values for adults, clear cut-off values for children and adolescents are missing. This highlights the critical need for age- and sex-specific cut-off values, given the observed changes in LCEA with age and sex in the current study. This would aid in a more consistent definition in epidemiological research and help create more insight into the development of these hip morphologies. Accurately assessing acetabular dysplasia and pincer morphology in pediatric populations is important for understanding their etiology. Critically, extending this research into late adolescence and early adulthood will aid

in thoroughly characterizing the developmental trajectory of acetabular coverage and, thus, the development of acetabular dysplasia and pincer morphology. This would provide valuable insights into skeletal maturation and could help identify at-risk individuals for developing hip problems in adulthood, such as hip pain, reduced function, and hip osteoarthritis.

In conclusion, acetabular coverage continues to increase throughout childhood and is associated with birth-assigned sex, weight status, and TCO. Age- and birth-assigned sex-specific cut-off values for LCEA are needed to improve the assessment of acetabular dysplasia and pincer morphology in the pediatric population. Future research should investigate the long-term implications of the observed developmental changes in LCEA.

Acknowledgements

The Generation R Study is conducted by the Erasmus MC, University Medical Center Rotterdam in close collaboration with the Erasmus University Rotterdam and the city of Rotterdam. We gratefully acknowledge the contribution of children and parents. The general design of Generation R Study is made possible by long-term financial support from Erasmus MC, University Medical Center Rotterdam, the Netherlands, Organization for Health Research and Development (ZonMw) and the Dutch Ministry of Health, Welfare and Sport.

Data sharing

This manuscript is based on data from the Generation R Study. The Generation R Study has an open policy in regard to collaboration with other research groups. Requests for collaboration regarding the data should primarily be pointed to Generation R, https://generationr.nl/researchers/collaboration/. The automated measurements are published open access, for collaboration regarding the measurements please contact the corresponding author.

Declaration of interests

This research is supported by the Dutch Arthritis Association [grant number 19-2-202, 2019] and the European Research Council (ERC) grant “HIPSTAR” ERC-adv: 101054778.

Part 3

Hip Morphology and Osteoarthritis

Chapter 10

Acetabular dysplasia and the risk of developing hip osteoarthritis within 4-8 years: An individual participant data meta-analysis of 18,807 hips from the World COACH consortium

F. Boel*, N.S. Riedstra*, M.M.A. van Buuren, H. Ahedi, V. Arbabi, N.K. Arden, S. J. Baart, S.M.A. Bierma-Zeinstra, F.M. Cicuttini, T.F. Cootes, K. Crossley, D.T. Felson, W.P. Gielis J. Heerey, G. Jones, S. Kluzek, N.E. Lane, C. Lindner, J. Lynch, J.B.J. van Meurs, A.E. Nelson, A. Mosler, M.C. Nevitt, E.H. Oei, J. Runhaar, J. Tang, H. Weinans, R. Agricola.

*Shared first authorship

Osteoarthritis and Cartilage. 2025 Mar;33(3):373-382.

ABSTRACT

Objective: To study the association between various radiographic definitions of acetabular dysplasia (AD) and incident radiographic hip osteoarthritis (RHOA), and to analyze in subgroups.

Methods: Hips free of RHOA at baseline and with follow-up within 4–8 years were drawn from the World COACH consortium. The Wiberg center edge angle (WCEA), acetabular depth width ratio (ADR), and the modified acetabular index (mAI) were calculated. AD was defined as WCEA ≤ 25°, and for secondary analyses as WCEA ≤ 20°, ADR ≤ 250, mAI ≥ 13°, and a combination. A logistic regression model with generalized mixed effects with 3 levels adjusted for age, biological sex, and body mass index (BMI) was used. Descriptive statistics stratified by age, biological sex and BMI were reported.

Results: A total of 18,807 hips from 9 studies were included. Baseline characteristics: age 61.84 (± 8.32) years, BMI 27.40 (± 4.49) kg/m^2, 70.1% women. 4766 hips (25.3%) had WCEA ≤ 25°. Within 4–8 years (mean 5.8 ±1.6) follow-up, 378 hips (2.0%) developed incident RHOA. We found an association between AD and RHOA (adjusted OR [aOR] 1.80 95% confidence interval [CI] 1.40–2.34). In secondary analyses, all other definitions of AD were also associated with incident RHOA (aOR ranging from 1.52 95% CI 1.19–1.94 to 1.96 95% CI 1.26–3.02). Descriptive statistics showed that the relative risk (RR) in AD hips to develop RHOA was higher compared to non-AD hips in age group 61–70 (RR 1.70), BMI<25 (RR 1.66), and in female hips (RR 1.73).

Conclusion: AD was consistently associated with incident RHOA. Explorative analyses show that AD hips in women and age group 61–70 years seem to be more at risk of developing RHOA compared to non-AD hips.

INTRODUCTION

There is no curative nonsurgical treatment available for hip osteoarthritis (OA)[197,198]. Therefore, prevention is critical, but there is a lack of knowledge on risk factors for the development of radiographic hip OA (RHOA). Identifying risk factors for this disease should be prioritized.

Subtle features of hip shape may predate the development of OA by many years and might therefore be a preventative target[199,200]. Acetabular dysplasia (AD) has previously been identified as a risk factor for developing RHOA[45,201]. AD is defined by insufficient coverage of the femoral head by the acetabulum[202]. Concentrated focal stress on a relatively small area of the acetabulum[202] is thought to lead to early mechanical failure of the cartilage, and to eventually cause hip OA[10,11,45,203].

A systematic review on hip morphology and OA found an association between AD and RHOA odds ratio [OR] 2.38, 95% confidence interval [CI]: 1.84 to 3.07)[11]. However, when analyzing individual studies, these have yielded conflicting results and highlight the need for robust analysis, avoiding inconsistencies in measurements and definitions[8,10,11,58,126]. Single cohorts are likely to be underpowered to determine whether specific high-risk subgroups are responsible for the associations found[11]. Likely due to the overall low number of included individuals and therefore decreased statistical power, existing prospective cohort studies include hips free of RHOA as well as those with doubtful RHOA at baseline, which may bias the presently known associations. Hips with doubtful RHOA already show mild radiographic changes (possible joint space narrowing [JSN] and signs of osteophytes), which may influence the radiographic measures of AD and represent the first signs of potential osteoarthritic changes[8,10,58].

Using an individual participant data (IPD) meta-analysis, we aimed to investigate the association between AD, defined by the Wiberg center edge angle (WCEA) ≤25° at baseline, and developing incident RHOA within 4–8 years follow-up. For secondary analyses, we investigated whether other measures of AD and other threshold values to quantify AD were associated with incident RHOA. Finally, we performed subgroup analyses stratified by age, biological sex, and body mass index (BMI) to assess potential high-risk subgroups.

METHODS

Study design and participants

Participants were drawn from the Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) consortium. The World COACH consortium is an international collaboration of all worldwide available prospective cohort studies with sequential pelvic or hip imaging. The consortium profile has previously been published in detail elsewhere[48].

For the present study, we included all cohorts with a follow-up anteroposterior (AP) pelvic radiograph within 4–8 years of a baseline radiograph, that also had an RHOA score available. This led to the inclusion of 9 cohorts (Cohort Hip and Cohort Knee [CHECK], Multi-center Osteoarthritis Study [MOST], Osteo Arthritis Initiative [OAI], Rotterdam Study-I [RS-I], Rotterdam Study-II [RS-II], Rotterdam Study-III [RS-III], the Chingford Study, The Johnston County Project [JoCo] and the Study of Osteoporotic Fractures [SOF]), and exclusion of two cohorts (Tasmanian Older Adults Cohort [TASOAC] and Femoroacetabular impingement and hip osteoarthritis cohort [FORCe]).

We included hips with known BMI, biological sex, and age at baseline. Next, we excluded hips without an original baseline RHOA score. We then excluded radiographs of insufficient quality for automated AD measurement calculation and all AP hip radiographs as they did not allow for constructing a horizontal reference line to adjust for pelvic rotation. Next, we included only the hips with an original RHOA score at follow-up and excluded all baseline hips with pincer morphology (acetabular overcoverage) as determined by a lateral center edge angle (LCEA) ≥40°. We chose to do the latter to compare hips with AD to a reference group with normal acetabular coverage, and because studies have found an association between pincer morphology and RHOA or total hip replacement (THR)[8,11]. Finally, we included only hips free of any signs of RHOA at baseline (any score=0). Studying a population completely free of RHOA at baseline allows for the determination of true predictors of RHOA, as existing osteophytes may affect the measurement of AD and bias the association between AD and incident RHOA. Furthermore, excluding hips with doubtful RHOA isolates the effect of AD on incidence RHOA rather than the effect of AD on progression in hips that likely already have some form of RHOA. This led to a total inclusion of 18,807 hips.

Radiographs

AP pelvic radiographs were taken at baseline and at follow-up between 4–8 years in each included cohort, according to cohort-specific protocols which have been published previously[204-208] (Supplementary material 1). To study the impact of the full-limb films from the MOST cohort on the associations between AD and RHOA, that are otherwise studied on pelvic radiographs, sensitivity analyses were performed excluding hips from the MOST cohort from the primary analysis.

Radiographic measurements

To avoid measurement variability across cohorts, for the present study all AD measurements were calculated uniformly on baseline radiographs. The bony outline of the proximal femur and acetabulum were automatically annotated on the AP pelvic radiographs with a point set using the BoneFinder® software (www.bone-finder.com; The University of Manchester, UK)[64]. This point set was used to perform automated radiographic measurements using a previously published Python script, which was adapted and validated for World COACH data, for which a detailed description can be found elsewhere[91,117]. The average of two trained reader's manual measurements were compared to automated morphological measurements. The average of the trained manual readers was considered the gold standard to which the automated method is compared. Intermethod interclass correlation coefficients (ICCs) range from 0.80 (0.60 – 0.90) for the acetabular depth width ratio (ADR) to 0.88 (0.70 – 0.95) for the WCEA[117].

Radiographic measurements to define AD are depicted (Figure 1). The amount of weight-bearing coverage of the femoral head by the acetabulum is measured by the WCEA[8,9,11]. AD was defined as a WCEA ≤ 25° in the primary analysis and in subgroup analyses, and additionally by WCEA≤20° for secondary analyses[10,126]. The ADR is a measure of depth of the acetabulum. AD was defined as an ADR ≤ 250 for secondary analyses[92]. The modified acetabular index (mAI) measures the inclination of the acetabular roof. AD was defined as an mAI ≥ 13° for secondary analyses[92]. In secondary analyses, the radiographic definitions of AD were studied individually and combined.

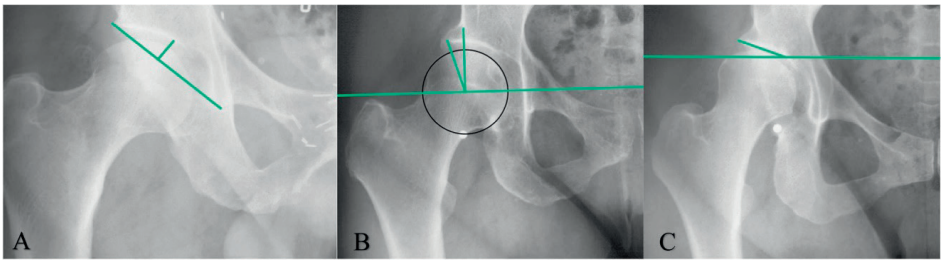


Figure 1. Anteroposterior in pelvic radiographs with three radiographic measurements to define AD. **A: The acetabular depth-width ratio (ADR):** The acetabular width was defined as a line across the length of the acetabular opening, extending from the lateral edge of the acetabulum to the pelvic teardrop. Next, the acetabular depth was determined by constructing a line perpendicular to the acetabular width, extending from the most medial point of the sourcil. The ADR is defined as the ratio of the depth to the width, multiplied by 1000. **B: The Center Edge Angle of Wiberg (WCEA):** To determine the center of the femoral head, a best-fitting circle is outlined around the femoral head based on the SSM points. The WCEA is then formed by a line drawn vertically through the center of the femoral head, perpendicular to the horizontal reference line, the second is drawn from the center of the femoral head to the most lateral weight-bearing part of the sourcil. **C: The modified acetabular index (mAI):** The mAI measures the acetabular roof inclination. The measure is modified, as the original acetabular index is applied to hips with an open triradiate cartilage. The mAI measures inclination from the medial sourcil to the lateral bony part of the acetabulum. Horizontal reference line (B+C): To correct for pelvic rotation, a horizontal reference line is calculated based on the average of 4 lines, between 1) both femoral head centers, 2) the most cranial points of the foramen obturator, 3) the most caudal points of the ischial tuberosity and 4) the most caudal points of the pelvic teardrop.

Radiographic Hip Osteoarthritis Grading

At baseline and follow-ups, all included radiographs had scores available by either the Kellgren and Lawrence (KL) classification (CHECK, Chingford, JoCo, MOST, RS-I, RS-II, RS-III)[39], the modified Croft classification (SOF)[209], or a modified OA score (OAI)[210].

The KL grading system defines OA severity in five grades(0−4) using a combination of osteophyte, JSN severity, sclerosis and bone deformity[39]. The modified Croft grading system defines OA severity in five grades(0−4) and is based on 5 radiographic features: JSN, osteophytes, subchondral sclerosis, cyst formation, and deformity[211]. The modified OA grades are based on the modified Croft grades and defines OA in 3 grades(0−2), where 0 marks hips free of RHOA, 1 defines doubtful RHOA and 2 is definite RHOA[210].

Original OA scores per cohort were defined as “free of RHOA” (any score 0), “doubtful RHOA” (any score 1), or “definite RHOA” (KL ≥2, Modified Croft ≥2, Modified OA=2, or THR)[10,212,213].

Outcome measurements

The outcome was incident score “definite RHOA” within 4–8 years follow-up. Additionally for secondary analyses RHOA was defined as an ordinal outcome “free of RHOA”, “doubtful RHOA” and “definite RHOA”.

Statistical analysis

All statistical analyses were performed in R version 4.1.1. Univariate differences in baseline characteristics between complete included and excluded cases were inspected. This means that we compared the included hips to the hips that were excluded because of OA score of 1 or 2 at baseline (Fig. 2). The associations between baseline AD, defined by the WCEA≤ 25°, and incident RHOA were estimated with mixed effects logistic regression models. Mixed effects were added to account for the potential clustering in the data within cohorts and participants. Random intercepts were determined on both participant and cohort level, with participants nested within the cohorts. The cohort was added as a level in this multi-level model to adjust for possible residual confounding by study differences. An example is the difference between an open population cohort (Chingford, JoCo, RS-I, RS-II, RS-III), and closed population cohort (CHECK, OAI, MOST, SOF). The results are expressed as adjusted OR (aOR) and unadjusted OR with 95% CI and were adjusted for baseline age, sex, and BMI. Additionally, a mixed model for ordinal data, namely RHOA classified as "free of RHOA", "doubtful RHOA", and "definite RHOA" was created using a forward build continuation ratio model to assess the impact of doubtful RHOA. Random effects were added to adjust for clustering, and the model was adjusted for baseline age, sex, and BMI. The ordinality assumption of the continuous ratio model was relaxed for AD, allowing the effect of AD to be different for each level of the outcome RHOA at follow-up. The results were presented as an effect plot of the marginal probabilities marginalized over the random effects for women, with mean baseline age and BMI and randomly selected left hip side. Secondary analyses were performed using the same model and 5 definitions of AD: 1) WCEA ≤ 20°, 2) ADR ≤ 250, 3) mAI ≥ 13°, 4) three combined measures (WCEA ≤ 25° and ADR ≤ 250 and mAI ≥ 13°), and 5) a pooled definition of any of the three measures (WCEA ≤ 25° or ADR ≤ 250 or mAI ≥ 13°). In the secondary analysis, when using a WCEA ≤ 20° as the predictor, hips with a WCEA > 20° and WCEA ≤ 25° were excluded from the reference group to compare AD hips to a clean population of hips completely free of AD. In the secondary analysis, when using any of the three measures for AD, the reference group contained hips free of AD and hips with only one or two measures of AD. We used descriptive statistics to explore whether specific subgroups may be more at risk to develop RHOA. Because of limited outcomes, it was not possible to perform subgroup analyses using logistic regression. We reported absolute risk (AR) and relative risk (RR) in AD and non-AD hips of developing RHOA stratified by age groups 40–50, 51–60, 61–70 and >70 years of age, by BMI by studying groups with a BMI > 25 and BMI ≤ 25, and by biological sex. The following packages in R were used: Logistic regression was performed using the lme4-package[214]. The continuation ratio model was created using the GLMMadaptive package[215]. The effect plot was created using the ggplot2-package[74].

RESULTS

Participants

The flow of hips from those available in World COACH to the current final study population is depicted (Figure 2). 18,807 hips free of any signs of RHOA at baseline were included. The mean interval between the baseline and follow-up radiograph across all cohorts was 5.8 ± 1.6 years. Baseline demographic data stratified per cohort are presented in Table 1. Our study population was younger than the excluded hips (61.84 ± 8.32 versus 64.56 ± 8.49 years, respectively); all other baseline characteristics and predictors were similar across included and excluded hips.

Acetabular dysplasia

At baseline, 4766 (25.3%) hips had AD defined by a WCEA ≤ 25°, 1164 (6.2%) according to a WCEA ≤ 20°, 5917 (31.5%) hips had an ADR ≤ 250 and 397 (2.1%) hips had an mAI ≥ 13°. The overlap between measures is illustrated in supplementary material 2.

Incident radiographic hip osteoarthritis

378 hips (2.0%) developed incident RHOA within 4–8 years follow-up. The incidence of RHOA at follow-up per cohort were: CHECK: 13.4% Chingford: 8.4%, JoCo: 6.4%, MOST 0.6%, OAI: 0.5%, SOF: 1.7%, RS-I: 0.9%, RS-II: 0.4%, RS-III: 2.2%.

Primary analysis: association between acetabular dysplasia and radiographic hip osteoarthritis

A significant association (aOR 1.80 (95% CI 1.40–2.34) between AD (WCEA ≤ 25°) and incident RHOA within 4–8 years was observed. The association remained statistically significant after adjustment for covariates.

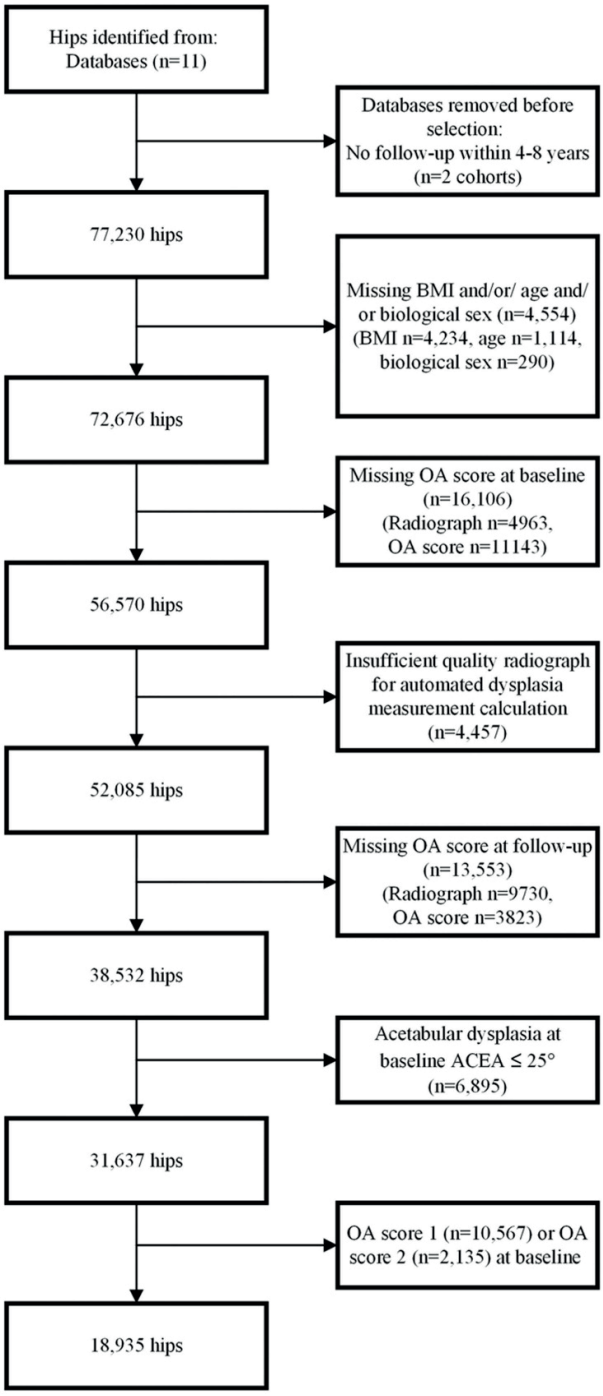


Figure 2. Flow of hips from consortium inclusion to final study population. OA: osteoarthritis. LCEA: lateral center edge angle. AP: anteroposterior. BMI: body mass index.

Table 1. Baseline characteristic of included hips, stratified per cohort.

	CHECK	Ching	JoCo	MOST	OAI	SOF	RS-I	RS-II	RS-III	Total incl.	Total excl.[a]
Hips, n (participants)	753 (457)	727 (464)	614 (442)	1776 (1078)	4552 (2618)	3547 (2399)	2460 (1660)	1810 (1056)	2568 (1463)	18,807 (11,637)	9708 (6772)
Age, mean (sd) years	55.65 (5.24)	53.09 (5.59)	58.21 (8.49)	60.84 (7.43)	59.96 (8.90)	70.07 (4.27)	64.84 (6.46)	62.75 (6.15)	56.13 (4.85)	61.84 (8.32)	64.56 (8.49)
BMI, mean (sd) kg/m^2	26.11 (4.02)	25.44 (4.09)	29.62 (6.01)	29.60 (5.03)	28.12 (4.60)	26.42 (4.37)	26.37 (3.53)	27.20 (3.86)	27.46 (4.08)	27.40 (4.49)	27.59 (4.73)
Men, n (%)	117 (15.5)	0 (0)	298 (48.5)	501 (28.2)	1,895 (41.6)	0 (0)	955 (38.8)	756 (41.8)	1109 (43.2)	5631 (29.9)	3158 (32.5)
WCEA ≤20°, n (%)	48 (6.4)	43 (5.9)	18 (2.9)	274 (15.4)	234 (5.1)	106 (3.0)	161 (6.5)	106 (5.9)	174 (6.8)	1164 (6.2)	510 (5.3)
WCEA ≤25°, n (%)	218 (29.07)	173 (23.8)	130 (21.2)	795 (44.8)	1,085 (23.8)	575 (16.2)	628 (25.5)	453 (25.0)	709 (27.6)	4766 (25.3)	2105 (21.7)
ADR ≤250, n (%)	248 (32.9)	212 (29.2)	159 (25.9)	715 (40.3)	1,273 (28.0)	955 (26.9)	857 (34.8)	651 (36.0)	847 (33.0)	5917 (31.5)	2694 (27.8)
mAI ≥ 13°, n (%)	14 (1.9)	26 (3.6)	11 (1.8)	72 (4.1)	84 (1.8)	54 (1.5)	50 (2.0)	35 (1.9)	51 (2.0)	397 (2.1)	191 (2.0)
OA score=2 follow-up, men/women (%/%)	101 (13.4)	61 (8.4)	39 (6.4)	10 (0.6)	21 (0.5)	60 (1.7)	22 (0.9)	7 (0.4)	57 (2.2)	77/301 (1.4/2.3)	2390 (24.6)

CHECK= Cohort Hip and Cohort Knee, MOST= Multi-center Osteoarthritis Study, OAI= Osteo Arthritis Initiative, RS-I= Rotterdam Study-I, RS-II=Rotterdam Study-II, RS-III= Rotterdam Study-III (RS-III), Ching= the Chingford Study, JoCo=The Johnston County Project, SOF= Study of Osteoporotic Fractures, WCEA= Wiberg Center Edge Angle, ADR= Acetabular Depth-Width Ratio, mAI= Modified Acetabular Index. OA score: 0= no RHOA, 1= Doubtful RHOA, 2= Definite RHOA. [a] Excluded hips are defined as all eligible hips for analysis but with OA score 1 or 2 at baseline.

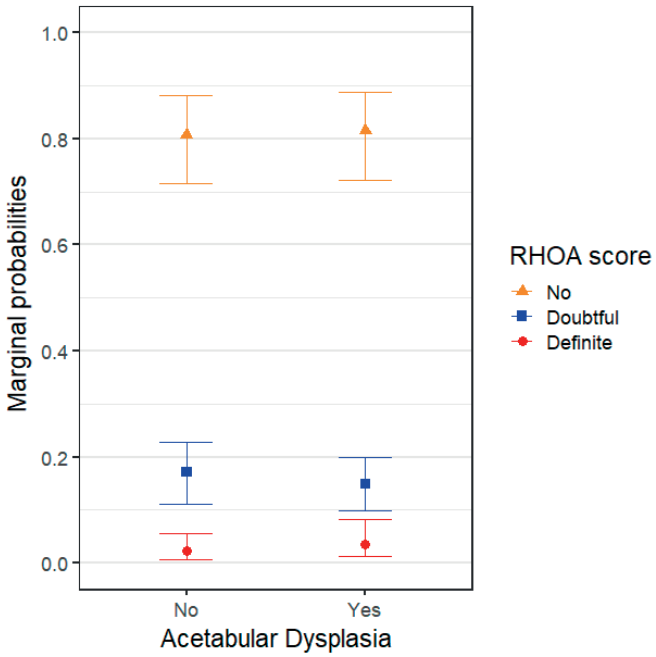


Figure 3. Marginal probabilities of RHOA score within 4–8 years for women aged 62 years and BMI of 27.4 kg/m^2 in hips with AD (WCEA ≤ 25°) or without AD. The probabilities were marginalized over the random effects, i.e., cohort and individual and the model was adjusted for age, BMI, biological sex, and hips side.

The effect plot of the marginal probabilities from the mixed model for ordinal data is shown in Figure 3. All marginal probabilities were calculated for hips free of RHOA, in women aged 62 years with a BMI of 27.4 kg/m^2 at baseline. The marginal probability for hips with AD to develop doubtful RHOA within 4–8 years is 0.15 (95% CI 0.10–0.20), compared to 0.17 (95% CI 0.11–0.23) for hips free of AD. The marginal probability for AD hips to develop definite RHOA within 4–8 years is 0.03 (95% CI 0.01–0.08), compared to 0.02 (95% CI 0.01–0.06) for AD-free hips.

Sensitivity analysis excluding MOST

The study population excluding MOST resulted in a total of 17,031 hips. A significant association was found (aOR 1.89 95%CI 1.45 −2.47) between hips with AD (WCEA ≤ 25°) and incident RHOA in the study population excluding hips from the MOST cohort.

Secondary analyses: association between various measures of acetabular dysplasia and radiographic hip osteoarthritis

Significant associations between AD defined by WCEA ≤ 20°, ADR ≤ 250 or either WCEA ≤ 25° or ADR ≤ 250 or mAI ≥ 13°) and incident RHOA within 4–8 years were observed. The associations remained statistically significant after adjusting for covariates. Because of a limited number of events (14 hips), it was not possible to calculate the association in the AD defined by mAI ≥ 13° group. All ORs are summarized in Table 2.

Table 2. Associations between various radiographic definitions of AD and incident RHOA.

Definition of AD	Hips with AD at baseline, n	Hips with incident RHOA at follow-up, n	Absolute risk (%)	Unadjusted OR (95% CI)[a]	Adjusted OR (95% CI)
WCEA ≤ 25°	4766	127	127/4766 (2.7)	**1.73 (1.33-2.25)**	**1.80 (1.40-2.34)**
WCEA ≤ 20°	1164	34	34/1164 (2.9)	**1.80 (1.28-2.52)**	**1.96 (1.26-3.02)**
ADR ≤ 250	5917	144	144/5917 (2.4)	**1.48 (1.15-1.90)**	**1.53 (1.19-1.96)**
mAI ≥ 13°	397	14	14/397 (3.5)	-[b]	-[b]
WCEA ≤ 25° & ADR ≤ 250 & mAI ≥ 13° [c]	351	14	14/351 (4.0)	-[b]	-[b]
WCEA ≤ 25° or ADR ≤ 250 or mAI ≥ 13° [d]	7480	176	176/7480 (2.4)	**1.47 (1.16-1.88)**	**1.52 (1.19-1.94)**

ORs were adjusted for age, BMI, biological sex, and hip side, and were accounted for by clustering cohort and individual. WCEA: Wiberg center edge angle. ADR: acetabular depth-width ratio. mAI: modified acetabular index. OR: odds ratio. CI: confidence interval. Significant associations are printed in bold. [a] The unadjusted odds ratios are calculated using the logistic regression model with generalized mixed effects with 3 levels (cohort, person and -hip side correlation) unadjusted for age, biological sex, and BMI. [b] Too few cases with both predictor and outcome to calculate an OR. [c] The reference group contained hips free of AD and hips with only 1 or 2 measures of AD. [d] The reference group contained hips free of any measure to define AD.

Subgroup analyses

Descriptive statistics stratified by age group, biological sex, and BMI are summarized in Table 3. The RR for hips with AD to develop RHOA was highest in age group 61–70, in hips with BMI < 25, and in women.

Table 3. Absolute and relative risk of hips with acetabular dysplasia (WCEA ≤ 25°) to develop incident radiographic hip osteoarthritis stratified by age group, BMI, and biological sex.

Strata	Total hips in group, n	Hips with AD (WCEA ≤ 25°), n	Hips with incident RHOA, n	Hips with AD and incident RHOA, n	Absolute Risk, % [a]	Relative Risk, % (95% CI) [b]
Age group (years)						
40-50	1753	526 (30.0)	35 (2.0)	11	0.6	1.07 (0.53-2.17)
51-60	6738	1921 (28.5)	159 (2.4)	57	0.8	1.40 (1.02-1.93)
61-70	7192	1691 (23.5)	128 (1.8)	44	0.6	1.70 (1.19-2.44)
70+	3124	628 (20.1)	56 (1.8)	15	0.5	1.45 (0.81-2.61)
BMI						
< 25	5874	1380 (23.5)	142 (2.4)	48	0.8	1.66 (1.18-2.34)
≥ 25	12,933	3386 (26.2)	236 (1.8)	79	0.6	1.42 (1.09-1.85)
Biological sex						
Men	5631	1369 (24.3)	77 (1.4)	14	0.2	0.69 (0.39-1.23)
Women	13,176	3397 (25.8)	301 (2.3)	113	0.9	1.73 (1.37-2.18)

AD: acetabular dysplasia. WCEA: Wiberg center edge angle. RHOA: Radiographic hip osteoarthritis. BMI: body mass index. [a] The absolute risk was calculated using the following equation: (number of hips with AD and RHOA/Total number of hips in subgroup). [b] The relative risk was calculated using the following equation: (number of hips with AD & RHOA/(number of hips with AD & RHOA + number of hips with AD only)) / (number of hips with RHOA without AD/ (number of hips with RHOA without AD + number of hips without AD or RHOA)).

DISCUSSION

This IPD meta-analysis on the association between AD and incident RHOA in a large prospective study of 18,807 hips free of any RHOA at baseline, demonstrated a significant association between AD defined by a WCEA ≤ 25° and incident RHOA within 4–8 years. Additionally, hips with AD were more likely to progress from being RHOA-free to definite RHOA rather than doubtful RHOA compared to non-AD hips. Secondary analyses showed that other measures of AD (WCEA ≤ 20°, ADR ≤ 250 and a combination of WCEA ≤ 25° and ADR ≤ 250) were also associated with an increased risk of developing RHOA. Descriptive statistics show that AD hips in women, individuals aged 61–70 and individuals with BMI < 25 have a higher RR to develop RHOA.

Several studies have shown that AD is associated with the development of RHOA. The strength of associations in prospective cohort studies ranged from aOR 1.56 (95% CI 1.09–2.24) to aOR 5.45 (95% CI 2.40–12.34)[8,10,11,52,126,216]. Conversely, a number of studies (case-control, prospective and cross-sectional) have failed to find such an association[58,217,218]. Our results support the finding that AD is associated with RHOA, although the association in the present study is not as strong as previously reported. This may be explained by the fact that the present study population only included hips free of any RHOA at baseline, whereas previous prospective cohort studies also included hips with doubtful RHOA at baseline, in which the stronger associations may reflect an association between AD and progression of RHOA, rather than incident RHOA[8,10,11,52,126,216]. Furthermore, publication bias may have played a role in selective publication of (strong) associations between AD and RHOA previously, and negative results may have been disfavored[219]. Time to follow-up as well as how AD and RHOA are measured and defined may also have contributed to variable strengths of associations in prospective studies or absence of an association in cross-sectional studies.

Although generalizable evidence is lacking, it has been hypothesized that AD leads to RHOA only in younger individuals[8]. Saberi et al. studied hips from RS-I and RS-II with an average age of 65 years at baseline and found that the magnitude of the association AD and RHOA was stronger in persons aged ≤65 years at baseline (OR 2.59 95% CI 1.62–4.16) compared to those aged > 65 years (OR 1.74 95% CI 0.90–3.37). On average, the population in this IPD meta-analysis is younger (61.84 years), which is likely because only hips completely free of RHOA were included at baseline, whereas the aforementioned study also included hips with doubtful RHOA at baseline. Our study lacked sufficient statistical power to stratify associations by age. However, the descriptive subgroup statistics showed that hips with AD aged 61–70 at baseline had an increased RR (1.70 95% CI 1.19–2.44) of developing incident RHOA, which was lower in other age groups although the CI overlaps (age 40–50 RR: 1.07 95% CI 0.53–2.17, age 51–60 RR: 1.40 95% CI 1.02–1.93, age 70+ RR: 1.45 95% CI 0.81–2.61). This finding suggests that younger individuals with AD may not be more at risk, but future studies with sufficient power should further analyze these associations.

The prevalence of AD defined by a WCEA ≤ 25° in our study population was similar in women (25.8%) compared to men (24.3%) in the study population. Although acetabular undercoverage was equally common in men and women in our study, we found that the RR is significantly lower in men with AD to develop RHOA (RR 0.69 95% CI 0.39–1.23) compared to women (RR 1.73 95% CI 1.37–2.18). It has been hypothesized that women have different joint alignment and thus joint load distribution than men. Estrogen metabolism, or pregnancy related pelvic instability may play a role in sex differences[145].

We hypothesized that a higher overall body weight may lead to higher joint load and therefore increase the risk of RHOA in overweight individuals with AD. Descriptive statistics show a slightly increased RR for AD hips in individuals who have a BMI<25 (RR 1.66 95% CI 1.18–2.34) compared to AD hips of individuals with BMI ≥25 (RR 1.42 1.09–1.85) in our study population, but CIs overlap meaning this is likely not significant. Interestingly, a recent study in children found a negative association between being overweight and developing dysplasia[106]. A previous study in the Rotterdam Study also reported low BMI as a risk factor for AD hips to develop RHOA[9].

The sensitivity analysis excluding the MOST cohort, as this contained long-limb radiographs rather than AP pelvic radiographs, yielded similar results to the primary analysis. Excluding MOST led to an aOR of 1.89 (95%CI 1.45 −2.47) for the association between AD and RHOA, compared to the primary analysis, which does include the MOST cohort of aOR 1.80 (95% CI 1.40–2.34). By including the hips from the MOST cohort in the primary analysis, we argue that the reported aORs contribute to generalizable results considering the added variation of a different radiographic view. Furthermore, the CIs largely overlap, from which we conclude that there are no statistically important differences between the study population, including and excluding the hips from the MOST cohort.

Quantification of AD may have impacted the previously reported associations between AD and RHOA[92]. In the present study, WCEA rather than LCEA was employed, as we argue that the weight-bearing surface, rather than the entire bony femoral head coverage, is under stress as a result of AD. Secondly, the threshold to define AD also vary in the literature. We used a threshold of WCEA ≤ 25° which indicates mild AD and should be kept in mind when interpreting the results. The association between AD when defined by WCEA ≤ 20° increased, which may indicate that more severe AD increases the risk of developing RHOA. Finally, we found that most studies only use acetabular coverage as a measure of AD, but for the present study we examined if acetabular depth or roof inclination influenced the reported associations. We found that both acetabular undercoverage as well as a shallow acetabulum were significantly associated with RHOA at follow-up in the present population. Whether acetabular roof inclination is also associated with RHOA could not be concluded from the present study, but future studies with long-term follow-up and therefore likely a higher incidence of RHOA may shed light on this measurement as a predictor.

A comprehensive definition of hip OA in epidemiological studies is still lacking[220]. Commonly used RHOA classification systems are the KL and (modified) Croft grading systems[39,209,213,220], for which good ICCs (κ = 0.55–0.92) have been reported in the World COACH cohorts[39,205,206,209,213,220,221]. The inevitable variability in RHOA grading was

corrected for in the logistic regression model by accounting for within-cohort effects. The incidence of RHOA in the present study of 2.01% (range per cohort 0.5–13.4%) is relatively low compared to similar studies[8,10,126,221,222]. Interestingly, the cohorts with the highest incidence of RHOA at follow-up (CHECK (22.6%), JoCo (10.9%) and Chingford (8.6%) were on average younger at baseline than the cohorts with the lowest incidence (OAI (0.5%), RS-II (0.4%) and RS-III (0.2%). The overall low incidence of incident RHOA is likely related to the exclusion of hips with doubtful RHOA at baseline.

The primary strength of the present study is the design. IPD meta-analysis created increased statistical power, reduced publication bias, and allowed for investigation of subgroup effects[223]. As RHOA is a heterogeneous disease, identifying subgroups for interventions is likely a promising way forward in clinical research. A second benefit of IPD meta-analysis compared to meta-analysis alone is that we were able to choose confounders for all included hips. This allowed us to correct for the same covariates across all cohorts and perform uniform analyses. IPD also helped improve data quality by combining studies with different follow-up and outcomes, to improve generalizability of findings[224,225]. A second strength is the use of multiple, uniform radiographic measurements to quantify AD. Although the WCEA proved to be the only measure of AD significantly associated with incident RHOA in our analysis, we argue that by additionally studying the ADR and mAI, we captured more of the AD characteristics compared to studies only employing a CEA[8,52]. A third strength is the uniformity of automated measurements, which removed variability compared to manual measurements and allowed for objective AD measurements across all cohorts[11].

This study was subject to limitations. The primary limitation is the subjective nature of original OA grading systems, as they rely on subjective assessment of OA features. We accounted for variability in OA scores per cohort in our statistical model and argue that, as these grading systems are still primarily used in a clinical setting, our study represents best current clinical practice. It should also be kept in mind when interpreting the results that RHOA does not equate clinical hip OA[221]. A second limitation is the variety of radiographic protocols per cohort, such as supine vs. weight-bearing radiographs. However, a recent study showed that for JSN, no difference in measurements between weightbearing or supine AP radiographs was found[226]. A horizontal reference line allowed for standardization of all other measurements, which reduced variability to a minimum. A third limitation is the lack of statistical power to perform logistic regression in subgroups. As AD has been shown to be a risk factor in younger individuals to develop RHOA, and the number of individuals ≤ 50 years of age was very limited in the present study[8]. The current results cannot be generalized to the young adult population (≤ 50 years). Prospective studies of younger populations are needed to study this further. A

limitation of IPD meta-analysis is that it may become prone to selection bias when IPD are only sought for a specific subset of studies. The World COACH cohorts however have been recruited based on a systematic literature search, which has been repeated recently[48]. Clinicians, researchers, and patients are also actively involved to help identify studies that should be included in the consortium. We therefore argue that publication bias was minimized in our study. We used definitions of AD only in one plane, thereby potentially neglecting anatomical abnormalities that may exist at different planes simultaneously[126]. We argue however, that by using multiple measurements to define AD, we were still able to capture a wide array of anatomical variability, in line with current clinical practice.

In future studies, identification of modifiable risk factors is essential for prevention of hip OA, as well as improving quality of life by advancing individualized care and identification of new treatments. Hip dysplasia is recognized as a potentially modifiable risk factor. It has been hypothesized that there are two distinct forms of hip dysplasia; developmental dysplasia of the hip (DDH) which is diagnosed during infancy, and AD, which is diagnosed later in life[227]. A recent study found demographic differences between patients diagnosed with DDH in infancy and adults with AD, supporting this hypothesis[27]. Examination of newborns for hip instability exemplifies prevention for hip OA in DDH hips, as the plastic hip joint can be stabilized to produce a congruent joint. This study showed that AD in the adult population was highly prevalent depending on the definition used, but the association with RHOA in general may be weaker than previously thought. It is therefore warranted to further our understanding of which individuals with AD specifically are at high risk of developing hip OA, and, assuming that two distinct forms exist, investigate whether one form is clinically more relevant.

In conclusion, this study demonstrated that AD is a risk factor for incident RHOA in hips free of RHOA at baseline. This IPD meta-analysis allowed for a robust analysis of the association between AD and RHOA, due to the large sample size, uniform measurements of AD across all baseline radiographs, and harmonized outcome of RHOA. Identification of modifiable risk factors is essential for prevention of hip OA in the future, as well as improving quality of life by advancing individualized care and identification of new treatments.

Ethical approval

This study involves human participants but was excepted from ethical approval (Erasmus MC Medical Ethics Review Committee) as it uses previously collected observational data for which the participants had originally given informed consent, and all cohort studies included in this consortium already had ethics approval from their respective

committees. Participants gave informed consent to participate in the study before taking part.

Patient and public involvement

Patient involvement is ongoing in the World COACH consortium as they co-determine and prioritize research questions to be answered within World COACH. World COACH researchers attend annual conferences for patients with OA in the Netherlands (Artrose Gezond) and engage in open dialog with OA patients to form research goals. The potential of the consortium to discover risk factors and potential treatment options are explained to patients, and both patients and public are encouraged to share ideas and questions they want answered through www.worldcoachconsortium.com.

Funding

The World COACH consortium has been funded through grants by the Dutch Arthritis Society (grant nr 21-1-205), the Dutch Research Council (NOW, Veni: 09150161910071), and Erasmus MC University Medical Center Rotterdam. CL is funded by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (223267/Z/21/Z). This research was funded in whole, or in part, by the Wellcome Trust [Grant number 223267/Z/21/Z]. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

NSR, FB and RA initiated the study. NSR, FB and RA worked on the conceptual design of the study. MMAvB and RA identified eligible cohorts and contacted cohort investigators for collaboration. MMAvB, RA, NSR, FB, HA, AM, KC, JH, SK, JAL, JVM, ABM, AEN, MN, JT and HW collected the existing cohort data. MMAvB, NSR, FB, JT and RA have worked on the database and on the harmonization process. NSR, FB, and RA have worked on statistical analyses. NSR and FB wrote the manuscript under supervision of RA. All authors critically reviewed and revised the manuscript and contributed to interpretation of the data. All authors read and approved the final version of the manuscript. NR acts as guarantor and accepts full responsibility for the finished work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Declaration of competing interests

GJ reports personal fees from Novartis outside the submitted work. SBZ reports consulting fees from Pfizer Infirst Healthcare and personal fees for being a Deputy Editor for Osteoarthritis and Cartilage outside the submitted work. CL and TC report a patent for an image processing apparatus and method for fitting a deformable shape model to

an image using random forest regression voting. CL reports licensing royalties for this patent from Optasia Medical outside the submitted work. AN is an associate editor for Osteoarthritis and Cartilage and is on the OARSI Board of Directors outside the submitted work. AM is on the Editorial Board for the British Journal of Sports Medicine and the Journal of Science and Medicine in Sport outside the submitted work. HW reports being a minority shareholder of Uplanner BV and Replasia BV outside the submitted work.

Acknowledgments

We would like to thank all participants of each of the cohort studies that are involved in World COACH. We gratefully acknowledge all international organisations that collaborated with the cohort studies in World COACH, as well as the OARSI for endorsing the World COACH consortium.

CHECK

The CHECK study was initiated by the Dutch Arthritis Society 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.

Chingford

We would like to thank all the participants of the Chingford Women Study, Professor Nigel Arden, Professor Tim Spector, Dr Deborah Hart, Mr Gem Lawson, Maxine Daniels and Alison Turner for their time and dedication and Arthritis Research UK for their funding support to the study and the Oxford NIHR Musculoskeletal Biomedical Research Unit for funding contributions.

JoCo-OA

Support for data from the Johnston County Osteoarthritis Project was provided in part by: the Center for Disease Control and Prevention (CDC) U01DP006266 and U01DP003206; Association of Schools of Public Health/ CDC S043, S1734, S3486; and National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases P60AR30701, P60AR049465, P60AR064166, and P30AR072580.

MOST

The MOST study was funded by the National Institutes of Health – National Institute on Aging grants AG19069 (Michael Nevitt, University of California, San Francisco) AG18820

(David Felson, Boston University) AG18947 (Cora Lewis, University of Alabama at Birmingham) and AG18832 (James Torner, University of Iowa).

OAI

The cohort, clinical data and image acquisitions used in these analyses were fund as the Osteoarthritis Initiative by the National Institutes of Health (NIH) through a Foundation for NIH public private partnership with GlaxoSmithKline, Merck & Co., Inc., Novartis Pharmaceuticals Corporation, and Pfizer.

Rotterdam Study

The Rotterdam Study is funded by Erasmus University Medical Center and Erasmus University, Rotterdam, The Netherlands Organisation for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

SOF

The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576.

TASOAC

The TASOAC study was supported by the National Health and Medical Research Council of Australia, Tasmanian Community Fund, Masonic Centenary Medical Research Foundation, Royal Hobart Hospital Research Foundation and Arthritis Foundation of Australia.

Data availability

Data are available upon reasonable request. Data may be obtained from a third party and are not publicly available. We encourage the use of data by third parties, although this is subject to approval by the steering committees of the World COACH consortium and the participating cohorts, as well as to legal boundaries regarding data ownership. A standardized data request form is available for which will be reviewed uniformly in order to consistently handle World COACH data requests.

SUPPLEMENTARY MATERIAL 1: COHORT-SPECIFIC OVERVIEW OF BASELINE AND FOLLOW-UP RADIOGRAPHS

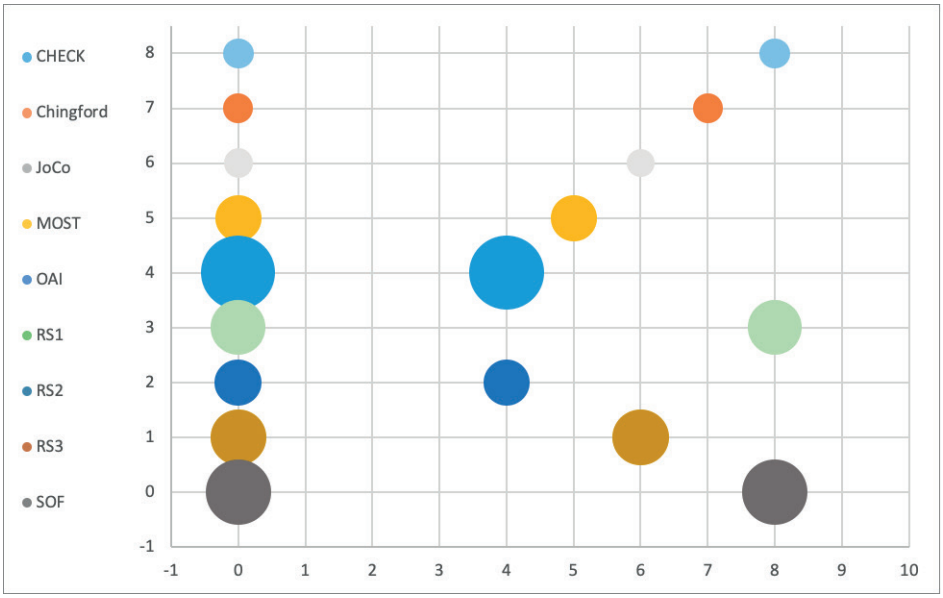


Figure 1. Radiographs per cohort at baseline and follow-up within 4-8 years. The size of the dot is proportionate to the number of included individuals at baseline and at each follow-up moment.

SUPPLEMENTARY MATERIAL 2: OVERLAP BETWEEN ACETABULAR DYSPLASIA MEASUREMENTS

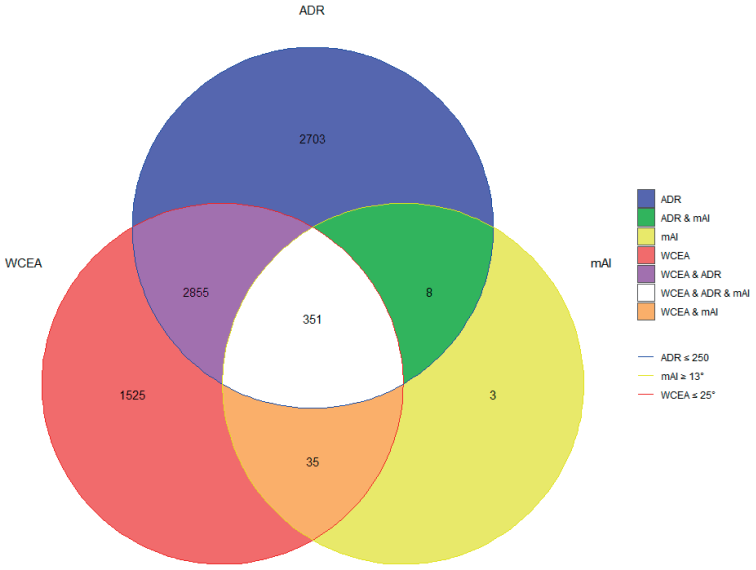


Figure 1. Venn diagram of all AD measures. ADR: acetabular depth-width ratio ≤ 250. mAI: modified acetabular index ≥ 13°. WCEA: Wiberg center edge angle ≤ 25°.

Table 1. Absolute and relative risk of hips with acetabular dysplasia to develop incident radiographic hip osteoarthritis in males.

Definition of AD	Total hips in group, n	Hips with AD, n (prevalence, %, 95% CI)	Hips with incident RHOA, n (prevalence, %, 95% CI)	Hips with AD and incident RHOA, n	Absolute Risk, % [c]	Relative Risk (95% CI) [d]
WCEA ≤ 25°	5631	1369 (24.3 (23.2-25.5))	77 (1.4 (1.1-1.7))	14	0.25	0.69 (0.39-1.23)
WCEA ≤ 20°	5631	300 (5.3 (4.8-5.9))	77 (1.4 (1.1-1.7))	4	0.07	0.97 (0.36-2.65)
ADR ≤ 250	5631	2239 (39.8 (38.5-41.1))	77 (1.4 (1.1-1.7))	35	0.62	1.26 (0.81-1.97)
mAI ≥ 13°	5631	105 (1.9 (1.5-2.3))	77 (1.4 (1.1-1.7))	1	0.02	0.69 (0.10-4.93)
WCEA ≤ 25° & ADR ≤ 250 & mAI ≥ 13° [a]	5631	100 (1.8 (1.4-2.2)	77 (1.4 (1.1-1.7))	1	0.02	0.73 (0.10-5.18)
WCEA ≤ 25° or ADR ≤ 250 or mAI ≥ 13° [b]	5631	2518 (44.7 (43.4-46.0))	77 (1.4 (1.1-1.7))	39	0.69	1.27 (0.81-1.98)

AD: acetabular dysplasia. WCEA: Wiberg center edge angle. ADR: acetabular depth-width ratio. mAI: modified acetabular index. CI: confidence interval. [a] The reference group contained hips free of AD and hips with only 1 or 2 measures of AD. [b] The reference group contained hips free of any measure to define AD. [c] The absolute risk was calculated using the following equation: (number of hips with AD and RHOA/Total number of hips in subgroup)*100%. [d] The relative risk was calculated using the following equation: (number of hips with AD & RHOA/ (number of hips with AD & RHOA + number of hips with AD only)) / (number of hips with RHOA without AD/ (number of hips with RHOA without AD + number of hips without AD and RHOA))

Table 2. Absolute and relative risk of hips with acetabular dysplasia to develop incident radiographic hip osteoarthritis in females.

Definition of AD	Total hips in group, n	Hips with AD, n (prevalence, %, 95% CI)	Hips with incident RHOA, n (prevalence, %, 95% CI)	Hips with AD and incident RHOA, n	Absolute Risk, % [c]	Relative Risk (95% CI) [d]
WCEA ≤ 25°	13,176	3397 (25.8 (25.0-26.5))	301 (2.3 (2.0-2.6))	113	0.86	1.73 (1.37-2.18)
WCEA ≤ 20°	13,176	864 (6.6 (6.1-7.0))	301 (2.3 (2.0-2.6))	30	0.23	1.58 (1.09-2.29)
ADR ≤ 250	13,176	3678 (27.9 (27.1-28.7))	301 (2.3 (2.0-2.6))	109	0.83	1.47 (1.16-1.85)
mAI ≥ 13°	13,176	292 (2.2 (2.0-2.5))	301 (2.3 (2.0-2.6))	13	0.10	1.99 (1.16-3.43)
WCEA ≤ 25° & ADR ≤ 250 & mAI ≥ 13° [a]	13,176	251 (1.9 (1.7-2.2)	301 (2.3 (2.0-2.6))	13	0.10	2.32 (1.35-3.99)
WCEA ≤25° or ADR ≤250 or mAI ≥ 13° [b]	13,176	4962 (37.7 (36.8-38.5))	301 (2.3 (2.0-2.6))	137	1.04	1.38 (1.11-1.73)

AD: acetabular dysplasia. WCEA: Wiberg center edge angle. ADR: acetabular depth-width ratio. mAI: modified acetabular index. CI: confidence interval. [a] The reference group contained hips free of AD and hips with only 1 or 2 measures of AD. [b] The reference group contained hips free of any measure to define AD. [c] The absolute risk was calculated using the following equation: (number of hips with AD and RHOA/Total number of hips in subgroup)*100% [d] The relative risk was calculated using the following equation: (number of hips with AD & RHOA/ (number of hips with AD & RHOA + number of hips with AD only)) / (number of hips with RHOA without AD/ (number of hips with RHOA without AD + number of hips without AD and RHOA))

Chapter 11

Severe pincer morphology is associated with incident hip osteoarthritis: prospective individual participant data meta-analysis from 18,935 hips from the World COACH consortium

F. Boel*, N.S. Riedstra*, M.M.A. van Buuren, H. Ahedi, V. Arbabi, N.K. Arden, S.M.A. Bierma-Zeinstra, F.M. Cicuttini, T.F. Cootes, K. Crossley, D.T. Felson, W.P. Gielis, J. Heerey, G. Jones, S. Kluzek, N.E. Lane, C. Lindner, J. Lynch, J. van Meurs, A. Mosler, A.E. Nelson, M.C. Nevitt, E.H. Oei, J. Runhaar, J. Tang, H. Weinans, R. Agricola.

*Shared first authorship

Under review.

ABSTRACT

Objective: To assess the relationship between pincer morphology and incident radiographic hip osteoarthritis (RHOA), and study specific subgroups.

Methods: Hips completely free of RHOA at baseline and with follow-up within 4-8 years were drawn from the World COACH consortium. The lateral center edge angle (LCEA) was calculated uniformly on all baseline radiographs. Moderate pincer morphology was defined as a LCEA≥40°, and severe pincer morphology as a LCEA≥45° in sensitivity analyses. The primary outcome was incident RHOA defined by a harmonized OA score. A logistic regression model with generalized mixed effects with 3 levels (within- cohort, -person and -hip side correlation) adjusted for age, biological sex, and BMI was employed. Descriptive statistics are reported for age, biological sex and BMI.

Results: 18,935 hips from 9 cohorts were included. 4,894 hips (25.8%) had moderate pincer morphology. Within 8 years (mean 6.0 ± 1.7 years), 352 hips (1.9%) developed RHOA. Moderate pincer morphology was not associated with RHOA (OR 1.15 (0.92-1.51), whereas severe pincer morphology was significantly associated (OR 1.50 95% CI 1.05-2.15). Moderate pincer morphology in groups aged 40-50 (RR 2.67, 95% CI 1.43-4.95) and BMI ≥25 (RR 1.23 95% CI 0.98-1.71) had a higher risk compared to non-pincer hips. Women (RR 1.20 95% CI 0.93-1.56) with pincer morphology may be more at risk than men (RR 0.95 95% CI 0.57-1.58)

Conclusion: The odds of developing RHOA within 8 years for hips with severe pincer morphology are 1.50 times higher than pincer-free hips, whereas moderate pincer morphology was not significantly associated with RHOA. Further research is necessary to uncover high risk subgroups of pincer morphology.

KEY MESSAGES

What is already known on this topic - Pincer morphology, characterized by acetabular overcoverage of the femoral head which leads to impingement, has been proposed as a risk factor for hip osteoarthritis. Evidence for the association between pincer morphology and hip osteoarthritis however, remains conflicting.

What this study adds - This study provides evidence that severe pincer morphology in hips free of radiographic hip osteoarthritis at baseline is significantly associated with incident radiographic hip osteoarthritis at 4-8 years follow-up.

How this study might affect research, practice or policy - Given the substantial impact of osteoarthritis on quality of life, elucidating the role of pincer morphology in disease risk may help clinicians to identify at-risk individuals and guide interventions targeting this modifiable risk factor.

INTRODUCTION

Osteoarthritis (OA) is a debilitating disease that significantly impacts quality of life[34]. It is therefore essential to identify risk factors for OA, which can potentially be targeted in prevention and treatment strategies[10,11,228]. Risk factors for hip OA that have been identified include age, biological sex, genetics, physical workload and hip shape[11,228-231].

Pincer morphology is a hip shape characterized by acetabular over coverage of the femoral head, and is associated with femoroacetabular impingement syndrome (FAIS), a motion related clinical disorder of the hip[45,84,203,232]. Pincer morphology may cause repeated abutment between the proximal femur and the acetabulum during repetitive and terminal motion of the hip[45]. It has been proposed that the repeated impingement leads to intra-articular damage (e.g., cartilage and labral pathology), and ultimately to hip OA[45,203].

Conflicting evidence has been reported on the association between pincer morphology and radiographic hip osteoarthritis (RHOA)[8,11,58,201,233]. A recent systematic review identified 9 prospective cohort studies but did not demonstrate an association between pincer morphology and RHOA, whereas cross-sectional studies showed that hips with OA were 3.7 times more likely to have a LCEA ≥ 40° [11]. However, substantial heterogeneity (I^2 60%) was observed between the results of the prospective studies, making it difficult to draw conclusions from this meta-analysis[11]. Furthermore, study populations and how pincer morphology is defined and measured varies significantly across studies, which may influence the reported associations.

Our aim is to perform an individual participant data (IPD) meta-analysis on the association between pincer morphology at baseline and the risk of developing RHOA within 4-8 years follow-up. Additionally, we will study this association in subgroups stratified by age, biological sex and BMI.

METHODS

Study design and participants

Participants were drawn from the Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) consortium. The World COACH consortium is a global collaboration of all available prospective cohort studies with prospective pelvic or hip imaging. The consortium profile has previously been published in detail elsewhere[48]. In this study we included all cohorts with a follow-up anteroposterior (AP) pelvic radiograph within

4-8 years of a baseline radiograph, and therefore included 9 cohorts (Cohort Hip and Cohort Knee (CHECK), Multi-center Osteoarthritis Study (MOST), Osteo Arthritis Initiative (OAI), Rotterdam Study-I (RS-I), Rotterdam Study-II (RS-II), Rotterdam Study-III (RS-III), the Chingford Study, The Johnston County Project (JoCo) and the Study of Osteoporotic Fractures (SOF)), and excluded two cohorts (Tasmanian Older Adults Cohort (TASOAC), Femoroacetabular impingement and hip osteoarthritis cohort (FORCe)).

All included hips needed to have known BMI, biological sex, and age at baseline. Next, hips without an original baseline RHOA score at baseline were excluded. All radiographs of insufficient quality for automated pincer morphology measurements and all AP hip radiographs were excluded as they did not allow for constructing a horizontal reference line to adjust for pelvic obliquity. Next, we excluded all hips lacking an original RHOA score at follow-up and excluded all baseline hips with AD as determined by a Wiberg center edge angle (WCEA) ≤ 25°. We chose to do this in order to compare the pincer hips to a clean reference group of hips with normal acetabular coverage. Furthermore, multiple studies have demonstrated a significant association between AD and RHOA[8,10,12]. Finally, we included only hips free of any signs of RHOA at baseline (any OA score=0). We chose to focus on a population of hips completely free of RHOA to identify the true predictors of this disease. This led to a total inclusion of 18,935 hips.

Radiographs

AP pelvic radiographs were obtained by cohorts at baseline and at follow-up between 4-8 years (Supplementary material 1). All radiographs were obtained based on a cohort-specific predetermined protocol established by each cohort. Detailed information about specific radiographic protocols, was previously published[48]. Five cohorts (CHECK, OAI, RS-I, RS-II, RS-III) had weight-bearing AP pelvic radiographs, one cohort (MOST) had weight-bearing full-limb radiographs, and three cohorts (the Chingford Study, JoCo, and SOF) had supine AP pelvic radiographs.

Radiographic measurements

Lateral Center Edge Angle

To avoid measurement variability across cohorts, uniform pincer morphology measurements were performed on all baseline radiographs. The bony outline of the proximal femur and acetabulum were annotated on the AP pelvic radiographs with a point set using the BoneFinder® software (www.bone-finder.com; The University of Manchester, UK)[64]. This point set was used to perform automated radiographic measurements using a previously published Python script, which was adjusted and validated on World COACH data[91,117].

The LCEA quantifies bony coverage of the femoral head by the acetabulum (Figure 1)[151]. Moderate pincer morphology was defined as a LCEA ≥ 40°. Sensitivity analyses with an LCEA threshold of ≥ 45° to define severe pincer morphology were performed to determine whether increased acetabular overcoverage influences the risk of developing RHOA.

Radiographic Hip Osteoarthritis Grading

Original OA scores per cohort were harmonized into "free of RHOA" (any score 0), "doubtful RHOA" (any score 1), or "definite RHOA" (KL ≥2, Modified Croft ≥2, Modified OA=2, or total hip replacement (THR)[183].

Outcome measurements

The primary outcome was "definite RHOA" defined by the harmonized RHOA score (OA score = 2) within 4-8 years follow-up from baseline. Additionally, RHOA was defined as an ordinal outcome "free of RHOA", "doubtful RHOA" and "definite RHOA" in secondary analyses.

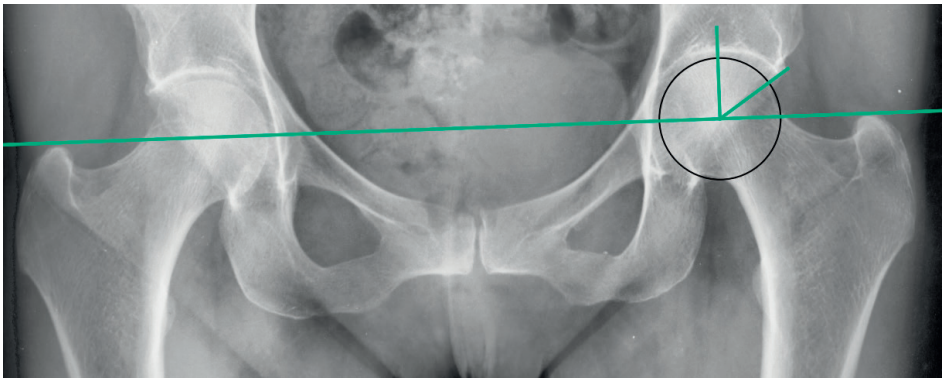


Figure 1. The lateral center edge angle (LCEA) is measured on an AP pelvic radiograph. The LCEA was constructed according to the following steps. A horizontal reference line is constructed to correct for pelvic tilt in the radiograph, and is based on the average of 4 lines, between 1) both femoral head centers, 2) the most cranial points of the foramen obturator, 3) the most caudal point of the ischial tuberosity and 4) the most caudal point of the pelvic teardrop. To determine the center of the femoral head, a best fitting circle is drawn around the femoral head based on the SSM points. The LCEA is then formed by two lines drawn from the center of the best fitting circle. The first line is drawn vertically through the center of the femoral head, perpendicular to the horizontal reference line. The second line is drawn from the center of the best fitting circle to the most lateral bony point of the acetabulum.

Statistical Analysis

All statistical analyses were performed in R version 4.1.1. Univariate differences in baseline characteristics between complete included and excluded cases were inspected, meaning the included hips were compared to the hips that were excluded because of an OA score of 1 or 2 at baseline (Figure 2). The association between baseline moder-

ate pincer morphology defined by LCEA ≥ 40° and incident RHOA was estimated using a one-stage logistic regression model with generalized mixed effects with 3 levels: hip side (left/right), individual and cohort. We corrected for the cohort in this multi-level model in order to adjust for possible residual confounding by study differences. The model accounted for the difference between open (Chingford, JoCo, RS-I, RS145 II, RS-III), and closed population cohorts (CHECK, OAI, MOST, SOF). The inclusion criteria for various population types vary notably, with a key distinction centered on enrollment characteristics. The results are expressed as adjusted (aOR) and unadjusted odds ratios (OR) with 95% confidence intervals (95% CI) and were adjusted for baseline age, modeled using splines with three degrees of freedom to account for non-linearity, biological sex, and BMI. A sensitivity analysis was performed using LCEA ≥ 45° to define severe pincer morphology. In the sensitivity analysis hips with a 40° ≤ LCEA < 45° were excluded from the reference group in order to compare pincer hips to a clean population of hips free of pincer morphology by any definition. The statistical significance threshold was set at $p<0.05$. Additionally, a continuation ratio model with ordinal outcome RHOA classified as "free of RHOA", "doubtful RHOA" and "definite RHOA" was created to assess the influence of doubtful RHOA as reference group. Random effects were added to adjust for clustering of cohorts and individual, and the model was adjusted for baseline age, sex, and BMI. Moderate pincer morphology was defined as LCEA ≥ 40°. The model was built in a forward fashion and a relaxed ordinality assumption for pincer morphology, allowing the effect of pincer morphology to be different for each level of the outcome RHOA within 4-8 years. The results were presented as an effect plot of the marginal probabilities with reference to the random effects for females, with mean baseline age and BMI and randomly selected left hip side. Because of limited outcomes, it was not possible to perform subgroup analyses using the same logistic regression model. We reported absolute risk (AR) and relative risk (RR) in moderate pincer morphology and non-pincer hips to develop RHOA stratified by age (40-50, 51-60, 61-70 and >70 years of age), by BMI (BMI > 25 and BMI ≤ 25), and by biological sex. The AR 95% CI was calculated based on the observed absolute risk (AR): AR ± 1.96*sqrt((AR(1-AR))/total number of individuals. The RR with corresponding 95% CI was determined by unconditional maximum likelihood estimation. Logistic regression was performed using the lme4-package[214]. The continuation ratio model was created using the GLMMadaptive package[215]. The effect plot was created using the ggplot2-package[74].

Equity, diversity, and inclusion statement

The current study includes participants with a variety of ethnic backgrounds, more women than men and individuals from 3 continents. Details are reported in the World COACH description paper[48]. The author team was gender balanced and includes both junior and senior researchers with a variety of academic backgrounds who were actively involved in the writing process. Our study includes individuals from marginalized communities.

RESULTS

Participants

The flow of World COACH hips to the current final study population is depicted (Figure 2). 18,935 hips were included for analysis. The average time between the baseline and follow-up radiograph across all cohorts is 6.0 ± 1.7 years. Baseline demographic data stratified per cohort are presented in Table 1. The excluded hips were on average slightly older (65.68 years vs 62.66 years at baseline) and had a higher prevalence of pincer morphology as defined by moderate and severe thresholds.

Pincer morphology

A total of 4,894 (25.8%) hips had moderate pincer morphology defined by LCEA ≥ 40° and 1,121 (5.9%) hips had severe pincer morphology defined by a threshold LCEA ≥ 45°. In females, 3,542 (26.6%) hips had moderate pincer morphology and 810 (6.1%) had severe pincer morphology. In males, 1,352 (24.1%) hips had moderate pincer morphology and 311 (5.5%) severe pincer morphology.

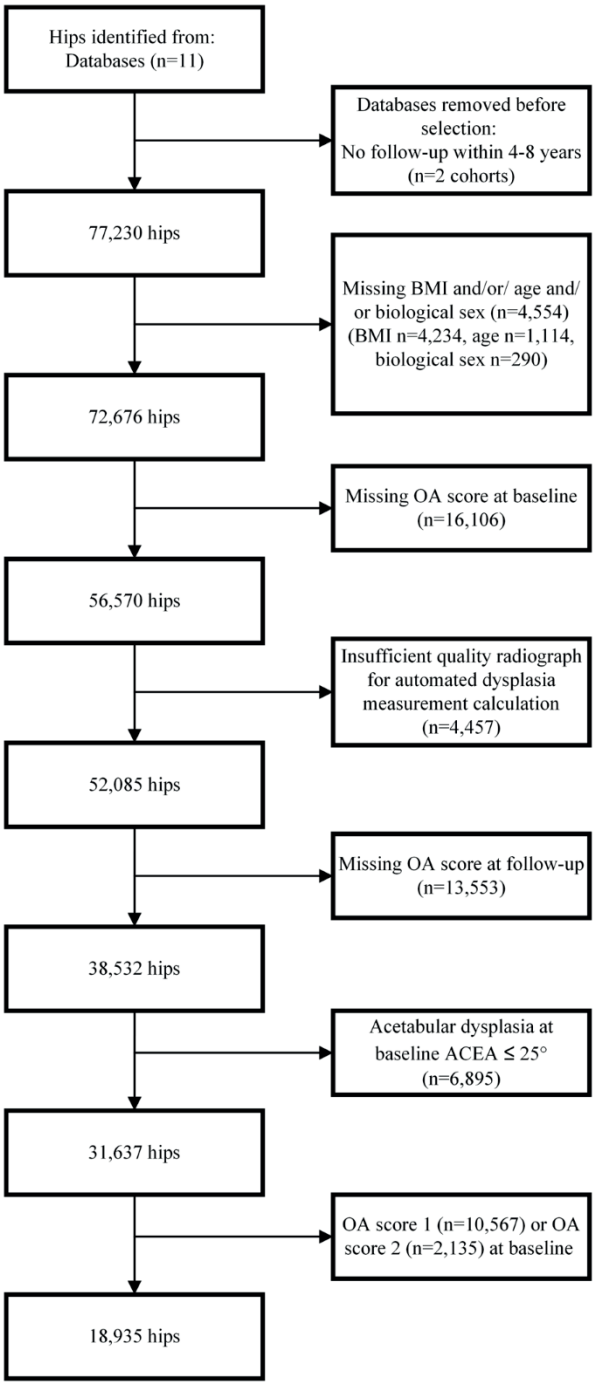


Figure 2. Flow of hips from consortium inclusion to final study population.

Table 1. Baseline characteristic of included hips, stratified per cohort

	CHECK	Ching	JoCo	MOST	OAI	RS-I	RS-II	RS-III	SOF	Total incl.	Total excl.*
Hips, n (%)	678	815	633	1,202	4,481	2,719	1,797	2,381	4,229	18,935	12,702
Age, mean (sd) years	55.64 (5.28)	53.62 (5.73)	59.25 (8.72)	61.09 (7.24)	60.66 (8.92)	65.21 (6.33)	63.00 (6.40)	56.34 (4.85)	70.38 (4.36)	62.66 (8.36)	65.68 (8.32)
BMI, mean (sd) kg/m^2	26.29 (4.23)	25.38 (3.97)	29.60 (6.02)	29.93 (5.13)	28.27 (4.64)	26.15 (3.45)	27.11 (3.83)	27.56 (4.17)	26.38 (4.31)	27.30 (4.49)	27.51 (4.75)
Men, n (%)	114 (16.8)	0 (0.0)	282 (44.5)	378 (31.4)	1,860 (41.5)	1,120 (41.2)	795 (44.2)	1,065 (44.7)	0 (0)	5,614 (29.6)	3,444 (27.1)
LCEA ≥ 40°, n (%)	143 (21.1)	261 (32.0)	149 (23.5)	221 (18.4)	1,014 (22.6)	887 (32.6)	440 (24.5)	522 (21.9)	1,257 (29.7)	4,894 (25.8)	5,099 (40.1)
LCEA ≥45°, n (%)	33 (4.9)	73 (9.0)	35 (5.5)	55 (4.6)	204 (4.6)	230 (8.5)	93 (5.2)	108 (4.5)	290 (6.9)	1,121 (5.9)	1,815 (14.3)
OAscore=2 follow-up, male/female (%/%)	19/ 63 (16.8/11.2)	0/72 (0.0/8.8)	18/36 (6.4/10.3)	3/4 (0.8/0.5)	3/10 (0.2/0.4)	1/11 (0.1/0.7)	2/4 (0.3/0.4)	36/17 (3.4/1.3)	0/53 (0.0/1.3)	82/270 (1.5/2.0)	3,025

CHECK= Cohort Hip and Cohort Knee, MOST= Multi-center Osteoarthritis Study, OAI= Osteo Arthritis Initiative, RS-I= Rotterdam Study-I, RS-II=Rotterdam Study-II, RS-III= Rotterdam Study-III (RS-III), Ching=the Chingford Study, JoCo=The Johnston County Project, SOF= Study of Osteoporotic Fractures, BMI= body mass index, LCEA= Lateral Center Edge Angle. OA score: 0= no RHOA, 1= Doubtful RHOA, 2= Definite RHOA. *Excluded hips are defined as all eligible hips for analysis but with OA score 1 or 2 at baseline

Incident radiographic hip osteoarthritis

Definite RHOA had developed in 352 hips (1.9 %) within 8 years follow-up. The distribution of RHOA incidence per cohort is 82 hips (12.1%) in CHECK, 72 hips (8.8%) in Chingford, 54 hips (8.5%) in JoCo, 7 hips (0.6%) in MOST, 13 hips (0.3%) in OAI, 12 hips (0.5%) in RS-I, 6 hips (0.4%) in RS-II, 53 hips (2.2%) in RS-III and 53 (1.9%) in SOF.

Association between pincer morphology and radiographic hip osteoarthritis

The association between moderate pincer morphology and incident RHOA within 8 years was 1.15 (0.92-1.43), p-value of 0.22. The association between severe pincer morphology and incident RHOA was 1.50 (1.05-2.14), with a p-value of 0.026. The associations between pincer 234 morphology and incident RHOA are summarized in Table 2.

Table 2. Associations between LCEA measures using two cut points to define pincer morphology and RHOA.

Definition pincer morphology	Hips with pincer morphology (%)	Hips with incident RHOA at follow-up (%)	Absolute risk (%)	Unadjusted OR (95% CI) *	p-value	Adjusted OR (95% CI)	p-value
Moderate (LCEA ≥ 40°)	4,894	101	101/4,894 (2.1)	1.19 (0.91-1.57)	0.21	1.15 (0.92-1.43)	0.22
Severe (LCEA ≥ 45°)	1,121	31	31/1,121 (2.8)	1.57 (1.10-2.24)	0.013	1.50 (1.05-2.14)	0.026

RHOA= radiographic hip osteoarthritis, OR= odds ratio, CI= confidence interval, LCEA= lateral center edge angle.

The marginal probability for hips with moderate pincer morphology (LCEA ≥ 40°) to develop doubtful RHOA within 4-8 years is 0.20 (95% CI 0.14-0.28), compared to 0.17 (95% CI 0.11-0.24) for hips free of pincer morphology. The marginal probability for moderate pincer hips (LCEA ≥ 40°) to develop definite RHOA within 4-8 years is 0.03 (95% CI 0.01-0.06), compared to 0.02 (95% CI 0.01-0.06) for pincer-free hips. The effect plot of the marginal probabilities from the continuation ratio model with ordinal outcome RHOA is shown in supplementary material 2.

Table 3. Absolute and relative risk of hips with pincer morphology (LCEA ≥ 40°) to develop incident radiographic hip osteoarthritis stratified by age group, BMI, and biological sex.

Strata	Total hips in group, n	Hips with pincer, n	Hips with incident RHOA, n	Hips with pincer, and incident RHOA, n	Absolute Risk, % (95% CI) *	Relative Risk, % (95% CI) **	p-value
Age (years)							
40-50	1,534	307 (20.0)	40 (2.6)	16	1.0 (0.5 – 1.6)	2.67 (1.43-4.95)	0.004
51-60	6,180	1,363 (22.1)	149 (2.4)	47	0.8 (0.5- 1.0)	1.63 (1.16-2.29)	0.007
61-70	7,557	2,056 (27.2)	110 (1.5)	26	0.3 (0.2- 0.5)	0.93 (0.54-1.28)	0.45
70+	3,664	1,168 (31.9)	53 (1.4)	12	0.3 (0.1- 0.5)	0.62 (0.33 to 1.17)	0.18
BMI							
< 25	6,094	1,600 (26.3)	125 (2.1)	31	0.5 (0.3- 0.7)	0.79 (0.62-1.38)	0.76
≥ 25	12,841	3,294 (25.7)	227 (1.8)	70	0.5 (0.4- 0.7)	1.23 (0.98-1.71)	0.078
Biological sex							
Male	5,614	1352 (24.1)	82 (1.5)	19	0.3 (0.2- 0.5)	0.95 (0.57-1.58)	1
Female	13,321	3,542 (26.6)	270 (2.0)	82	0.6 (0.5- 0.7)	1.20 (0.93-1.56)	0.16

LCEA= lateral center edge angle, RHOA= radiographic hip osteoarthritis, CI= confidence interval, BMI= body mass index. *The absolute risk was calculated using the following equation: (number of hips with pincer morphology and RHOA/Total number of hips in subgroup) **The relative risk was calculated using the following equation: (number of hips with pincer & RHOA/(number of hips with pincer & RHOA + number of hips with pincer only)) / (number of hips with RHOA without pincer morphology/ (number of hips with RHOA without pincer morphology + number of hips without pincer morphology or RHOA)).

Sensitivity analysis excluding the MOST cohort.

The study population excluding the MOST cohort comprised a total of 17,733 hips. Of all hips in the study population, only 7 hips develop RHOA within 8 years in the MOST cohort. No hips with pincer morphology develop RHOA. The non-significant association between hips with moderate pincer morphology (LCEA ≥ 40°) and incident RHOA was 1.19 (95% CI 0.90-1.56) in the remaining study population (n=17,733) when hips from the MOST cohort were excluded.

Subgroup analyses

Descriptive statistics stratified by age group, biological sex, and BMI are summarized in Table 3. The RR for moderate pincer hips to develop RHOA was highest in age group 40-50 (RR 2.67 (95% CI 1.43-4.95), p-value 0.004), in hips with BMI ≥25 (RR 1.23 (95% CI 0.98-1.71), p-value 0.078), and in female hips (RR 1.20 (95%CI 0.93-1.56), p-value 0.16).

DISCUSSION

This first IPD meta-analysis in a large prospective consortium of 18,935 hips completely free of RHOA at baseline, did not find significant association between moderate pincer morphology defined by LCEA ≥ 40° and incident RHOA within 8 years. However, severe pincer morphology (LCEA ≥ 45°) was significantly associated with RHOA. Hips with moderate pincer morphology may also be more likely to progress to doubtful RHOA within this follow-up compared to non-pincer hips, although no conclusions on clinical significance can be drawn. Subgroup statistics point in the direction that hips with moderate pincer morphology in younger individuals (aged 40-50) and with higher baseline BMI (≥ 25) are more at risk of developing RHOA compared to non-pincer hips. Additionally, hips in females with moderate pincer morphology were slightly more at risk to develop RHOA within 8 years compared to hips in males.

Several previous studies have been unable to establish an association between moderate pincer morphology and RHOA[11]. For instance, a large prospective study of over 4,000 hips with 9.2 years of follow-up from the Rotterdam Study found no significant association between pincer morphology, defined as an LCEA ≥ 40°, and RHOA[8] . Similarly, in the CHECK cohort, which included 1,002 hips with 10 years of follow-up, no overall association was observed; however, the presence of hip pain at baseline did appear to modify this relationship, as acetabular overcoverage increased the risk of developing RHOA in such cases[46]. In the primary analyses, we also did not find a significant association between moderate pincer morphology and RHOA. Nevertheless, prior results from the Rotterdam Study demonstrated that pincer morphology increased the risk of developing RHOA specifically in hips that were completely free of RHOA at baseline (KL grade 0)[8]. Furthermore, a cross-sectional study by Faber et al. reported that pincer morphology was associated with an increased risk of JSN, providing additional evidence that pincer morphology may pose a risk for developing RHOA[58]. Interestingly, in our secondary analyses, severe pincer morphology (LCEA ≥ 45°) was significantly associated with the development of RHOA, further emphasizing the complexity of this relationship.

In the present study, the average BMI was 27.4 kg/m^2. Conflicting evidence excists on the relationship between increased BMI and hip osteoarthritis, although a systematic review suggested that the risk of hip OA increases with BMI and a dose–response relationship exists[234]. In our study, in a subgroup op people with a baseline BMI ≥25kg/m^2, descriptive statistics indicated that pincer morphology hips, - compared to non-pincer hips, had a higher RR (1.23 305 vs 0.79) of developing hip OA. Importantly however, their confidence intervals overlap.

Our study population consists mostly of female hips (70%), but the incidence of pincer morphology was similar in female and male hips (26.6% and 24.1% respectively). This is in line with previous findings[30]. Research shows that women have greater pelvic obliquity and less vertical center of mass displacement compared to men, which may influence biomechanics of the hip joint, and could potentially lead to a higher RHOA risk in female hips with pincer morphology[235]. Unfortunately, it was not possible to perform regression analyses in subgroups by biological sex in the present study, as only 19 male hips with pincer developed RHOA.

It is possible that the definition of pincer morphology has a direct impact on its association with RHOA[30]. This is illustrated by the significant association between severe pincer morphology defined by LCEA ≥ 45° and incident RHOA, which was not present in the current population when moderate pincer morphology was defined by LCEA ≥ 40°. Most studies have relied on a LCEA ≥ 40° to define pincer morphology, but based on results from the present study with almost 19,000 hips, we argue that this threshold may be too low to be clinically relevant. A recent study of 6,807 individuals from the UK Biobank found a prevalence in the general population of pincer morphology defined by a LCEA ≥ 45°, of 8.1% in females and 8.9% in males[58]. This is similar to the prevalence in this study (LCEA ≥ 45° 6.1% in female hips and 5.5% in male hips). In the excluded hips from the present study, a prevalence of 14.3% of hips with LCEA ≥ 45° was found. These hips were only excluded from analysis because they were not free of RHOA at baseline. It may be that these hips had already developed RHOA as a result of acetabular overcoverage. This is further supported by the increased relative risk in the younger subgroups as pincer morphology may be a considerable risk factor for more rapid development of RHOA. On the other hand, the LCEA might also be influenced by the presence of RHOA, for example due to (subtle) acetabular osteophytes which potentially causes falsely positive classification of pincer morphology which is why we excluded these hips. Subsequent studies should aim to conduct sensitivity analyses employing this threshold, which may elucidate a more clinically relevant study population in the search for modifiable risk factors for RHOA.

Importantly, the definition of pincer morphology as a static concept (defined only by radiological excessive acetabular coverage) differs from the dynamic concept of FAI syndrome with pincer morphology. The definition of FAI syndrome as stated by the 2016 Warwick Agreement, does not only pertain to radiological findings, but to a triad of radiological signs, clinical signs (hip impingement tests, limited range of motion) and symptoms (motion or position related pain in the hip or groin)[84]. Clinical signs and symptoms have not yet been harmonized in the World COACH consortium but including those might enhance the predictive ability and identify hips with pincer morphology which are at higher risk of developing RHOA as this has been shown previously for patients with FAI syndrome with cam morphology in the CHECK cohort.

This study has several strengths. The first is the inclusion of hips completely free of any signs of RHOA at baseline, which differs from some previous prospective studies[8,10,11]. This allowed us to study associations that were unbiased by pre-existing doubtful RHOA. Though previous prospective studies generally correct for baseline RHOA grade in statistical models, we believe risk factors are best demonstrated when RHOA-free hips are followed until a subset develops disease. Furthermore, LCEA measurements may be affected by the presence of osteophytes as it is possible that spurious osteophytes are mistaken for pincer hips. We were able to rule out the presence of osteophytes at baseline as all included hips were completely free of RHOA. Second, IPD meta-analysis study design is a significant strength. By collecting, pooling and analyzing original cohort data, we achieved increased statistical power allowing for subgroup and sensitivity analyses. Our results confirm a robust estimate of the risk pincer morphology poses to RHOA-free hips within 8 years. This could inform the clinical approach to patients with severe pincer morphology, including future treatment and preventative strategies for hip OA. Finally, we used uniform automated measurements. Using a validated algorithm to quantify acetabular coverage of all hips on baseline radiographs reduces variability and bias in predictor measure

This study is subject to several limitations. First, it has been suggested that pincer morphology potentially only leads to RHOA when mixed with other shape features, or specific subtypes of pincer morphology which were not captured by the LCEA only[96]. The LCEA however, is presently the most commonly used and reliable measurement of pincer morphology[236]. Furthermore, a recent study compared radiographs to computed tomography (CT) scans and found similar sensitivity and specificity in defining pincer morphology when comparing radiographs to CT scans[237]. A second limitation is that we included both supine and weight bearing radiographs, which may influence RHOA grading. However, a study comparing the joint space width (JSW) on weight bearing and supine radiographs found that how the radiograph was obtained does not significantly

impact JSW measurements[226]. In the present study we adjusted for age, biological sex and BMI as these are important in hip OA research and were harmonized variables in the World COACH consortium[34]. Importantly, unmeasured confounders that were not incorporated may be important in the association of pincer morphology and hip OA and should be studied in future work. Finally, we only studied RHOA, which may differ from clinically relevant hip OA where symptoms are taken into account. Elucidating the association between pincer morphology and a clinical definition of hip OA should be prioritized in future research.

Modifiable risk factors are essential for preventing hip osteoarthritis in the general population as well as athletes. Our study shows that severe pincer morphology only explains a small subset of individuals at risk for hip OA. However, we argue that severe pincer morphology is a potentially modifiable risk factor for hip OA for at least three reasons. First, physical therapy might increase strength and stability of the joint. Second, activity modification might help avoid excessive joint-loading. Third, surgical interventions might help improve the joint shape and could potentially aid in preventing osteoarthritis, although this is presently unknown. Prevention of hip osteoarthritis can improve overall quality of life and aid in relieving the substantial and increasing societal burden of this disease[34].

To the best of our knowledge, our IPD meta-analysis is the first study of its kind to investigate the relationship between pincer morphology and the risk of developing RHOA. Severe pincer morphology defined by a LCEA ≥ 45° is significantly associated with incident RHOA in a population of RHOA-free hips at baseline. This study offers new insight into a potentially modifiable risk factor for RHOA in specific subgroups, which contributes to discovering targets for prevention and treatment of hip osteoarthritis in the future.

Clinical implications

Pincer morphology, characterized by acetabular over-coverage, is not currently identified as a risk factor for radiographic hip osteoarthritis (RHOA) in existing literature, potentially due to variability in measurement methods, thresholds, and study populations, as well as reader variability in manual measurements. However, this study provides robust evidence that severe pincer morphology, defined by a LCEA of 45° or more, is significantly associated with incident RHOA, whereas a LCEA of 40° is not. By including only hips free of RHOA at baseline and accounting for variations in defining the outcome, this analysis avoids inconsistencies seen in prior research. Since osteoarthritis significantly impacts quality of life, understanding how pincer morphology increases the risk of disease could help identify high-risk individuals and inform strategies to mitigate this modifiable risk factor.

Acknowledgments

We would like to thank all participants of each of the cohort studies that are involved in World COACH. We gratefully acknowledge all international organisations that collaborated with the cohort studies in World COACH, as well as the OARSI for endorsing the World COACH consortium.

CHECK

The CHECK study was initiated by the Dutch Arthritis Society 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.

Chingford

We would like to thank all the participants of the Chingford Women Study, Professor Nigel Arden, Professor Tim Spector, Dr Deborah Hart, Mr Gem Lawson, Maxine Daniels and Alison Turner for their time and dedication and Arthritis Research UK for their funding support to the study and the Oxford NIHR Musculoskeletal Biomedical Research Unit for funding contributions.

JoCo-OA

Support for data from the Johnston County Osteoarthritis Project was provided in part by: the Center for Disease Control and Prevention (CDC) U01DP006266 and U01DP003206; Association of Schools of Public Health/ CDC S043, S1734, S3486; and National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases P60AR30701, P60AR049465, P60AR064166, and P30AR072580.

MOST

The MOST study was funded by the National Institutes of Health – National Institute on Aging grants AG19069 (Michael Nevitt, University of California, San Francisco) AG18820 (David Felson, Boston University) AG18947 (Cora Lewis, University of Alabama at Birmingham) and AG18832 (James Torner, University of Iowa).

OAI

The cohort, clinical data and image acquisitions used in these analyses were fund as the Osteoarthritis Initiative by the National Institutes of Health (NIH) through a Foundation

for NIH public private partnership with GlaxoSmithKline, Merck & Co., Inc., Novartis Pharmaceuticals Corporation, and Pfizer.

Rotterdam Study
The Rotterdam Study is funded by Erasmus University Medical Center and Erasmus University, Rotterdam, The Netherlands Organisation for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

SOF
The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576.

TASOAC
The TASOAC study was supported by the National Health and Medical Research Council of Australia, Tasmanian Community Fund, Masonic Centenary Medical Research Foundation, Royal Hobart Hospital Research Foundation and Arthritis Foundation of Australia.

Funding
The World COACH consortium has been funded through grants by the Dutch Arthritis Society (grant nr 21-1-205), the Dutch Research Council (NOW, Veni: 09150161910071), and Erasmus MC University Medical Center Rotterdam. CL is funded by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (223267/Z/21/Z). This research was funded in whole, or in part, by the Wellcome Trust [Grant number 223267/Z/21/Z]. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Patient and public involvement
Continuous patient engagement is a fundamental aspect of the World COACH consortium. Together, patients actively participate in shaping and prioritizing research inquiries within the consortium. World COACH researchers actively attend annual conferences, such as Artrose Gezond in the Netherlands, fostering open dialogues with osteoarthritis (OA) patients to collaboratively define research objectives. Patients are informed about

the consortium's capacity to identify risk factors and explore treatment options, and a platform at www.worldcoachconsortium.com encourages both patients and the public to contribute their ideas and questions for future research.

Data sharing statement

Data are available upon reasonable request. Data may be obtained from a third party and are not publicly available. We encourage the use of data by third parties, although this is subject to approval by the steering committees of the World COACH consortium and the participating cohorts, as well as to legal boundaries regarding data ownership. A standardized data request form is available for which will be reviewed uniformly in order to consistently handle World COACH data requests.

Competing interests

GJ reports personal fees from Novartis outside the submitted work. SBZ reports consulting fees from Pfizer Infirst Healthcare and personal fees for being a Deputy Editor for Osteoarthritis and Cartilage outside the submitted work. CL and TC report a patent for an image processing apparatus and method for fitting a deformable shape model to an image using random forest regression voting. CL reports licensing royalties for this patent from Optasia Medical outside the submitted work. AN is an associate editor for Osteoarthritis and Cartilage and is on the OARSI Board of Directors outside the submitted work. AM is on the Editorial Board for the British Journal of Sports Medicine and the Journal of Science and Medicine in Sport outside the submitted work. HW reports being a minority shareholder of Uplanner BV and Replasia BV outside the submitted work.

Patient and public involvement

Patients and public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Patient and public involvement section for further details.

Ethical approval

This study involves human participants but was excepted from ethical approval (Erasmus MC Medical Ethics Review Committee) as it uses previously collected observational data for which the participants had originally given informed consent, and all cohort studies included in this consortium already had ethics approval from their respective committees. Participants gave informed consent to participate in the study before taking part.

Contributorship

NSR, FB and RA initiated the study. NSR, FB and RA worked on the conceptual design of the study. MMAvB and RA identified eligible cohorts and contacted cohort investigators for collaboration. MMAvB, RA, NSR, FB, HA, AM, KC, JH, SK, JAL, JVM, ABM, AEN, MN, JT and HW collected the existing cohort data. MMAvB, NSR, FB, JT and RA have worked on the database and on the harmonisation process. NSR, FB, and RA have worked on statistical analyses. NSR and FB wrote the manuscript under supervision of RA. All authors critically reviewed and revised the manuscript and contributed to interpretation of the data. All authors read and approved the final version of the manuscript. NR acts as guarantor and accepts full responsibility for the finished work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

SUPPLEMENTARY MATERIAL 1: COHORT-SPECIFIC OVERVIEW OF BASELINE AND FOLLOW-UP RADIOGRAPHS

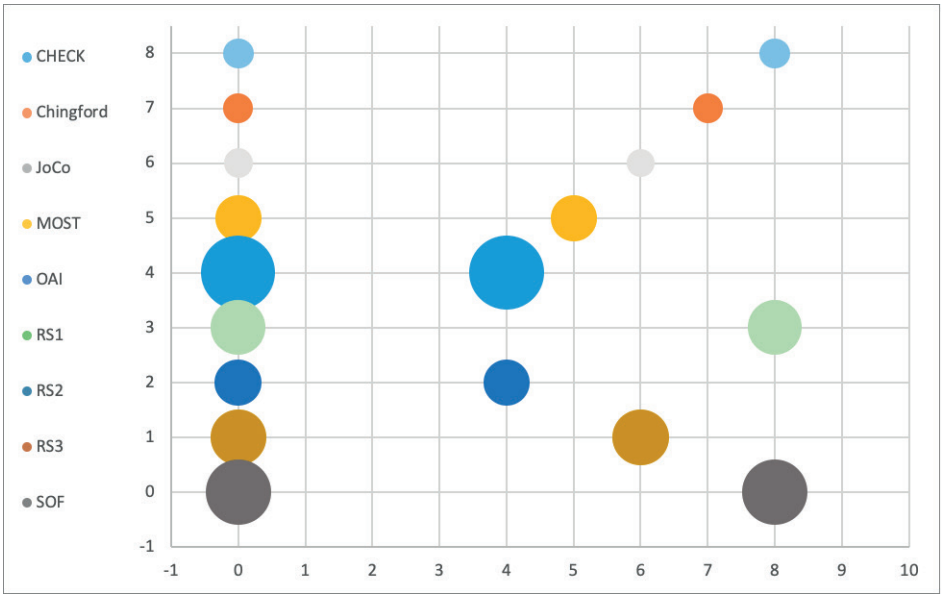


Figure 1. Radiographs per cohort at baseline and follow-up within 4-8 years. The size of the dot is proportionate to the number of included individuals at baseline and at each follow-up moment.

SUPPLEMENTARY MATERIAL 2: EFFECT PLOT OF THE MARGINAL PROBABILITIES OF RHOA WITHIN 4-8 YEARS

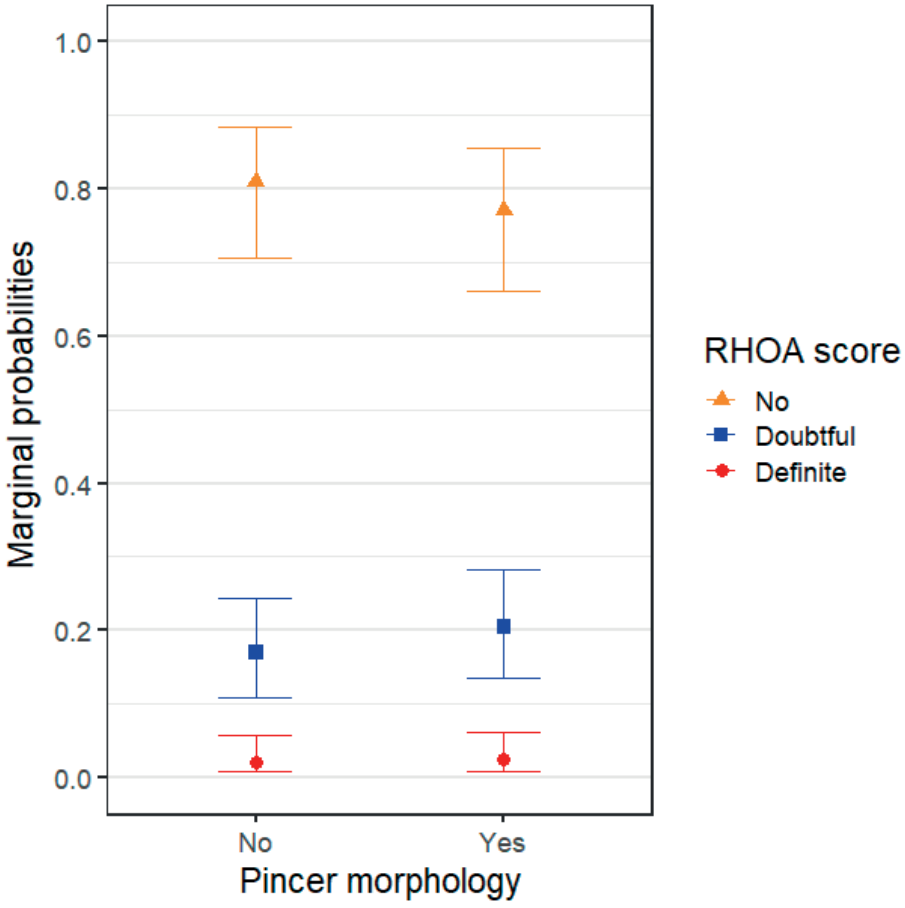


Figure 1. Effect plot of the marginal probabilities with 95% confidence intervals of RHOA within 4-8 years for females aged 63 years and BMI of 27 kg/m^2 in hips with pincer morphology (LCEA ≥ 40°) or without pincer morphology. The probabilities were marginalized over the random effects (cohort and individual), and adjusted for baseline age, BMI, biological sex, and hips side.

Chapter 12

General discussion

This thesis investigated acetabular dysplasia and pincer morphology to gain insight into their prevalence, development and contribution to the development of hip osteoarthritis. The first part described the development and validation of an open-access automated method for quantifying hip morphology, which can be used in children and adults and on DXA images and radiographs. In the second part, the prevalence and determinants of acetabular dysplasia and pincer morphology among children in the general population were studied. The third part investigated the role of acetabular dysplasia and pincer morphology in the development of hip osteoarthritis. The main findings and limitations of each study were described in the previous chapters. This chapter will discuss the broader implications of these findings, address remaining challenges, and offers future perspectives.

QUANTIFYING HIP MORPHOLOGY

Hip morphology is often quantified on anteroposterior (AP) radiographs using radiographic measurements. However, the definitions and methods used to quantify hip morphology in research are inconsistent and often poorly described. Additionally, morphology measurements are performed by hand or using various types of (semi-) automated software. These inconsistencies make it challenging to compare between different observers and studies[53-55]. Automated, validated and open-access methods for determining radiographic morphology measurements could aid in resolving these challenges. It would also allow for fast determination of multiple measurements, making it feasible to perform large population studies. While some automated methods are used, they are not open-source and can lack transparency on how the measurements are obtained. Chapters 2 and 3 described the development and validation of such an automated method and found the method to be comparable or better than manual measurements. Additionally, chapter 4 showed DXA images to be a reliable alternative to pelvic radiographs for quantifying hip morphology.

While automated methods offer significant advantages, it is crucial to acknowledge their current limitations. Firstly, as part of the automated method, landmarks are placed on the contour of the acetabulum and proximal femur using an automatic system based on Random Forest Regression Voting in the Constrained Local Model framework[64]. This automatic system is trained for specific populations and may not be generalizable to all hips. For instance, the automatic system used in chapters 2, 6, 7, and 8 was built specifically for this age group and, therefore, will not perform well on hips at a different stage of skeletal maturity. Secondly, if the automatic system is applied to hips with a hip shape not included in the training set, the landmarks will probably not be placed correctly.

Similarly, if there are artifacts in the image, this might result in the failure of the automatic landmark placement. Correct landmark placement is essential since the resulting radiographic morphology measurements are only as good as the landmark placement. A human observer performing manual measurements could easily overcome these difficulties related to the correct recognition of landmarks. Lastly, assumptions are made in the automated calculation of the radiographic morphology measurements (chapter 2). While these assumptions will hold true for most hips, this could lead to incorrect measurements in cases that were not considered while making these assumptions. Therefore, performing a visual check of a random sample and extreme measurements within your dataset is always important. This is why the developed methods in part 1 of this thesis contain an option for visualization of all measurements.

There are also limitations related to the use of two-dimensional (2D) imaging. Specifically, the developed methods (chapters 2 and 3) do not yet account for all variations in projected hip shape caused by patient positioning on the 2D imaging[238-241]. This makes it difficult to distinguish the true anatomical shape variation of the hip from shape variation caused by patient positioning. While the automated method included a correction for the pelvic obliquity (chapter 2), there is no correction for pelvic tilt, pelvic rotation, or position of the femur. Including these corrections might improve the morphological measurements, especially when comparing results from imaging made in different positions, for instance, supine and weight-bearing radiographs. However, while determining and correcting measurements for pelvic obliquity is relatively straightforward, correcting the others can be more challenging. Different methods have been proposed to determine pelvic tilt and rotation on AP radiographs[242-245]. How the determined tilt and rotation should be used to correct radiographic morphology measurements on AP imaging is still unclear. Similarly, the positioning of the patient's legs can drastically influence the projected femur shape and, therefore, the morphology measurements[246-248]. A better understanding of the relationship between the projected hip shape and the true anatomical hip shape will aid in creating more accurate measurements on AP radiographs. Indeed, the inherent limitations of 2D imaging have led some researchers to advocate for a shift away from 2D methods altogether, favoring the use of 3D imaging for studying hip morphology.

The use of three-dimensional (3D) imaging could resolve the problems encountered when using conventional 2D imaging. However, in clinical practice and most research, slices through or projections of the hip joint are used to quantify hip shape using similar radiographic morphology measurements as 2D imaging[249,250]. Additionally, obtaining 3D imaging is more time-consuming and costly, and, in the case of computer tomography (CT), involves a higher radiation burden. The advantage of 3D imaging is that the po-

sitioning of the femur and pelvis can be standardized after segmenting the hip joint, eliminating the influence of patient positioning. This would eliminate the influence of patient positioning on the measurements. Additionally, the hip joint could be studied in multiple planes at once. There is a need to develop radiographic measurements further to use the full 3D imaging potential. An example could be assessing the full 3D acetabular coverage instead of just the lateral and anterior center edge angles[249]. For these 3D measurements, it still holds that automated, validated, open-access, insightful methods are the way to move forward.

Although quantifying hip morphology using radiological morphology measurements on 2D and 3D imaging is common practice, it does reduce the complex hip shape to specific characteristics. In these measurements, the full hip shape is no longer considered. One method of capturing the full hip shape is statistical shape modeling (SSM). SSM is a method to capture variation in shape within a specific population, which can be applied to 2D and 3D imaging.

While acetabular dysplasia and pincer morphology are hip morphologies related to the acetabulum, cam morphology is a change in the proximal femur morphology characterized by extra bone formation at the anterolateral femoral head-neck junction resulting in a non-spherical femoral head[251]. Cam morphology is often quantified on AP pelvic radiographs using an alpha angle greater than 60 degrees. A recent study using a SSM of just the femoral head-neck junction on pelvic radiographs showed distinct variations in the shape of the femoral head-neck junction, indicating potential sub-types of cam morphology, despite all being associated with an alpha angle greater than 60 degrees[252]. Interestingly, only a few of these distinct variations were also associated with the development of radiographic hip osteoarthritis. This illustrates the limitations of reducing the complex hip shape to specific characteristics, like the alpha angle, to quantify and study hip shape as a risk factor for radiographic hip osteoarthritis.

Another promising application of SSM is the ability to study the change in hip shape over time. This could be used to model the developing hip by focusing on normal growth and abnormal development like acetabular dysplasia and pincer morphology, as well as the change in hip shape as a predictor for the development of osteoarthritis or a result of osteoarthritis. When we better understand the developing hip and the development of an abnormal hip shape, we can also better our understanding of the etiology, including (modifiable) risk factors. Additionally, subtle changes in hip shape in the adult hip, possibly compared to healthy hips of people in the same age category, could be a way to move towards early detection of osteoarthritis. For example, SSM could potentially identify early cartilage degeneration by detecting subtle changes in joint space on an AP

pelvic radiograph. A major challenge of using SSM is the interpretation of shape modes and, therefore, the found associations, which can be quite difficult. Shape modes are often interpreted subjectively, introducing possible reader bias. A second challenge is that the shape modes resulting from an SSM are unique to the specific SSM used. This means that comparing results from studies using different SSMs is impossible. Therefore, the aim should be to create universal SSMs that are available open-access, allowing comparison between studies. With that, since the studied shape modes are the same, the shape modes only need to be interpreted once.

Future research should investigate the hip shape beyond just the "standard" radiographical morphological measurements. The use of 2D and 3D imaging should be explored further, including the capability of 3D imaging to quantify the full 3D shape and techniques that combine both, like 2D biplanar imaging, which can translate into 3D models.

Defining hip morphology

Based on the quantified hip shape, different hip morphologies, such as acetabular dysplasia and pincer morphology, can be defined on imaging (chapters 2, 3 and 4).

The center edge angle is most often used in literature to define acetabular dysplasia and pincer morphology. Two variations of the center edge angle exist: the lateral center edge angle (LCEA) and the Wiberg center edge angle (WCEA). Generally, the LCEA measures the bony coverage of the femoral head by the acetabulum, while the WCEA measures the weight-bearing coverage. The weight-bearing portion of the acetabulum is also referred to as the sourcil, which is the radiographic presentation of the compressive stress in the hip joint[253]. While the most lateral bony point and the most lateral weight-bearing point of the acetabulum can be the same, this is not always the case. The most lateral bony point can also be more lateral than the most lateral weight-bearing point, resulting in a higher LCEA than WCEA in these hips, see Figure 1. Especially in dysplastic hips, the sourcil is shown to end more medial and superior than the bony acetabulum[253]. However, the terms LCEA and WCEA are often used interchangeably in literature, and a clear definition of the measurement used within a specific study is often missing. Therefore, it is essential to describe the measurement used within research clearly.

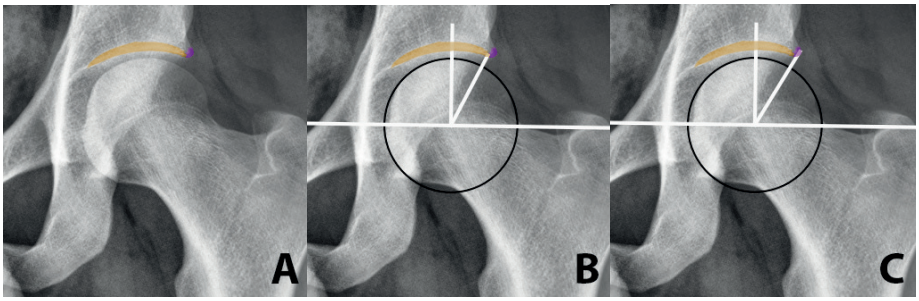


Figure 1. The Wiberg center edge angle (WCEA) and the lateral center edge angle (LCEA) as measured on an anteroposterior pelvic radiograph of the same person. The pelvic radiograph is cropped to the left hip. **A:** The acetabular sourcil, the weight-bearing part of the acetabulum, is indicated in orange. The additional bony coverage is indicated in purple. **B:** The WCEA determined using the most lateral point of the sourcil. **C:** The LCEA determined using the most lateral bony point of the acetabulum.

On top of this, multiple cut-off values are used to define acetabular dysplasia and pincer morphology based on the CEA. In adults, both 20° and 25° are often used to define acetabular dysplasia. Sometimes ≤ 20° is defined as definite acetabular dysplasia, and 20°- 25° is defined as borderline. Similarly, an LCEA higher than 35°, 38°, 40° or 45° is commonly used in literature to define pincer morphology. Additionally, both the LCEA and WCEA are used to define acetabular dysplasia. Since the LCEA can result in higher values than the WCEA, using the LCEA will lead to a lower prevalence of acetabular dysplasia than using the WCEA in the same population. All these factors combined make comparing research findings from different studies challenging.

In the pediatric population, the cut-off value gets even more unclear (chapter 5). With the developing hip joint, the CEA and other morphological measurements change with the maturation of the hip[33,68,192]; see Figure 2. Additionally, the average values of the morphological measurements for age are different in boys and girls. This is likely because, on average, girls are more skeletally mature than boys of the same age. Therefore, girls' hips are further in the skeletal development than boys of the same age. Consequently, using adult cut-off values in pediatric populations may not be appropriate. However, there are no clear cut-offs available, especially in late childhood and adolescence. This calls for more personalized cut-off values, taking at least age and biological sex into account. Alternatively, skeletal age-specific cut-offs could be used to account for any biological sex differences related to skeletal maturity. However, this would mean that the skeletal maturity needs to be determined for each hip in order to apply these cut-off values, making implementation of skeletal age-specific cut-off values more challenging.

Different methods could be used to find appropriate cut-off values. However, a gold standard is needed for determining cut-off values for diagnostic tests[254]. These methods

do not apply to hip morphology since a clear gold standard is lacking, especially in childhood and adolescent populations. A method often employed for hip morphology cut-offs is population-based cut-offs set at two standard deviations (SD) from the population mean or the 2.5 and 97.5 percentiles[193,255]. This predefines that the prevalence of the morphology of interest is 2.5% within this population. Additionally, values outside 2 SD or the 2.5 and 97.5 percentiles represent the extremities of the range in the population and are not necessarily pathological. Therefore, I do not think this is an appropriate method of determining cut-off values for the pediatric population. More effective methods might be based on a longitudinal population cohort, like Generation R, where one can work backward from adulthood (fully developed hips). By tracing the developmental trajectories of hip measurements in individuals who ultimately have acetabular dysplasia or pincer morphology at skeletal maturity and comparing these trajectories to those of individuals who have normal hip morphology, insight can be gained into how these morphologies develop and guide the definition of age-appropriate cut-off values. Additionally, clinically relevant factors could aid in this process, such as the development of hip pain, performed interventions, and the development of hip osteoarthritis.

Standardizing the definitions and cut-off values for acetabular dysplasia and pincer morphology, particularly in pediatric populations, is crucial for comparison across research findings and advancing our understanding of these conditions and their implications.

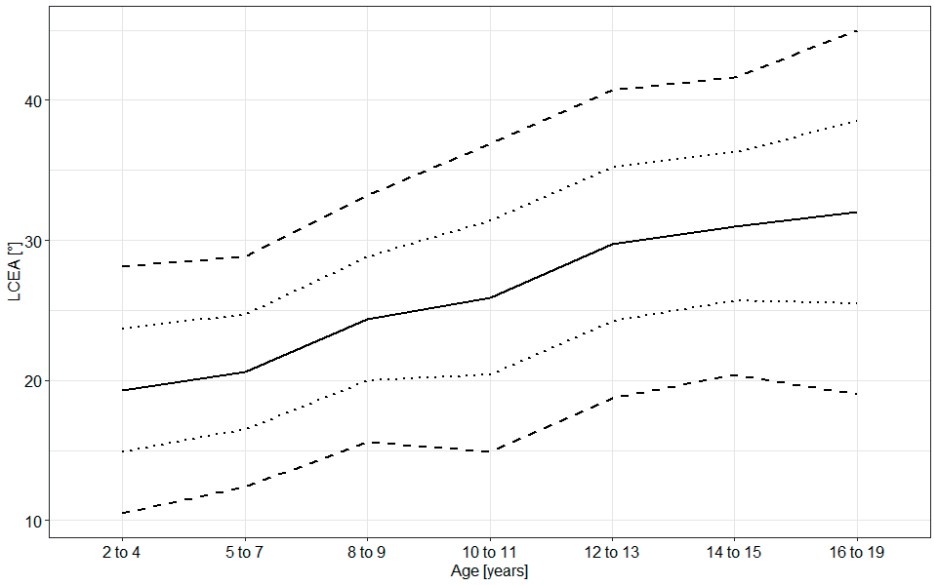


Figure 2. Change of acetabular coverage as defined by the lateral center edge angle (LCEA) with age, the mean is presented as a solid line, one standard deviation (SD) as a dotted line and two SD as a dashed line. Based on data from Monazzam et al. 2018 [33].

DEVELOPMENT OF THE HIP

Chapters 6 and 8 showed that acetabular dysplasia might develop during skeletal maturation of the hip. The prevalence in early adolescents (6.4%, 95% CI 5.6 – 7.1) was higher than the prevalence of developmental dysplasia of the hip (DDH) from a recent meta-analysis(1.4%, 95% CI 0.86 – 2.3)[16]. Additionally, common risk factors associated with DDH were not associated with acetabular dysplasia in early adolescence (chapter 8), indicating that not all dysplastic hips found in adolescents are residual dysplasia from DDH or missed DDH diagnoses. However, no data is currently available on DDH diagnoses and treatment within the Generation R population. Therefore, residual dysplasia from DDH cannot be excluded definitively. Additionally, further diagnostic evaluation of DDH in the Netherlands is based on the presence of one or more risk factors, so it might be that the early adolescents identified in chapter 8 did not have any of the studied risk factors and were thus missed within the screening program. These risk factors include a positive family history of DDH or early-onset hip osteoarthritis (before age 50 years), a breech position after 32 weeks of gestation, a breech presentation at birth, or abnormal findings on physical examination such as limited abduction < 70°, an abduction difference ≥ 20°, or a difference in leg length or knee height[20]. Of these risk factors, only a breech presentation at birth was studied in chapter 8. Systematic reviews with meta-analyses have shown only a small or non-significant difference in late-diagnosed DDH between selective and universal ultrasound diagnostic evaluation[256-259]. Therefore, the selective screening performed in the Netherlands might not only explain the difference between the prevalence of DDH and acetabular dysplasia found in chapters 6 and 8. This could indicate that not all dysplastic hips found in adolescents in chapters 6 and 8 are missed DDH diagnoses. Previous studies also found differences between DDH and adolescent or adult-diagnosed acetabular dysplasia[27,28]. Therefore, I hypothesize that acetabular dysplasia can also have an onset during childhood or adolescence.

Pincer morphology is thought to start to develop around age 12 [33]. We found a prevalence of 3.1% in early adolescents aged 13 years from the general population (chapter 7). While the reported prevalence of pincer morphology in adults varies widely (3.0% to 74%)[30], 3.1% is on the lowest side of the range. This might indicate that more people in our study population might still develop pincer morphology. Analysis of the Generation R cohort at the next stage, around age 18 years, would allow investigation of this hypothesis.

A challenge for chapters 6, 7 and 8 was choosing the correct definition of acetabular dysplasia and pincer morphology. As described previously, clear cut-off values for this age range are missing. In chapters 6 and 8, the choice was made to characterize

acetabular dysplasia using LCEA < 20°. The ossification center of the acetabular lip, situated at the most lateral part of the acetabulum, is not fully fused yet within this group of 13-year-olds, and therefore, making an accurate distinction between the most lateral bony and most lateral weight-bearing point of the acetabulum could not always be done accurately. This is why the LCEA was selected above the WCEA since the LCEA could be performed more consistently within this age group. The cut-off of 20°, often used in adults, was chosen since no established cut-off value for this age group is available, and this allowed us to compare our findings to existing literature. However, both choices can be debated and were expert opinion-based, due to a lack of evidence. The selection of the LCEA with this specific cut-off could have led to an underestimation of the prevalence of acetabular dysplasia within this age group. Therefore, it could also have influenced the associations found in chapter 8. Similarly, the cut-off of 40°, often used in adults, was chosen to define pincer morphology in chapter 7. This choice could again have resulted in an underestimation of pincer morphology in this age group.

Another limitation of chapters 6 through 9 is the possible selection bias within Generation R. At the start of the cohort, 9,778 mothers with a delivery date from March 2002 until January 2006 were included[47]. Throughout the follow-up period, loss to follow-up and withdrawal from the study occurred, see Figure 3. While this is a regular occurrence for longitudinal studies, it could have introduced selection bias. For instance, more participants with underlying diseases, such as congenital hip defects like DDH, could have dropped out of the study. One could imagine that taking part in the Generation R study on top of doctor visits and treatment would be too much of a commitment for parents and children, resulting in withdrawal from the cohort. Creating a link between the Generation R database and general practice databases, for instance, the Rijnmond Primary Care Database (RPCD)[260], could help us provide insight into the prevalence of different diseases within the population and the possible occurrence of selection bias. Creating such a datalink could also aid in providing information on congenital hip defects within the study population. Additionally, if it is known if a participant has had DDH and treatment for DDH in the past, a distinction could be made between DDH and childhood-onset acetabular dysplasia.

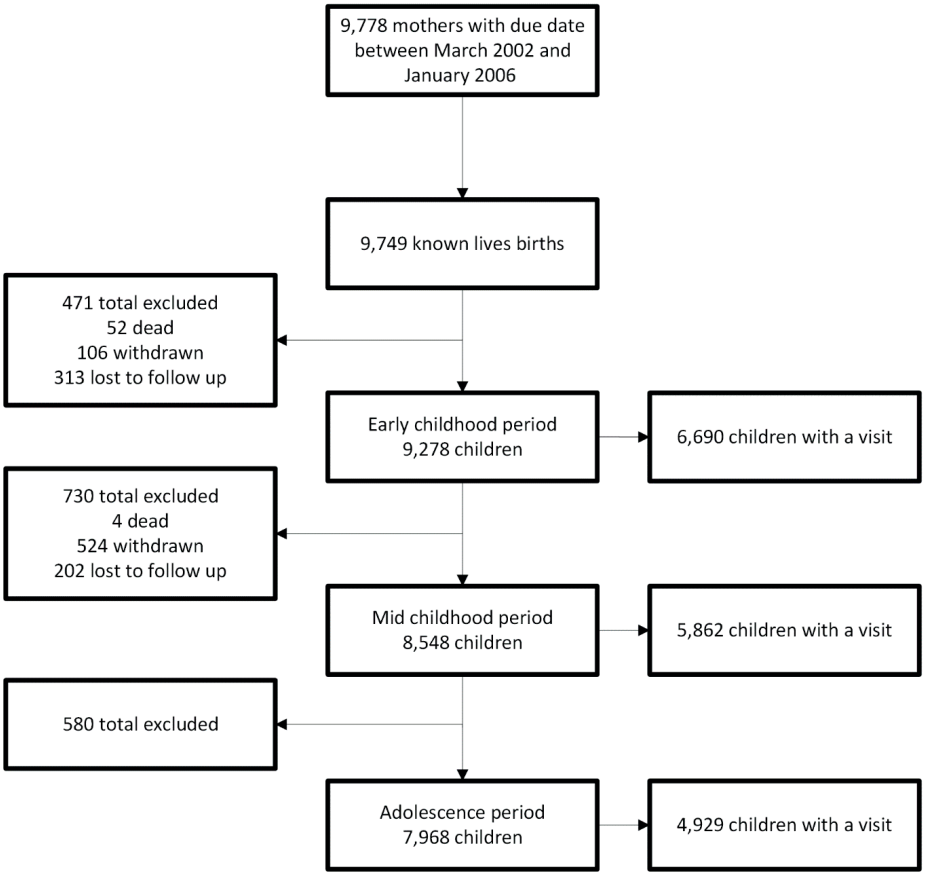


Figure 3. Enrolment, follow-up rates and center visits for measurements in the Generation R study. The early childhood period is around age 6 years, the mid childhood period around age 9 years and the adolescence period around age 13 years. The figure is based on data from Kooijman et al. 2016 [47] and chapter 8.

Skeletal growth, and thus the development of the hip, is influenced by hormonal, genetic, mechanical, and environmental factors. It was determined that while being overweight was protective for having acetabular dysplasia in adolescence, sports participation was not associated (chapter 8). It should be taken into account that sports participation was measured crudely and based on self-reported variables. In a similar cross-sectional study within Generation R at the time point of age 9, both BMI and each hour increase in physical activity were found to be protective for acetabular dysplasia[106]. Physical activity was defined by hours per week walking or cycling to school or hours per week spent playing outside. This might indicate a dose-response relationship between physical activity and acetabular dysplasia, and perhaps also the development of the acetabulum. Therefore, mechanical loading during development may play a role in

developing the acetabulum and acetabular dysplasia or pincer morphology; however, the relationship needs to be studied further to understand it better.

While we studied some possible risk factors for acetabular dysplasia in childhood (chapters 8 and 9), much remains unknown. Additionally, no determinants for pincer morphology have been studied yet. While several possible risk factors are of interest for future research, I would recommend a focus on possible easily modifiable risk factors for acetabular dysplasia or pincer morphology, since this could lead to more easily implementable preventative strategies. One such factor could be physical activity. Furthering the understanding of the relationship between physical activity and the development of the hip, and thus the development of acetabular dysplasia and pincer morphology, could be the first step. This could entail the time spent playing sports, the intensity, the type of sports played, and daily physical activity like traveling to school, which could be measured using the Short QUestionnaire to ASsess Health-enhancing physical activity (SQUASH)[261]. Additionally, the impact of the timing of physical activity, i.e. related to the developmental timeline of the hip, should be studied further.

Chapter 9 also showed that acetabular coverage, and thus acetabular dysplasia, can improve during growth. Conversely, this also shows that pincer morphology can develop during growth. Currently, it is difficult to distinguish normal and abnormal development. Reference growth charts could aid in our understanding of normal and abnormal development. These growth charts could be developed based on the Generation R data, using the different study phases throughout childhood and adolescence. This would include phases around age 6, 9, 13 and 18 years. Additionally, a new phase around age 21 is being prepared, making it possible to extend the growth chart to young adulthood. However, the development of the hip should not only be measured by the acetabular coverage as determined using the LCEA. This measure alone might not properly reflect the development of the hip joint, as discussed in the previous section "Quantifying hip morphology". More insight is needed on the correlation between the radiographic definitions of acetabular dysplasia and pincer morphology and pathological acetabular dysplasia and pincer morphology.

Defining normal and abnormal development of the hip joint and uncovering the etiology of acetabular dysplasia and pincer morphology can aid in clinical decision-making in the future. For instance, in children with a mild dysplastic acetabulum, knowing the predicted growth of the acetabulum could aid in the decision between watchful waiting and intervening right away.

HIP OSTEOARTHRITIS

The development of hip osteoarthritis

As discussed in the introduction of this thesis, hip morphology is considered an important risk factor for developing hip osteoarthritis due to its influence on local biomechanics. However, the association between acetabular dysplasia and pincer morphology and the development of hip osteoarthritis reported in literature can vary significantly[11]. In chapters 10 and 11, this relationship was investigated further using individual participant data (IPD) meta-analysis in the World COACH consortium. Both acetabular dysplasia and pincer morphology were associated with the development of radiographic hip osteoarthritis (RHOA) within 4-8 years.

Dysplastic hips had 1.80 times higher odds of developing RHOA compared to hips with a normal acetabulum (chapter 10). These odds are lower than those from a recent meta-analysis (2.38, 95% CI 1.84-3.07)[11]. This could be partly influenced by the included population with a relatively high mean age at baseline of 61 years since acetabular dysplasia is thought to be a risk for early onset hip osteoarthritis[8]. Additionally, chapter 10 and 11 exclusively included hips free of any radiographic signs of osteoarthritis at baseline. This differs from almost all other studies, which include hips with doubtful RHOA at baseline, potentially confounding the association between acetabular dysplasia and RHOA. Lastly, the range of follow-up time could be of influence, since the association between acetabular dysplasia and incident RHOA seems to diminish with longer follow-up[12].

Similarly, pincer hips had 1.50 times higher odds of developing RHOA within 4-8 years compared to hips with normal acetabular coverage of the femoral head (chapter 11). This is the first prospective cohort study that has found an association between pincer morphology and RHOA. Multiple factors could be related to this fact, namely, the selected study population or the strict threshold of an LCEA ≥ 45° used to define pincer morphology.

Even though there was an association between both morphologies and incident RHOA, not all hips with one of these morphologies will develop hip osteoarthritis. Future studies should aim to explore this relationship further. There might be specific characteristics within individuals with either acetabular dysplasia or pincer morphology that cause those individuals to develop hip osteoarthritis, while others with the same morphology will not. Possible characteristics of influence could be related to the biomechanics of the hip joint, like movement patterns, specific hip shape variations within acetabular dysplasia or pincer morphology, or more indirect factors like BMI and physical activity. Gaining more insight into the specific population at risk of development of hip osteo-

arthritis could aid in preventative efforts. This is especially important since no curative treatment is currently available, and the prevalence of hip osteoarthritis is predicted to keep increasing[36,37].

Quantifying hip osteoarthritis

A major challenge in hip osteoarthritis research is the lack of a universal definition[220], hindering comparisons across studies. Chapters 10 and 11 addressed this within the World COACH consortium by creating a harmonized RHOA score based on the original RHOA scores (Kellgren and Lawrence, modified Croft and OARSI individual features) of the parent cohorts[48]. However, the reliability of these different RHOA scores can vary widely and is highly dependent on the experience of the reader[220]. Beyond reliability issues, manual RHOA scoring is time-consuming. Automating RHOA scores offers a solution for faster, more consistent, and objective grading across cohorts. Such an automated method would require validation to ensure reliability and accurate categorization of osteoarthritis presence and severity. Open-access availability of this automated measurement would further promote widespread use and comparability in future research. Given its established reliability in epidemiological and clinical studies and the best correlation with pain compared to other radiographic osteoarthritis features[220], minimal joint space presents a promising starting point for developing automated RHOA assessment.

FUTURE PERSPECTIVES

Hip morphology

Advancing the understanding of the etiology of hip morphology requires moving beyond the current methods. While 3D imaging is increasingly used, its full capabilities are not yet utilized. Future research should focus on the development and validation of automated, open-access methods for quantifying 3D hip morphology. Transparency, explainability to physicians, and interpretability are crucial for the adoption of these methods. While designing new studies, the costs and benefits of 2D and 3D imaging must be carefully considered. The incorporation of 3D imaging can be quite costly, so this will not always be feasible. However, 3D reconstruction based on 2D biplanar imaging could be a viable alternative that should be explored further. Validation of this approach before implementation is essential for reliable results. This approach would expand the excess to 3D hip morphology analyses.

While quantifying hip morphology in 3D has a lot of potential, I do think there is much knowledge to be gained from existing 2D radiographs. Improving hip morphology measurements on 2D imaging by eliminating limitations caused by patient position

would result in more reliable results. Better understanding the relationship between the projected hip shape and the true anatomical hip shape would aid in developing corrections for patient positioning beyond the pelvic obliquity. One approach would be the use of digitally reconstructed radiographs (DRRs), which are simulated 2D radiographic images generated from 3D imaging[262]. CT images are commonly used for this purpose; however, magnetic resonance images (MRI) could also be used. After segmentation of the proximal femur and the pelvis, DRRs could be created with the pelvis and femur in various positions simulating various patient positions. Subsequently, the influence of positioning on the 2D hip morphology measurements can be assessed, and corrections can be developed. This approach will not only improve the reliability of 2D hip morphology measurements, but also strengthen the understanding of the relationship between 2D and 3D hip shape.

Acetabular dysplasia and pincer morphology

The work presented in chapter 9 should be extended by incorporating additional time points, additional hip morphology measurements, and 3D imaging to gain a better understanding of the development of abnormal hip morphology like acetabular dysplasia and pincer morphology. Analyzing longitudinal data from studies like Generation R will enable the creation of reference growth charts for hip development analogous to those used for height and weight. These reference charts will allow a comparison of individual hip development to the general population, facilitating the detection of abnormal growth trajectories. Within Generation R, a subset of participants received an MRI around ages 9, 13 and 18. Analyzing these data would allow analysis of the development of the hip joint in 3D, further enhancing the understanding of hip joint development.

Additionally, longitudinal studies employing latent growth models or spatiotemporal SSMs can provide valuable insights into how hip morphology changes over time. Such studies can help to identify typical and atypical growth patterns, predict future changes, and might eventually be employed to predict the results of planned interventions. For example, understanding how hip morphology develops during adolescence could help with developing preventative strategies to prevent hip problems later in life.

Developing appropriate cut-off values for acetabular dysplasia and pincer morphology in pediatric populations is also crucial. I would suggest the use of longitudinal data while considering the correlation between radiographic definitions and pathological findings. Clinical aspects such as hip pain, stiffness, diagnoses, treatment, and the development of hip osteoarthritis should be considered in this process. Defining clear cut-off values will help to standardize the definition of acetabular dysplasia and pincer morphology in

research. This would allow for study of their etiology, and widespread adoption would allow for comparison of research findings.

A better understanding of acetabular dysplasia and pincer morphology can aid in the development of preventative measures. Since acetabular dysplasia and pincer morphology are related to higher odds of developing hip osteoarthritis, prevention of these morphologies would also aid in the prevention of hip osteoarthritis. Additionally, if the development of (abnormal) hip morphology is better understood, this could result in the construction of prediction models that can aid in clinical decision-making. For instance, if a child around age 6 gets referred for acetabular dysplasia, a decision needs to be made between treatment and watchful waiting. If the expected growth trajectory of the dysplastic hip can be predicted, this could aid in the decision-making.

Hip osteoarthritis

While chapters 10 and 11 showed that hips with acetabular dysplasia or pincer morphology have higher odds of developing RHOA, not everyone with acetabular dysplasia or pincer morphology will develop RHOA. This association should, therefore, be explored further. A first step could be to study whether there are specific shape variations or baseline characteristics within individuals with either acetabular dysplasia or pincer morphology that are associated with the development of RHOA. This involves creating subpopulations of people with acetabular dysplasia or pincer morphology to analyze shape variations related to the acetabulum of the femoral head and characteristics such as BMI, biological sex, and genetics within these specific populations. This approach could help identify specific risk factors within these subpopulations, leading to better identification of people at risk of developing RHOA. Understanding the relationship between hip morphology and hip osteoarthritis development could inform prevention and early intervention strategies, improving patient care and potentially delaying or preventing the need for treatments like joint replacement surgery.

Automated RHOA scoring, particularly of joint space width, is a crucial next step for research within World COACH. This would allow for uniform RHOA scores throughout World COACH, and beyond. While artificial intelligence offers a promising approach, the measurements must remain understandable and interpretable for clinical use. Automated RHOA scores have the potential to be more time-efficient, accurate, and reproducible than manual scores. The development of an automated score would not only facilitate efficient assessment but also reduce inter- and intra-observer variability. Additionally, I think such a method should be published open-access so that it can be externally validated and utilized by other research groups.

A step that could aid in the development of an automated RHOA measurement would be to investigate the correlation between (minor) radiographic changes and changes observed on MRI, as defined by scoring hip osteoarthritis with MRI (SHOMRI) scores. This can offer valuable insights into the features that automated RHOA scores should prioritize. SSM can help with describing radiographic changes with regard to shape. However, other image features, such as density, should also be considered. Linking these radiographic changes to MRI findings allows for a better understanding of the development and progression of hip osteoarthritis. This will not only contribute to the development of automated scoring systems, but will also aid in the early detection of RHOA on pelvic radiographs.

Finally, given the association between acetabular dysplasia and early-onset hip osteoarthritis, establishing younger cohorts for research is essential. Studying the relationship between hip morphology, particularly acetabular dysplasia, and RHOA in younger individuals can help further the understanding of early disease development.

Chapter 13

Summary

The hip joint begins its development in utero and continues to develop through childhood. Abnormal hip morphology can arise throughout this development. This thesis focused on abnormal morphology of the acetabulum, namely acetabular dysplasia and pincer morphology. Acetabular dysplasia is characterized by a shallow acetabulum with insufficient femoral head coverage. It is crucial to distinguish between developmental dysplasia of the hip (DDH), which develops during the perinatal period and infancy, and acetabular dysplasia, which is diagnosed later in life. The development timeline of these cases of late-diagnosed acetabular dysplasia remains unclear. Pincer morphology, conversely, involves femoral head overcoverage by the acetabulum. Although pincer morphology has been studied extensively in adults, there is a lack of knowledge on the development of pincer morphology during childhood.

Hip morphology is an important risk factor for the development of hip osteoarthritis. However, literature reports varying findings, and some studies find no relationship. Therefore, the true impact of acetabular dysplasia and pincer morphology on the development of hip osteoarthritis remains unclear.

This thesis aimed to understand the prevalence, development, and contribution to the risk of hip osteoarthritis of acetabular dysplasia and pincer morphology. To study acetabular dysplasia and pincer morphology, part 1 focused on the development and validation of an automated method for quantifying hip morphology. Part 2 aimed to study the prevalence and determinants of hip morphology in children and early adolescents, a critical yet understudied period for hip development. Part 3 aimed to further study the association between hip morphology, specifically acetabular dysplasia and pincer morphology, and radiographic hip osteoarthritis.

PART 1: QUANTIFYING HIP MORPHOLOGY

In Chapter 2, we presented our in-house developed automated method to determine radiographic hip morphology measurements on dual-energy x-ray absorptiometry (DXA) images. The acetabular depth-width ratio, acetabular index, alpha angle, Wiberg and lateral center edge angle, extrusion index, neck-shaft angle, and triangular index could be automatically determined, and the measurements performed comparable to or better than manual measurements, except for the acetabular index. The presented method allows for fast and reproducible calculation of radiographic measurements of hip morphology.

Chapter 3 aimed to validate this method externally in the adult population from the Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH). The automated morphological measurements were a reliable alternative to manual measurements performed by trained observers.

Traditionally, these radiographic hip morphology measurements are determined on pelvic radiographs. However, DXA images are increasingly used to study hip morphology. In Chapter 4, we compared hip morphology measurements performed on DXA images and pelvic radiographs to see if these imaging modalities can be used interchangeably. Participants from the Rotterdam Study who underwent DXA imaging and pelvic radiography on the same day were included. We found that DXA images and pelvic radiographs can both reliably be used to study hip morphology.

PART 2: HIP MORPHOLOGY IN CHILDREN: PREVALENCE AND DETERMINANTS

Utilizing the validated methods from Part I, Part II investigated the prevalence and possible risk factors associated with hip morphology in a general population of children and early adolescents utilizing the Generation R cohort. In Chapter 5, we performed a systematic review of the literature on the prevalence of acetabular dysplasia between the ages of two and eighteen. Additionally, we described the radiological measurements used to diagnose acetabular dysplasia on the imaging. This systematic review highlighted the lack of knowledge on the prevalence of acetabular dysplasia in children after the age of 2 years, especially in the general population. The Wiberg center edge angle and the acetabular index were used to quantify hip morphology. However, the definition of the measurements was inconsistent or even lacking.

Chapter 6 aimed to describe the prevalence of acetabular dysplasia in 3,986 early adolescents of the general population. Additionally, differences in prevalence related to birth-assigned sex, ethnicity, and skeletal maturity, defined by the triradiate cartilage status, were studied. Acetabular dysplasia was defined as a lateral center edge angle <20° as determined on the right hip DXA images using the methods described in part 1. The included participants had a mean age of 13.6 ± 0.3 years, and 47% were male. The prevalence of acetabular dysplasia was 6.4% (95% CI 5.6-7.1%) and was higher in participants with an open triradiate cartilage than those with a closed triradiate cartilage. No statistically significant difference in the prevalence of acetabular dysplasia was found related to categories of birth-assigned sex and ethnicity.

Chapter 7 mirrored chapter 6 and aimed to describe the prevalence of pincer morphology within this same population of early adolescents of the general population and the birth-assigned sex-specific prevalence of pincer morphology. Pincer morphology was defined by a lateral center edge angle ≥ 40°. The lateral center edge angle was automatically determined on the right hip DXA images using the validated method from part 1. The prevalence of pincer morphology was 3.1% (95% CI 2.6-3.6%) and was similar in males and females.

In Chapter 8, we aimed to further our understanding of acetabular dysplasia in early adolescents. We investigated whether known risk factors for DDH are also associated with acetabular dysplasia in the same population of 3,986 early adolescents from the general population studied in chapters 6 and 7. We also investigated BMI and physical activity since these were shown to be associated with acetabular dysplasia in childhood. Acetabular dysplasia was defined as a lateral center edge angle <20°, similar to chapter 6. We found that known risk factors for DDH and sports participation were not associated with acetabular dysplasia in early adolescents, while weight status over time was. Being overweight between the ages of 6 and 13 years, compared to participants of normal weight, was protective for acetabular dysplasia, aOR 0.39 (95% CI 0.18 – 0.85).

In Chapter 9, we continued the work of chapter 8 and aimed to study the development of the acetabulum during childhood. We investigated how acetabular coverage changes over time and its associations with birth-assigned sex, weight status, triradiate cartilage orientation, and head-shaft angle at age 9, to provide insights into the developmental trajectory of the acetabulum. The change of acetabular coverage from age 9 to age 13 was studied in 516 participants. Acetabular coverage was automatically determined on all DXA images using the lateral center edge angle, as described in part 1. We found that the acetabular coverage changed throughout childhood. Birth-assigned sex, weight status, and the triradiate cartilage orientation were associated with this change, while the head-shaft angle was not.

PART 3: HIP MORPHOLOGY AND OSTEOARTHRITIS: CONNECTING HIP SHAPE TO MORBIDITY

In chapters 10 and 11, we performed an individual participant data meta-analysis on the association between hip morphology and incident radiographic hip osteoarthritis within the World COACH consortium. Chapter 10 studied the association between acetabular dysplasia, automatically determined using the methods from part 1, and incident radiographic hip osteoarthritis in 18,807 hips free of any signs of radiographic hip

osteoarthritis at baseline. We demonstrated an association (OR 1.80 95% CI 1.40-2.34) between acetabular dysplasia defined by a Wiberg center edge angle ≤ 25° and incident radiographic hip osteoarthritis within 4-8 years. Additional measures of acetabular dysplasia (Wiberg center edge angle ≤ 20°, acetabular depth-width ratio ≤ 250, acetabular index ≥ 13°, or a combination) were also associated with an increased risk of developing radiographic hip osteoarthritis.

Chapter 11 studied the association between pincer morphology and incident radiographic hip osteoarthritis in 18,935 hips free of any signs of radiographic hip osteoarthritis at baseline. Pincer morphology was automatically determined on pelvic radiographs using the methods from part 1. In contrast to findings from earlier prospective cohort studies, we found a significant association between pincer morphology defined by a lateral center edge angle ≥ 45° and incident radiographic hip osteoarthritis (OR 1.50, 95% CI 1.05-2.15), but not when pincer morphology was defined by lateral center edge angle ≥ 40° (OR 1.15, 95% CI 0.87-1.51).

Appendices

NEDERLANDSE SAMENVATTING

De ontwikkelen van het heupgewricht begint in de baarmoeder en deze ontwikkeling gaat door tot het kind volledig is uitgegroeid. Afwijkende heupmorfologie kan gedurende deze ontwikkeling ontstaan. Dit proefschrift richt zich op de afwijkende morfologie van het acetabulum, namelijk acetabulaire dysplasie en pincermorfologie.

Acetabulaire dysplasie wordt gekenmerkt door een ondiep acetabulum met onvoldoende bedekking van de femurkop. Het is essentieel om onderscheid te maken tussen developmental dysplasia of the hip (DDH), wat zich ontwikkelt tijdens de perinatale periode en de babytijd, en acetabulaire dysplasie, wat later in het leven wordt vastgesteld. Het ontwikkelingstraject van deze laat gediagnosticeerde gevallen van acetabulaire dysplasie blijft onduidelijk. Pincermorfologie, daarentegen, is de overmatige bedekking van de femurkop door het acetabulum. Hoewel pincermorfologie uitgebreid is bestudeerd bij volwassenen, is er een gebrek aan kennis over de ontwikkeling van pincermorfologie gedurende de kindertijd.

Heupmorfologie is een belangrijke risicofactor voor de ontwikkeling van heupartrose. De literatuur rapporteert echter uiteenlopende bevindingen, en sommige studies vinden zelfs geen verband. Hierdoor blijft de werkelijke impact van acetabulaire dysplasie en pincermorfologie op de ontwikkeling van heupartrose onduidelijk.

Het doel van dit proefschrift was inzicht te krijgen in de prevalentie, ontwikkeling en bijdrage aan het risico op heupartrose van acetabulaire dysplasie en pincermorfologie. Om acetabulaire dysplasie en pincermorfologie te bestuderen, richtte Deel 1 zich op de ontwikkeling en validatie van een geautomatiseerde methode voor het kwantificeren van heupmorfologie. Deel 2 had als doel de prevalentie en determinanten van heupmorfologie bij kinderen en jonge adolescenten te bestuderen, een kritieke maar onderbelichte periode voor heupontwikkeling. Deel 3 was gericht op het verder bestuderen van de associatie tussen heupmorfologie en radiografische heupartrose.

Deel 1: Het Kwantificeren van Heupmorfologie

In Hoofdstuk 2 presenteerden we de intern ontwikkelde geautomatiseerde methode voor het bepalen van radiografische heupmorfologie metingen op dual-energy x-ray absorptiometrie (DXA) beelden. De acetabulaire depth-width ratio, acetabulaire index, alfahoek, Wiberg en lateral center edge angle, extrusion index, neck-shaft angle en triangular index konden automatisch worden bepaald. De automatische metingen presteerden vergelijkbaar met of beter dan handmatige metingen, met uitzondering van de

acetabulaire index. Deze methode zorgt voor een snelle en reproduceerbare bepaling van de radiografische metingen van heupmorfologie.

Hoofdstuk 3 richtte zich op het extern valideren van deze methode in de volwassen populatie van de Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH). De geautomatiseerde morfologische metingen bleken een betrouwbaar alternatief voor handmatige metingen.

Traditioneel worden deze radiografische heupmorfologie metingen bepaald op bekkenfoto's. DXA beelden worden echter steeds vaker gebruikt om heupmorfologie te bestuderen. In Hoofdstuk 4 vergeleken we heupmorfologie metingen uitgevoerd op DXA beelden en bekkenfoto's om te bepalen of DXA beelden een goed alternatief zijn voor bekkenfoto's voor het onderzoeken van heupmorfologie. Deelnemers van de Rotterdam Studie bij wie op dezelfde dag een DXA beeld en een bekkenfoto zijn gemaakt, werden geïncludeerd. We vonden dat zowel DXA beelden en bekkenfoto's betrouwbaar kunnen worden gebruikt om heupmorfologie te onderzoeken.

Deel 2: Heupmorfologie in kinderen: Prevalentie en Determinanten

Met behulp van de gevalideerde methoden uit Deel 1 onderzocht Deel 2 de prevalentie en mogelijke risicofactoren voor heupmorfologie in een algemene populatie van kinderen en jonge adolescenten uit het Generation R cohort. In Hoofdstuk 5 voerden we een systematische review van de literatuur uit over de prevalentie van acetabulaire dysplasie bij kinderen tussen de leeftijd van twee en achttien jaar. Daarnaast beschreven we de radiologische metingen die worden gebruikt om acetabulaire dysplasie op beeldvorming vast te stellen. Deze systematische review benadrukte het gebrek aan kennis over de prevalentie van acetabulaire dysplasie bij kinderen na de leeftijd van 2 jaar. De Wiberg center edge angle en de acetabulaire index werden gebruikt om acetabulaire dysplasie te kwantificeren. De definitie van de metingen binnen de verschillende studies was echter inconsistent of ontbrak volledig.

Hoofdstuk 6 had als doel de prevalentie van acetabulaire dysplasie te beschrijven bij 3.986 jonge adolescenten uit de algemene bevolking. Daarnaast werden verschillen in prevalentie met betrekking tot geboortegeslacht, etniciteit en skeletale uitrijping, gedefinieerd door de status van het triradiate kraakbeen, bestudeerd. Acetabulaire dysplasie werd gedefinieerd als een lateral center edge angle < 20°, zoals bepaald op de DXA beelden van de rechterheup met behulp van de methoden beschreven in Deel 1. De geïncludeerde deelnemers hadden een gemiddelde leeftijd van 13,6 ± 0,3 jaar en 47% was man. De prevalentie van acetabulaire dysplasie was 6,4% (95% CI 5,6-7,1%) en was hoger bij deelnemers met een open triradiate kraakbeen dan bij deelnemers met

een gesloten triradiate kraakbeen. Er werd geen statistisch significant verschil gevonden in de prevalentie van acetabulaire dysplasie met betrekking tot de categorieën van geboortegeslacht en etniciteit.

Hoofdstuk 7 had als doel de prevalentie van pincermorfologie te beschrijven binnen dezelfde populatie van jonge adolescenten uit de algemene bevolking. Daarnaast werd de prevalentie van pincermorfologie per categorie van geboortegeslacht bekeken. Pincermorfologie werd gedefinieerd door een lateral center edge angle ≥ 40°. De lateral center edge angle werd automatisch bepaald op de DXA beelden van de rechterheup met behulp van de gevalideerde methode uit Deel 1. De prevalentie van pincermorfologie was 3,1% (95% CI 2,6-3,6%) en was vergelijkbaar bij mannen en vrouwen.

In Hoofdstuk 8 onderzochten we of bekende risicofactoren voor DDH ook geassocieerd zijn met acetabulaire dysplasie in dezelfde populatie van 3.986 jonge adolescenten uit de algemene bevolking die in Hoofdstuk 6 en 7 werden bestudeerd. We onderzochten ook BMI en fysieke activiteit, aangezien deze geassocieerd zijn met acetabulaire dysplasie op 9-jarige leeftijd. Acetabulaire dysplasie werd gedefinieerd als een lateral center edge angle < 20°, vergelijkbaar met Hoofdstuk 6. We vonden dat bekende risicofactoren voor DDH en sportdeelname niet geassocieerd waren met acetabulaire dysplasie bij jonge adolescenten, terwijl gewichtsstatus in de kindertijd dat wel was. Overgewicht tussen de leeftijd van 6 en 13 jaar in vergelijking met deelnemers met een normaal gewicht was beschermend voor acetabulaire dysplasie, aOR 0,39 (95% CI 0,18 - 0,85).

In Hoofdstuk 9 breidden we het werk van Hoofdstuk 8 uit en wilden we de ontwikkeling van het acetabulum tijdens de kindertijd bestuderen. We onderzochten hoe de acetabulaire bedekking van de femurkop in de loop van de tijd verandert en de associatie met geboortegeslacht, gewichtsstatus, triradiate kraakbeenoriëntatie en head-shaft angle op 9-jarige leeftijd, om inzicht te geven in het ontwikkelingstraject van het acetabulum. De verandering van acetabulaire bedekking van 9 tot 13 jaar werd bestudeerd bij 516 deelnemers. Acetabulaire bedekking werd automatisch bepaald op alle DXA beelden met behulp van de lateral center edge angle, zoals beschreven in Deel 1. We vonden dat de acetabulaire bedekking veranderde gedurende de kindertijd. Geboortegeslacht, gewichtsstatus en de triradiate kraakbeenoriëntatie waren geassocieerd met deze verandering, maar de head-shaft angle niet.

Deel 3: Heupmorfologie en artrose

In hoofdstuk 10 en 11 voerden we een meta-analyse van individuele deelnemersgegevens uit naar de associatie tussen heupmorfologie en het ontwikkelen van radiografische heupartrose binnen het World COACH-consortium. Hoofdstuk 10 onderzocht

het verband tussen acetabulaire dysplasie, automatisch bepaald met behulp van de methoden uit Deel 1, en het ontwikkelen van radiografische heupartrose in 18.807 heupen zonder tekenen van radiografische heupartrose op baseline. We vonden een associatie tussen acetabulaire dysplasie, gedefinieerd door een Wiberg center edge angle ≤ 25°, en het ontwikkelen van radiografische heupartrose binnen 4-8 jaar (OR 1,80, 95% CI 1,40-2,34). Andere definities van acetabulaire dysplasie (Wiberg center edge angle ≤ 20°, acetabulaire depth-width ratio ≤ 250, acetabulaire index ≥ 13° of een combinatie) waren ook geassocieerd met een verhoogd risico op het ontwikkelen van radiografische heupartrose.

Hoofdstuk 11 bestudeerde de associatie tussen pincermorfologie en het ontwikkelen van radiografische heupartrose in 18.935 heupen zonder tekenen van radiografische heupartrose op baseline. Pincermorfologie werd automatisch bepaald op bekkenfoto's met behulp van de methoden uit Deel 1. In tegenstelling tot resultaten van eerdere prospectieve cohortstudies, vonden we een significante associatie tussen pincermorfologie, gedefinieerd door een lateral center edge angle ≥ 45°, en het ontwikkelen van radiografische heupartrose (OR 1,50, 95% CI 1,05-2,15), maar niet wanneer pincermorfologie werd gedefinieerd door een lateral center edge angle ≥ 40° (OR 1,15, 95% CI 0,87-1,51).

REFERENCES

1. R.S. Watanabe. Embryology of the human hip. Clin Orthop Relat Res. 1974(98):8-26.
2. M.C. Lee, C.P. Eberson. Growth and development of the child's hip. Orthop Clin North Am. 2006;37(2):119-32, v.
3. R.S. Siffert. Patterns of deformity of the developing hip. Clin Orthop Relat Res. 1981(160):14-29.
4. S. Kim, H. Yoon, H. Han, S. Cho, Y. Seo. Ossification of the Triradiate Cartilage and Posterior Acetabulum. J Korean Soc Radiol. 2019;80(3):503.
5. M.R. Safran, A.J. Money, Z.D. Vaughn, D.A. Brcka, R. Stone McGaver, M.R. Giveans, et al. Acetabular Rim Ossification Variants Are Found in Almost 20% of Patients Presenting With Hip Pain. Arthrosc Sports Med Rehabil. 2023;5(4):100742.
6. C. Birkenmaier, G. Jorysz, V. Jansson, B. Heimkes. Normal development of the hip: a geometrical analysis based on planimetric radiography. J Pediatr Orthop B. 2010;19(1):1-8.
7. N. Herregods, J. Jaremko, L. Jans. Hip and Pelvis. In: P. Simoni, M.P. Aparisi Gómez, editors. Essential Measurements in Pediatric Musculoskeletal Imaging. Cham: Springer International Publishing; 2023. p. 139-60.
8. F. Saberi Hosnijeh, M.E. Zuiderwijk, M. Versteeg, H.T. Smeele, A. Hofman, A.G. Uitterlinden, et al. Cam Deformity and Acetabular Dysplasia as Risk Factors for Hip Osteoarthritis. Arthritis Rheumatol. 2017;69(1):86-93.
9. M. Reijman, J.M. Hazes, H.A. Pols, B.W. Koes, S.M. Bierma-Zeinstra. Acetabular dysplasia predicts incident osteoarthritis of the hip: the Rotterdam study. Arthritis Rheum. 2005;52(3):787-93.
10. R. Agricola, M.P. Heijboer, R.H. Roze, M. Reijman, S.M. Bierma-Zeinstra, J.A. Verhaar, et al. Pincer deformity does not lead to osteoarthritis of the hip whereas acetabular dysplasia does: acetabular coverage and development of osteoarthritis in a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage. 2013;21(10):1514-21.
11. N.C. Casartelli, N.A. Maffiuletti, P.L. Valenzuela, A. Grassi, E. Ferrari, M.M.A. van Buuren, et al. Is hip morphology a risk factor for developing hip osteoarthritis? A systematic review with meta-analysis. Osteoarthritis Cartilage. 2021;29(9):1252-64.
12. N.S. Riedstra, R. Vinge, J. Herfkens, D. Eygendaal, S.M.A. Bierma-Zeinstra, J. Runhaar, et al. Acetabular dysplasia and the risk of developing hip osteoarthritis at 2,5,8, and 10 years follow-up in a prospective nationwide cohort study (CHECK). Semin Arthritis Rheum. 2023;60:152194.
13. I.O. Engesaeter, T. Lehmann, L.B. Laborie, S.A. Lie, K. Rosendahl, L.B. Engesaeter. Total hip replacement in young adults with hip dysplasia: age at diagnosis, previous treatment, quality of life, and validation of diagnoses reported to the Norwegian Arthroplasty Register between 1987 and 2007. Acta Orthop. 2011;82(2):149-54.
14. M. de Hundt, F. Vlemmix, J.M. Bais, E.K. Hutton, C.J. de Groot, B.W. Mol, M. Kok. Risk factors for developmental dysplasia of the hip: a meta-analysis. Eur J Obstet Gynecol Reprod Biol. 2012;165(1):8-17.
15. C.L. Talbot, R.W. Paton. Screening of selected risk factors in developmental dysplasia of the hip: an observational study. Arch Dis Child. 2013;98(9):692-6.
16. Z. Tao, J. Wang, Y. Li, Y. Zhou, X. Yan, J. Yang, et al. Prevalence of developmental dysplasia of the hip (DDH) in infants: a systematic review and meta-analysis. BMJ Paediatr Open. 2023;7(1).
17. C.L. Ortiz-Neira, E.O. Paolucci, T. Donnon. A meta-analysis of common risk factors associated with the diagnosis of developmental dysplasia of the hip in newborns. Eur J Radiol. 2012;81(3):e344-51.

18. L.B. Laborie, I.O. Engesaeter, T.G. Lehmann, D.M. Eastwood, L.B. Engesaeter, K. Rosendahl. Screening strategies for hip dysplasia: long-term outcome of a randomized controlled trial. Pediatrics. 2013;132(3):492-501.
19. L.B. Laborie, T.J. Markestad, H. Davidsen, K.R. Bruras, S.M. Aukland, J.A. Bjorlykke, et al. Selective ultrasound screening for developmental hip dysplasia: effect on management and late detected cases. A prospective survey during 1991-2006. Pediatr Radiol. 2014;44(4):410-24.
20. M. Boere-Boonekamp, A. Broerse, J. Deurloo, C. Lanting, A. Konijnendijk. JGZ-richtlijn Heupdysplasie: Nederlands Centrum Jeugdgezondheid; 2018 [cited 2024 Oct 14]. Available from: https://www.jgzrichtlijnen.nl/alle-richtlijnen/richtlijn/heupdysplasie.
21. R. Graf, C. Tschauner, W. Klapsch. Progress in Prevention of Late Developmental Dislocation of the Hip by Sonographic Newborn Hip Screening - Results of a Comparative Follow-up-Study. J Pediatr Orthop B. 1993;2(2):115-21.
22. T. Terjesen, J. Horn. Management of late-detected DDH in children under three years of age: 49 children with follow-up to skeletal maturity. Bone Jt Open. 2020;1(4):55-63.
23. I.O. Engesaeter, L.B. Laborie, T.G. Lehmann, J.M. Fevang, S.A. Lie, L.B. Engesaeter, K. Rosendahl. Prevalence of radiographic findings associated with hip dysplasia in a population-based cohort of 2081 19-year-old Norwegians. Bone Joint J. 2013;95-B(2):279-85.
24. S. Jacobsen. Adult hip dysplasia and osteoarthritis. Studies in radiology and clinical epidemiology. Acta Orthop Suppl. 2006;77(324):1-37.
25. R. Raveendran, J.L. Stiller, C. Alvarez, J.B. Renner, T.A. Schwartz, N.K. Arden, et al. Population-based prevalence of multiple radiographically-defined hip morphologies: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage. 2018;26(1):54-61.
26. S. Humphry, T. Hall, M.A. Hall-Craggs, A. Roposch. Predictors of Hip Dysplasia at 4 Years in Children with Perinatal Risk Factors. JB JS Open Access. 2021;6(1).
27. C.B. Lee, A. Mata-Fink, M.B. Millis, Y.J. Kim. Demographic differences in adolescent-diagnosed and adult-diagnosed acetabular dysplasia compared with infantile developmental dysplasia of the hip. J Pediatr Orthop. 2013;33(2):107-11.
28. E.L. Sink, B.F. Ricciardi, K.D. Torre, C.T. Price. Selective ultrasound screening is inadequate to identify patients who present with symptomatic adult acetabular dysplasia. J Child Orthop. 2014;8(6):451-5.
29. M.M. Hadeed, J.M. Cancienne, F.W. Gwathmey. Pincer Impingement. Clin Sports Med. 2016;35(3):405-18.
30. P. van Klij, J. Heerey, J.H. Waarsing, R. Agricola. The Prevalence of Cam and Pincer Morphology and Its Association With Development of Hip Osteoarthritis. J Orthop Sports Phys Ther. 2018;48(4):230-8.
31. I. Zaltz, B.T. Kelly, I. Hetsroni, A. Bedi. The crossover sign overestimates acetabular retroversion. Clin Orthop Relat Res. 2013;471(8):2463-70.
32. J.J. Nepple, C.L. Lehmann, J.R. Ross, P.L. Schoenecker, J.C. Clohisy. Coxa profunda is not a useful radiographic parameter for diagnosing pincer-type femoroacetabular impingement. J Bone Joint Surg Am. 2013;95(5):417-23.
33. S. Monazzam, J.D. Bomar, J.R. Dwek, H.S. Hosalkar, A.T. Pennock. Development and prevalence of femoroacetabular impingement-associated morphology in a paediatric and adolescent population: a CT study of 225 patients. Bone Joint J. 2013;95-B(5):598-604.
34. D.J. Hunter, S. Bierma-Zeinstra. Osteoarthritis. Lancet. 2019;393(10182):1745-59.

35. D. Pereira, B. Peleteiro, J. Araujo, J. Branco, R.A. Santos, E. Ramos. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. Osteoarthritis Cartilage. 2011;19(11):1270-85.
36. S.L. James, D. Abate, K.H. Abate, S.M. Abay, C. Abbafati, N. Abbasi, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1789-858.
37. L.B. Murphy, C.G. Helmick, T.A. Schwartz, J.B. Renner, G. Tudor, G.G. Koch, et al. One in four people may develop symptomatic hip osteoarthritis in his or her lifetime. Osteoarthritis Cartilage. 2010;18(11):1372-9.
38. Conservatieve behandeling van artrose in heup of knie: Federatie Medisch Specialisten; 2018 [cited 2024 Oct 10]. Available from: https://richtlijnendatabase.nl/richtlijn/artrose_in_heup_of_knie/diagnostiek_heup-_of_knieartrose.html.
39. J.H. Kellgren, J.S. Lawrence. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494-502.
40. L. Gala, J.C. Clohisy, P.E. Beaule. Hip Dysplasia in the Young Adult. J Bone Joint Surg Am. 2016;98(1):63-73.
41. M. Armand, J. Lepisto, K. Tallroth, J. Elias, E. Chao. Outcome of periacetabular osteotomy: joint contact pressure calculation using standing AP radiographs, 12 patients followed for average 2 years. Acta Orthop. 2005;76(3):303-13.
42. J.A. Hipp, N. Sugano, M.B. Millis, S.B. Murphy. Planning acetabular redirection osteotomies based on joint contact pressures. Clin Orthop Relat Res. 1999(364):134-43.
43. D.A. Michaeli, S.B. Murphy, J.A. Hipp. Comparison of predicted and measured contact pressures in normal and dysplastic hips. Med Eng Phys. 1997;19(2):180-6.
44. M. Beck, M. Kalhor, M. Leunig, R. Ganz. Hip morphology influences the pattern of damage to the acetabular cartilage: femoroacetabular impingement as a cause of early osteoarthritis of the hip. J Bone Joint Surg Br. 2005;87(7):1012-8.
45. R. Ganz, M. Leunig, K. Leunig-Ganz, W.H. Harris. The etiology of osteoarthritis of the hip: an integrated mechanical concept. Clin Orthop Relat Res. 2008;466(2):264-72.
46. N.S. Riedstra, F. Boel, M. van Buuren, D. Eygendaal, S. Bierma-Zeinstra, J. Runhaar, R. Agricola. Pincer Morphology Is Not Associated With Hip Osteoarthritis Unless Hip Pain Is Present: Follow-Up Data From a Prospective Cohort Study. Arthritis Care Res (Hoboken). 2024;76(5):644-51.
47. M.N. Kooijman, C.J. Kruithof, C.M. van Duijn, L. Duijts, O.H. Franco, I.M.H. van, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016;31(12):1243-64.
48. M.M.A. van Buuren, N.S. Riedstra, M.A. van den Berg, F. Boel, H. Ahedi, V. Arbabi, et al. Cohort profile: Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) - an international consortium of prospective cohort studies with individual participant data on hip osteoarthritis. BMJ Open. 2024;14(4):e077907.
49. D.T. Felson, R.C. Lawrence, P.A. Dieppe, R. Hirsch, C.G. Helmick, J.M. Jordan, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med. 2000;133(8):635-46.
50. F. Guilak. Biomechanical factors in osteoarthritis. Best Pract Res Clin Rheumatol. 2011;25(6):815-23.
51. F. Saberi Hosnijeh, M. Kavousi, C.G. Boer, A.G. Uitterlinden, A. Hofman, M. Reijman, et al. Development of a prediction model for future risk of radiographic hip osteoarthritis. Osteoarthritis Cartilage. 2018;26(4):540-6.

52. G.E. Thomas, A.J. Palmer, R.N. Batra, A. Kiran, D. Hart, T. Spector, et al. Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women. A 20 year longitudinal cohort study. Osteoarthritis Cartilage. 2014;22(10):1504-10.
53. V.V. Mascarenhas, M.O. Castro, P.D. Afonso, P. Rego, M. Dienst, R. Sutter, et al. The Lisbon Agreement on femoroacetabular impingement imaging-part 2: general issues, parameters, and reporting. Eur Radiol. 2021;31(7):4634-51.
54. M.P. Reiman, R. Agricola, J.L. Kemp, J.J. Heerey, A. Weir, P. van Klij, et al. Consensus recommendations on the classification, definition and diagnostic criteria of hip-related pain in young and middle-aged active adults from the International Hip-related Pain Research Network, Zurich 2018. Br J Sports Med. 2020;54(11):631-41.
55. J.A. Hanson, A.L. Kapron, K.M. Swenson, T.G. Maak, C.L. Peters, S.K. Aoki. Discrepancies in measuring acetabular coverage: revisiting the anterior and lateral center edge angles. J Hip Preserv Surg. 2015;2(3):280-6.
56. A. Aldieri, M. Terzini, A.L. Audenino, C. Bignardi, U. Morbiducci. Combining shape and intensity dxa-based statistical approaches for osteoporotic HIP fracture risk assessment. Comput Biol Med. 2020;127:104093.
57. B.G. Faber, D. Baird, C.L. Gregson, J.S. Gregory, R.J. Barr, R.M. Aspden, et al. DXA-derived hip shape is related to osteoarthritis: findings from in the MrOS cohort. Osteoarthritis Cartilage. 2017;25(12):2031-8.
58. B.G. Faber, R. Ebsim, F.R. Saunders, M. Frysz, J.S. Gregory, R.M. Aspden, et al. Cam morphology but neither acetabular dysplasia nor pincer morphology is associated with osteophytosis throughout the hip: findings from a cross-sectional study in UK Biobank. Osteoarthritis Cartilage. 2021;29(11):1521-9.
59. A.V. Pavlova, F.R. Saunders, S.G. Muthuri, J.S. Gregory, R.J. Barr, K.R. Martin, et al. Statistical shape modelling of hip and lumbar spine morphology and their relationship in the MRC National Survey of Health and Development. J Anat. 2017;231(2):248-59.
60. J.H. Waarsing, R.M. Rozendaal, J.A. Verhaar, S.M. Bierma-Zeinstra, H. Weinans. A statistical model of shape and density of the proximal femur in relation to radiological and clinical OA of the hip. Osteoarthritis Cartilage. 2010;18(6):787-94.
61. K. Yoshida, R.J. Barr, S. Galea-Soler, R.M. Aspden, D.M. Reid, J.S. Gregory. Reproducibility and Diagnostic Accuracy of Kellgren-Lawrence Grading for Osteoarthritis Using Radiographs and Dual-Energy X-ray Absorptiometry Images. J Clin Densitom. 2015;18(2):239-44.
62. F.A. Mettler, Jr., W. Huda, T.T. Yoshizumi, M. Mahesh. Effective doses in radiology and diagnostic nuclear medicine: a catalog. Radiology. 2008;248(1):254-63.
63. B.F. Wall, D. Hart. Revised radiation doses for typical X-ray examinations. Report on a recent review of doses to patients from medical X-ray examinations in the UK by NRPB. National Radiological Protection Board. Br J Radiol. 1997;70(833):437-9.
64. C. Lindner, S. Thiagarajah, J.M. Wilkinson, O.C. arc, G.A. Wallis, T.F. Cootes. Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE Trans Med Imaging. 2013;32(8):1462-72.
65. G. Van Rossum, F.L. Drake. Python 3 Reference Manual: Scotts Valley, CA: CreateSpace; 2009.
66. J.W. Walker. Computing Angle Between Vectors. 2014 [updated 5 June 2016]. Available from: https://www.jwwalker.com/pages/angle-between-vectors.html.
67. A. Al-Sharadqah, N. Chernov. Error analysis for circle fitting algorithms. Electron J Stat. 2009;3:886-911.

68. D. Tönnis. Congenital dysplasia and dislocation of the hip in children and adults: Springer-Verlag Berlin and Heidelberg GmbH & Co. KG.; 1987.
69. R. Agricola, M.P. Heijboer, A.Z. Ginai, P. Roels, A.A. Zadpoor, J.A. Verhaar, 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.
70. P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020;17:261-72.
71. S. Van der Walt, J.L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J.D. Warner, N. Yager, et al. scikit-image: image processing in Python. PeerJ. 2014;2:e453.
72. R.C. Gonzalez, R.E. Woods. Opening and Closing. Digital Image Processing. Fourth Edition ed: Pearson Education Limited; 2018. p. 644-8.
73. M. Gamer, J. Lemon. irr: Various Coefficients of Interrater Reliability and Agreement. 2019.
74. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.; 2016.
75. A.E. Nelson, J.L. Stiller, X.A. Shi, K.M. Leyland, J.B. Renner, T.A. Schwartz, et al. Measures of hip morphology are related to development of worsening radiographic hip osteoarthritis over 6 to 13 year follow-up: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage. 2016;24(3):443-50.
76. B.G. Faber, R. Ebsim, F.R. Saunders, M. Frysz, G. Davey Smith, T. Cootes, et al. Deriving alpha angle from anterior-posterior dual-energy x-ray absorptiometry scans: an automated and validated approach. Wellcome Open Res. 2021;6:60.
77. H. Archer, S. Reine, A. Alshaikhsalama, J. Wells, A. Kohli, L. Vazquez, et al. Artificial intelligence-generated hip radiological measurements are fast and adequate for reliable assessment of hip dysplasia : an external validation study. Bone Jt Open. 2022;3(11):877-84.
78. G.M. Schwarz, S. Simon, J.A. Mitterer, S. Huber, B.J. Frank, A. Aichmair, et al. Can an artificial intelligence powered software reliably assess pelvic radiographs? Int Orthop. 2023;47(4):945-53.
79. C. Stotter, T. Klestil, C. Roder, P. Reuter, K. Chen, R. Emprechtinger, et al. Deep Learning for Fully Automated Radiographic Measurements of the Pelvis and Hip. Diagnostics (Basel). 2023;13(3).
80. J. Jensen, O. Graumann, S. Overgaard, O. Gerke, M. Lundemann, M.H. Haubro, et al. A Deep Learning Algorithm for Radiographic Measurements of the Hip in Adults-A Reliability and Agreement Study. Diagnostics (Basel). 2022;12(11).
81. G.E. Thomas, A. Kiran, R.N. Batra, D. Hart, T. Spector, A. Taylor, et al., The association between hip morphology and end-stage osteoarthritis at 12-year follow up, Osteoarthritis Cartilage 20 (2012) S204.
82. A. Hoch, P. Schenk, T. Jentzsch, S. Rahm, P.O. Zingg, FAI morphology increases the risk for osteoarthritis in young people with a minimum follow-up of 25 years, Arch. Orthop. Trauma Surg. 141 (2021) 1175–1181.
83. M. Harris-Hayes, N.K. Royer, Relationship of acetabular dysplasia and femoroacetabular impingement to hip osteoarthritis: a focused review, PM&R 3 (11) (2011) 1055–67.e1.
84. D.R. Griffin, E.J. Dickenson, J. O'donnell, T. Awan, M. Beck, J.C. Clohisy, et al., The Warwick Agreement on femoroacetabular impingement syndrome (FAI syndrome): an international consensus statement, Br. J. Sports Med. 50 (19) (2016) 1169–76.
85. S.D. Walter, M. Eliasziw, A. Donner, Sample size and optimal designs for reliability studies, Stat. Med. 17 (1) (1998 Jan 15) 101–110.
86. I.Ø. Engesæter, L.B. Laborie, T.G. Lehmann, F. Sera, J. Fevang, D. Pedersen, et al., Radiological findings for hip dysplasia at skeletal maturity. Validation of digital and manual measurement techniques, Skeletal Radiol. 41 (7) (2012 Jul) 775–85.

87. M. Umer, A. Thambyah, W. Tan, S.D. De, Acetabular morphometry for determining hip dysplasia in the Singaporean population, J. Orthop. Surg. 14 (1) (2006) 27–31.
88. P. van Klij, M.P. Reiman, J.H. Waarsing, M. Reijman, W.M. Bramer, J.A.N. Verhaar, et al., Classifying cam morphology by the alpha angle: a systematic review on threshold values, Orthop J Sports Med 8 (8) (2020 Aug 10) 2325967120938312.
89. K.K. Gosvig, S. Jacobsen, H. Palm, S. Sonne-Holm, E. Magnusson, A new radiological index for assessing asphericity of the femoral head in cam impingement, The Journal of Bone & Joint Surgery British 89 (10) (2007) 1309–16.
90. P.N. Ramkumar, J.M. Karnuta, H.S. Haeberle, S.W. Sullivan, D.H. Nawabi, A.S. Ranawat, et al., Radiographic indices are not predictive of clinical outcomes among 1735 patients indicated for hip arthroscopic surgery: a machine learning analysis, Am. J. Sports Med. 48 (12) (2020) 2910–8.
91. F. Boel, S. de Vos-Jakobs, N.S. Riedstra, C. Lindner, J. Runhaar, S.M.A. Bierma-Zeinstra, et al., Automated radiographic hip morphology measurements: An open-access method. Osteoarthr Imaging. 2024 Jun;4(2):100181.
92. G.P. Wilkin, M.M. Ibrahim, K.M. Smit, P.E. Beaulé, A contemporary definition of hip dysplasia and structural instability: toward a comprehensive classification for acetabular dysplasia, J. Arthroplasty 32 (9) (2017) S20–7.
93. M. Tannast, M.S. Hanke, G. Zheng, S.D. Steppacher, K.A. Siebenrock, What are the radiographic reference values for acetabular under-and overcoverage? Clin. Orthop. Relat. Res. 473 (4) (2015) 1234–46.
94. J.D. Harris, B.J. Gerrie, K.E. Varner, D.M. Lintner, P.C. McCulloch, Radiographic prevalence of dysplasia, cam, and pincer deformities in elite ballet, Am. J. Sports Med. 44 (1) (2016) 20–7.
95. H.P. Nötzli, T.F. Wyss, C.H. Stoecklin, M.R. Schmid, K. Treiber, J. Hodler, The contour of the femoral head-neck junction as a predictor for the risk of anterior impingement, J Bone Joint Surg Br 84 (4) (2002 May) 556–60.
96. M. van Buuren, N.K. Arden, S. Bierma-Zeinstra, W.M. Bramer, N.C. Casartelli, D.T. Felson, et al., Statistical shape modeling of the hip and the association with hip osteoarthritis: a systematic review, Osteoarthritis Cartilage 29 (2020) 607–18.
97. N. Arevalo, N. Santamaria, E. Diez, J. Gredilla Molinero, M. Grande Barez, Imaging Findings of Developmental Dysplasia of the Hip in Adults, European Congress of Radiology - ECR, 2016.
98. A.S. Nicholls, A. Kiran, T.C.B. Pollard, D.J. Hart, C.P.A. Arden, T. Spector, et al., The association between hip morphology parameters and nineteen-year risk of end-stage osteoarthritis of the hip: a nested case-control study, Arthritis Rheum. 63 (11) (2011 Nov) 3392–400.
99. W. Yang, Q. Ye, S. Ming, X. Hu, Z. Jiang, Q. Shen, et al., Feasibility of automatic measurements of hip joints based on pelvic radiography and a deep learning algorithm, Eur. J. Radiol. 132 (2020 Nov) 109303.
100. S. Tontanahal, V. Madhuri, Reproducibility of radiographic measurements made in the active stages of legg-calvé-perthes disease: evaluation of a prognostic indicator and an interim outcome measure, J. Pediatr. Orthop. 41 (10) (2021) e938–9.
101. J. Powell, R.F. Gibly, L.W. Faulk, P. Carry, S.W. Mayer, C.M. Selberg, Can EOS imaging substitute for conventional radiography in measurement of acetabular morphology in the young dysplastic hip? J. Pediatr. Orthop. 40 (6) (2020 Jul) 294–9.
102. M.E. Air, J.R. Harrison, J.T. Nguyen, B.T. Kelly, E.A. Bogner, P.J. Moley, Correlation of measurements of the prearthritic hip between plain radiography and computed tomography, Pharm. Manag. PM R 11 (2) (2019 Feb) 158–66.

103. S. Lerch, A. Kasperczyk, T. Berndt, O. Rühmann, Ultrasound is as reliable as plain radiographs in the diagnosis of cam-type femoroacetabular impingement, Arch. Orthop. Trauma Surg. 136 (10) (2016 Oct) 1437–43.
104. N.H. Mast, F. Impellizzeri, S. Keller, M. Leunig, Reliability and agreement of measures used in radiographic evaluation of the adult hip, Clin. Orthop. Relat. Res. 469 (1) (2011 Jan) 188–99.
105. M. Tannast, K.A. Siebenrock, S.E. Anderson, Femoroacetabular impingement: radiographic diagnosis—what the radiologist should know, Am. J. Roentgenol. 188 (6) (2007) 1540–52.
106. W.K. Chung, S. De Vos-Jakobs, F. Rivadeneira, S.M. Bierma-Zeinstra, J.H. Waarsing. The association of BMI and physical activity on acetabular dysplasia in children. Osteoarthritis Cartilage. 2021;29(1):50-8.
107. M. Frysz, B.G. Faber, R. Ebsim, F.R. Saunders, C. Lindner, J.S. Gregory, et al. Machine Learning-Derived Acetabular Dysplasia and Cam Morphology Are Features of Severe Hip Osteoarthritis: Findings From UK Biobank. J Bone Miner Res. 2022;37(9):1720-32.
108. J. Zheng, M. Frysz, B.G. Faber, H. Lin, R. Ebsim, J. Ge, et al. Comparison between UK Biobank and Shanghai Changfeng suggests distinct hip morphology may contribute to ethnic differences in the prevalence of hip osteoarthritis. Osteoarthritis Cartilage. 2023.
109. D.J. Brenner, E.J. Hall. Computed tomography--an increasing source of radiation exposure. N Engl J Med. 2007;357(22):2277-84.
110. J. Damilakis, T.G. Maris, A.H. Karantanas. An update on the assessment of osteoporosis using radiologic techniques. Eur Radiol. 2007;17(6):1591-602.
111. G. Solomou, J. Damilakis. Radiation Exposure in Bone Densitometry. Semin Musculoskelet Radiol. 2016;20(4):392-8.
112. N.B. Smith, A.G. Webb. Introduction to medical imaging : physics, engineering, and clinical applications. United Kingdom: Cambridge University Press; 2011.
113. A. Bazzocchi, F. Ponti, U. Albisinni, G. Battista, G. Guglielmi. DXA: Technical aspects and application. Eur J Radiol. 2016;85(8):1481-92.
114. K.L. Welton, M.K. Jesse, M.J. Kraeutler, T. Garabekyan, O. Mei-Dan. The Anteroposterior Pelvic Radiograph: Acetabular and Femoral Measurements and Relation to Hip Pathologies. J Bone Joint Surg Am. 2018;100(1):76-85.
115. M.A. Ikram, G. Brusselle, M. Ghanbari, A. Goedegebure, M.K. Ikram, M. Kavousi, et al. Objectives, design and main findings until 2020 from the Rotterdam Study. Eur J Epidemiol. 2020;35(5):483-517.
116. J.H. Kellgren, M.R. Jeffrey, J. Ball. The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications; 1963.
117. F. Boel, N.S. Riedstra, J. Tang, D.F. Hanff, H. Ahedi, N. Arden, et al. Reliability and Agreement of Manual and Automated Morphological Radiographic Hip Measurements. Osteoarthr Cartil Open. 2024;6(3):100510.
118. S. Delaunay, R.G. Dussault, P.A. Kaplan, B.A. Alford. Radiographic measurements of dysplastic adult hips. Skeletal Radiol. 1997;26(2):75-81.
119. R. Agricola, J.H. Waarsing, G.E. Thomas, A.J. Carr, M. Reijman, S.M. Bierma-Zeinstra, et al. Cam impingement: defining the presence of a cam deformity by the alpha angle: data from the CHECK cohort and Chingford cohort. Osteoarthritis Cartilage. 2014;22(2):218-25.
120. S.S. Mangiafico. rcompanion: Functions to Support Extension Education Program Evaluation. version 2.4.36 ed2024.
121. D. Datta. blandr: a Bland-Altman Method Comparison package for R. Version 0.6.0, 2024-06-09 ed2017.

122. T. Nishimura, H. Watanabe, N. Taki, I. Kikkawa, K. Takeshita. Standard radiographic values for the acetabulum in Japanese adolescents: a cross-sectional study. BMC Musculoskelet Disord. 2023;24(1):257.
123. R. Bouttier, J. Morvan, B. Mazieres, A.C. Rat, L.E. Ziegler, P. Fardellone, et al. Reproducibility of radiographic hip measurements in adults. Joint Bone Spine. 2013;80(1):52-6.
124. B.J. Hansen, M.D. Harris, L.A. Anderson, C.L. Peters, J.A. Weiss, A.E. Anderson. Correlation between radiographic measures of acetabular morphology with 3D femoral head coverage in patients with acetabular retroversion. Acta Orthop. 2012;83(3):233-9.
125. R. Leide, A. Bohman, D. Wenger, S. Overgaard, C.J. Tiderius, C. Rogmark. Hip dysplasia is not uncommon but frequently overlooked: a cross-sectional study based on radiographic examination of 1,870 adults. Acta Orthop. 2021 Oct;92(5):575-80.
126. J. Herfkens, M.M.A. van Buuren, N.S. Riedstra, J.A.N. Verhaar, V.V. Mascarenhas, R. Agricola. Adding false-profile radiographs improves detection of developmental dysplasia of the hip, data from the CHECK cohort. J Hip Preserv Surg 2022; 9:3–9.
127. F.S. Song, J.J. McCarthy, G.D. MacEwen, K.E. Fuchs, S.E. Dulka. The incidence of occult dysplasia of the contralateral hip in children with unilateral hip dysplasia. J Pediatr Orthop 2008; 28:173–6.
128. D. Tönnis. Normal values of the hip joint for the evaluation of X-rays in children and adults. Clin Orthop Relat Res 1976; 119:39 47.
129. M.D. LaPrade, H.P. Melugin, R.F. Hale, D.P. Leland, C.D. Bernard, R.J. Sierra, et al. Incidence of Hip Dysplasia Diagnosis in Young Patients With Hip Pain: A Geographic Population Cohort Analysis. Orthop J Sports Med. 2021 Mar;9(3):2325967121989087.
130. D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009; 62:1006–12.
131. M.L. Rethlefsen, S. Kirtley, S. Waffenschmidt, A.P. Ayala, D. Moher, M.J. Page, et al.; PRISMA-S Group. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021; 10:39.
132. W.M. Bramer, G.B. de Jonge, M.L. Rethlefsen, F. Mast, J. Kleijnen. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc 2018; 106:531–41.
133. W.M. Bramer, P. Bain. Updating search strategies for systematic reviews using endnote. J Med Libr Assoc 2017; 105:285–9.
134. W.M. Bramer. Reference checking for systematic reviews using endnote. J Med Libr Assoc 2018; 106:542–6.
135. W.M. Bramer, D. Giustini, G.B. de Jonge, L. Holland, T. Bekhuis. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc 2016; 104:240–3.
136. W.M. Bramer, J. Milic, F. Mast. Reviewing retrieved references for inclusion in systematic reviews using endnote. J Med Libr Assoc 2017; 105:84–7.
137. G. Wells, B. Shea, D. O'Connell, J. Peterson, V. Welch, M. Losos, P. Tugwell. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. The Ottawa Hospital Research Institute; 2018.
138. I. Akel, M. Songur, S. Karahan, G. Yilmaz, H.G. Demirkiran, Y. Tumer. Acetabular index values in healthy Turkish children between 6 months and 8 years of age: a cross-sectional radiological study. Acta Orthop Traumatol Turc 2013; 47:38–42.
139. A.I. Tugrul, G. Yilmaz, B.K. Aydin, I. Akel, F. Durgut, H. Senaran. Center-edge angle values in healthy children between 5 and 14 years old in Turkey. Acta Orthop Traumatol Turc 2020; 54:15–9.

140. Y.Y. Shi, T.J. Liu, Q. Zhao, L.J. Zhang, S.J. Ji, E.B. Wang. The normal centre-edge angle of Wiberg in the Chinese population: a population-based cross-sectional study. J Bone Joint Surg Br 2010; 92:1144–7.
141. C.H. Chang, Y.T. Chiang, L. Chen, K.N. Kuo. The influence of health policy on early diagnosis and surgical incidence of developmental dysplasia of the hip. PLoS One 2018; 13:e0200995.
142. W. Szulc. The frequency of occurrence of congenital dysplasia of the hip in Poland. Clin Orthop Relat Res 1991; 272:100–2.
143. C. Broadhurst, A.M.L. Rhodes, P. Harper, D.C. Perry, N.M.P. Clarke, A. Aarvold. What is the incidence of late detection of developmental dysplasia of the hip in England?: a 26-year national study of children diagnosed after the age of one. Bone Joint J 2019; 101-B:281–7.
144. K.J. Donnelly, K.W. Chan, A.P. Cosgrove. Delayed diagnosis of developmental dysplasia of the hip in Northern Ireland: can we do better? Bone Joint J 2015; 97-B:1572–6.
145. R.T. Loder, E.N. Skopelja. The epidemiology and demographics of hip dysplasia. ISRN Orthop 2011; 2011:238607.
146. C. Dezateux, K. Rosendahl. Developmental dysplasia of the hip. Lancet. 2007;369:1541-52.
147. S. de Vos-Jakobs, F. Boel, W.M. Bramer, S.M.A. Bierma-Zeinstra, R. Agricola. Prevalence and radiological definitions of acetabular dysplasia after the age of 2 years: a systematic review. J Pediatr Orthop B. 2024 Jul 1;33(4):334-9.
148. C.J. Kruithof, M.N. Kooijman, C.M. van Duijn, O.H. Franco, J.C. de Jongste, C.C. Klaver, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014 Dec;29(12):911-27.
149. Statistics Netherlands. Nieuwe indeling bevolking naar herkomst [Internet]. Centraal Bureau voor de Statistiek. 2022. [cited 2022 Feb 16]. Available from: https://www.cbs.nl/nl-nl/longread/statistische-trends/2022/nieuwe-indeling-bevolking-naar-herkomst?onepage=true
150. F. Boel, J. Wortel, M.M.A. van Buuren, F. Rivadeneira, J.B.J. van Meurs, J. Runhaar, et al. DXA images vs. pelvic radiographs: Reliability of hip morphology measurements. Osteoarthritis Cartilage. 2025 Feb;33(2):283-92.
151. G. Wiberg. Studies on Dysplastic Acetabula and Congenital Subluxation of the Hip Joint with Special Reference to the Complication of Osteo-Arthritis. Journal of the American Medical Association. 1940;115:81-81.
152. R.M. Acheson. The Oxford method of assessing skeletal maturity. Clin Orthop. 1957;10:19-39.
153. C. Palazzo, J.F. Ravaud, A. Papelard, P. Ravaud, S. Poiraudeau. The Burden of Musculoskeletal Conditions. PLoS ONE. 2014;9:e90633.
154. S.M. Freiman, M.T. Schwabe, L. Fowler, J.C. Clohisy, J.J. Nepple. Prevalence of Borderline Acetabular Dysplasia in Symptomatic and Asymptomatic Populations: A Systematic Review and Meta-analysis. Orthop J Sports Med. 2022;10:23259671211040455.
155. M.B. Millis, M. McClincy. Periacetabular osteotomy to treat residual dysplasia in adolescents and young adults: indications, complications, results. J Child Orthop. 2018;12:349-57.
156. S.D. Steppacher, M. Tannast, R. Ganz, K.A.0 Siebenrock. Mean 20-year follow-up of Bernese periacetabular osteotomy. Clin Orthop Relat Res. 2008;466:1633-44.
157. Collaborators GBDO. Global, regional, and national burden of osteoarthritis, 1990-2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol 2023; 5(9): e508-22.
158. S. Glyn-Jones, A.J. Palmer, R. Agricola, A.J. Price, T.L. Vincent, H. Weinans, et al. Osteoarthritis. Lancet 2015; 386(9991): 376-87.
159. R. Agricola, M.M.A. van Buuren, J.L. Kemp, H. Weinans, J. Runhaar, S.M.A. Bierma-Zeinstra. Femoroacetabular impingement syndrome in middle-aged individuals is strongly associated with the

development of hip osteoarthritis within 10-year follow-up: a prospective cohort study (CHECK). Br J Sports Med 2024; 58(18): 1061-7.
160. C. Pascual-Garrido, D.J. Li, G. Grammatopoulos, E.L. Yanik, J.C. Clohisy, ANCHOR Group. The Pattern of Acetabular Cartilage Wear Is Hip Morphology-dependent and Patient Demographic-dependent. Clin Orthop Relat Res 2019; 477(5): 1021-33.
161. A.S. Nicholls, A. Kiran, T.C. Pollard, D.J. Hart, C.P. Arden, T. Spector, et al. The association between hip morphology parameters and nineteen-year risk of end-stage osteoarthritis of the hip: a nested case-control study. Arthritis Rheum 2011; 63(11): 3392-400.
162. N.S. Riedstra, M.M.A. Van Buuren, F. Boel, H. Ahedi, V. Arbabi, N. Arden, et al. Pincer Morphology Is A Risk Factor For Developing Radiographic Hip Osteoarthritis; Data From The World Coach Consortium. Osteoarthritis and Cartilage 2023; 31: S249-50.
163. R. Agricola, M.P. Heijboer, A.Z. Ginai, P. Roels, A.A. Zadpoor, J.A.N. Verhaar, 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.
164. P. van Klij, M.P. Heijboer, A.Z. Ginai, J.A.N. Verhaar, J.H. Waarsing, R. Agricola. 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(9): 532-8.
165. E. von Elm, D.G. Altman, M. Egger, S.J. Pocock, P.C. Gotzsche, J.P. Vandenbroucke, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-7.
166. S. Dorai-Raj. binom: Binomial Confidence Intervals for Several Parameterizations. 2022.
167. L.B. Laborie, T.G. Lehmann, I.O. Engesaeter, D.M. Eastwood, L.B. Engesaeter, K. Rosendahl. Prevalence of radiographic findings thought to be associated with femoroacetabular impingement in a population-based cohort of 2081 healthy young adults. Radiology 2011; 260(2): 494-502.
168. C.M. Larson, A. Moreau-Gaudry, B.T. Kelly, J.W. Byrd, J. Tonetti, S. Lavallee, et al. Are normal hips being labeled as pathologic? A CT-based method for defining normal acetabular coverage. Clin Orthop Relat Res 2015; 473(4): 1247-54.
169. J.J. Nepple, C.L. Lehmann, J.R. Ross, P.L. Schoenecker, J.C. Clohisy. Coxa profunda is not a useful radiographic parameter for diagnosing pincer-type femoroacetabular impingement. J Bone Joint Surg Am 2013; 95(5): 417-23.
170. V.V. Mascarenhas, M.O. Castro, P.A. Rego, R. Sutter, L.M. Sconfienza, A. Kassarjian, et al. The Lisbon Agreement on Femoroacetabular Impingement Imaging-part 1: overview. Eur Radiol. 2020;30(10):5281-97.
171. S. de Vos-Jakobs, F. Boel, D. Chen, J.C.M. van Haasteren, J.J. Tolk, S.M.A. Bierma-Zeinstra, R. Agricola. Prevalence of acetabular dysplasia in early adolescents in a general population: the Generation R study. *Submitted.*
172. A. Niklasson, K. Albertsson-Wikland. Continuous growth reference from 24th week of gestation to 24 months by gender. BMC Pediatr. 2008;8:8.
173. T.J. Cole, T. Lobstein. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes. 2012;7(4):284-94.
174. E.J. Troe, H. Raat, V.W. Jaddoe, A. Hofman, E.A. Steegers, F.C. Verhulst, et al. Smoking during pregnancy in ethnic populations: the Generation R study. Nicotine Tob Res. 2008;10(8):1373-84.
175. D.H. Heppe, C. Medina-Gomez, J.C. de Jongste, H. Raat, E.A. Steegers, A. Hofman, et al. Fetal and childhood growth patterns associated with bone mass in school-age children: the Generation R Study. J Bone Miner Res. 2014;29(12):2584-93.

176. R. Meinert, T. Frischer, J. Kuehr. Assessing the effect of time-varying covariates in cross-sectional studies. J Clin Epidemiol. 1994;47(9):983-91.
177. I.R. White, P. Royston, A.M. Wood. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377-99.
178. R.C. Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021.
179. S. van Buuren, K. Groothuis-Oudshoorn. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011;45(3):1-67.
180. J. Fox, S. Weisberg. An R Companion to Applied Regression. Third ed: Sage, Thousand Oaks CA; 2019.
181. L.B. Laborie, S.A. Lie, K. Rosendahl. Radiographic markers of hip dysplasia in young adults: predictive effect of factors in early life. BMC Musculoskelet Disord. 2023;24(1):119.
182. M. Kaya, T. Suzuki, M. Emori, T. Yamashita. Hip morphology influences the pattern of articular cartilage damage. Knee Surg Sports Traumatol Arthrosc. 2016;24(6):2016-23.
183. N.S. Riedstra, F. Boel, M.M.A. van Buuren, H. Ahedi, V. Arbabi, N. Arden, et al. Acetabular dysplasia and the risk of developing hip osteoarthritis within 4-8 years: An individual participant data meta-analysis of 18,807 hips from the World COACH consortium. Osteoarthritis Cartilage. 2025 Mar;33(3):373-82.
184. N.S. Riedstra, F. Boel, M.M.A. van Buuren, H. Ahedi, V. Arbabi, N.K. Arden, S.M.A. Bierma-Zeinstra, F.M. Cicuttini, T.F. Cootes, K. Crossley, D.T. Felson, W.P. Gielis, J. Heerey, G. Jones, S. Kluzek, N.E. Lane, C. Lindner, J. Lynch, J. van Meurs, A. Mosler, A.E. Nelson, M.C. Nevitt, E.H. Oei, J. Runhaar, J. Tang, H. Weinans, R. Agricola. Pincer morphology is associated with 2 incident hip osteoarthritis: prospective 3 individual participant data from 18,935 4 hips from the World COACH consortium. *Under review.*
185. F. Boel, S. de Vos-Jakobs, D. Chen, J. Runhaar, S.M.A. Bierma-Zeinstra, R. Agricola. Known risk factors of developmental dysplasia of the hip are not associated with early adolescent acetabular dysplasia: findings from a prospective open population cohort.
186. A.J. Teichtahl, Y. Wang, S. Smith, A.E. Wluka, M. Zhu, D. Urquhart, et al. Bone geometry of the hip is associated with obesity and early structural damage--a 3.0 T magnetic resonance imaging study of community-based adults. Arthritis Res Ther. 2015;17(1):112.
187. E.N. Novais, S.J. Shefelbine, K.P. Kienle, P.E. Miller, G. Bowen, Y.J. Kim, S.D. Bixby. Body Mass Index Affects Proximal Femoral but Not Acetabular Morphology in Adolescents Without Hip Pathology. J Bone Joint Surg Am. 2018;100(1):66-74.
188. P. Van Geel, W. Cools, A. Laumen. Evolution of head-shaft angle and neck-shaft angle in childhood. Acta Orthop Belg. 2021;87(3):435-41.
189. Chen D, F. Boel, P. van Klij, M.M.A. van Buuren, S.M.A. Bierma-Zeinstra, R. Agricola. The prevalence of primary cam morphology in early adolescents from the general population: A population-based study (Generation R). Osteoarthritis Cartilage. [published online: February 17, 2025]. 10.1016/j.joca.2025.02.775
190. J. Pinheiro, D. Bates, R.C. Team. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-166 ed2024.
191. D. Lüdecke. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software. 2018;3(26):772.
192. P. Than, T. Sillinger, J. Kranicz, A. Bellyei. Radiographic parameters of the hip joint from birth to adolescence. Pediatr Radiol. 2004;34(3):237-44.

193. L.B. Laborie, I.O. Engesaeter, T.G. Lehmann, F. Sera, C. Dezateux, L.B. Engesaeter, K. Rosendahl. Radiographic measurements of hip dysplasia at skeletal maturity--new reference intervals based on 2,038 19-year-old Norwegians. Skeletal Radiol. 2013;42(7):925-35.
194. M.D. Taussig, K.P. Powell, H.A. Cole, S.K. Nwosu, T. Hunley, S.E. Romine, et al. Prevalence of Hypertension in Pediatric Tibia Vara and Slipped Capital Femoral Epiphysis. J Pediatr Orthop. 2016;36(8):877-83.
195. D. Ke, D. Lu, G. Cai, J. Zhang, X. Wang, K. Suzuki. Accelerated skeletal maturation is associated with overweight and obesity as early as preschool age: a cross-sectional study. BMC Pediatr. 2020;20(1):452.
196. A.B. Sopher, A.M. Jean, S.K. Zwany, D.M. Winston, C.B. Pomeranz, J.J. Bell, et al. Bone age advancement in prepubertal children with obesity and premature adrenarche: possible potentiating factors. Obesity (Silver Spring). 2011;19(6):1259-64.
197. D.J. Hunter, L. March, M. Chew. Osteoarthritis in 2020 and beyond: a Lancet Commission. Lancet. 2020 Nov 28;396(10264):1711-2.
198. M. Hiligsmann, C. Cooper, N. Arden, M. Boers, J.C. Branco, M. Luisa Brandi, et al. Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Semin Arthritis Rheum. 2013 Dec;43(3):303-13.
199. W.H. Harris. Etiology of osteoarthritis of the hip. Clin Orthop Relat Res. 1986 Dec;(213):20-33.
200. J.C. Baker-LePain, N.E. Lane. Relationship between joint shape and the development of osteoarthritis. Curr Opin Rheumatol. 2010 Sep;22(5):538-43.
201. R. Agricola, M.P. Heijboer, S.M.A. Bierma-Zeinstra, J.A.N. Verhaar, H. Weinans, J.H. Waarsing. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). Ann Rheum Dis. 2013 Jun;72(6):918-23.
202. S. Jacobsen, S. Sonne-Holm. Hip dysplasia: a significant risk factor for the development of hip osteoarthritis. A cross-sectional survey. Rheumatology (Oxford). 2005 Feb;44(2):211-8.
203. R. Ganz, J. Parvizi, M. Beck, M. Leunig, H. Nötzli, K.A. Siebenrock. Femoroacetabular impingement: a cause for osteoarthritis of the hip. Clin Orthop Relat Res. 2003 Dec; (417):112-20.
204. J. Wesseling, M. Boers, M.A. Viergever, W.K. Hilberdink, F.P. Lafeber, J. Dekker, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. Int J Epidemiol. 2016 Feb;45(1):36-44.
205. J.M. Jordan, G.F. Linder, J.B. Renner, J.G. Fryer. The impact of arthritis in rural populations. Arthritis Care Res. 1995 Dec;8(4):242-50.
206. N.A. Segal, M.C. Nevitt, K.D. Gross, J. Hietpas, N.A. Glass, C.E. Lewis, et al. The Multicenter Osteoarthritis Study: opportunities for rehabilitation research. PM R. 2013 Aug;5(8):647-54.
207. M.C. Nevitt, D.T. Felson, G. Lester. The Osteoarthritis Initiative. Protocol for the cohort study.;V 1.1 6.21.06 (National Institute of Arthritis, Musculoskeletal and Skin Diseases.) Date accessed: 12-09-23.
208. A. Hofman A, S Darwish Murad, C.M. van Duijn, O.H. Franco, A. Goedegebure, M.A. Ikram, et al. The Rotterdam Study: 2014 objectives and design update. Eur J Epidemiol. 2013 Nov;28(11):889-926.
209. P. Croft, C. Cooper, C. Wickham, D. Coggon. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol. 1990 Sep;132(3):514-22.
210. G.B. Joseph, J.F. Hilton, P.M. Jungmann, J.A. Lynch, N.E. Lane, F. Liu, et al. Do persons with asymmetric hip pain or radiographic hip OA have worse pain and structure outcomes in the knee opposite the more affected hip? Data from the Osteoarthritis Initiative. Osteoarthritis Cartilage. 2016 Mar;24(3):427-35.

211. N.E. Lane. We Challenged the Kellgren and Lawrence Radiographic Scoring Method and Came Up With Some Interesting Epidemiology for Osteoarthritis of the Hip. HSS J. 2023 Nov;19(4):402-6.
212. A.G. Culvenor, C.N. Engen, B.E. Øiestad, L. Engebretsen, M.A. Risberg. Defining the presence of radiographic knee osteoarthritis: a comparison between the Kellgren and Lawrence system and OARSI atlas criteria. Knee Surg Sports Traumatol Arthrosc. 2015 Dec;23(12):3532-9.
213. N.E. Lane, M.C. Nevitt, M.C. Hochberg, Y.Y. Hung, L. Palermo. Progression of radiographic hip osteoarthritis over eight years in a community sample of elderly white women. Arthritis Rheum. 2004 May;50(5):1477-86.
214. D. Bates, M. Mächler, B. Bolker, S. Walker. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Soft. 2015 Oct;67(1):1-48.
215. D. Rizopoulos. Generalized Linear Mixed Models using Adaptive Gaussian Quadrature.
216. M.C. Castaño-Betancourt, J.B.J. Van Meurs, S.M.A. Bierma-Zeinstra, F. Rivadeneira, A. Hofman, H. Weinans, et al. The contribution of hip geometry to the prediction of hip osteoarthritis. Osteoarthritis Cartilage. 2013 Oct;21(10):1530-6.
217. D.R. Cooperman, R. Wallensten, S.D. Stulberg. Acetabular dysplasia in the adult. Clin Orthop Relat Res. 1983 May;(175):79-85.
218. S. Jacobsen, S. Sonne-Holm, K. Søballe, P. Gebuhr, B. Lund. Joint space width in dysplasia of the hip: a case-control study of 81 adults followed for ten years. J Bone Joint Surg Br. 2005 Apr;87(4):471-7.
219. A.S. Nair. Publication bias - Importance of studies with negative results. Indian J Anaesth. 2019 Jun;63(6):505-7.
220. M. Reijman, J.M. Hazes, B.W. Koes, A.P. Verhagen, S.M.A. Bierma-Zeinstra. Validity, reliability, and applicability of seven definitions of hip osteoarthritis used in epidemiological studies: a systematic appraisal. Ann Rheum Dis. 2004 Mar;63(3):226-32.
221. N.E. Lane, P. Lin, L. Christiansen, L.R. Gore, E.N. Williams, M.C. Hochberg, 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 Feb;43(2):400-4.
222. T. Iidaka, S. Muraki, H. Oka, C. Horii, H. Kawaguchi, K. Nakamura, et al. Incidence rate and risk factors for radiographic hip osteoarthritis in Japanese men and women: a 10-year follow-up of the ROAD study. Osteoarthritis Cartilage. 2020 Feb;28(2):182-8.
223. R.D. Riley, T.P.A. Debray, D. Fisher, M. Hattle, N. Marlin, J. Hoogland, et al. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning. Stat Med. 2020 Jul 10;39(15):2115-37.
224. S. Sud, J. Douketis. The devil is in the details...or not? A primer on individual patient data meta-analysis. Evid Based Med. 2009 Aug;14(4):100-1.
225. S. Bujkiewicz, J.R. Thompson, A.J. Sutton, N.J. Cooper, M.J. Harrison, D.P. Symmons, et al. Multivariate meta-analysis of mixed outcomes: a Bayesian approach. Stat Med. 2013 Sep 30;32(22):3926-43.
226. F.S. Bessa, B.T. Williams, E.M. Polce, B. Maheshwer, J.C. Williams, S.J. Nho, et al. No Differences in Hip Joint Space Measurements Between Weightbearing or Supine Anteroposterior Pelvic Radiographs. Arthroscopy. 2020 Nov;36(11):2843-8.
227. R. Wynne-Davies. Acetabular dysplasia and familial joint laxity: two etiological factors in congenital dislocation of the hip. A review of 589 patients and their families. J Bone Joint Surg Br. 1970 Nov;52(4):704-16.

228. C. Cooper, H. Inskip, P. Croft, L. Campbell, G. Smith, M. McLaren, et al. Individual risk factors for hip osteoarthritis: obesity, hip injury, and physical activity. Am J Epidemiol. 1998 Mar 15;147(6):516-22.
229. C. Palazzo, C. Nguyen, M.M. Lefevre-Colau, F. Rannou, S. Poiraudeau. Risk factors and burden of osteoarthritis. Ann Phys Rehabil Med. 2016 Jun;59(3):134-8.
230. V.K. Srikanth, J.L. Fryer, G. Zhai, T.M. Winzenberg, D. Hosmer, G. Jones. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthritis Cartilage. 2005 Sep;13(9):769-81.
231. A.M. Lievense, S.M.A. Bierma-Zeinstra, A.P. Verhagen, R.M. Bernsen, J.A.N. Verhaar, B.W. Koes. Influence of sporting activities on the development of osteoarthritis of the hip: a systematic review. Arthritis Rheum. 2003 Apr 15;49(2):228-36.
232. H.P. Dijkstra, S. Mc Auliffe, C.L. Ardern, J.L. Kemp, A.B. Mosler, A. Price, et al. Oxford consensus on primary cam morphology and femoroacetabular impingement syndrome: part 1-definitions, terminology, taxonomy and imaging outcomes. Br J Sports Med. 10.1136/bjsports-2022-106085
233. R. Agricola, J.H. Waarsing, N.K. Arden, A.J. Carr, S.M. Bierma-Zeinstra, G.E. Thomas, et al. Cam impingement of the hip: a risk factor for hip osteoarthritis. Nat Rev Rheumatol. 2013 Oct;9(10):630-4.
234. L. Jiang, J. Rong, Y. Wang, F. Hu, C. Bao, X. Li, et al. The relationship between body mass index and hip osteoarthritis: a systematic review and meta-analysis. Joint Bone Spine. 2011 Mar;78(2):150-5.
235. L.K. Smith, J.L. Lelas, D.C. Kerrigan. Gender differences in pelvic motions and center of mass displacement during walking: stereotypes quantified. J Womens Health Gend Based Med. 2002 Jun;11(5):453-8.
236. S. Monazzam, J.D. Bomar, K. Cidambi, P. Kruk, H. Hosalkar. Lateral center-edge angle on conventional radiography and computed tomography. Clin Orthop Relat Res. 2013 Jul;471(7):2233-7.
237. M.A. Röling, N.M.C. Mathijssen, R.M. Bloem. Diagnostic sensitivity and specificity of dynamic three-dimensional CT analysis in detection of cam and pincer type femoroacetabular impingement. BMC Musculoskelet Disord. 2020 Jan 16;21(1):37.
238. A. Henebry, T. Gaskill. The effect of pelvic tilt on radiographic markers of acetabular coverage. Am J Sports Med. 2013;41(11):2599-603.
239. R.T. Li, E. Hu, H. Gould, N. Valentin, M.J. Salata, R.W. Liu. Does Pelvic Rotation Alter Radiologic Measurement of Anterior and Lateral Acetabular Coverage? Arthroscopy. 2019;35(4):1111-6 e1.
240. M. Tannast, S. Fritsch, G. Zheng, K.A. Siebenrock, S.D. Steppacher. Which radiographic hip parameters do not have to be corrected for pelvic rotation and tilt? Clin Orthop Relat Res. 2015;473(4):1255-66.
241. S. Monazzam, M. Agashe, H.S. Hosalkar. Reliability of overcoverage parameters with varying morphologic pincer features: comparison of EOS(R) and radiography. Clin Orthop Relat Res. 2013;471(8):2578-85.
242. S. Jacobsen, S. Sonne-Holm, K. Soballe, P. Gebuhr, B. Lund. Hip dysplasia and osteoarthrosis: a survey of 4151 subjects from the Osteoarthrosis Substudy of the Copenhagen City Heart Study. Acta Orthop. 2005;76(2):149-58.
243. Hunter J. Muir, J. Benson, J. Napenas, J. Vigdorchik. Correcting for pelvic rotation on anteroposterior pelvic radiographs. Orthopaedic Proceedings. 2020;102-B(SUPP_2):54-54.
244. J.M. Muir, J. Vincent, J. Schipper, V.D. Gobin, M. Govindarajan, K. Fiaes, J. Vigdorchik. A Novel Method for Correcting Pelvic Tilt on Anteroposterior Pelvic Radiographs. Cureus. 2019;11(12):e6274.

245. K.A. Siebenrock, D.F. Kalbermatten, R. Ganz. Effect of pelvic tilt on acetabular retroversion: a study of pelves from cadavers. Clin Orthop Relat Res. 2003(407):241-8.
246. J.D. O'Connor, M. Rutherford, J.C. Hill, D.E. Beverland, N.J. Dunne, A.B. Lennon. Effect of combined flexion and external rotation on measurements of the proximal femur from anteroposterior pelvic radiographs. Orthop Traumatol Surg Res. 2018;104(4):449-54.
247. J. Tatka, D. Delagrammaticas, B.R. Kemler, S.I. Rosenberg, A.W. Brady, A.R. Bryniarski, et al. A new understanding of radiographic landmarks of the greater trochanter that indicate correct femoral rotation for measurement of femoral offset. Arthroplasty. 2022;4(1):21.
248. S. Monazzam, J.D. Bomar, M. Agashe, H.S. Hosalkar. Does femoral rotation influence anteroposterior alpha angle, lateral center-edge angle, and medial proximal femoral angle? A pilot study. Clin Orthop Relat Res. 2013;471(5):1639-45.
249. H.D. Aitken, J.E. Goetz, W.M. Sailer, D.J.L. Rivas, C.L. Farnsworth, V.V. Upasani, J.B. Holt. Three-dimensional assessment of subchondral arc and hip joint coverage angles in the asymptomatic young adult hip. Clin Anat. 2024.
250. P.E. Beaule, E. Zaragoza, K. Motamedi, N. Copelan, F.J. Dorey. Three-dimensional computed tomography of the hip in the assessment of femoroacetabular impingement. J Orthop Res. 2005;23(6):1286-92.
251. K. Ito, M.A. Minka, 2nd, M. Leunig, S. Werlen, R. Ganz. Femoroacetabular impingement and the cam-effect. A MRI-based quantitative anatomical study of the femoral head-neck offset. J Bone Joint Surg Br. 2001;83(2):171-6.
252. J. Tang, F. Boel, M.M.A. van Buuren, N.S. Riedstra, J. Runhaar, S. Bierma-Zeinstra, R. Agricola. The different subtypes of cam morphology as defined by statistical shape modeling and their relationship with the development of hip osteoarthritis: A nationwide prospective cohort study (CHECK) with 10 years follow-up. Osteoarthritis Cartilage. 2024;32(12):1647-54.
253. C.L. Maddock, S. Noor, A. Kothari, C.S. Bradley, S.P. Kelley. Reliability of the sourcil method of acetabular index measurement in developmental dysplasia of the hip. J Child Orthop. 2019;13(2):167-71.
254. M. Hassanzad, K. Hajian-Tilaki. Methods of determining optimal cut-point of diagnostic biomarkers with application of clinical data in ROC analysis: an update review. BMC Med Res Methodol. 2024;24(1):84.
255. J.C. Boyd. Defining laboratory reference values and decision limits: populations, intervals, and interpretations. Asian J Androl. 2010;12(1):83-90.
256. T. Cheok, T. Smith, K. Wills, M.P. Jennings, J. Rawat, B. Foster. Universal screening may reduce the incidence of late diagnosis of developmental dysplasia of the hip : a systematic review and meta-analysis. Bone Joint J. 2023;105-B(2):198-208.
257. H.W. Jung, W.Y. Jang. Effectiveness of different types of ultrasonography screening for developmental dysplasia of the hip: A meta-analysis. Medicine (Baltimore). 2020;99(50):e23562.
258. L.B. Laborie, K. Rosendahl, A. Dhouib, P. Simoni, P. Toma, A.C. Offiah. The effect of selective ultrasound screening on the incidence of late presentation of developmental hip dysplasia-a meta-analysis. Pediatr Radiol. 2023;53(10):1977-88.
259. D. Shorter, T. Hong, D.A. Osborn. Screening programmes for developmental dysplasia of the hip in newborn infants. Cochrane Database Syst Rev. 2011;2011(9):CD004595.
260. Rijnmond Primary Care Database; [cited 2024 Dec 19] Available from: https://www.rijnmondgezond.nl/.

261. G.C. Wendel-Vos, A.J. Schuit, W.H. Saris, D. Kromhout. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol. 2003 Dec;56(12):1163-9.
262. M.J. van der Bom, M.E. Groote, K.L. Vincken, F.J. Beek, L.W. Bartels. Pelvic rotation and tilt can cause misinterpretation of the acetabular index measured on radiographs. Clin Orthop Relat Res. 2011;469(6):1743-9.
263. C. Sayon-Orea, C. Moreno-Iribas, J. Delfrade, M. Sanchez-Echenique, P. Amiano, E. Ardanaz, et al. Inverse-probability weighting and multiple imputation for evaluating selection bias in the estimation of childhood obesity prevalence using data from electronic health records. BMC Med Inform Decis Mak. 2020; 20(1): 9.
264. W.M. van der Wal and R.B. Geskus. ipw: An R Package for Inverse Probability Weighting. Journal of Statistical Software. 2011; 43(13): 1-23.

LIST OF ABBREVIATIONS

2D	two-dimensional
3D	three-dimensional
AA	alpha angle
AD	acetabular dysplasia
ADR	acetabular depth-width ratio
AI	acetabular index
aOR	adjusted odd ratio
AP	anteroposterior
AR	absolute risk
ASM	automatic search model
BMI	body mass index
CEA	center edge angle
CHECK	Cohort Hip and Cohort Knee
CI	confidence interval
CT	computed tomography
DDH	developmental dysplasia of the hip
DRR	digitally reconstructed radiograph
DXA	dual-energy x-ray absoptiometry
EI	extrusion index
FAIs	femoroacetabular impingement syndrome
FNI	femoral neck isthmus
FORCe	Femoroacetabular impingement and hip osteoarthritis cohort
HRLP	horizontal reference line of the pelvis
HSA	head-shaft angle
ICC	intraclass correlation coefficient
IOTF	international obesity task force BMI cut-offs
IPD	individual participant data
JoCo	Johnston County Project
JSN	joint space narrowing
JSW	joint space width
KL	Kellgren-Lawrence

LA	lateral point of the bony acetabulum
LCEA	lateral center edge angle
LGP	longitudinal growth plate
mAI	modified acetabular index
MeSH	medical subject headings
MICE	multiple imputation by chained equations
MOST	Multi-center Osteoarthritis Study
MRI	magnetic resonance images
MS	medial point of the sourcil
NA	not applicable
NOS	Newcastle-Ottawa scale
NSA	neck-shaft angle
OA	osteoarthritis
OAI	Osteo Arthritis Initiative
OARSI	osteoarthritis research society international
OR	odds ratio
PRISMA	preferred reporting items for systematic reviews and meta-analyses
r	radius
RHOA	radiographic hip osteoarthritis
ROI	region of interest
RPCD	Rijnmond Primary Care Database
RR	relative risk
RS	Rotterdam Study
SD	standard deviation
SOF	Study of Osteoporotic Fractures
SSM	statistical shape model
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TASOAC	Tasmanian Older Adults Cohort
TC	triradiate cartilage
TCO	triradiate cartilage orientation
TD	teardrop
TGP	greater trochanter growth plate
THR	total hip replacement

TI	triangular index
TIR	triangular index ratio
WCEA	center edge angle of Wiberg
World COACH	Worldwide Collaboration on OsteoArthritis prediCtion for the Hip

DANKWOORD

In dit dankwoord wil ik mijn oprechte dank uitspreken aan iedereen die heeft bijgedragen aan dit proefschrift. Niet alleen aan het proefschrift en het onderzoek zelf, maar ook aan alles eromheen. Een aantal personen wil ik in het bijzonder bedanken.

Beste **Dr. Agricola**, beste Rintje, dank voor alles wat ik de afgelopen jaren van je heb geleerd, en vooral ook het vertrouwen om mijzelf als onderzoeker te ontwikkelen. Dankjewel dat je me naast het werk in Generation R ook betrokken hebt in het World COACH consortium. Jouw fijne begeleiding en enthousiasme voor onderzoek zullen me altijd bijblijven, net als de vele mooie ervaringen tijdens de congressen.

Beste **Dr. Runhaar**, beste Jos, wat ben ik blij dat ook jij onderdeel bent geworden van mijn supervisie team. Jouw kritische blik en vragen daagden me telkens uit om verder te denken. Ik probeer nu op dezelfde manier mijn studenten aan het denken te zetten. Dank voor de fijne begeleiding, ik kijk terug op mooie herinneringen, zoals de IWOAI.

Beste **Prof. dr. Bierma-Zeinstra**, beste Sita, hartelijk dank voor het vertrouwen en de kans om aan dit avontuur te beginnen. Wat begon als een onderzoek plek voor twee jaar, groeide uit tot een breed PhD traject waarin ik ook mijn technische kant kon ontwikkelen. Jouw steun en begeleiding waren daarin onmisbaar.

Mijn dank gaat ook uit naar de andere leden van de promotiecommissie, **Dr. M. van Middelkoop**, **Dr. J.F. Felix** en **Prof. dr. J.N. Doornberg**, voor hun tijd en aandacht als lid van de beoordelingscommissie. Daarnaast wil ik de opponenten tijdens de verdediging van dit proefschrift danken voor hun deelname en tijd.

Alle coauteurs wil ik bedanken voor hun ideeën, samenwerking en waardevolle feedback. In het bijzonder het World COACH team (**Michiel, Jinchi, Noor en Myrthe**) voor de gezellige samenwerking en mooie belevenissen, onder andere tijdens de vele OARSI congressen. Daarnaast wil ik natuurlijk ook het HIPSTAR team bedanken voor de samenwerking en de vele discussies die we samen hebben gevoerd.

Lieve collega's - **Abi, Annika, Brechtje, Britt, Derek, Delong, Eline, Erin, Floris, Irene, Jinchi, Josh, Lichelle, Mariska, Mark, Merel, Michiel, Mirthe, Myrthe, Nadine, Niels, Ninka, Noor, Tjerk, Wouter en Zhongcan** - bedankt voor de hulp, discussie en alle

leuke momenten die we samen hebben mogen beleven. Ik heb enorm genoten van alle potjes tafelvoetbal, koffiemomenten en lunchgesprekken, en vooral alle gezellige uitjes.

Annet en **Simone**, dank voor jullie praktische hulp en het altijd beschikbaar zijn bij vragen en oplossingen. Jullie zijn een onmisbare steun voor mij en vele andere onderzoekers.

Lieve vrienden en familie, bedankt voor jullie belangstelling, hulp en steun of het nu ging om stoom afblazen, successen vieren, nieuwe ideeën opdoen of gewoon samenzijn (met een biertje). Lieve handbalteams en -vrienden , dank voor alle gezellige trainingen, wedstrijden, toernooien en uitjes. Lieve **Maartje**, na jaren samen strijden op het veld ben ik blij dat jij als paranimf naast me staat tijdens de verdediging.

Lieve Entropietjes, **Derek, Lisa** en **Sybren**, wat hebben we samen al veel beleefd. En dat begon allemaal met een zak spekjes tijdens wiskunde in ons eerste studiejaar. Ik kijk uit naar alle avonturen die we (samen) nog gaan beleven.

Lieve **Anne, Pauline, Jacqueline, Paul,** en **Lien**, dank voor alle interesse in mijn proefschrift, maar bovenal alle steun en de mooie momenten samen.

Lieve **pap** en **mam**, op jullie onvoorwaardelijke steun mag ik altijd rekenen. Zonder jullie was ik nooit gekomen waar ik nu ben. Ik ben ontzettend dankbaar voor alles wat jullie voor mij hebben gedaan. **Anne, Guus, Wouter, Marco, Esmée en Judith** - ook geldt dit natuurlijk voor jullie.

Tot slot, lieve **Thijs**, altijd ben jij mijn steun en toeverlaat. Jij hebt het meest alle pieken en dalen van dit traject van dichtbij meegemaakt en me altijd aangemoedigd als ik dat nodig had. We hebben al een zoveel moois samen beleefd, en ik kijk ernaar uit om nog veel meer avonturen met jou te delen.

PHD PORTFOLIO

Description	Year	ECTS*
Courses		
Biostatistics I	2021	4.5
Biomedical Writing for PhD candidates	2021	1.5
Statistical Shape Modeling	2022	4.5
Repeated Measurements	2022	1.7
Re-registration BROK	2022	0.5
Scientific Integrity	2022	0.3
Logistic Regression	2022	1.4
Data Science in Epidemiology	2022	0.7
Topics in Medical Decision-making	2023	1.4
Presentation Skills	2023	1
Conferences		
Oral presentations		
Science Day Dept. Of Orthopedics and Sports Medicine *"Why you should collaborate with a Technical Physician"*	2022	1
SMWJC *"Asymmetric sports played during childhood are associated with acetabular dysplasia at adolescence"*	2023	1
SMWJC - Star paper session *"Type of sport played during childhood is not associated with cam morphology at adolescence"*	2023	1
Science Day Dept. Of Orthopedics and Sports Medicine *"The HIPSTAR Chronicles: Designing and Tweaking Your Research Plan"*	2024	1
OARSI World Congress - Pitch	2024	0.9
Combined NOV/NOF congress *"DXA images are a reliable alternative to pelvic radiographs for performing hip morphology measurements"*	2024	1
International workshop on osteoarthritis imaging *"The presence of hip pain does not modify the association between hip morphology and incident radiographic hip OA within 5-8 years"* *"DXA images are a reliable alternative to pelvic radiographs for performing hip morphology measurements"*	2024	2

Poster presentations		
Biomedical Science PhD Day	2022	1
OARSI World Congress	2023	1
Sophia Research Day	2023	1
NOV congress	2023	1
OARSI World Congress	2024	1
Sophia Research Day	2024	1

Teaching activities		
Teaching assistant - Basic Course on R	2023	1
Scientific internship for medicine master student education	2021-2025	2
Supervision Masterstudent - Jolien van Haasteren	2021	2
Supervision Masterstudent - Julia Wortel	2022	2
Supervision Masterstudent - Casper Donkervoort	2023	1
Supervision Masterstudent - Tom Lansink	2024	1

Other		
Chair – Research meetings on hip related Generation R research	2021-2023	1
Journal club department of Orthopedics and Sports Medicine	2021-2024	2.5
Presentations various research meetings	2021-2024	1
Attendence OARSI World Congress	2022	1
Organisation TiiM Conference - NVvTG	2022	8
Attendence ESOC	2023	1
Reviewer – Osteoarthritis and Cartilage	2023	0.3
Reviewer – BMC Musculoskeletal Disorders	2024	0.4

*1 ECTS (European Credit Transfer System) equals a 28-hour workload.

LIST OF PUBLICATIONS

Peer reviewed publications

J. Tang, M.M.A. van Buuren, N.S. Riedstra, **F. Boel**, J. Runhaar, S.M.A. Bierma-Zeinstra et al. Cam morphology is strongly and consistently associated with development of radiographic hip osteoarthritis throughout 4 follow-up visits within 10 years. Osteoarthritis and Cartilage. 2023 Dec;31(12):1650-1656. Epub 2023 Aug 18. doi: 10.1016/j.joca.2023.08.006

N.S. Riedstra, **F. Boel**, M.M.A. van Buuren, D. Eygendaal, S.M.A Bierma-Zeinstra, J. Runhaar et al. Pincer morphology is not associated with hip osteoarthritis unless hip pain is present: Follow-Up Data From a Prospective Cohort Study. Arthritis Care & Research. 2024 May;76(5):644-651. Epub 2023 Dec 21. doi: 10.1002/acr.25285

M.M.A. van Buuren, N.S. Riedstra, M.A. van den Berg, **F. Boel**, H. Ahedi, V. Arbabi et al. Cohort profile: Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) - an international consortium of prospective cohort studies with individual participant data on hip osteoarthritis. BMJ open. 2024 Apr 18;14(4):e077907. e077907. doi: 10.1136/bmjopen-2023-077907

F. Boel, S. de Vos-Jakobs, N.S. Riedstra, C. Lindner, J. Runhaar, S.M.A. Bierma-Zeinstra et al. Automated radiographic hip morphology measurements: An open-access method. Osteoarthritis imaging. 2024 Jun;4(2):100181. doi: 10.1016/j.ostima.2024.100181

S. de Vos-Jakobs, **F. Boel**, W.M. Bramer, S.M.A. Bierma-Zeinstra, R. Agricola. Prevalence and radiological definitions of acetabular dysplasia after the age of 2 years: a systematic review. Journal of Pediatric Orthopaedics. 2024 Jul 1;33(4):334-339. Epub 2023 Aug 7. doi: 10.1097/bpb.0000000000001113

F. Boel*, N.S. Riedstra*, J. Tang , D.F. Hanff, H. Ahedi, V. Arbabi, et al. Reliability and agreement of manual and automated morphological radiographic hip measurements. Osteoarthritis and Cartilage Open. 2024 Sept;6(3):100510. doi: 10.1016/j.ocarto.2024.100510

F. Boel, J. Wortel, M.M.A. van Buuren, F. Rivadeneira, J.B.J. van Meurs, J. Runhaar et al. DXA images vs. pelvic radiographs: Reliability of hip morphology measurements. Osteoarthritis and Cartilage. 2024 Oct 24;33(2):283-292. Epub 2024 Oct 24. doi: 10.1016/j.joca.2024.10.010

J. Tang, M.M.A. van Buuren, **F. Boel**, N.S. Riedstra, M.A. van den Berg, J. Runhaar et al. The association between cam morphology and hip pain in males and females within 10 years: A national prospective cohort study (CHECK). Seminars in Arthritis and Rheumatism. 2024 Dec;69:152539. doi: 10.1016/j.semarthrit.2024.152539

J. Tang, **F. Boel**, M.M.A. van Buuren, N.S. Riedstra, J. Runhaar, S.M.A. Bierma-Zeinstra et al. The different subtypes of cam morphology as defined by statistical shape modeling and their relationship with the development of hip osteoarthritis: A nationwide prospective cohort study (CHECK) with 10 years follow-up. Osteoarthritis and Cartilage. 2024 Dec;32(12):1647-1654. doi: 10.1016/j.joca.2024.08.003

N.S. Riedstra*, **F. Boel***, M.M.A. van Buuren, H. Ahedi, V. Arbabi, N. Arden et al. Acetabular dysplasia and the risk of developing hip osteoarthritis within 4-8 years: An individual participant data meta-analysis of 18,807 hips from the World COACH consortium. Osteoarthritis and Cartilage. 2025 Mar;33(3):373-382. 120756. doi: 10.1016/j.joca.2024.12.001

D. Chen, **F. Boel**, P. van Klij, M.M.A. van Buuren, S.M.A. Bierma-Zeinstra, R. Agricola. The prevalence of primary cam morphology in early adolescents from the general population: a population-based study (Generation R). Osteoarthritis and Cartilage. 2025 May;33(5):633-637. doi: 10.1016/j.joca.2025.02.775.

D. Chen, **F. Boel**, S. de Vos-Jakobs, P. van Klij, M.M.A. van Buuren, S.M.A. Bierma-Zeinstra, R. Agricola. The prevalence of pincer morphology in early adolescents from the general population: a population-based study (Generation R). Arthritis care & research. 2025 Jul. doi: 10.1002/acr.25601. Epub ahead of print.

M.A. van den Berg, **F. Boel**, M.M.A. van Buuren, N. S. Riedstra, J. Tang, H. Ahedi et al. Hip morphology-based osteoarthritis risk prediction models: development and external validation using individual participant data from the World COACH Consortium. Arthritis care & research. 2025 Aug. doi: 10.1002/acr.25629. Epub ahead of print.

Submitted for publication

N.S. Riedstra*, **F. Boel***, M.M.A. van Buuren, H. Ahedi, V. Arbabi, N.K. Arden et al. Severe pincer morphology is associated with incident hip osteoarthritis: prospective individual participant data meta-analysis from 18,935 hips from the World COACH consortium. *Under review.*

J. Tang, **F. Boel**, M.M.A. van Buuren, N.S. Riedstra, M.A. van den Berg, J. Runhaar et al. Triangular Index Ratio as An Alternative Method to the Alpha Angle for Defining the Presence of Cam Morphology: Data from the CHECK Cohort. *Under review.*

J. Tang; **F. Boel**, M.M.A. van Buuren, N.S. Riedstra, M.A. van den Berg, H. Ahedi et al. Cam Morphology and the Risk of Developing Radiographic Hip Osteoarthritis within 8 Years: An Individual Participant Data Meta-analysis from the World COACH Consortium. *Under review.*

S. de Vos-Jakobs, **F. Boel**, D. Chen, J.C.M. van Haasteren, J.J. Tolk, S.M.A. Bierma-Zeinstra, R. Agricola. Prevalence of acetabular dysplasia in early adolescents in a general population: the Generation R study. *Submitted.*

F. Boel, S. de Vos-Jakobs, D. Chen, J. Runhaar, S. Bierma-Zeinstra, R. Agricola. Known risk factors for developmental dysplasia of the hip are not associated with early adolescent acetabular dysplasia: findings from a prospective open population cohort. *Submitted.*

F. Boel, S. de Vos-Jakobs, D. Chen, F. Rivadeneira, S.M.A. Bierma-Zeinstra, J. Runhaar[±], R. Agricola[±]. Hip shape is associated with the change in acetabular coverage in early adolescence: a longitudinal study. *Submitted.*

**Shared first author. [±]Shared last author*

CURRICULUM VITAE

Fleur Désirée Emile Maria Boel is op 29 februari 1996 geboren te Breda. In 2014 behaalde zij haar Gymnasium diploma op De Nassau in Breda. Direct daarna begon zij haar bachelor Klinische Technologie aan de Technische Universiteit Delft, de Erasmus Universiteit Rotterdam en de Universiteit Leiden. Fleur was onderdeel van de eerste lichting van de studie waardoor zij betrokken was bij het opzetten van de studievereniging S.V.K.T. Variscopic. In het tweede jaar van haar bachelor werd zij commissaris interne betrekkingen van Variscopic, waarbij zij zich onder andere bezig hield met het opstellen van de statuten van de vereniging. Na het behalen van haar bachelor diploma Klinische Technologie in 2017 begon zij aan haar master Technical Medicine binnen de track Imaging & Interventions. Tijdens haar master behaalde zij ook het diploma voor Coördinerend Deskundige in de Stralingsbescherming. Fleur liep haar stages in het Leids Universitair Medisch Centrum, waar ze haar klinische ervaring opdeed en zich bezig heeft gehouden met vier verschillende onderzoeksvragen. Daar werd haar interesse voor onderzoek en het inzetten van beeldanalyse gewekt. Dit heeft zich uiteindelijk verder uitgebreid tijdens haar afstudeeronderzoek bij de afdeling Interventie Radiologie in het Leids Universitair Medisch Centrum. In 2020 heeft Fleur succesvol haar master Technical Medicine afgerond. Al tijdens de laatste twee jaar van haar masteropleiding wist Fleur dat ze verder wilde met onderzoek in de vorm van PhD traject. In maart 2021 deed de mogelijkheid zich voor om een PhD traject te starten bij de afdeling Orthopedie en Sportgeneeskunde van het Erasmus Universitair Medisch Centrum en deze mogelijkheid greep zij dan ook vol enthousiasme aan. Onder begeleiding van Prof. dr. Sita Bierma-Zeinstra, Dr. Rintje Agricola en Dr. Jos Runhaar heeft zij onderzoek gedaan naar hoe afwijkende heup morfologie ontstaat gedurende de groei en hoe dit een risicofactor is voor heupartrose. In 2025 zal ze haar loopbaan als onderzoeker vervolgen als Postdoc waar ze het werk van haar PhD zal voorzetten.