

Novel Indices for the Diagnosis of Infant Hip Dysplasia using Two and Three Dimensional  
Ultrasound

by

Myles Garnet Mabee

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biomedical Engineering  
University of Alberta

© Myles Garnet Mabee, 2014

## **Abstract**

Infant developmental dysplasia of the hip (DDH) is a common condition in which the ball and socket joint of the hip is not formed properly. Current diagnosis techniques are controversial and lack firm agreement between physicians. Diagnosis is based on a joint application of physical examination and imaging of the hip, generally by two dimensional (2D) ultrasound. 2D ultrasound techniques are based upon the Graf method, which has many limitations such as high inter/intra- observer variability and large inter-scan variability. Understanding of DDH may be improved by adding secondary indices such as the acetabular radius of curvature discussed in this thesis. With recent improvements in 3D ultrasound, there may be added benefit in using this technology to image infant hips for the purpose of a more reliable diagnosis. The 3D inStability angle introduced in the thesis has high reliability and demonstrates preliminary diagnostic utility for the identification of dysplastic hips.

## PREFACE

The research projects, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name “Assessment of developmental hip dysplasia by 3D ultrasound – correlation with clinical and imaging findings at initial infant screening and follow-up”, Study ID MS6\_Pro00032107.

A version of Chapter 2 has been accepted for publication in *Radiology* – Jaremko JL., Mabee M, Swami VG, et al. Potential for change in the US diagnosis of hip dysplasia solely caused by changes in probe orientation: Patterns of alpha-angle variation revealed by using three-dimensional US. *Radiology* 2014 (in press). © Radiological Society of North America.

A version of Chapter 3 has been accepted for publication in *Ultrasound in Medicine and Biology* – Cheng E, Mabee M, Swami VG, et al. Ultrasound Quantification of Acetabular Rounding in Hip Dysplasia: Reliability and Correlation to Treatment Decisions in a Retrospective Study. *Ultrasound in Medicine and Biology* 2014 (in press).

A version of Chapter 4 has been accepted for publication in *Ultrasonic Imaging* – Mabee M, Dulai S., Thompson R.B., Jaremko J.L., Reproducibility of acetabular landmarks and a standardized system obtained from 3D hip ultrasound. *Ultrasonic Imaging* 2014 (in press).

## ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor – Dr. Jacob Jaremko. I never thought I would end up in a graduate degree, but I am certainly glad that I was provided this opportunity. I am incredibly grateful for the enthusiasm which he brings to this project, especially given his many (**MANY**) other commitments and incredibly tight schedule. He has always been able to provide me with guidance and assistance whenever I needed it, even answering emails with time stamps far too late into the night and again the next morning at 8 AM; all of this from a man I have seen yawn twice in two years and who I have never seen drink ANY coffee. So thank you again, I truly appreciate everything I have learned and accomplished these past few years.

I would like express my gratitude to Dr. Richard Thompson who served as co-supervisor and a member of my supervisory committee. Your guidance and input into these projects was invaluable, and I appreciate the advice you have provided me.

The PhD students who I have gotten to know in my degree have been a valuable source of information and knowledge for me, particularly Kelvin Chow and June Cheng-Baron, thank you for all of the programming advice and general problem solving talks.

My friends and family also get an honorable mention for bearing with me when I get talking about my research, the nodding and pretending to know what I'm talking about makes me happier than you all know.

A special thanks to my parents for being incredibly supportive, even if they think that 'one degree' was enough.

Finally, I would like to thank the staff of the Radiology and Diagnostic Imaging department, spending my Wednesdays scanning with you all has been incredibly pleasant, especially Jan Riddell and Elizabeth Fehlauer. Thanks again Elizabeth for keeping everyone organized!

## TABLE OF CONTENTS

Preface.....	iii
Acknowledgements .....	iv
Table of Contents .....	v
List of Tables.....	x
List of Figures .....	xi
List of Abbreviations.....	xiii
List of Symbols.....	xiv
Chapter 1 Introduction .....	1
1.1 Ultrasound .....	1
1.1.1 The Ultrasound Probe.....	2
1.1.2 Reflection and Transmission.....	4
1.1.3 Refraction .....	5
1.1.4 Attenuation and Resolution .....	6
1.1.5 3D Ultrasound .....	7
1.2 Hip Dysplasia.....	9
1.2.1 Definition.....	9
1.2.2 Pelvis Anatomy .....	10
1.2.3 Diagnosis.....	11
1.2.4 Treatment.....	13
1.3 Thesis Scope .....	15

1.4 References.....	16
Chapter 2 Potential for change in ultrasound diagnosis of hip dysplasia due solely to changes in probe orientation: patterns of alpha angle variation revealed using 3D ultrasound .....	20
2.1 Introduction.....	20
2.2 Materials and Methods .....	22
2.2.1 Imaging .....	22
2.2.2 Imaging Processing .....	23
2.2.3 Statistical Analysis.....	28
2.3 Results .....	28
2.3.1 Section Orientation .....	28
2.3.2 Variability of Alpha Angle with Section Orientation.....	28
2.3.3 Variability of Alpha Angle Between and within Observers .....	31
2.3.4 Normal versus Dysplastic Hips .....	32
2.4 Discussion .....	34
2.5 CONCLUSION.....	37
2.6 References .....	38
Chapter 3 Ultrasound Quantification of Acetabular Rounding in Hip Dysplasia: Reliability and Correlation to Treatment Decisions .....	40
3.1 Introduction.....	40
3.2 Materials and Methods .....	41
3.2.1 Patients .....	41
3.2.2 Index Measurement.....	41

3.2.3 Statistics.....	43
3.3 Results.....	44
3.4 Discussion .....	50
3.4.1 Limitations.....	52
3.5 Conclusions .....	54
3.6 References.....	54
Chapter 4 Reproducibility of acetabular landmarks and a standardized coordinate system obtained from 3D hip ultrasound.....	56
4.1 Introduction.....	56
4.2 Materials and Methods .....	57
4.2.1 Patients .....	57
4.2.2 Imaging.....	58
4.2.3 Image Processing.....	58
4.2.4 Statistical Analysis.....	60
4.3 Results.....	61
4.3.1 Variability of Landmarks .....	62
4.3.2 Variability of Central Plane Orientation.....	63
4.3.3 Variability of Slice Selection .....	64
4.3.4 Inter-Observer Variability of Alpha Angle .....	65
4.3.5 Inter-scan Variability of Alpha Angle.....	67
4.3.6 Comparison of 3D and 2D Alpha Angles.....	69
4.3.7 Alpha Angles in Normal vs. Dysplastic Hips .....	69

4.4 Discussion .....	70
4.5 Conclusion.....	73
4.6 References.....	73
Chapter 5 A novel index in the diagnosis of infant hip dysplasia using 3D ultrasound: The inStability Index.....	75
5.1 Introduction.....	75
5.2 Methods.....	75
5.2.1 Patients.....	75
5.2.2 Imaging.....	76
5.2.3 Image Processing.....	77
5.2.4 Statistical Analysis.....	82
5.3 Results .....	82
5.3.1 Variability of inStability Index.....	82
5.3.2 Normal vs. Dysplastic Hips .....	83
5.3.3 Receiver Operating Characteristic Curves.....	85
5.3.4 Correlation of inStability and Alpha Angles.....	85
5.3.5 Acetabulum Fusion.....	86
5.4 Discussion .....	87
5.5 Conclusion.....	91
5.6 References .....	91
Chapter 6 Discussion and Conclusions.....	93
6.1 Discussions .....	93

6.1.1 Thesis Overview.....	93
6.1.2 Problems with 2D Ultrasound and Potential Benefits of 3D.....	93
6.1.3 Review of Quantitative Results.....	95
6.1.4 Evaluation of Clinical Utility of 3D Ultrasound.....	97
6.1.5 Limitations.....	98
6.1.5 Future Directions.....	99
6.2 Conclusions.....	100
6.3 References.....	101
Chapter 7 Complete Bibliography.....	102

## LIST OF TABLES

TABLE 2.1 CATEGORIES OF HIP DYSPLASIA BY ALPHA ANGLE.....	21
TABLE 3.1 INTER AND INTRA-OBSERVER INTRA-CLASS CORRELATION COEFFICIENT (ICC) AND COEFFICIENT OF VARIANCE (COV) FOR ALPHA ANGLE, COVERAGE, ACETABULAR RADIUS OF CURVATURE (AROC), AND ARC LENGTH.....	44
TABLE 3.2 INTER- AND INTRA-OBSERVER VARIABILITY OF SUBJECTIVE ASSESSMENT OF ACETABULAR ROUNDING .....	45
TABLE 3.3 PEARSON CORRELATION COEFFICIENTS BETWEEN FOUR ULTRASOUND INDICES .....	46
TABLE 4.1 RELIABILITY OF 3D HIP ULTRASOUND LANDMARK LOCALIZATION AND CENTRAL PLANE ORIENTATION. ....	61
TABLE 4.2 RELIABILITY OF ALPHA ANGLE -- 2D VS. 3D ULTRASOUND .....	62
TABLE 5.1 RMS VALUES FOR ACETABULUM SURFACE MODELS.....	87

## LIST OF FIGURES

FIGURE 1.1 ULTRASOUND TRANSDUCER COMPONENTS.....	3
FIGURE 1.2 MECHANICAL LOCALIZERS USED IN 3D ULTRASOUND. ....	9
FIGURE 1.3 ILLUSTRATION OF HIP DYSPLASIA.. ....	10
FIGURE 1.4 INFANT HIP RADIOGRAPH WITH IMPORTANT BONES LABELLED .....	11
FIGURE 1.5 STANDARD GRAF CORONAL PLANE.....	13
FIGURE 1.6 INFANT SECURED IN A PAVLIK HARNESS.....	14
FIGURE 2.1 GENERATION OF A 3D HIP SURFACE MODEL FROM 3D US IMAGES. ....	24
FIGURE 2.2 MODEL DEMONSTRATING THE CENTRAL PLANE AND ROTATED PLANES .....	26
FIGURE 2.3 EFFECT OF PROBE TILT IN TWO AXES.. ....	29
FIGURE 2.4 US SCANS OF THE RIGHT HIP OF A 19-DAY-OLD GIRL SHOW FINDINGS CROSSING TWO GRAF DIAGNOSTIC CATEGORIES BECAUSE OF PROBE TILT. ....	31
FIGURE 2.5 US SCANS IN LEFT HIP IN A 3-MONTH-OLD GIRL SHOW FINDINGS CROSSING THREE GRAF DIAGNOSTIC CATEGORIES BECAUSE OF PROBE TILT. ....	31
FIGURE 2.6 US IN LEFT HIP IN A 6-DAY-OLD GIRL SHOWS THAT A DYSPLASTIC HIP CAN APPEAR NORMAL DEPENDING ON PROBE TILT.....	33
FIGURE 2.7 POINT ESTIMATES SHOW DIFFERENCES IN ALPHA ANGLES ACCORDING TO CLINICAL DIAGNOSIS (DX). ....	34
FIGURE 3.1 ULTRASOUND TRACINGS CONDUCTED BY OBSERVERS ON CUSTOM MATLAB SOFTWARE. ....	43
FIGURE 3.2 EXAMPLE OF A 2D HIP ULTRASOUND WITH INDISTINCT ACETABULAR MARGIN THAT WAS DIFFICULT TO TRACE.....	45
FIGURE 3.3 ACETABULAR RADIUS OF CURVATURE (AROC) VS. ALPHA ANGLE FOR ONE SET OF READS CONDUCTED BY AN OBSERVER. ....	47
FIGURE 3.4 MEAN VALUES OF ALPHA ANGLE, COVERAGE, ACETABULAR RADIUS OF CURVATURE AND ARC LENGTH FOR EACH CLINICAL DIAGNOSTIC GROUP.....	48

FIGURE 3.5 TYPICAL ULTRASOUND APPEARANCE IN EACH CLINICAL GROUP.....	49
FIGURE 3.6 RECEIVER OPERATING CHARACTERISTIC CURVES FOR ALPHA ANGLE, COVERAGE, ACETABULAR RADIUS OF CURVATURE 'AROC' AND ARC LENGTH. ....	50
FIGURE 4.1 COMPARISON OF INFANT HIPS BETWEEN MODALITIES.....	57
FIGURE 4.2 IMAGES FROM 3D ULTRASOUND OF THE RIGHT HIP IN A 101 DAY OLD BOY DIAGNOSED CLINICALLY WITH NORMAL HIPS, SHOWING CLEARLY DEFINED LANDMARK POINTS. ....	59
FIGURE 4.3 ULTRASOUND IMAGES OF A GIRL DIAGNOSED WITH DDH.....	60
FIGURE 4.4 INTER-OBSERVER VARIATION IN SELECTION OF LANDMARK LOCATIONS AND CENTRAL PLANE ORIENTATIONS. . ....	63
FIGURE 4.5 SELECTION OF BEST CENTRAL SLICE.. ....	65
FIGURE 4.6 EFFECT OF INTER-OBSERVER VARIABILITY IN PROCESSING 3D ULTRASOUND .....	67
FIGURE 4.7 INTER-SCAN VARIABILITY.....	68
FIGURE 4.8 DATA PROCESSING WORKFLOW AND RELIABILITY FOR 2D AND 3D ULTRASOUND. ....	69
FIGURE 5.1 ACETABULUM LANDMARKS, TRACING TECHNIQUES, AND SURFACE MODEL.. ....	80
FIGURE 5.2 SURFACE NORMALS AND INSTABILITY ANGLE.....	80
FIGURE 5.3 ILLUSTRATION OF THE RMS BETWEEN SURFACE MODELS.....	81
FIGURE 5.4 ALPHA AND INSTABILITY ANGLE BY DIAGNOSTIC CATEGORY. ....	84
FIGURE 5.5 PYRAMID GRAPH FOR ALPHA AND INSTABILITY ANGLE BY DIAGNOSTIC CATEGORY.....	84
FIGURE 5.6 ROC CURVE FOR ALPHA AND INSTABILITY ANGLES.....	85
FIGURE 5.7 CORRELATION OF ALPHA ANGLE TO INSTABILITY ANGLE.....	86
FIGURE 6.1 COMMON ERRORS OF 2D ULTRASOUND SCANNING.....	94

## LIST OF ABBREVIATIONS

DDH	Developmental Dysplasia of the Hip
CDH	Congenital Dislocation of the Hip
2D	Two Dimensional
3D	Three Dimensional
US	Ultrasound
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
A-mode	Amplitude Mode
B-mode	Brightness Mode
PZT	Lead zirconate titanate
PVDF	Polyvinylidene difluoride
AROC	Acetabular Radius of Curvature

## LIST OF SYMBOLS

$\alpha$	Alpha Angle
$\beta$	Beta Angle
$\varphi$	inStability Index/Angle
Z	impedance
$\lambda$	wavelength
c	Speed of sound
v	Speed
R	Reflection coefficient
T	Transmission coefficient
$\rho$	density
I	Intensity
$\mu_{us}$	Ultrasound intensity attenuation coefficient

# CHAPTER 1

## Chapter 1 ..

### INTRODUCTION

#### 1.1 ULTRASOUND

Sound waves are longitudinal (compressional) waves which travel through a physical medium (usually liquid or tissue for ultrasound), in contrast to transverse waves which oscillate at  $90^\circ$  to the direction of travel [1]). Ultrasound in its most basic definition is any sound wave which has a frequency greater than the upper limit of the human hearing range (typically in the 1 – 10 MHz range [1, 2]). Ultrasound is much different than other imaging modalities in that it does not use ionizing radiation (x-ray, computed tomography), which has lasting negative effects for the individuals it is used upon, or radio frequency energy (like MRI).

The first detailed experiments which indicated that inaudible sound existed were performed by Lazzaro Spallanzani, who found that bats use ultrasound to echo-locate their surroundings and identify food sources. Spallanzani showed that blindfolded bats were able to navigate without sight, but would run into obstacles when their mouths were covered [3]. Bats emit a call into their environment and then listen for the echo, with the echo helping to locate and discern where objects in its surroundings are. Initial applications of using ultrasound and its reflections were similar, these first applications tried to detect objects for nautical navigation purposes [2]. By 1942 the use of ultrasound as a medical diagnostic tool was beginning, when Dr Karl Dussik attempted to locate tumors in the brain through transmissions of ultrasound beams in the head [3]. Ultrasound continued to develop with the use of A-mode (amplitude) scanning;

A-mode scanning consists of simple one dimensional representation of the echoes along a single line plotted as a function of depth, tissue structures can also be depicted or tracked over time (motion mode or M-mode) [1, 2, 4, 5]. Brightness mode (named because the strength of the echoes are converted to brightness on the display), which is usually shortened to just B-mode, was introduced next and added an extra dimension of information over A-mode. By the 1960s many different variations and methods for ultrasound had been introduced, such as: continuous wave, pulsed Doppler, and real-time B-mode [1, 3-5]. Color Doppler entails the velocity information inside of the body and was introduced in the 1980s. It allowed color-coding of the moving parts (usually blood flow) over standard B-mode images [1, 2, 4, 5]. A recent advance which has wide clinical implications is 3D ultrasound (discussed in further detail in section 1.1.5).

All of these advancements have allowed ultrasound to remain an integral piece of any clinician's diagnostic toolset. The imaging technique consists of three main steps: (1) generating a sound pulse for emission, (2) receiving the sound pulse echoes, and (3) transforming the received echoes into an image that is discernible to a human. Ultrasound probes are designed cleverly to take care of steps (1) and (2), but a computer is required to take all of the information and convert it in step (3).

#### 1.1.1 THE ULTRASOUND PROBE

The ultrasound probe is the source of the emitted sound waves as well as the destination of the received echoes. The transducer is made of piezoelectric ceramic crystals (PZT – lead zirconate titanate) and typically the plastic polyvinylidene difluoride (PVDF) [1, 2], which convert electrical energy into mechanical energy during transmission and mechanical energy into electrical energy during receiving. Other pieces of the transducer include: damping block, matching layer, resonance transducers, as well as cables and plastic cases/housing shells [1, 2, 4], see Fig 1.1.

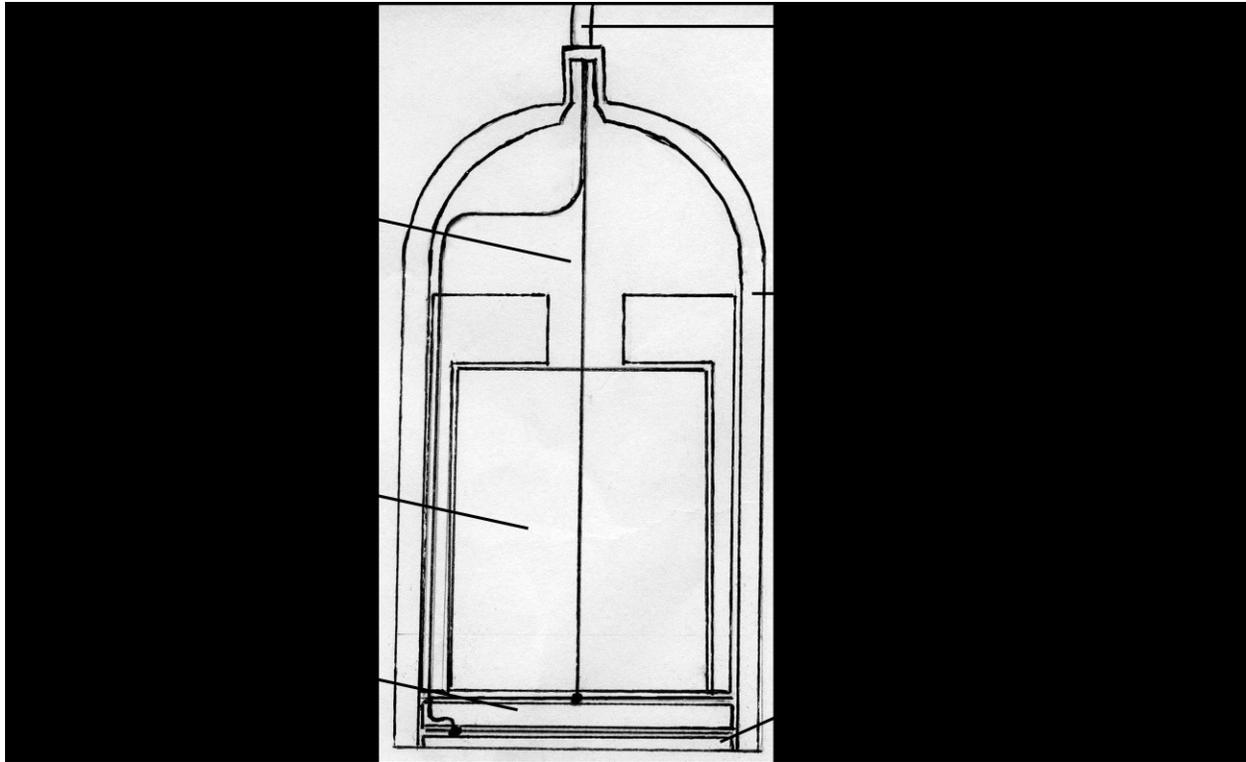


Figure 1.1 Ultrasound Transducer Components. Image was adapted from [1].

Resonance transducers are essential to produce sound waves of adequate amplitude to penetrate human tissue; these pieces of the transducer convert the initial force created by the piezoelectric crystals into something with greater amplitude [1]. Damping blocks aid in lessening the purity of the resonance frequency of the device and create a broadband frequency spectrum that helps establish better resolution. Matching layers in transducers provide the interface between the transducer's piezoelectric element and the damping block behind and (on the opposite side) the interface between the element and the patient's tissue [1, 2]. This layer is needed as it helps to minimize the acoustic impedance differences between the transducer and the tissue, which allows for better penetration and maximizes power transmission into the body; most manufactures minimize impedance differences by constructing the damping block out of material with the same acoustic impedance as that of the piezoelectric element, which results in almost all of the energy being absorbed into the tissue [1, 2]. The element to tissue interface

impedance is minimized by creating a layer by means of impedance matching. Impedance matching involves covering the head of a transducer with a layer which has an impedance of the square root of the product of the element impedance and tissue impedance (see eq 1.1), this layer is generally a quarter-wavelength in thickness for narrow bandwidth operation and from several small layers for broadband transmission [1, 2].

$$Z_{match} = \sqrt{Z_{Piezo}Z_{tissue}} \quad (1.1)$$

Finally, good physical contact and coupling gel is used to fully minimize the remaining acoustic impedance differences. The transducer is then arranged into arrays, meaning the piezoelectric crystal is segmented into different pieces which may act independently from one another. Transducer arrays generally come in two main types: sequential, and phased. Sequential arrays typically consist of 64-512 elements with a width of around  $\frac{1}{2} \lambda$ , these arrays fire simultaneously and run parallel to each other, causing each element to create a single scan line and the image will be the exact shape of the array [1, 4]. Phase arrays typically consist of 64 – 128 elements, with all transducer elements activated with time delays to produce an ultrasound beam at a given direction. By adding time delays, the beam may be steered or directed to regions of interest without moving the actual probe, while during the receive mode all elements are active to detect returning echoes [1, 4].

### 1.1.2 REFLECTION AND TRANSMISSION

Ultrasound images of the tissues within the body are able to be visualized due to their differences in density and compressibility. The differences in density result in different acoustic impedances of materials. The relationship between density and impedance is defined in equation 1.2 below [1, 2, 4, 6].

$$Z = \rho \times c \quad (1.2)$$

Where  $Z$  = impedance of material

$\rho$  = density of material

$c$  = speed of sound in material

As mentioned earlier, differences in impedance result in lower transmission of ultrasound energy, as the rest of the energy is reflected or scattered. The transmitted and reflected intensity of the ultrasound energy is related by equation 1.3 [1, 2, 4, 6].

$$R_I = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2} \text{ and } T_I = 1 - R_I \quad (1.3)$$

Where  $R_I$  = Reflected Intensity

$Z_2$  = Impedance of object 2

$Z_1$  = Impedance of object 1

$T_I$  = Transmitted Intensity

Equation 1.3 highlights why impedance matching is important in the ultrasound transducer. When large differences in impedance exist between two surfaces, most energy is reflected and not transmitted. Large impedance differences are why bones appear as bright white structures on ultrasound as almost all of the energy makes it back to the transducer to be converted back to electrical energy. For example, the reflection coefficient of intensity between muscle and bone is 0.41 (41% of ultrasound energy is reflected), and that of muscle and air is even higher (resulting in almost 100% reflection).

### 1.1.3 REFRACTION

When a beam of ultrasound passes from one medium to another of different density, its speed will change and, depending on if the wave is incident upon the next medium, the direction of propagation will also change. The frequency of the ultrasound wave is constant on either side of the interface; however, since the speed is changing, the wavelength of the beam must also change by the same proportionality (as shown in eq 1.4)

$$\frac{\lambda_2}{\lambda_1} = \frac{v_2}{v_1} \quad (1.4)$$

Where  $\lambda$  = wavelength

$v$  = speed

When a wavefront reaches a new medium, part of the wave will reside in the old medium while another part will reside in the new medium; the wavecrests must be separated by a distance representative of the two different size wavelengths in each medium. The only way this situation can occur is if a discontinuity in the direction of propagation occurs [2]. The exact refraction and change in direction of propagation that will occur can be calculated from Snell's law, as in equation 1.5 [1, 2, 4].

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{v_1}{v_2} \quad (1.5)$$

Where  $\theta$  = angle of incident or refraction

$v$  = speed in the medium

Refraction in ultrasound can cause distortions in the image, in the same way in which something is partially submerged in water appears to be displaced a different distance than it is in reality. This can also cause losses in resolution.

#### 1.1.4 ATTENUATION AND RESOLUTION

When ultrasound travels through a single homogenous medium, its intensity decays exponentially with respect to the distance it has travelled, as represented by equation 1.6 [2].

$$I(x) = I(0)e^{\mu_{US}x} \quad (1.6)$$

Where  $I$  = intensity

$x$  = distance travelled by ultrasound

$\mu_{US}$  = ultrasound intensity attenuation coefficient

The *ultrasound intensity attenuation coefficient* characterizes the losses from scattering of the wave due to minute changes in the uniform nature of the medium, as well as absorption due to heat in the tissue/medium. The rate of attenuation is highly dependent on the frequency and wavelength of the beam. The higher the frequency of the ultrasound beam, the faster the attenuation and the lower the distance of penetration. For muscle, blood, and most soft tissues

the attenuation is linear with respect to frequency; however, in water, bone, and a few other materials the attenuation increases with the square of frequency [1, 2].

This would suggest that picking the lowest frequency would be ideal for using ultrasound in imaging; however, there is often a need to visualize small objects within the body, and visualizing small objects requires fine resolution. In order to increase both lateral (perpendicular to the ultrasound beam) and axial (parallel to the ultrasound beam) resolution, a higher frequency beam is required. The selection of frequency and transducer is based on this trade off: if you are attempting to image something deep inside the body and require high penetration, a lower frequency is essential; but, if you are attempting to image a small structure, then a higher frequency would be required.

#### 1.1.5 3D ULTRASOUND

Ultrasound images can be confusing without orientation, and it is difficult to interpret the underlying complexities of the 3D geometry of organs and tissues in only two dimensions. Consequentially it is also difficult to orient the transducer to create a true proper image in the required plane [7]. Since any 2D image is a limited sample of the 3D body structure obtained by a handheld probe, it is difficult to reproducibly create a representative 2D image of a given body structure at follow up exams, or in research for longitudinal studies [5]. The patient's own internal anatomy or orientation may restrict the image angle, resulting in an impossible optimal plane. Guided therapeutic procedures are susceptible to error because the process of quantifying and monitoring the body's changes over the time of the procedure are limited by the 2D restrictions of the conventional exam [7]. All of these issues are diminished with the use of 3D ultrasound. The main benefit of the added dimension is the ability to reduce variability and to allow physicians to see the complex 3D anatomy of tissues, organs, and structures. Unlike MRI and CT which are generally slow to acquire a full set of images in a stack of parallel planes and have lengthy scan times, ultrasound is very quick (10-60 images per second) and is not limited to a stack of parallel planes for data acquisition [7]. Ultrasound has no contraindications,

contrast agents, or ionizing agents which makes a 3D ultrasound invaluable for patients ineligible for CT, MRI or other 3D imaging modalities. 3D ultrasound conceptually ought to be (and this thesis is a step towards proving it is) less operator dependent than 2D ultrasound and should have excellent reproducibility, since the ideal plane will exist within the acquired volume data set, meaning it can simply be manipulated to acquire this ideal orientation [8-10]. 3D ultrasound also introduces many orientations and planes which may be impossible to acquire using conventional 2D, due to patient position obstructions or other factors. The archived data may also be reviewed post exam by a clinician who may want to review more of the patient data or in a different way than anticipated at the time of scanning.

3D ultrasound exists in many different forms for acquisition of data, such as: freehand acquisition, 2D matrix arrays, and mechanical localizers [1]. Freehand acquisition is typically the cheapest of acquisition techniques and can be further segmented into using acoustic positioners, articulated arm positioners, or magnetic field sensors. 2D matrix arrays were first introduced in the 1990s and are capable of acquiring 3D images in real time, which is referred to as 4D imaging [5]. The array is a stationary square of elements which use a built in beam-former to electronically sweep the region of interest in two orthogonal directions (azimuth and elevation) which create a pyramidal shape of acquisition [1, 5]. 2D arrays have evolved from sparse arrays to fully sampled arrays in recent years, and in some cases are capable of producing clinically useful 4D images making it likely that they are the future of 3D ultrasound imaging. Mechanical localizer transducers are a typical 2D transducer probe in which the third dimension is obtained by the mechanical movement of the transducer in a precise predefined manner. As the transducer is moved, the 2D images are acquired at preset intervals so that the image sequences build an appropriate volume of interest [1, 7]. The movement is generally guided through one of three processes: linear, fan (tilt), or rotation (see Fig 1.2). The 3D data sets gathered in this thesis were acquired through a 'tilt style' transducer.

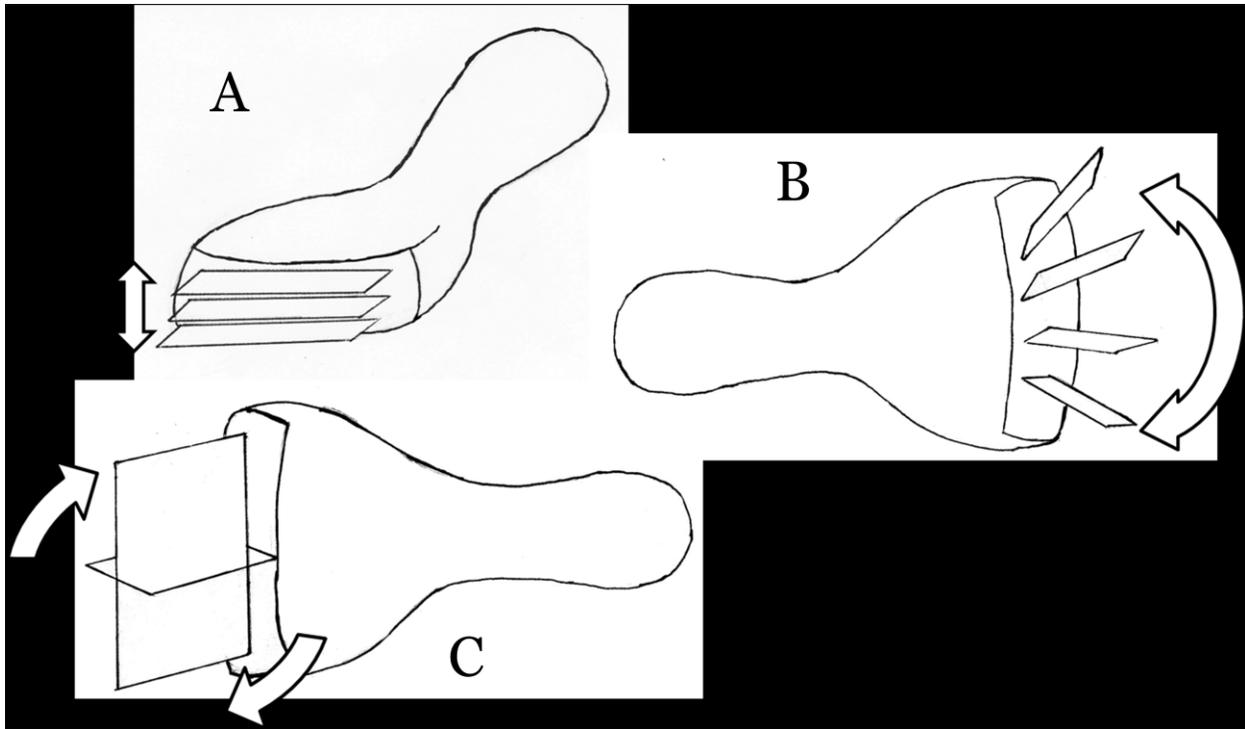


Figure 1.2 Mechanical Localizers used in 3D Ultrasound. A represents a linear transducer, B is a tilt style, and C is a rotational transducer. Image adapted from [3].

## 1.2 HIP DYSPLASIA

### 1.2.1 DEFINITION

Developmental dysplasia of the hip (DDH) is a common condition with prevalence varying from 1.6-28.5/1000 infants [10-12]. It is characterized by a shallow hip joint in which the acetabulum is smoothed and does not contain the femoral head, see Fig 1.3.

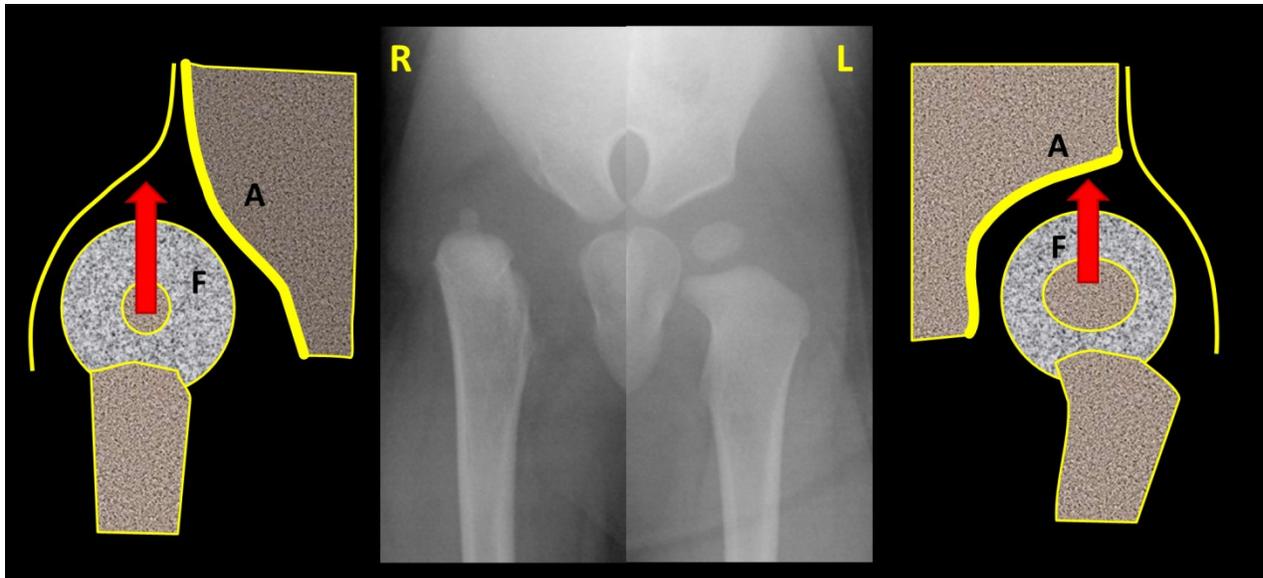


Figure 1.3 Illustration of Hip Dysplasia. On the right hip, the force from weight-bearing is transmitted upward from the ground to the femoral head (F) and the head tends to slide out of the socket created by the acetabulum (A), requiring soft tissue structure such as the labrum to take heavy loads to preserve stability; in the left hip, the force is stabilized by acetabulum anatomy within the well defined socket.

DDH includes hips that are unstable, subluxed, dislocated (luxated) and/or have a malformed acetabulum [13]. As the name implies, DDH is a developmental process where the results of the clinical exams may change and progress. Also it is important to note that the commonly used misnomer Congenital dislocation of the hip (CDH) is sometimes used interchangeably with DDH, but is more appropriately used for children who have “permanently displaced hips” which were diagnosed late and occurred prenatally, usually after they had started to walk [13, 14].

### 1.2.2 PELVIS ANATOMY

The hip joint is based on a ball (the femoral head) and socket (the acetabulum) joint. During embryonic development, the femoral head and acetabulum begin from the same block of cartilage and then develop as a result of genetics and loading [13]. Hip dysplasia is based on the interaction of these two bones, while several others, such as the os ischium, are important

landmarks for visualization in ultrasound screening. Select bones can be seen identified in Fig 1.4.

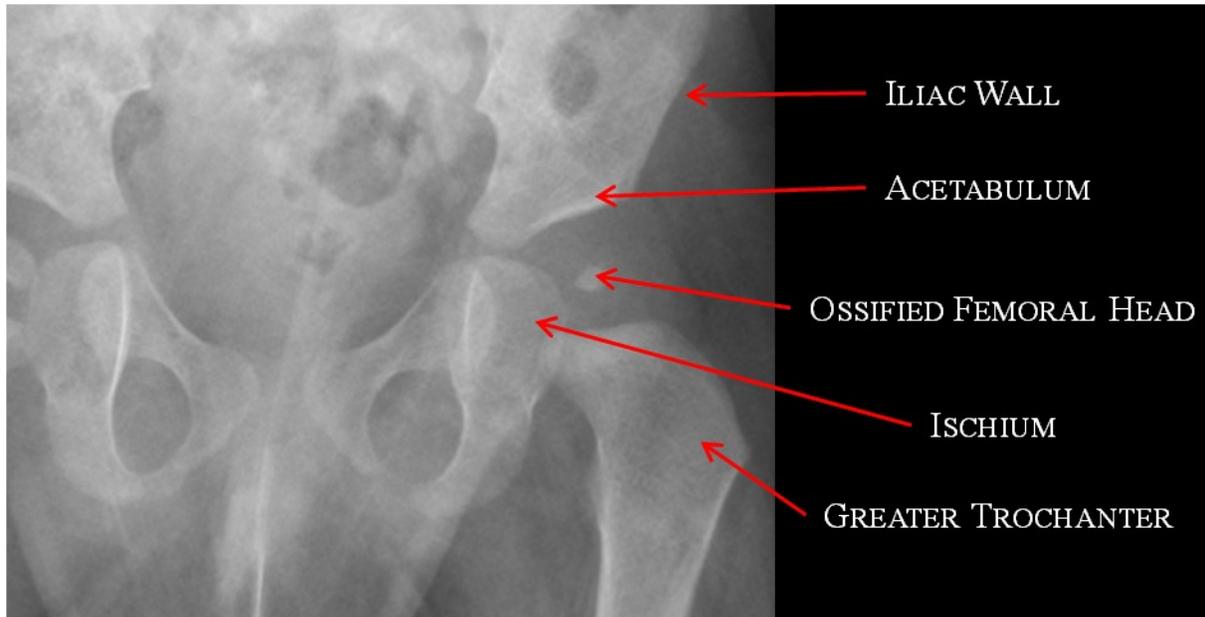


Figure 1.4 Infant Hip Radiograph with important bones labeled.

The infant and adult pelvis are very different as many of the bones are not fully formed. Especially during the final 4 weeks of pregnancy, mechanical forces have a role in bone remodeling [13]. When the femoral head is not properly contained in the acetabulum, the socket does not grow and remodel correctly, and this is a main factor in the acetabulum becoming more shallow [13]. The labrum begins to take on extra forces and stresses in order to compensate for the acetabulum's lack of stability.

### 1.2.3 DIAGNOSIS

Risk factors for DDH include positive family history, race, firstborn child, breech, and female gender [11, 15-17]. The higher incidence in girls may be due to their extra susceptibility to the maternal hormone relaxin, which can cause ligamentous laxity with the resultant instability of the hip [13, 18]. Asymmetric thigh creases, gluteal folds, apparent limb length discrepancy,

and restricted motion are significant but do not mean that the infant has hip dysplasia. As a result the examiner should look for these signs, but not rely upon them for diagnosis [16].

Screening consists of initial applications of the Barlow and Ortolani tests [13], which test the stability of the hip within the socket. A positive test is indicated by the presence of a 'clunking' noise or sensation. For the Ortolani test, a clunk represents the femoral head being reduced into the acetabulum as the hip is adducted while lifting the leg anteriorly when the hip is flexed 90°. Clunking in the Barlow test is indicative of the femoral head exiting the acetabulum posteriorly when pressure is applied on the knee posteriorly. By 8-12 weeks old, the Barlow and Ortolani maneuvers are generally no longer positive regardless of that status of the femoral head due to the decrease in capsule laxity and increasing in muscle tightness [15, 16]. Additionally, some malformations of the acetabulum and/or severe dysplasia cannot be detected by these clinical exams and must be visualized with an imaging technique. Thus, after or instead of the initial clinical tests the infant may be sent for an ultrasound scan to determine if the hip is dysplastic. Other diagnostic imaging performed may be an x-ray, which is of limited utility until the infant's age is above 6 months [15], since the infant's bones are not completely ossified and therefore are unable to be visualized fully. Although all of these methods may be used by a clinician to diagnose DDH in early infancy, there is no consensus on how to provide a definitive diagnosis. This is especially true during the first 8 weeks of life, as it has been suggested that abnormal clinical and sonographic findings may resolve spontaneously within this age group [12, 18, 19]. Additionally, the diagnostic utility of imaging has yet to be fully established in DDH as they are fraught with false-positive and false-negative results [18, 20]. Despite this, real-time ultrasonography has been established as an accurate method for imaging the hip during the first few months of life [12, 15, 21]. Clinical ultrasound tests generally follow the principles laid out by Graf in his 1984 paper [21], as well as dynamic testing in the Harcke method [22]. ACR-AIUM guidelines recommend that hospitals perform imaging in two views, coronal view (in the standard plane) and a transverse view of the flexed hip; the transverse view should be performed

with and without stress to assess stability of the hip [15]. In the Graf method, the standard coronal plane image must be achieved; this standard plane must contain a flat iliac wing, clear labrum, round femoral head, and the presence of the os ischium. DDH is generally quantified on ultrasound by the acetabular alpha ( $\alpha$ ), and beta ( $\beta$ ) angles, as well as total femoral head coverage (Fig 1.5).

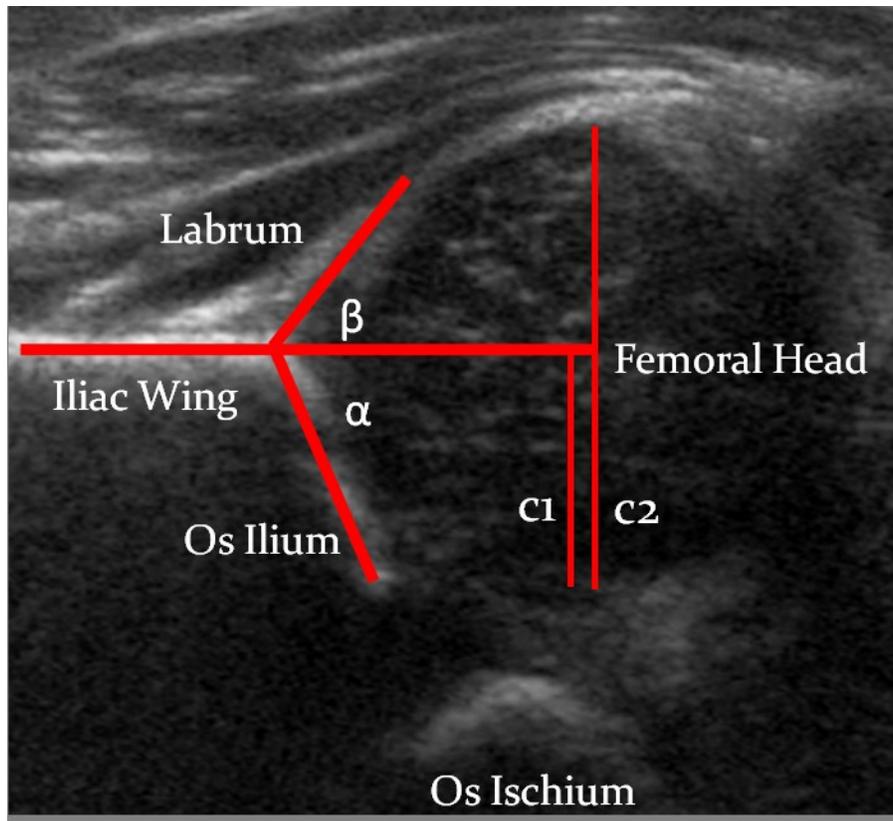


Figure 1.5 Standard Graf Coronal Plane shown with the required standard landmarks in a high quality image are shown (flat iliac wing, os ilium, os ischium, round femoral head), as well as alpha angle ( $\alpha$ ), beta angle ( $\beta$ ), and coverage (measured as C1 line length/C2 line length)

#### 1.2.4 TREATMENT

Untreated DDH is associated with premature osteoarthritis [23] and can account for up to 30% of hip replacements in patients younger than 60 years old [24]. Some advocate treatment within the first 8 weeks [12]; however, due to the spontaneous resolving nature of

some cases, others suggest surveillance or no follow up [18]. If there are abnormal ultrasound findings or clinical exam issues past 8 weeks, then a referral to an orthopedic surgeon is strongly recommended. Orthopedic surgeons will begin treatment with a Pavlik harness if the infant is less than 6 months old [25] or another type of flexion-abduction-external rotation harness, splint, or cast [13]. The Pavlik harness, named for Arnold Pavlik, has been in use since 1950 and achieves suitable flexion of the hips as well as nonviolent unforced abduction of the hips with simultaneous facilitation of movement in the hip joint. This results in spontaneous centralization of the femoral head in the acetabulum [26]. An example of an infant within the Pavlik harness can be seen in Fig 1.6.



Figure 1.6 Infant secured in a Pavlik Harness.

There are varying opinions on the exact duration of flexion harness treatment in DDH, with recommendations ranging from 21 days [27] to 6 months [28]. The Pavlik harness boasts a 95% success rate overall [29], with most failures occurring in dislocated hips or older patients [30]. Graf category III hips ( $\alpha < 43^\circ$ , subluxed hip) are successful in 97% of cases, while Graf IV ( $\alpha < 43^\circ$ , dislocated hip) category successes occur only 50% of the time [25]. Complications from the Pavlik harness are minimal with the main risk being avascular necrosis (AVN) due to hyperabduction, but this affects only 1% of patients [12]. When harness treatment fails, or the patient is older than 6 months, treatment is moved onto surgery and/or the infant is placed into a spica cast [25]. The less invasive corrective method is a closed reduction which aims to release the adductor tendon. Once the tendon is released pressure is taken off the soft surfaces of the hip and the femoral head remains in the socket more easily. The hip is then physically manipulated into the best possible position within the socket and the child is put into a spica cast to keep the newly aligned positions while the joint heals, all while under anesthesia [31]. The more invasive form of reduction is referred to as an open reduction. This entails surgically removing obstacles to reduction, such as the labrum, while at the same time lengthening tendons and tightening the capsule [32]. Like the closed reduction, the infant is placed into a spica cast to retain positioning while the joint heals after surgery [33]. In older patients or more severe cases, osteotomies of the femur or pelvis may be required, with a range of bone reforming techniques used, depending on the exact bone deficiency [13, 14, 34-38].

### 1.3 THESIS SCOPE

As mentioned, there is currently no consensus on the method to screen for DDH. Clinical and diagnostic imaging exams lack sensitivity and are fraught with false-positive results. DDH is a condition for which screening could be highly appropriate as the condition is common and treatments are highly successful if detected in infancy, while the morbidity from the potential hip osteoarthritis which may develop later in life is substantial. The most common current

screening modality, 2D ultrasound, has high inter/intra observer and scan variability. 3D ultrasound probes can now acquire images rapidly enough for scans to be feasible on even the most uncooperative squirming infants. This thesis contains methods intended to illustrate the current intrinsic flaws of 2D ultrasound techniques on DDH (chapter 2), while proposing alternative indices to improve on current methods (chapter 3), as well as novel techniques to generate 3D indices intended to quantify the severity of DDH (chapters 4 and 5). The final chapter (chapter 6) of this thesis contains information on future directions as well as general discussions and conclusions formed from the rest of the chapters.

#### 1.4 REFERENCES

- [1] P. Hoskins, K. Martin and A. Thrush. *Diagnostic Ultrasound : Physics and Equipment* 2010[University of Alberta JW Scott Health Sciences RC 78.7 U4 D516 2010].
- [2] A. B. Wolbarst. *Physics of Radiology / Anthony Brinton Wolbarst ; with Illustrations by Gordon Cook* 2005[Cross Cancer Institute RC 78.7 D53 W65 2005].
- [3] D. Kane, W. Grassi, R. Sturrock and P. V. Balint. A brief history of musculoskeletal ultrasound: 'from bats and ships to babies and hips'. *Rheumatology (Oxford)* 43(7), pp. 931-933. 2004.
- [4] D. L. Hykes, W. R. Hedrick and D. E. Starchman. *Ultrasound Physics and Instrumentation / David L. Hykes, Wayne R. Hedrick, Dale E. Starchman* 1992[University of Alberta JW Scott Health Sciences RC 78.7 U4 H99 1992].
- [5] K. Karadayi, R. Managuli and Y. Kim. Three-dimensional ultrasound: From acquisition to visualization and from algorithms to systems. *IEEE REVIEWS IN BIOMEDICAL ENGINEERING* pp. 23-39. 2009.
- [6] J. T. Bushberg. *The Essential Physics of Medical Imaging / Jerrold T. Bushberg ... [Et Al.]* 2012[University of Alberta JW Scott Health Sciences RC 78.7 D53 E78 2012].
- [7] A. Fenster and D. B. Downey. 3-D ultrasound imaging: A review. *IEEE ENGINEERING IN MEDICINE AND BIOLOGY MAGAZINE* 15(6), pp. 41-51. 1996.
- [8] D. H. Pretorius, N. Borok, M. S. Coffler and T. R. Nelson. Three-dimensional ultrasound in obstetrics and gynecology. *Radiol. Clin. North Am.* 39(3), pp. 499-+. 2001.
- [9] A. Abuhamad. Automated multiplanar imaging: A novel approach to ultrasonography. *J. Ultrasound Med.* 23(5), pp. 573-576. 2004.

- [10] L. Mercier, T. Lango, F. Lindseth and D. L. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *ULTRASOUND IN MEDICINE AND BIOLOGY* 31(4), pp. 449-471. 2005.
- [11] D. Shorter, T. Hong and D. A. Osborn. Cochrane review: Screening programmes for developmental dysplasia of the hip in newborn infants. *Evid Based. Child. Health.* 8(1), pp. 11-54. 2013.
- [12] S. K. Storer and D. L. Skaggs. Developmental dysplasia of the hip. *Am. Fam. Physician* 74(8), pp. 1310-1316. 2006.
- [13] M. B. Ozonoff. *Pediatric Orthopedic Radiology / M.B. Ozonoff* 1992[University of Alberta JW Scott Health Sciences RD 732.3 C48 O99 1992].
- [14] R. B. Salter, G. Hansson and G. H. Thompson. Innominate osteotomy in the management of residual congenital subluxation of the hip in young adults. *Clin Orthop Relat Res* (182), pp. 53-68. 1984.
- [15] American Institute of Ultrasound in Medicine. AIUM practice guideline for the performance of an ultrasound examination for detection and assessment of developmental dysplasia of the hip. *J. Ultrasound Med.* 32(7), pp. 1307-1317. 2013.
- [16] C. Dezateux and K. Rosendahl. Developmental dysplasia of the hip. *Lancet* 369(9572), pp. 1541-1552. 2007.
- [17] C. E. Bache, J. Clegg and M. Herron. Risk factors for developmental dysplasia of the hip: Ultrasonographic findings in the neonatal period. *J. Pediatr. Orthop. B* 11(3), pp. 212-218. 2002.
- [18] A. Roposch, L. Q. Liu, F. Hefti, N. M. P. Clarke and J. H. Wedge. Standardized diagnostic criteria for developmental dysplasia of the hip in early infancy. *Clinical Orthopaedics and Related Research* (12), pp. 3451. 2011.
- [19] V. Bialik, G. M. Bialik, S. Blazer, P. Sujov, F. Wiener and M. Berant. Developmental dysplasia of the hip: A new approach to incidence. *Pediatrics* 103(1), pp. 93-99. 1999.
- [20] D. Shorter, T. Hong and D. A. Osborn. Cochrane review: Screening programmes for developmental dysplasia of the hip in newborn infants. *Evidence-Based Child Health* 8(1), pp. 11. 2013.
- [21] R. Graf. Fundamentals of sonographic diagnosis of infant hip dysplasia. *J. Pediatr. Orthop.* 4(6), pp. 735-740. 1984.
- [22] H. T. Harcke and L. E. Grissom. Performing dynamic sonography of the infant hip. *AJR Am. J. Roentgenol.* 155(4), pp. 837-844. 1990.
- [23] K. Bin, J. M. Laville and F. Salmeron. Developmental dysplasia of the hip in neonates: Evolution of acetabular dysplasia after hip stabilization by brief pavlik harness treatment. *Orthop. Traumatol. Surg. Res.* 100(4), pp. 357-361. 2014.

- [24] O. Furnes, S. A. Lie, B. Espehaug, S. E. Vollset, L. B. Engesaeter and L. I. Havelin. Hip disease and the prognosis of total hip replacements: A review of 53 698 primary total hip replacements reported to the norwegian arthroplasty register 1987-99. *J BONE JOINT SURG (BR)* 83B(4), pp. 579-586. 2001.
- [25] A. K. Mostert, N. J. Tulp and R. M. Castelein. Results of pavlik harness treatment for neonatal hip dislocation as related to graf's sonographic classification. *J. Pediatr. Orthop.* 20(3), pp. 306-310. 2000.
- [26] A. Pavlik. Stirrups as an aid in the treatment of congenital dysplasias of the hip in children. by arnold pavlik, 1950. *J. Pediatr. Orthop.* 9(2), pp. 157-159. 1989.
- [27] Y. Lefevre, J. M. Laville and F. Salmeron. Early short-term treatment of neonatal hip instability with the pavlik harness. *Rev. Chir. Orthop. Reparatrice Appar. Mot.* 93(2), pp. 150-156. 2007.
- [28] M. Mafalda Santos and G. Filipe. Treatment of congenital hip dislocation using pavlik's harness. long term results. *Rev. Chir. Orthop. Reparatrice Appar. Mot.* 83(1), pp. 41-50. 1997.
- [29] J. P. Cashman, J. Round, G. Taylor and N. M. Clarke. The natural history of developmental dysplasia of the hip after early supervised treatment in the pavlik harness. A prospective, longitudinal follow-up. *J. Bone Joint Surg. Br.* 84(3), pp. 418-425. 2002.
- [30] K. K. White, D. J. Sucato, S. Agrawal and R. Browne. Ultrasonographic findings in hips with a positive ortolani sign and their relationship to pavlik harness failure. *J. Bone Joint Surg. Am.* 92(1), pp. 113-120. 2010.
- [31] P. R. Fleissner Jr, C. J. Ciccarelli, R. E. Eilert, F. M. Chang and G. L. Glancy. The success of closed reduction in the treatment of complex developmental dislocation of the hip. *J. Pediatr. Orthop.* 14(5), pp. 631-635. 1994.
- [32] D. Olszewski C. and L. Karol A. The medial ludloff open reduction in developmental dysplasia of the hip before the age of walking. *OPER TECHNIQ ORTHOP* 23(3), pp. 109-114. 2013.
- [33] W. N. Sankar, C. R. Young, A. G. Lin, S. A. Crow, K. D. Baldwin and C. F. Moseley. Risk factors for failure after open reduction for DDH: A matched cohort analysis. *J. Pediatr. Orthop.* 31(3), pp. 232-239. 2011.
- [34] T. Jackson, J. Watson, J. LaReau and B. Domb. Periacetabular osteotomy and arthroscopic labral repair after failed hip arthroscopy due to iatrogenic aggravation of hip dysplasia. *Knee Surgery, Sports Traumatology, Arthroscopy* 22(4), pp. 911-914. 2014.
- [35] M. Fu, S. Xiang, Z. Zhang, G. Huang, J. Liu, X. Duan, Z. Yang, P. Wu and W. Liao. The biomechanical differences of rotational acetabular osteotomy, chiari osteotomy and shelf procedure in developmental dysplasia of hip. *BMC Musculoskelet. Disord.* 15pp. 47-2474-15-47. 2014.
- [36] K. Perry I., R. Trousdale T. and R. Sierra J. Hip dysplasia in the young adult: An osteotomy solution. *BONE JOINT J (BR)* 95-B(11), pp. 21-25. 2013.

[37] M. Akiyama, Y. Nakashima, M. Oishi, T. Sato, M. Hirata, D. Hara and Y. Iwamoto. Risk factors for acetabular retroversion in developmental dysplasia of the hip: Does the pemberton osteotomy contribute? *JOURNAL OF ORTHOPAEDIC SCIENCE* 19(1), pp. 90-96. 2014.

[38] C. Erturk, M. A. Altay and U. E. Isikan. A radiological comparison of salter and pemberton osteotomies to improve acetabular deformations in developmental dysplasia of the hip. *JOURNAL OF PEDIATRIC ORTHOPAEDICS-PART B* 22(6), pp. 527-532. 2013.

# CHAPTER 2

## Chapter 2 ..

### POTENTIAL FOR CHANGE IN ULTRASOUND DIAGNOSIS OF HIP DYSPLASIA DUE SOLELY TO CHANGES IN PROBE ORIENTATION: PATTERNS OF ALPHA ANGLE VARIATION REVEALED USING 3D ULTRASOUND<sup>1</sup>

#### 2.1 INTRODUCTION

The most widely used index for evaluation of DDH is the acetabular alpha angle, measured by using the Graf method [1] on a standardized two-dimensional (2D) coronal image (see Fig 1.5). The Graf ultrasound criteria are intended to standardize the plane in which imaging is performed. With the Graf method, a hip with an alpha angle greater than 60° is classified as normal, a hip with an alpha angle from 50° to 59° is classified as having mild dysplasia, a hip with an alpha angle from 43° to 49° is classified as having moderate dysplasia, and a hip with an alpha angle of less than 43° is classified as having severe dysplasia (this information is summarized in Table 2.1).

---

<sup>1</sup>A version of this Chapter has been accepted for publication in *Radiology* – Jaremko JL., Mabee M, Swami VG, et al. Potential for change in the US diagnosis of hip dysplasia solely caused by changes in probe orientation: Patterns of alpha-angle variation revealed by three-dimensional US. *Radiology* 2014 (in press). © Radiological Society of North America.

**Table 2.1 Categories of Hip Dysplasia by Alpha Angle**

	Normal	Mild Dysplasia	Moderate Dysplasia	Severe Dysplasia
Alpha Angle	60°	50-59°	43-49°	<43°

Diagnosing dysplasia requires combining clinical and imaging assessment, and, in current practice, quantifying the imaging component relies heavily upon the value of the alpha angle. The alpha angle is subject to wide variability of two main types: interscan and interobserver [2-6]. Interobserver variability is easily quantified: two people measuring the alpha angle differently on the same 2D US image. Inter scan error is more difficult to analyze. Because the US probe is handheld, each scan inevitably is obtained with the probe and patient oriented differently, which, because of the complex three-dimensional (3D) acetabular shape, results in images showing different alpha angles. Quantification of alpha-angle variability due to probe orientation in live patients or in dysplastic hips has not been well established in the literature. This variability is important because it could lead to misdiagnosis of DDH. Missed cases of DDH are particularly concerning because DDH is usually easily treated conservatively if diagnosed in infancy [7], whereas a later diagnosis is associated with more invasive treatments and lower success rates [8]. To avoid this problem, clinical practice often includes follow-up US until initially borderline hips definitely appear normal, which comes with associated costs and parental anxiety, or routine over treatment until hips normalize. The poor diagnostic accuracy of current US of the hip may account for its lack of beneficial effect on patient care, with rates of late DDH diagnosis or surgery not reduced by using universal or selective US, compared with rates of clinical screening alone, despite higher treatment rates [9].

By using 3D ultrasound, the entire acetabular shape can be acquired in a single scan, and the effect of differences in 2D probe orientation can be explored by viewing sections cut along

different planes through the 3D data. Compared with obtaining repeated 2D measurements, this approach has the advantage of holding all variables constant save for the 2D section orientation, allowing the effects of this orientation to be evaluated independently. The purpose of this study was to use 3D US to quantify the range of interscan variability in alpha angle caused by changes in 2D US probe orientation.

## 2.2 MATERIALS AND METHODS

This prospective study was approved by the University of Alberta Health Research Ethics Board (Biomedical Panel, Pro00032107). Imaging was performed at a tertiary pediatric hospital from October 2012 to November 2013. At the first routine clinical 2D US of the hip in each patient, written informed consent was obtained from a parent to add 3D US of the hip at the same visit. The imaging indication was clinical suspicion of DDH because of laxity at examination, asymmetric hip creases, or other risk factors such as a family history of a positive finding for DDH. Because dysplasia can be unilateral or bilateral, we included each hip separately. Considering only the 3D scans showing the entire range of planes necessary for this study, we had 56 hips in 35 patients, and 26 (74%) of them were female. Patients underwent US at the mean age of 41.7 days (range, 4-112 days; for female infants, mean age was 38.7 [range, 6 – 112 days]; for male infants, mean age was 50.2 days [range, 4 -111 days]). They received routine care at a pediatric orthopedic clinic from one or more of five clinic surgeons who were blinded to 3D US images and findings. We observed clinical care for at least 6 months to classify each imaged hip as normal at the first orthopedic assessment (category 0; n = 20), questionably abnormal initially but with findings that resolved spontaneously at follow-up imaging and clinical examination (category 1; n = 23), or dysplastic and requiring treatment by using a Pavlik harness and/or surgery (category 2; n = 13).

### 2.2.1 IMAGING

All imaging was performed by using platforms (Philips iU22; Philips Healthcare, Andover, Mass). We performed conventional 2D US in both hips by using a 12-MHz linear transducer, including static coronal imaging in the Graf standard plane and axial dynamic imaging to assess for hip stability by using our usual protocol. Two-dimensional US images were interpreted by a pediatric radiologist, with results and images made available to referring clinicians. In addition, two study team members, including a radiologist, technologist, or medical or graduate student trained by the study radiologists, used a high-resolution 13-MHz 3D linear (13VL5; Philips Healthcare, Andover, Mass) transducer in the coronal orientation to obtain a 3D US data set at each hip. With this transducer, we performed a 3.2 second automated sweep through a range of  $\pm 15$  degrees to generate a 3D data set of 256 US sections that were 0.13 mm thick, each containing 411 X 192 pixels measuring 0.11 X 0.20 mm. These 3D scans were not released for use in clinical treatment.

### 2.2.2 IMAGING PROCESSING

Images were analyzed off-line by using custom software (MATLAB R2010a-2012; MathWorks, Natick, Mass) that allowed viewing of 3D US images followed by markup of points, lines, and curves replicating the functionality of standard cardiac magnetic resonance imaging workstation analysis of features such as ventricular contours. The major obstacle to performing this analysis was the proprietary 3D data file format. A key function of the custom software was to extract manufacturer-specific pixel spacing information from the 3D US Digital Imaging and Communications in Medicine (DICOM) data. For each hip, the relevant anatomic contours were traced (acetabular rim, femoral head cartilage surface, os ischium) on the original 3D US data set by using a customized semi-automated interactive interface that allowed contours traced on selected sections to be interpolated to intervening sections, then reviewed and corrected by using a nudging tool. A 3D model then was generated for each structure by creating a patch surface connecting the traced and interpolated contours (Fig 2.1).

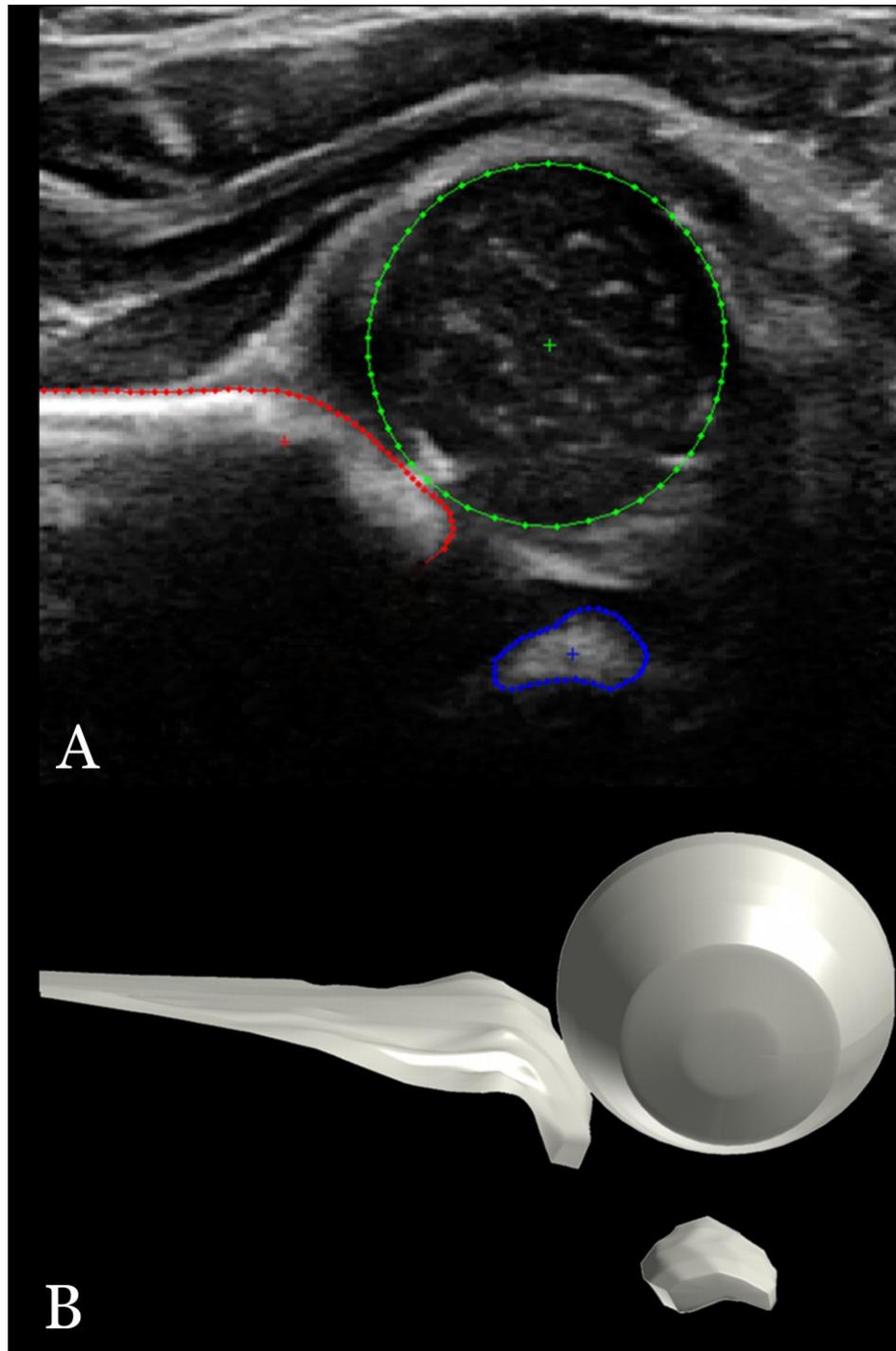


Figure 2.1 Generation of a 3D hip surface model from 3D US images. (a) The surfaces of the femoral head, ilium, and os ischium are traced on sections of 3D data at regular intervals and interpolated into the intervening sections, with adjustments made manually by nudging contours that do not match underlying anatomy into optimized position. (b) Resulting surface model demonstrating the complex 3D anatomy of the acetabulum (structure at left).

The training process involved initial combined review of anatomy on multi-plane and 3D-reformatted computer tomographic scans of a normal infant pelvis and a physical model of this pelvis generated by a 3D printer. This process was followed by a preliminary trials in which surface models from 3D US were traced by both observers in 15 hips in eight patients ranging from normal to severely dysplastic, separate from the final study, and consensus review of models produced by two observers to ensure anatomically appropriate appearance. Tracings in the final study were reviewed and approved by a radiologist.

As a reference orientation, we required an initial central plane located in the middle of the acetabulum in each patient. We generated this central plane in a simple way from two landmarks located on the acetabulum bone. The observer located the corner point of the acetabular angle in the most anterior and posterior sections in which this angle could be defined, points A and P on Figure 2.2.

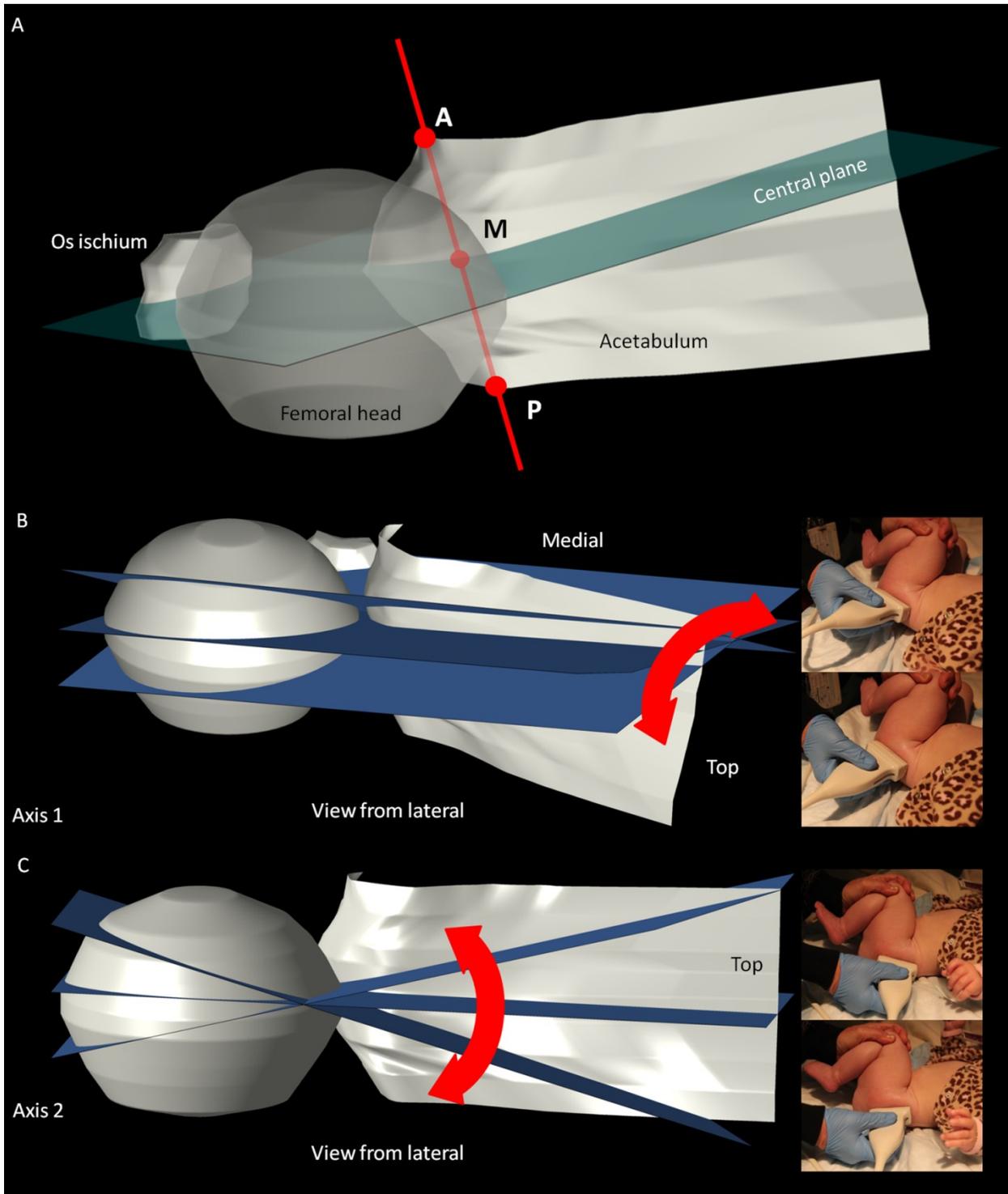


Figure 2.2 Model demonstrating the central plane and rotated planes. (a) Central plane is obtained through the middle of the acetabulum as the plane perpendicular to the line joining the corner points of the most anterior (A) and posterior (P) positions in which an acetabular angle can be drawn at the midpoint (M) of this line. (b) Planes were

cut at angles rotated from the central plane at  $5^\circ$  intervals in axis 1, representing probe tilt around a craniocaudal axis (ie, rotating the base of the transducer from the table toward the ceiling). (c) Planes were similarly cut at  $5^\circ$  intervals around a transverse axis (axis 2), representing probe tilt as if twisting the transducer around a line joining the two hips.

This central plane then was generated automatically perpendicular to the midpoint of line *A-P* (Fig 2.2). Three-dimensional data were resampled along this plane to produce the 2D US image that would result from a 2D probe placed in exactly that location and orientation. From this reference central plane, 2D US images were extracted along orientations rotated in  $5^\circ$  increments from  $-50^\circ$  to  $50^\circ$  away, as if slowly rotating a 2D US probe. Images were produced by rotation around axis 1 oriented cranio-caudally (Fig 2.2b) and axis 2 oriented transversely between hips (Fig 2.2c).

Each of the resulting 21 images obtained at  $5^\circ$  increments of probe orientation for each axis were reviewed for quality by two observers (a radiologist with dual pediatric and musculoskeletal fellowship training and 10 years of imaging experience, and the author of this thesis). Image quality was scored on the basis of the number of major Graf standard plane criteria met (maximum, four): flat horizontal iliac wing, labrum visible, os ischium present, and midportion of femoral head visible. As a nearly spherical structure, the midpoint of the femoral head is best determined by observing that the image saved is that in which the head has the largest possible radius. We rated the femoral head acceptable on an image if it was plausible from review of that image that the middle one-third of the femoral head was included, regardless of whether other 3D US images of the region (which would not be available to a radiologist reporting 2D US findings) might show a larger femoral head radius. Only 18% (428 of 2352) of the images reviewed had a quality score of 4 from both observers. On each of these, the observers separately measured the alpha angle between acetabular roof and iliac lines, and these observers were blinded to the alpha angle measured on other images of that hip, as well as to each other's work and to clinical data. For intraobserver variability, the alpha angle was

remeasured by the same observer on each of the 118 images obtained in this fashion in a subset of 15 hips, 1 month after the initial reading session, and this observer was blinded to the previously measured alpha angles.

### 2.2.3 STATISTICAL ANALYSIS

Descriptive statistics were recorded as the mean  $\pm$  standard deviation. We computed differences in alpha angles between observers and between observations at various plane orientations. We characterized variability in alpha angles as the upper limit of agreement in these tests [10]. For interobserver variability (assessment of the same image by different observers), this was the same as the repeatability coefficient (1.96 multiplied by the standard deviation), as also measured by Gwynne Jones et al [5], whereas for interscan variability, this was the mean difference plus 1.96 multiplied by the standard deviation (i.e., the upper limit of the 95% confidence interval [CI] for the alpha angle). Interobserver variability was assessed by using mean difference and intraclass correlation coefficients calculated by using a two-way mixed effects model. We used univariate analysis of variance to compare means in dysplastic and normal hips. Statistics were calculated with software (SPSS 20; SPSS, Chicago, Ill).

## 2.3 RESULTS

### 2.3.1 SECTION ORIENTATION

In all 56 hips, at least one image had an acceptable quality score of 4. The mean  $\pm$  standard deviation of simulated probe orientations providing images with a quality score of 4 was  $24^\circ \pm 10^\circ$  in axis 1 (maximum,  $45^\circ$ ) and  $23^\circ \pm 8.4^\circ$  in axis 2 (maximum,  $40^\circ$ ). The image with the central plane had a quality score of 4 in 38 (68%) of 56 hips and was within a  $10^\circ$  rotation from an image with a quality score of 4 in 50 (89%) of 56 hips.

### 2.3.2 VARIABILITY OF ALPHA ANGLE WITH SECTION ORIENTATION

The alpha angles recorded at each hip by the radiologist across all images with a quality score of 4 varied by a mean of  $6.9^\circ \pm 4.8^\circ$  in different probe orientations along axis 1 (range,

0-21°),  $10.8^\circ \pm 7.1^\circ$  in axis 2 (range, 0-31°); combining both axes the angles varied by a mean of  $13.9^\circ \pm 7.1^\circ$  (upper limit of 95% CI, 28°). Ranges were similar for observer 2. The proportion of this variability occurring for alpha angles greater than 60° may not be clinically relevant because those angles are considered normal. When we considered any measured alpha angle greater than 60° to being equal to 60°, the remaining variability was  $5.5^\circ \pm 6.9^\circ$  (range 0-24°; upper limit of 95% CI, 19°) across the two axes. For nearly all hips and all axes, changes in virtual probe orientation produced 2D images showing visually substantial variation in the acetabular shape corresponding to the numerical variation in alpha angle. For example, the hip in Figure 2.3 had alpha angles measured at 52° - 70° on images with a quality score of 4, depending on the section orientation, an 18° range that was slightly higher than the mean range of 13.9°.

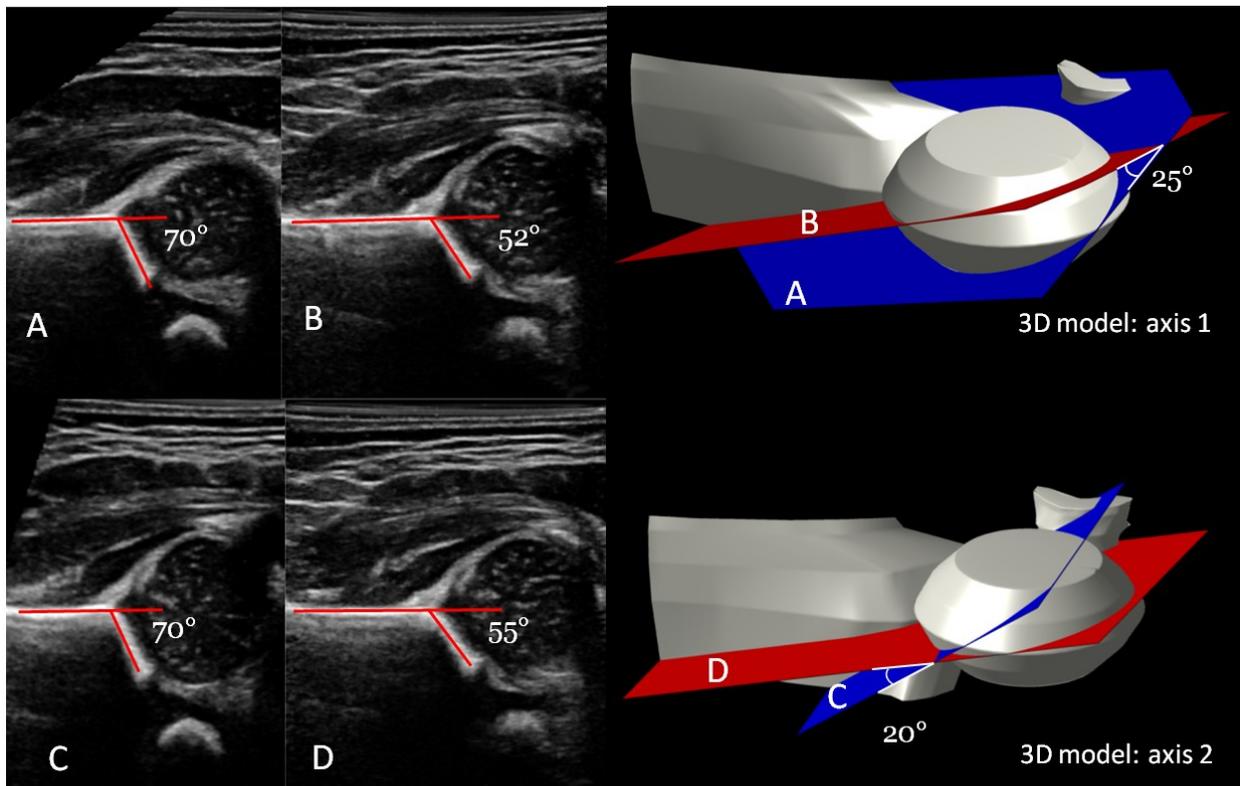


Figure 2.3 Effect of probe tilt in two axes. Top: A and B, 3D US images obtained on cut planes 25° apart along axis 1. Top right, Illustration of acetabular model shows cut planes A and B. Bottom: C and D, US images obtained on cut planes 20° apart along axis 2. Bottom right, Illustration of 3D acetabular model shows cut planes C and D.

Image C shows a trace of curvature at the junction of the iliac wall and acetabular roof, rather than this being the junction of two perfectly straight lines; the os ischium is more clearly defined in A and C than in the other images, and the femoral head is perhaps best seen in B or D, but, in our experience, each image potentially could be accepted in real-world clinical practice. These images demonstrate an overall range of alpha angle variation of  $18^\circ$  ( $52^\circ$ – $70^\circ$ ), which was typical in our data. In this hip, the images were in one of two possible diagnostic categories (normal or mildly dysplastic). The hip was assessed as having findings inconclusive for dysplasia at initial clinical and 2D US examination and was assessed as normal at follow-up without treatment (category 1). Note that the 3D US images were not released to clinicians and had no role in treatment.

To assess the clinical significance of this variation, we considered whether observed differences in alpha angle caused by simulated probe orientation led to a change in the Graf diagnostic category. Considering only 3D US images with a quality score of 4, 26 (46%) of 56 hips had images with findings within only one Graf category; 20 (36%) of 56 hips had images with findings in two Graf categories (Fig 2.4); nine (16%) of 56 hips had images with findings in three Graf categories (Fig 2.5); and one (2%) had images with findings in all four Graf categories.

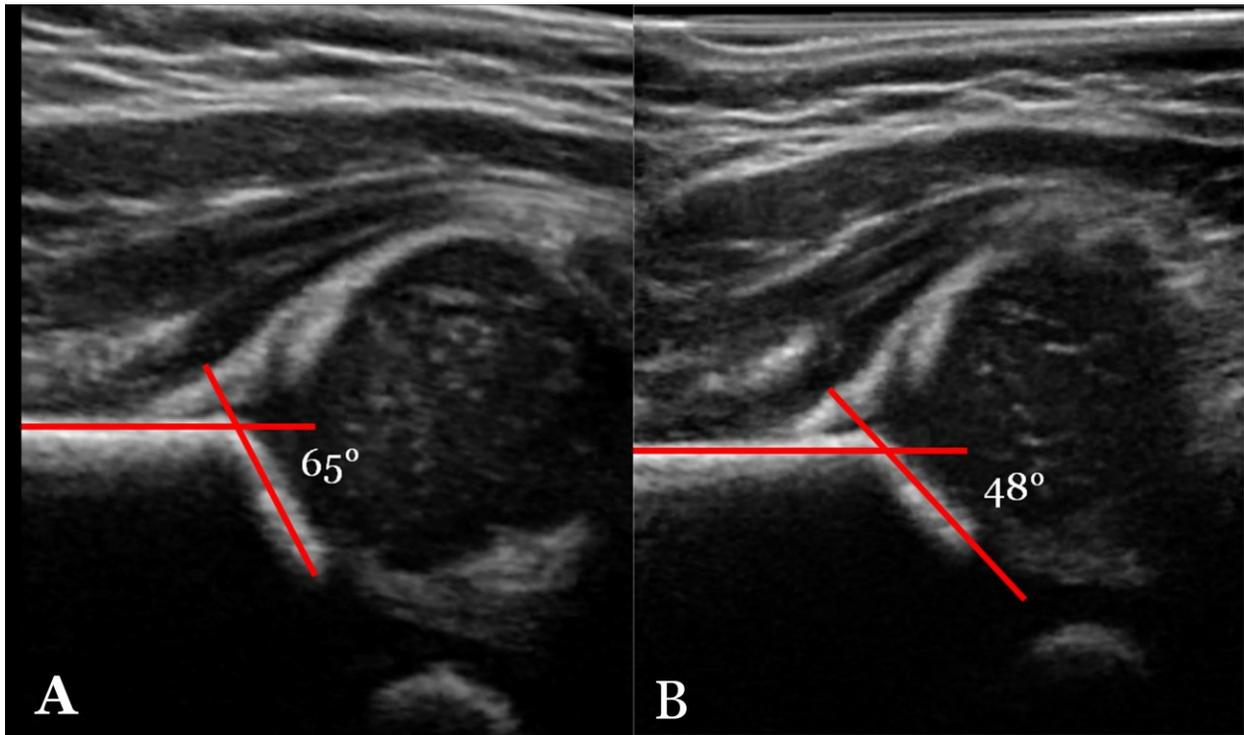


Figure 2.4 US scans in the right hip of a 19-day-old girl show findings crossing two Graf diagnostic categories because of probe tilt. The images extracted along rotated cut planes from a single 3D US scan show alpha angles ranging from 48° (image B) to 65° (image A) (ie, moderately dysplastic in B to normal in A), depending solely on probe tilt. Clinically, findings in this hip were considered borderline (category 1) at initial visit and normalized eventually at follow-up

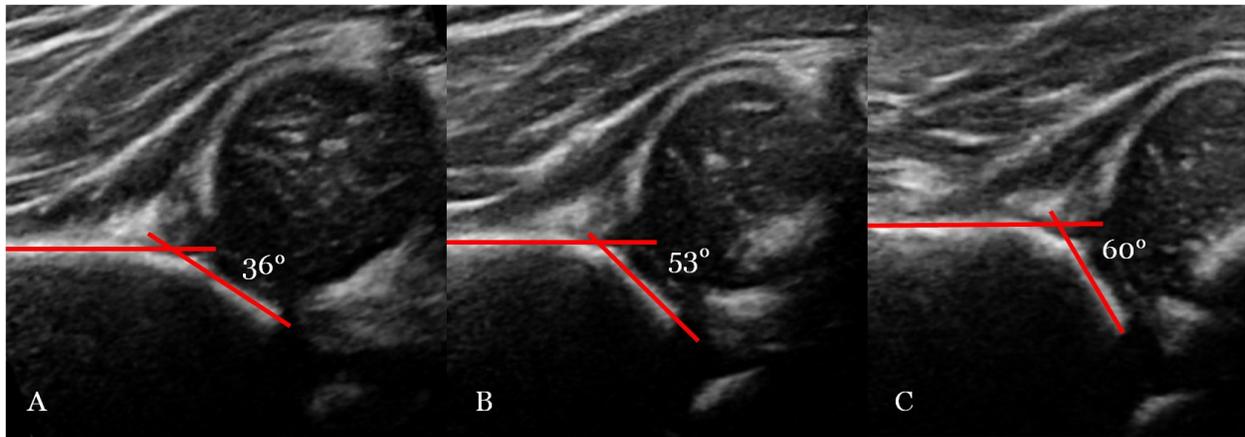


Figure 2.5 US scans in left hip in a 3-month-old girl show findings crossing three Graf diagnostic categories because of probe tilt. The images from cut planes in a single 3D US scan across a 25° range of probe tilt show a grossly abnormal hip with severe acetabular rounding and alpha angles varying from 36° to 60°. Note that the os ischium is seen poorly on the image best showing the midportion of the femoral head (a) and is more clearly visible on image c, where the femoral head is obscured partly by artifact from the greater trochanter, suggesting that the femoral head is probably subluxed. Pavlik harness treatment was unsuccessful, and treatment included spica casting. Reporting the finding as a normal 60° alpha angle in this case would have led to incorrect characterization of this obviously dysplastic hip.

Changes in diagnosis were significantly more common at younger ages: Findings on images in 18 (72%) of 25 hips scanned in infants younger than 31 days were classified in two or more Graf categories versus findings on images in just nine (29%) of 31 hips scanned in older infants ( $P = 0.002$ ).

### 2.3.3 VARIABILITY OF ALPHA ANGLE BETWEEN AND WITHIN OBSERVERS

Four hundred twenty-eight images had a quality score of 4 and were of acceptable quality in 56 hips. In these images, interobserver variation in alpha angle was a mean of  $0.1^\circ$   $\pm$   $5.1^\circ$  (repeatability coefficient,  $10^\circ$ ; coefficient of variation, 0.048). The intraclass correlation coefficient was 0.89. Alpha angles recorded by the two observers were within  $5^\circ$  and  $10^\circ$  of each other in 73% and 97% of cases, respectively. Intraobserver variability was lower, with the mean difference between alpha angles of  $1.8^\circ$   $\pm$   $3.7^\circ$  (repeatability coefficient,  $7.4^\circ$ ) and measurements within  $5^\circ$  and  $10^\circ$  of each other in 82% and 96% of cases, respectively.

### 2.3.4 NORMAL VERSUS DYSPLASTIC HIPS

We observed clinically normal hips that appeared dysplastic in some planes (Fig 2.3) and clinically dysplastic hips that had normal  $60^\circ$  alpha angles in certain planes, either with the acetabular contours appearing grossly abnormal (Fig 2.5) or essentially normal (Fig 2.6).

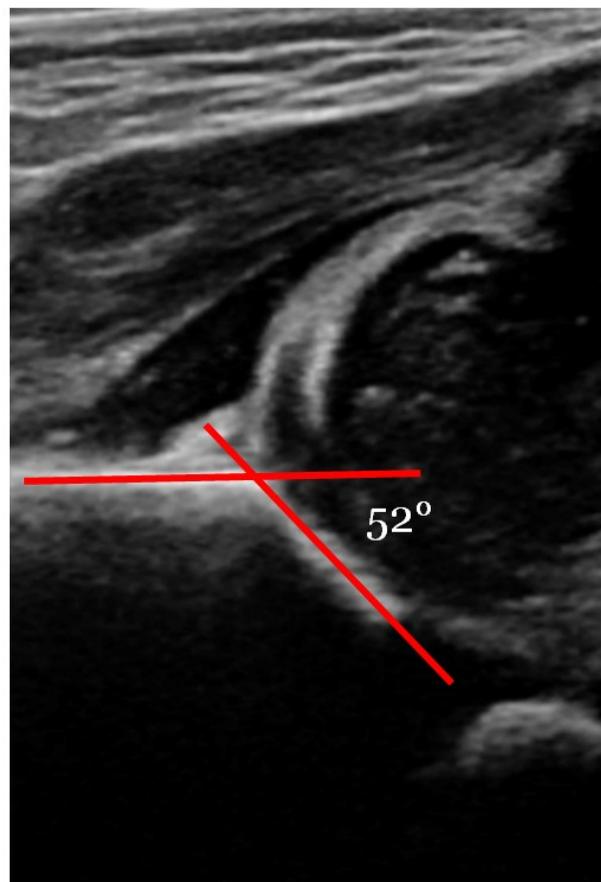
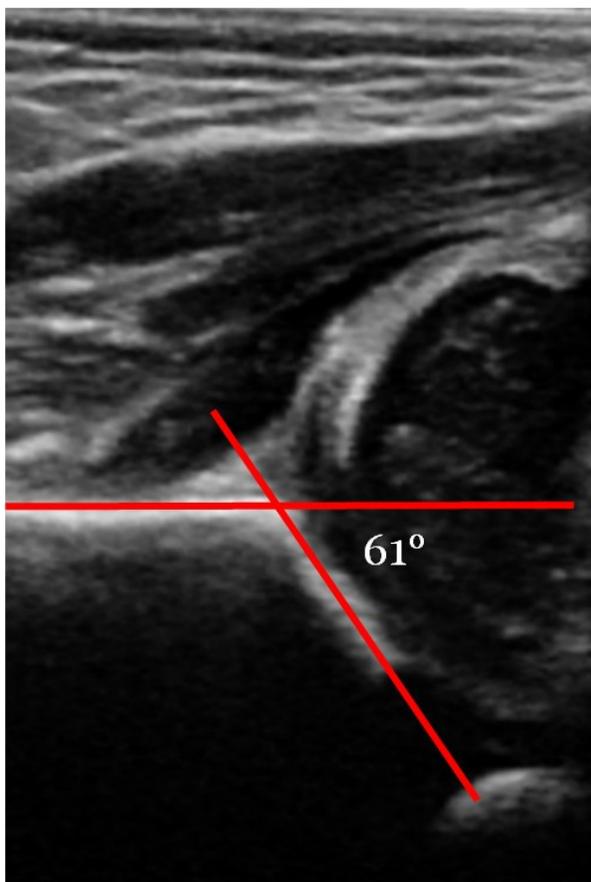


Figure 2.6 US in left hip in a 6-day-old girl shows that a dysplastic hip can appear normal, depending on probe tilt. The images from cut planes in a single 3D US scan show alpha angle variation from 52° to 61°. Other images lacking only the os ischium had alpha angles as low as 48°. Image at left is within normal limits. The hip was judged clinically dysplastic and normalized with treatment with a Pavlik harness.

The alpha angles we observed at 3D US, averaged from acceptable images across all simulated probe orientations in both axes, clearly indicated ( $P < 0.001$ ) clinical diagnostic categories: 3D US images in dysplastic hips requiring treatment (category 2;  $n = 13$ ) had a mean alpha angle of 49.5° versus that in images in initially borderline hips requiring follow-up but no treatment (category 1;  $n = 23$ ; mean angle, 63.2°) and that in images in hips not requiring treatment or follow-up (category 0;  $n = 20$ ; mean angle, 71.4°) (Fig 2.7).

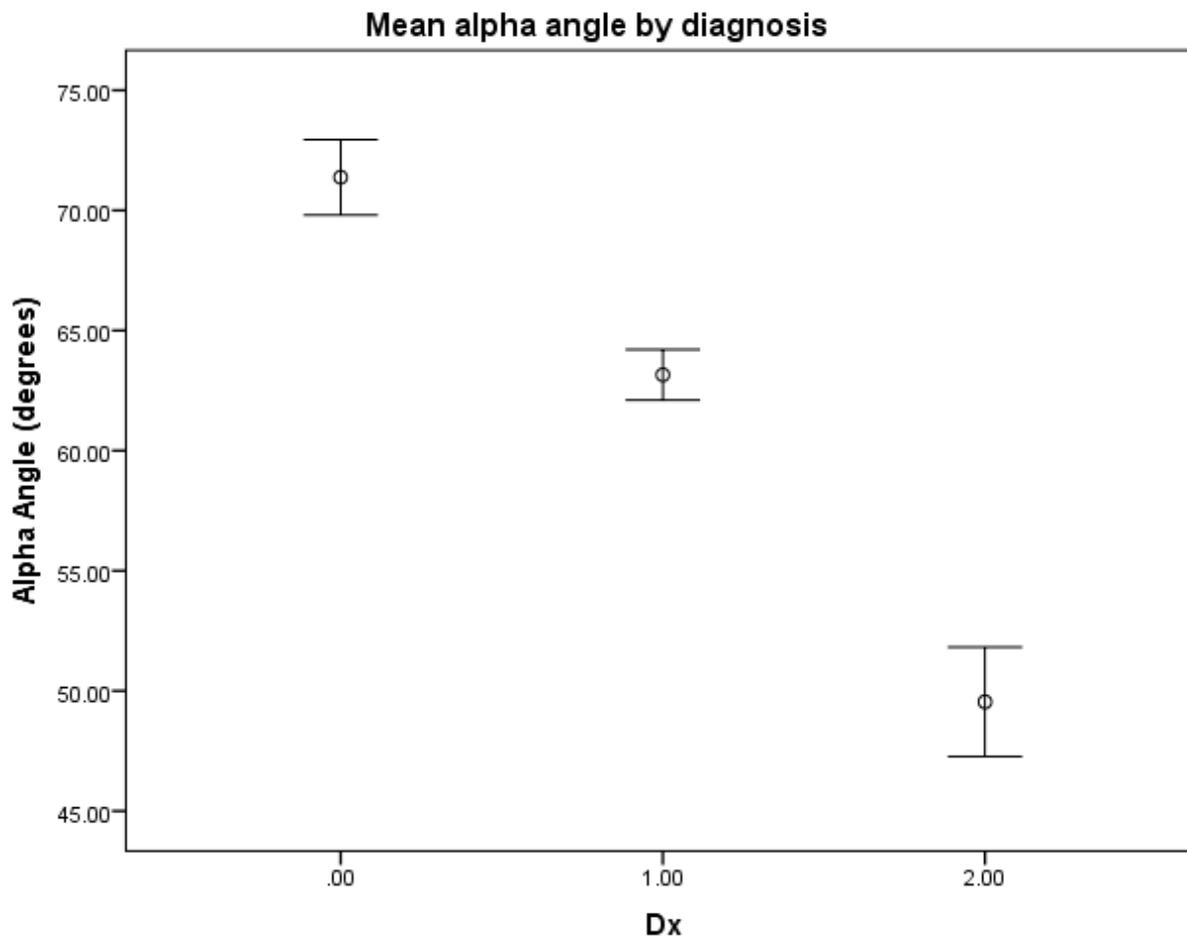


Figure 2.7 Point estimates show differences in alpha angles according to clinical diagnosis (Dx). Points = mean values, error bars = 95% CI of the mean for the alpha angles in all acceptable (quality score 4) images in each hip across all studied probe orientations, 0.0 = normal, 1.0 = hips with initially inconclusive findings but that normalized at follow-up, and 2.0 = dysplastic hips. Mean values differed significantly ( $P < .001$ ) between each category. Note that the normal hips had mean alpha angles substantially higher than the traditional threshold of  $60^\circ$  for detection of DDH, whereas even the hips that were rated as having inconclusive findings on the basis of 2D US and clinical assessment but that ultimately were determined to be normal had mean alpha angles slightly greater than  $60^\circ$  when all acceptable 3D US images were considered.

## 2.4 DISCUSSION

We used normal 3D US technology in infants with normal and dysplastic hips to examine the potential alpha angle inter scan variation solely caused by US probe orientation. Our most important finding was that changing probe orientation altered the alpha angle enough to lead to a change in more than one-half of the imaged hips overall and in nearly three quarters of hips imaged in patients younger than 1 month. We also found that a high-quality image satisfying Graf criteria could be obtained at orientations varying by an average of  $24^\circ$ , or in some patients by as much as  $45^\circ$ , in each axis. This finding means that the sonographer often has a relatively wide range of probe positions in which acceptable images can be generated and must rely on skill and experience to select the best image to save within this range. The interscan variability of alpha angles measured on 2D US images in those different probe orientations was substantial (upper limit of 95% CI:  $19^\circ$ ,  $28^\circ$ ) and higher than interobserver variability (repeatability coefficient,  $10^\circ$ ). DDH diagnosis from the alpha angle measured on a standardized 2D US image is based on the assumption that this is a unique clinically representative image (i.e., that if any trained individual performs US in that hip that day, any image obtained containing appropriate landmarks will have similar acetabular contours and alpha angle). The risk of misdiagnosis and overtreatment because of incorrect probe angulation is a recognized hazard of 2D US, with Graf et al [11] even advocating use of a probe guide-frame system to avoid tilt. Still, to our knowledge,

investigations in only one previous study directly assessed the effect of changes in US transducer angle on alpha angle measurements [12]. They used ex vivo 2D US in five infant cadaver hips in a water bath and showed that the alpha angle in a normal hip could vary from  $48^\circ$  to  $65^\circ$  (i.e. from moderately dysplastic to normal) by changing probe position. Our 3D US study results help confirm these findings in vivo and demonstrate that dysplastic hips can be shown to have normal  $60^\circ$  alpha angles in probe orientations producing images that still meet Graf standard plane criteria. The range of alpha angle variability in our study compares well with data in that study, with an upper limit of agreement of  $19^\circ$  versus a range of  $17^\circ$  in their study [12].

The  $19^\circ$  upper limit of agreement of interscan variability reported here is greater than the approximately  $8^\circ$  reported in a larger clinical study [4]. This finding is expected because we included all US images in each hip that met Graf standards, whereas in clinical practice, sonographers are trained to move the probe through a range of orientations to select the best image of the hip, with the steepest possible alpha angle, deepest acetabular fossa, and largest femoral head radius of all possible images meeting Graf standards. A sonographer typically would not select an image at the margins of the acceptable range of probe orientations. However, in a squirming, crying infant, even the most experienced sonographer has a limited opportunity to truly to assess all possible imaging planes with the probe. The hazard for the interpreting radiologist or surgeon is that the image may be of diagnostic quality but not actually represent the best image of the acetabulum that could have been obtained. Our study results demonstrate the wide possible range of interscan variability in alpha angle measurement. Careful sonographer training and experience, and thorough radiologist assessment of other factors such as acetabular rounding, coverage, and dynamic hip stability, are critical to minimize the practical effect of this variability, particularly in the youngest infants. This is especially true given the relatively high superimposed interobserver variability in alpha angle measurement once an image is obtained. The reliability of US diagnosis of hip dysplasia is likely considerably greater in clinical practice, when the alpha angle, acetabular coverage, acetabular morphology,

and hip stability are assessed together, than when the alpha angle alone is used for diagnosis. Testing this reliability was outside the scope of our study, which focused on the limitations of basing diagnosis primarily on the alpha angle.

We did not assess acetabular coverage formally in this study. Our preliminary testing of this index showed an even wider range of variation with plane orientation than the alpha angle, as expected, because this index involves the 3D geometric relationship between two bones (femoral head and acetabulum) rather than only the shape of the acetabulum.

Although the alpha angles differed significantly between dysplastic and normal hips in our study, it is evident from our analysis that there are intrinsic flaws in using the 2D Graf standard plane as a diagnostic method for DDH. Simply by rotating the US probe, the sonographer could change the diagnostic category for more than one-half of the hips in this study by at least one, and often two, levels (eg, from normal to moderately dysplastic). Because the image on screen still appears diagnostic, and there is no record of the position of the 2D US probe, this error invisibly degrades the accuracy and reliability of 2D US assessment of DDH. Conceptually, 3D US of DDH could reduce or eliminate this error by removing dependence on the specific plane of image acquisition.

Three-dimensional US ideally would allow firm diagnosis of every hip as normal or dysplastic at the initial scan. Although our study was not designed or powered to define these diagnostic thresholds in 3D US, hips classified in different diagnostic categories showed substantially different mean alpha angles across all 3D US imaging planes. We expect that diagnostic accuracy of US for DDH could be improved in the future by considering this additional acetabular shape information available from 3D US, both visually and by measuring alpha angles and novel indexes being developed (discussed later in this thesis).

Our study had limitations. As a pilot study, the sample size was relatively small, and some of our earlier scans included a range of the acetabulum too narrow to be included; however, our data set demonstrates the full range from normal to severely dysplastic hips. The

clinical diagnoses and treatment decisions were made by one of five orthopedic surgeons on the basis of their own experience and preferences, and they were not validated by any external reference standard. However, this situation represents typical clinical practice, and our conclusions regarding the variability of US measurements do not depend on the specific diagnosis in each patient. The follow-up interval was relatively short at 6 months, but this is the most clinically important period because the decision to begin treating DDH is generally within weeks of initial assessment. The probe movements we assessed in 3D data were limited to pure rotation in one of two planes, and we did not consider combinations of two rotations or rotation and translation. These assessments would have been challenging to perform mathematically, and although they might have revealed further increased variation in alpha angles, they would not alter our conclusion that this variation is an important hazard of 2D US assessment of DDH.

We found that 3D US of the infant hip is feasible and images can be analyzed quantitatively. After an initial flurry of studies in the 1990s [13-16], 3D US was not pursued in hip dysplasia assessment. Those early reports describe slow manual sweeps to produce scans and cumbersome post-processing. The probe we used can help obtain a 3D scan in 3.2 seconds with less difficulty in position than in 2D US. In our experience, almost all infants are able to stay still for this brief time, especially if distracted by toys, sucrose, a warmed blanket, and warmed US gel. Computer analysis time per hip is less than 10 minutes. As technology advances, 3D scanning and analysis times probably will continue to shorten, potentially further improving the practical application of this technology in the future.

## 2.5 CONCLUSION

By using 3D ultrasound data sets, the 2D acetabular alpha angle was shown to vary by an average of  $5.5^{\circ} \pm 6.9^{\circ}$  (mean  $\pm$  SD; range  $0-24^{\circ}$ ; upper limit of 95% CI,  $19^{\circ}$ ) simply by changing the apparent transducer probe orientation. This variation caused the diagnostic category to change in more than one-half of the imaged hips overall and in nearly three quarters

of hips imaged in patients younger than 1 month. Some hips which were classified as normal were able to appear dysplastic, and hips classified as dysplastic were able to appear normal. These results draw attention to the inadequacies of using current 2D ultrasound methods and indices (specifically relying upon the alpha angle for diagnosis), and it is evident that new methods and techniques are needed to resolve current shortfalls.

## 2.6 REFERENCES

- [1] R. Graf. Fundamentals of sonographic diagnosis of infant hip dysplasia. *J. Pediatr. Orthop.* 4(6), pp. 735-740. 1984.
- [2] E. A. Simon, F. Saur, M. Buerge, R. Glaab, M. Roos and G. Kohler. Inter-observer agreement of ultrasonographic measurement of alpha and beta angles and the final type classification based on the graf method. *Swiss Med. Wkly.* 134(45-46), pp. 671-677. 2004.
- [3] E. A. Roovers, M. M. Boere-Boonekamp, T. S. Geertsma, G. A. Zielhuis and A. H. Kerkhoff. Ultrasonographic screening for developmental dysplasia of the hip in infants. reproducibility of assessments made by radiographers. *J. Bone Joint Surg. Br.* 85(5), pp. 726-730. 2003.
- [4] M. Zieger. Ultrasound of the infant hip. part 2. validity of the method. *Pediatr. Radiol.* 16(6), pp. 488-492. 1986.
- [5] D. P. Gwynne Jones, A. G. Vane, G. Coulter, P. Herbison and J. D. Dunbar. Ultrasound measurements in the management of unstable hips treated with the pavlik harness: Reliability and correlation with outcome. *J. Pediatr. Orthop.* 26(6), pp. 818-822. 2006.
- [6] C. Morin, S. Zouaoui, A. Delvalle-Fayada, P. M. Delforge and H. Lecllet. Ultrasound assessment of the acetabulum in the infant hip. *Acta Orthop. Belg.* 65(3), pp. 261-265. 1999.
- [7] J. P. Cashman, J. Round, G. Taylor and N. M. Clarke. The natural history of developmental dysplasia of the hip after early supervised treatment in the pavlik harness. A prospective, longitudinal follow-up. *J. Bone Joint Surg. Br.* 84(3), pp. 418-425. 2002.
- [8] S. K. Storer and D. L. Skaggs. Developmental dysplasia of the hip. *Am. Fam. Physician* 74(8), pp. 1310-1316. 2006.
- [9] C. Dezateux and K. Rosendahl. Developmental dysplasia of the hip. *Lancet* 369(9572), pp. 1541-1552. 2007. . DOI: 10.1016/S0140-6736(07)60710-7.
- [10] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476), pp. 307-310. 1986.
- [11] R. Graf, M. Mohajer and F. Plattner. Hip sonography update. quality-management, catastrophes - tips and tricks. *Med. Ultrason.* 15(4), pp. 299-303. 2013.

[12] A. Falliner, H. J. Hahne, J. Hedderich, J. Brossmann and J. Hassenpflug. Comparable ultrasound measurements of ten anatomical specimens of infant hip joints by the methods of graf and terjesen. *Acta Radiol. (2)*, pp. 227. 2004.

[13] E. O. Gerscovich, A. Greenspan, M. S. Cronan, L. A. Karol and J. P. McGahan. Three-dimensional sonographic evaluation of developmental dysplasia of the hip: Preliminary findings. *Radiology 190(2)*, pp. 407-410. 1994.

[14] R. Graf and K. Lercher. [Experiences with a 3-D ultrasound system in infant hip joints]. *Ultraschall Med. 17(5)*, pp. 218-224. 1996.

[15] C. Sohn, G. P. Lenz and M. Thies. 3-dimensional ultrasound image of the infant hip. *Ultraschall Med. 11(6)*, pp. 302-305. 1990.

[16] U. von Jan, H. M. Overhoff and D. Lazovic. 3-D visualization of the newborn's hip joint using ultrasound and automatic image segmentation. *Stud. Health Technol. Inform. 77pp. 1170-1174*. 2000.

# Chapter 3

## Chapter 3 ..

### ULTRASOUND QUANTIFICATION OF ACETABULAR ROUNDING IN HIP DYSPLASIA: RELIABILITY AND CORRELATION TO TREATMENT DECISIONS<sup>2</sup>

#### 3.1 INTRODUCTION

The problems of current ultrasonographic diagnosis methods in infant hip dysplasia, such as the high inter observer and inter scan variation [1] of the alpha angle [2] or the high rate of “abnormalities” which spontaneously resolve upon follow up imaging [3, 4, 5], are convincing evidence that improvements would be beneficial. There would be direct improvement by including more useful secondary ultrasound indices which can entail information on the acetabulum, especially given ultrasounds ability to fully visualize the hip [6]. One potentially helpful step would be to quantify acetabular rounding. Rounding is a visually obvious aspect of acetabular deformity that is currently only informally assessed by radiologists and pediatric orthopedic surgeons, with no numeric index of rounding available. Therefore, we developed semi-quantitative and quantitative indices of acetabular rounding and assessed their reliability. Since rounding is visually obvious and already subjectively assessed in treatment decisions, it is not possible to blind observers to its presence and difficult to isolate its clinical impact. We

---

<sup>2</sup> A version of this chapter has been accepted for publication *Ultrasound in Biology and Medicine* – Cheng E, Mabee M, Swami VG, et al. Ultrasound Quantification of Acetabular Rounding in Hip Dysplasia: Reliability and Correlation to Treatment Decisions in a Retrospective Study. *Ultrasound in Biology and Medicine* 2014 (in press).

focused on creating novel indices to measure rounding and noted whether they were correlated to the decision to treat for hip dysplasia.

## 3.2 MATERIALS AND METHODS

### 3.2.1 PATIENTS

This retrospective image and chart review was approved by our institutional ethics committee, with requirement for consent waived. Consecutive infants referred to a tertiary pediatric orthopedic clinic by primary care physicians or neonatologists for possible hip dysplasia with available ultrasound imaging and clinical follow-up averaging 32.2 months (range: 11.0 – 86.6 months) were included. We used the initial ultrasound of the hip of clinical concern in each patient (selecting one side randomly if concern was bilateral), saving the image most closely matching the Graf standard plane and removing patient identifying information. Age at ultrasound averaged 6.8 weeks (range: 2 days – 8.1 months). Hips were categorized from clinical records as: (0) normal and patient discharged (n=22); (1) initially indeterminate but hip normalized at follow-up without treatment (n=5); (2) dysplastic hip treated with Pavlik harness or similar external brace (n=30); (3) dysplastic hip treated by surgical reduction and spica casting and/or osteotomy (n=33).

### 3.2.2 INDEX MEASUREMENT

We performed several reading exercises. Although acetabular morphology was unavoidably visually obvious in each image, for each reading exercise the observers were blinded to each other's results, the clinical treatment group, and calculated index values.

A semi-quantitative “rounding grade” in which the rounding of the acetabular margin was characterized as “minimal” (0), “mild” (1), “moderate” (2) and “severe” (3) was recorded for each hip by three observers: a board-certified pediatric radiologist with pediatric and musculoskeletal imaging fellowships (JJ), an engineering graduate student (MM – author of this thesis), and a medical student with six months dedicated hip ultrasound imaging experience

(EC). Each observer repeated the exercise after a week's time to assess intra-observer variability.

Quantitative indices were also measured. Using custom software (MATLAB, 2012; The MathWorks, Natick USA) on each image, users drew standard lines along the iliac wing (baseline), acetabular roof, femoral head diameter and 5 points tracing the acetabular curve. The acetabular curve was defined as the distance between the point where the bony margin began to curve away from iliac wing baseline (A in Fig. 3.1) and where its slope became equal to the acetabular roof line (B in Fig. 3.1). After  $\beta$ -spline smoothing, the arc length between A-B was calculated. Lines perpendicular to endpoints A and B and their intersection point (O) were generated. Acetabular radius of curvature (AROC) was the average distance between endpoints and intersection point (lines OA and OB, Fig. 3.1). Traditional coverage and alpha angle indices were also measured. Acetabular coverage was defined as the percentage of femoral head diameter located medial to the inferior projection of the baseline ( $c1/c2$ , Fig. 3.1) and alpha angle the angle between the iliac wing and acetabular roof ( $\alpha$ , Fig. 3.1). Two observers, JJ and MM, each made all measurements on all 90 hips in random order. Each reader also repeated measurements in a random subset of 20 hips after a week's time.

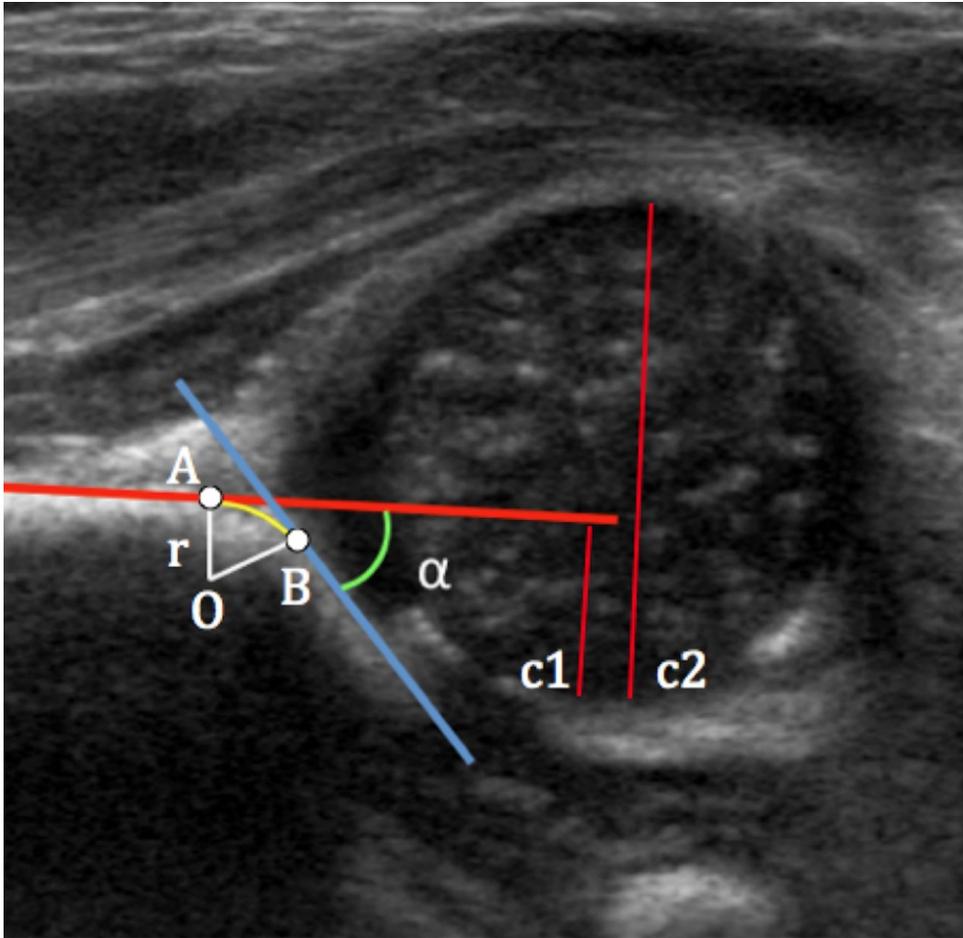


Figure 3.1 Ultrasound tracings conducted by observers on custom MATLAB software. The red line represents the iliac wing (baseline) and blue line the acetabular roof. Acetabular arc length was defined as the distance from A-B and acetabular radius of curvature (AROC) the average distance between O-B and O-B. The ratio of the lengths  $c1/c2$  was used to calculate acetabular coverage and the alpha angle is noted by  $\alpha$ .

### 3.2.3 STATISTICS

Each reading generated five indices per hip: rounding grade, alpha angle, coverage, acetabular radius of curvature (AROC) and arc length. For rounding grade, a categorical index, the intra-observer and inter-observer Cohen's kappa ( $\kappa$ ) and percent agreement were computed for each reader pair. For the four continuous indices, inter-observer coefficients of variation (CoV) and intraclass correlation coefficients (ICC) were calculated using  $n=90$ , and intra-observer CoV calculated using  $n = (20 \times 2)$  repeat readings. Bivariate Pearson correlation

coefficients were computed between indices. One-way analysis of variance (ANOVA) was performed to determine whether index means differed significantly between clinical groups, using student's t-test or if Levene's test of homogeneity was not satisfied, the Mann-Whitney U test. Receiver operating characteristic curves and the area under these were calculated for certain variables to investigate their predictive value. All statistics were calculated using SPSS (IBM Corp., v. 21, 2013).

### 3.3 RESULTS

Inter- and intra-observer reliability statistics for traditional (alpha angle, coverage) and novel (arc length, AROC) acetabular rounding indices are summarized in Table 3.1. Alpha angle was most consistently measured with a CoV ~ 8%. CoV was substantially higher for AROC and arc length (15-39%), although ICC values were comparable for AROC (0.835) and alpha angle (0.895). There was difficulty in reliably drawing the acetabular curve whenever the acetabular line was ill-defined, as seen in Fig. 3.2.

**Table 3.1 Inter and intra-observer intra-class correlation coefficient (ICC) and coefficient of variance (CoV) for alpha angle, coverage, acetabular radius of curvature (AROC), arc length**

	Inter-observer (Observer 1 vs. 2)		Intra-observer (Observer 1)	Intra-observer (Observer 2)
	ICC	CoV (%)	CoV (%)	CoV (%)
<b>Alpha Angle</b>	0.895	8.20	7.93	9.10
<b>Coverage</b>	0.929	21.1	10.8	15.1
<b>AROC</b>	0.835	33.8	29.4	32.9
<b>Arc Length</b>	0.743	28.7	15.1	39.0

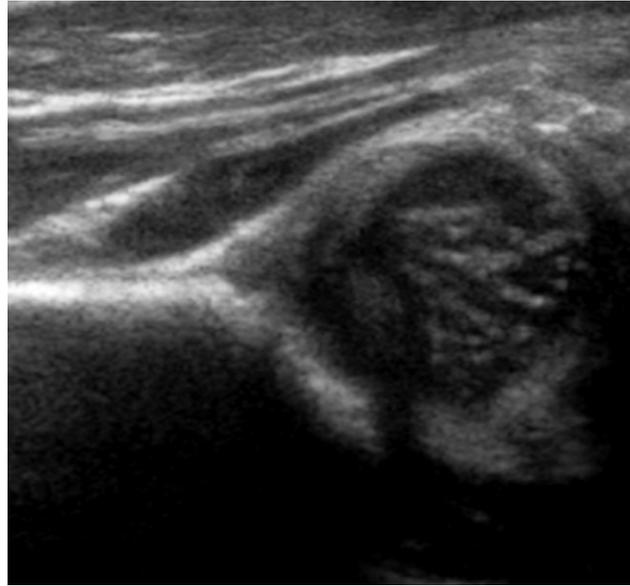


Figure 3.2 Example of a 2D hip ultrasound with indistinct acetabular margin that was difficult to trace.

There was moderate to substantial intra-observer agreement on rounding grade, with kappa 0.54-0.63 (table 3.2). However, inter-observer variability of the rounding grade was poor (kappa 0.30-0.37, agreement 48-52%). Rounding grade did correlate closely with AROC: the area under a receiver operating characteristic curve prepared using AROC values to predict the rounding grade was 0.91.

**Table 3.2 Inter- and intra-observer variability of subjective assessment of acetabular rounding (acetabular rounding grade) through calculation of Kappa ( $\kappa$ ) and % agreement (% A). A, B, C denotes the three readers who assigned semi-quantitative rounding grades to all n=90 hips twice with a one week delay between reads. Rounding grades were as follows: 0=minimal, 1=mild, 2=moderate and 3=severe.**

Comparison	Intra-observer						Inter-observer					
	A-A		B-B		C-C		A-B		A-C		B-C	
	$\kappa$	% A	$\kappa$	% A	$\kappa$	% A	$\kappa$	% A	$\kappa$	% A	$\kappa$	% A
Grades 0, 1, 2, 3	0.54	65.6	0.63	72.2	0.61	71.1	0.30	47.8	0.32	47.8	0.37	52.2
Grade 0-1 vs. 2-3							0.681	84.4	0.61	80.0	0.65	82.0

Arc length and AROC, both indices of acetabular rounding, were highly correlated ( $r=0.94$ ). Correlation between these and traditional DDH indices (alpha angle and coverage) was negative (i.e., dysplastic hips had low alpha angles and large radii of curvature) and poor to moderate (Table 3.3). AROC correlated nonlinearly to alpha angle: hips with high alpha angles (i.e., normal) had sharp acetabular edges with small AROC, while hips with low alpha angles (dysplastic) could have small or large AROC (Fig. 3.3; only the lower left half of the chart contains data points).

**Table 3.3 Pearson correlation coefficients between four ultrasound indices (alpha angle, coverage, acetabular radius of curvature [AROC] and arc length) of hip dysplasia**

	Coverage	Arc Length	AROC
Alpha Angle	0.556**	-0.400**	-0.588**
Arc Length	-0.134	1	0.935**
AROC	-0.203	0.935**	1

\*\*Correlation is significant at the 0.01 level (two tailed)

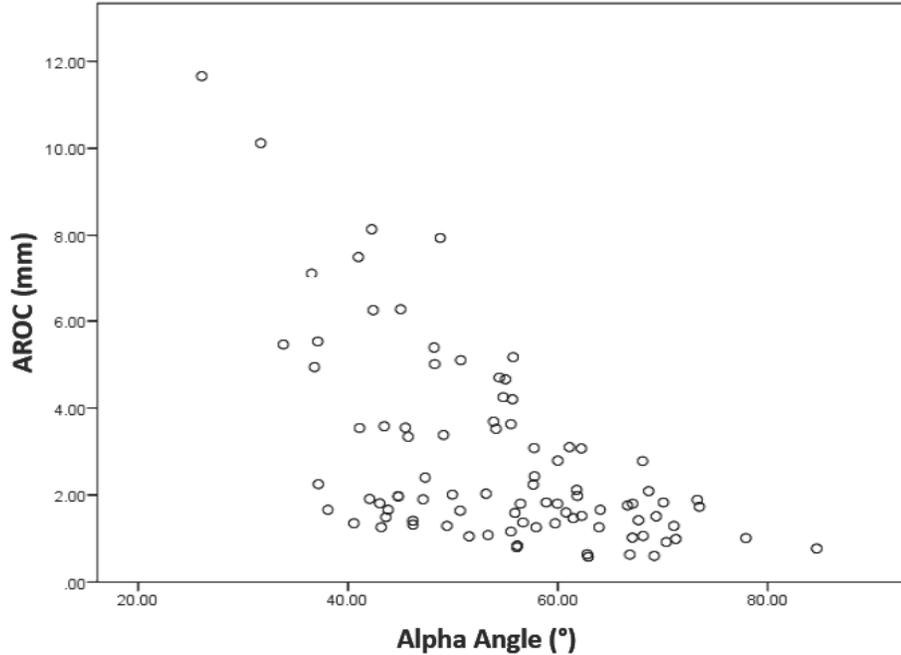


Figure 3.3 Acetabular radius of curvature (AROC) vs. alpha angle for one set of reads conducted by an observer.

Most hips with moderate to severe rounding ultimately received treatment, but many of the treated hips did not have rounded acetabulae. For example, rounding grade  $> 1$  predicted eventual treatment with specificity 0.93, but sensitivity of just 0.58; at 47% prevalence of treatment this gave a positive predictive value 0.95, negative predictive value 0.49. There was substantial agreement between the quantitative AROC and the semi-quantitative rounding grade on whether a hip was at least moderately dysplastic: at  $\text{AROC} > 3.0$  mm, agreement with rounding grade of 2 or 3 had  $\text{kappa} = 0.70$ , 80% concordance.

At ANOVA, all four ultrasound indices clearly distinguished between hips that did not (groups 0-1) or did (groups 2-3) eventually undergo treatment for DDH (Fig. 3.4, Fig. 3.5). Mean alpha angle declined linearly from groups 0-3, while mean coverage only declined significantly in group 3. Mean AROC and arc length both rose significantly between groups 0/1

(no treatment; mean AROC 1.6 mm and 1.6 mm) and groups 2/3 (treated; mean AROC 3.1 mm and 3.4 mm;  $p=0.007$ ). Receiver operating characteristic curves were prepared for use of each index as a predictor of the clinical decision to treat (Fig. 3.6). The areas under the receiver operating characteristic curves were 0.84 and 0.80 for traditional indices of alpha angle and coverage, and 0.76 and 0.69 for the acetabular radius of curvature and arc length, respectively.

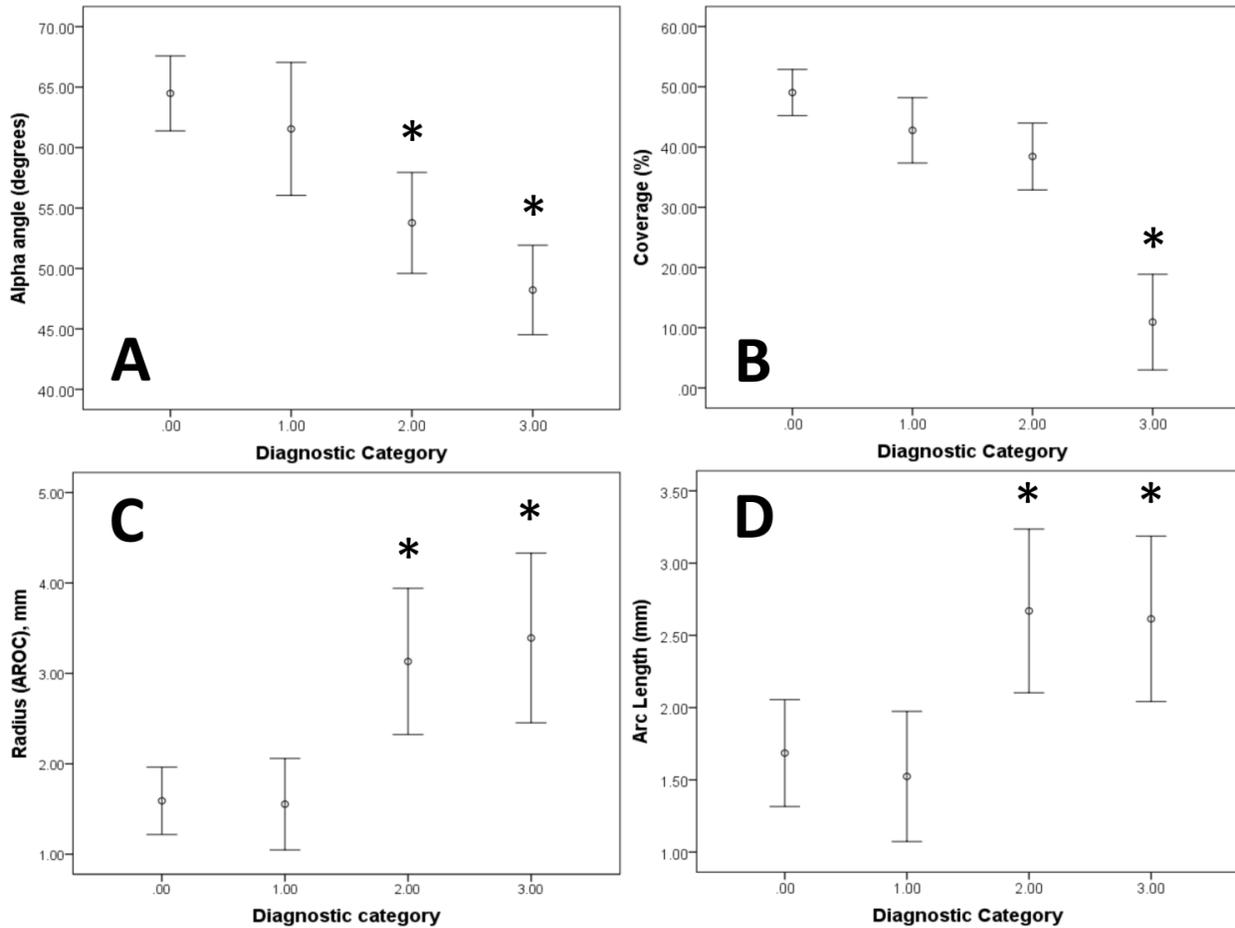


Figure 3.4 Mean values of alpha angle (A), coverage (B), acetabular radius of curvature (AROC; C) and arc length (D) for each clinical diagnostic group: 0=normal; 1=initially indeterminate, resolved at follow-up without treatment; 2=dysplastic, treated with Pavlik harness or similar; 3=dysplastic, treated surgically. \*= statistically significant difference in means from group 0 (normal). Error bars represent 95% confidence interval of each mean. Note that groups 0 and 1 required no treatment while groups 2 and 3 were treated.

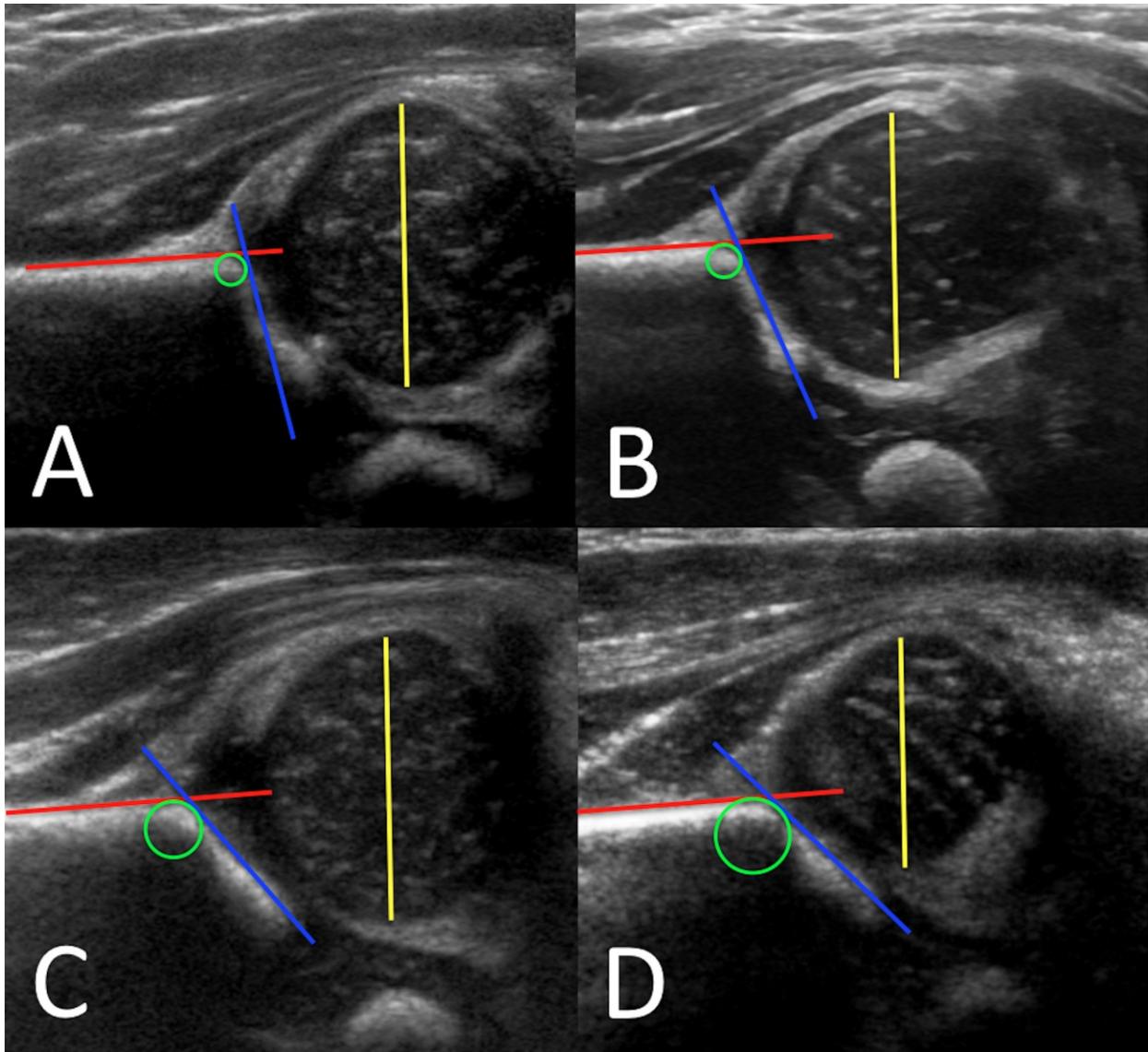


Figure 3.5 Typical ultrasound appearance in each clinical group. (A) group 0 (normal, no treatment): mean alpha angle  $65^{\circ}$ , small acetabular curve with mean radius of curvature 1.6 mm; (B) group 1 (indeterminate at initial ultrasound, resolved later without treatment): mean alpha angle  $61^{\circ}$ , mean radius 1.6 mm; (C) group 2 (Pavlik harness treatment): mean alpha angle  $54^{\circ}$ , mean radius 3.1 mm; (D) group 3 (surgical treatment): mean alpha angle  $48^{\circ}$ , mean radius 3.4 mm. The circles drawn on each image show the increasing size of the calculated acetabular radius of curvature from category 0-3.

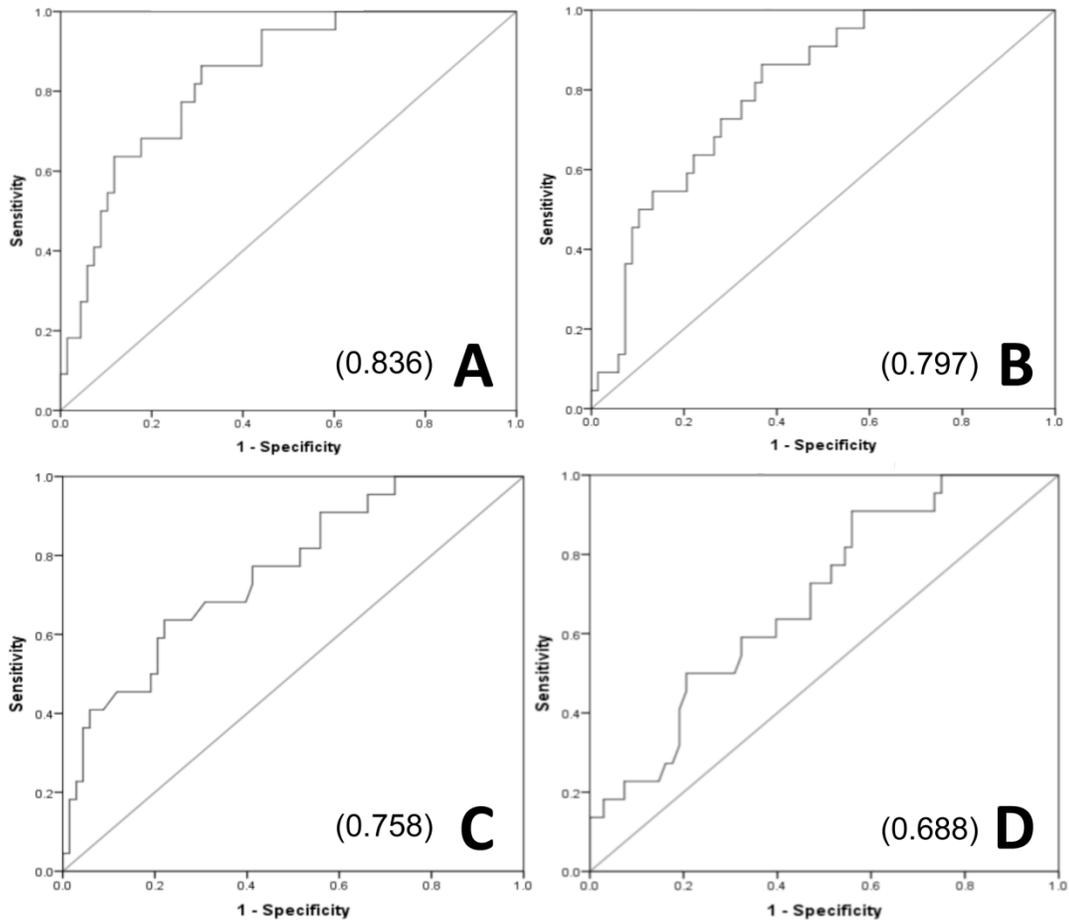


Figure 3.6 Receiver operating characteristic curves for alpha angle (A), coverage (B), acetabular radius of curvature ‘AROC’ (C) and arc length (D). The areas under the receiver operating characteristic curves are listed in brackets beside each respective letter. Reference lines are also shown in each panel.

### 3.4 DISCUSSION

There have been recent advances in the literature pertaining to measuring different components of acetabular morphology in pediatric hips [7]. However, this is the first study we know of to quantify acetabular rounding on hip ultrasound. Quantification is needed as subjective assessment of rounding is highly variable; even assigning a rounding grade semi-quantitatively in this study led to unacceptably high inter-observer variability. Given the tiny size of the structure being measured, we had low expectations. We found that acetabular

rounding was measurable with moderate reliability, although coefficients of variation were relatively high, CoV~30%. These CoV are larger than for alpha angle and coverage, undoubtedly because the acetabular curve is very small, radius averaging 1-3 mm, and difficult to trace with sub-millimeter precision. Tracing the acetabular curve also used 5 plotted points, compared to 2 points for each line from which alpha angle and coverage were calculated. Having more user input inevitably increased variability. Variability was greatest when the acetabular curve was indistinct and therefore difficult to accurately trace (Fig. 3.2). Sonographers are trained to prioritize saving images with crisp and straight iliac wing and acetabular roof lines to best measure the alpha angle and the planes in which these images are captured may not be optimal to capture a well defined acetabular curve outline. This could be avoided prospectively by instructing sonographers to also ensure a sharply defined acetabular curve is well seen, which may at times require saving an extra image in a slightly different plane.

The poor-to-moderate correlations between traditional ultrasound indices of DDH (alpha angle, coverage) and our novel indices of rounding (AROC, arc length) imply that the latter measure a related but different component of hip morphology. The high correlation between AROC and arc length ( $r=0.94$ ) suggests it is only necessary to measure one of these indices. AROC is preferred due to its higher ICC comparable to arclength (AROC 0.835, arclength 0.743), and higher area under the receiver operating characteristic curve predicting treatment (nearly as high as for the alpha angle 0.84 vs 0.76 AROC).

Differences in acetabular rounding between normal and dysplastic hips could be usefully quantified despite technical challenges in measurement. Specifically, AROC clearly distinguished between hips that ultimately went on to treatment for DDH and those that did not, regardless of the alpha angle. This was possible since the difference in AROC between normal and dysplastic hips (mean 1.6 vs. 3.3 mm) was large enough to overcome variability in AROC measurement, which was proportionately substantial (CoV~30%) but represented only a

tiny absolute distance (~0.5 mm). Based on our results, a threshold of AROC > 2.5-3 mm may be appropriate, but validating an exact threshold would require more extensive study.

Rounding and alpha angle had a nonlinear correlation, in that acetabular edges were sharp in all hips with normal alpha angles, but could be sharp (small AROC) or rounded (large AROC) in dysplastic hips. This suggests the depth of the acetabular socket and the roundness of its edge represent different aspects of dysplasia, only partly dependent on each other. Quantification of both features on ultrasound may ultimately improve prognostic accuracy. It is possible, for example, that for a given alpha angle, greater acetabular rounding (higher AROC) might be associated with greater likelihood of failure of conservative treatment. Now that we have established a method of measuring acetabular rounding that is easily performed and sufficiently reliable to make clinically useful distinctions, future prospective studies could apply this technique to further investigate diagnostic and prognostic significance of acetabular rounding in DDH.

It is important to note that AROC is best considered supplementary to existing ultrasound indices of DDH. Ideally, a clinician would combine all available information from ultrasound (indices including alpha angle, coverage and AROC, and dynamic testing) with clinical features (e.g., age, risk factors, stability on examination) to best predict patient outcomes, plan and monitor therapy.

#### 3.4.1 LIMITATIONS

This study had limitations, chiefly related to its retrospective nature. First, we were limited to use of the saved images from each ultrasound study and could not control image quality directly. Sonographers were focused on obtaining a Graf plane image rather than one with a well-defined acetabular curve, which limited reliability of retrospective tracing. Second, we limited our outcome measure to the decision to treat DDH within our follow-up period (average 32 months) rather than assessing radiographic outcomes. This is because in our retrospective study, radiographs were not consistently available for all patients, and the

meaning of correlation between initial rounding and ultimate radiographic outcome would have been unclear due to the various intervening treatments. Note that most decisions regarding treatment occur promptly within weeks to months of DDH diagnosis [4] . This study was not designed to evaluate very long term clinical outcomes; a task we believe will be better performed in future prospective study.

Third and most importantly, treatment decisions made by the attending pediatric orthopedic surgeon already incorporated visual assessment of acetabular rounding (in addition to other features including alpha angle, acetabular coverage, stability on dynamic ultrasound and clinical examination, clinical risk factors and patient age). Thus, the clinical outcome was not an independent gold standard but was partly based on some of our input indices. In this setting, it would be circular reasoning to conclude that acetabular rounding, as an imaging feature, correlates to treatment: “We think rounding requires treatment. Those with rounding were treated. Therefore measurement of rounding correlates with treatment.” However, this was not our hypothesis. We know that our surgeons treat severe rounding. Therefore, when we tested whether our specific technique for measuring rounding correlated closely to treatment, it was as an indirect way of determining whether the measurement accurately quantifies the clinically important essence of the feature that is already visually obvious to the surgeon. A low correlation to treatment decisions would not have meant that rounding is not important, but that our method of measuring the rounding is likely flawed. Acetabular rounding on ultrasound anecdotally forms part of surgical decision-making, but to an unknown extent since this visual finding has not been quantified until now. Our study results should be interpreted as describing current clinical practice as to who is selected for treatment rather than evaluating whether this is the most appropriate group of patients to be treated. For example, it remains unclear whether an abnormal AROC alone should be treated; this needs to be studied prospectively. The natural history of AROC in untreated hips is also not clear and could be studied in an appropriate prospective cohort.

Finally, this type of measurement lends itself to ultrasound measurements in three dimensions. By measuring the acetabulum curvature in two dimensions you can begin to understand more fully how stable the femoral head is in the ball and socket joint; once additional slices have been added (above and below) to the original two dimensional view, you can perform the same (or similar) measurement upon the entire acetabular curvature and end up with a complete picture of rounding which should be more precise than any of the current 2D methods. This type of measurement is discussed in chapter 5 of this thesis.

### 3.5 CONCLUSIONS

Subjective or semi-quantitative assessment of acetabular rounding has inter-observer variability that is unacceptably high. Acetabular rounding could be more reliably quantified on hip ultrasound by measuring the acetabular radius of curvature (AROC). Although this index had relatively high inter-observer variability compared to traditional indices (alpha angle and coverage), it was still sufficiently reliably measured to detect significant differences in acetabular rounding between infant hips that did or did not go on to receive treatment for hip dysplasia. Acetabular rounding was only moderately correlated to the alpha angle, and greater rounding was strongly associated with a decision to treat hip dysplasia. In light of this, measuring the radius of acetabular curvature may prove useful clinically, or in prospective studies, especially where rounding is the most significant feature of hip dysplasia, or additionally in aiding in creation of novel 3D indices.

### 3.6 REFERENCES

- [1] E. Bar-On, S. Meyer, G. Harari and S. Porat. Ultrasonography of the hip in developmental hip dysplasia. *J. Bone Joint Surg. Br.* 80(2), pp. 321-324. 1998.
- [2] R. Graf. Fundamentals of sonographic diagnosis of infant hip dysplasia. *J. Pediatr. Orthop.* 4(6), pp. 735-740. 1984.

- [3] A. Roposch, L. Q. Liu, F. Hefti, N. M. P. Clarke and J. H. Wedge. Standardized diagnostic criteria for developmental dysplasia of the hip in early infancy. *Clinical Orthopaedics and Related Research* (12), pp. 3451. 2011.
- [4] S. K. Storer and D. L. Skaggs. Developmental dysplasia of the hip. *Am. Fam. Physician* 74(8), pp. 1310-1316. 2006.
- [5] V. Bialik, G. M. Bialik, S. Blazer, P. Sujov, F. Wiener and M. Berant. Developmental dysplasia of the hip: A new approach to incidence. *Pediatrics* 103(1), pp. 93-99. 1999.
- [6] American Institute of Ultrasound in Medicine. AIUM practice guideline for the performance of an ultrasound examination for detection and assessment of developmental dysplasia of the hip. *J. Ultrasound Med.* 32(7), pp. 1307-1317. 2013.
- [7] O. Y. Yavuz, I. Uras, B. A. Tasbas, M. H. Ozdemir, M. Kaya and M. Komurcu. A new measurement method in graf technique: Prediction of future acetabular development is possible in physiologically immature hips. *J. Pediatr. Orthop.* 2014.

# CHAPTER 4

## Chapter 4 ..

### REPRODUCIBILITY OF ACETABULAR LANDMARKS AND A STANDARDIZED COORDINATE SYSTEM OBTAINED FROM 3D HIP ULTRASOUND<sup>3</sup>

#### 4.1 INTRODUCTION

Recalling the problems illustrated in chapter 2 of traditional 2D ultrasound images, I have shown that they offer only a partial view of the complex 3D shape of the acetabulum, and that the interscan variability in the 2D image orientations and the resulting calculated alpha angles exist because ultrasound probes are hand-held. No two images of the acetabulum will be exactly alike, because of the different probe orientations they were acquired at. Also shown in chapter 2, was that this variability can be substantial, with potential for misclassification of DDH in half of infants and up to three-quarters of neonates.

Three dimensional (3D) ultrasound probes can now acquire images rapidly enough to image even a non-compliant infant. 3D ultrasound can demonstrate all relevant anatomy (iliac wall, femoral head, and os ischium) for diagnosing DDH, in a more comprehensive way than 2D ultrasound (Figure 4.1).

---

<sup>3</sup> A version of Chapter 4 has been accepted for publication in *Ultrasonic Imaging* – Mabee M., Dulai S., Thompson R.B., Jaremko J.L., Reproducibility of acetabular landmarks and a standardized system obtained from 3D hip ultrasound. *Ultrasonic Imaging* 2014 (in press).

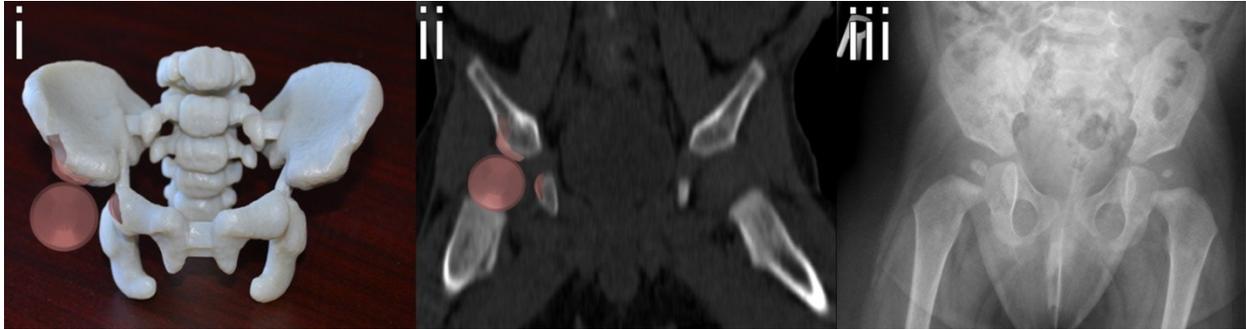


Figure 4.1 Comparison between modalities. A model of a normal infant hip acetabulum and cartilaginous femoral head, generated from 3D ultrasound (red), superimposed on a physical model of an infant pelvis generated via 3D printer from high resolution CT scan data (i), and on a single coronal image from the source CT scan in the same patient (ii). (iii) Comparison with a radiograph in a different infant emphasizes that the cartilaginous femoral head is much larger than the ossified core visible radiographically. Note that the images are in different patients and are only representative for visualization purposes.

Since the entire acetabular shape is acquired in a single 3D data set, differences in probe orientation may be less important than in 2D ultrasound. For reproducible comparisons between scans and modalities, it is essential to consistently establish the location of the acetabulum in the 3D ultrasound data set. We have developed a method to reorient the ultrasound image of the hip along a consistent axis using two acetabular landmarks visible in 3D ultrasound, giving a standardized central plane. The purpose of this chapter is to showcase the reliability of this technique. Our primary outcomes were the inter-observer reliability of landmark selection to generate the central plane, and of the associated variation in orientation of the central plane. The chapter also shows the inter-observer variability of alpha angle measurements on 3D vs on 2D ultrasound.

## 4.2 MATERIALS AND METHODS

### 4.2.1 PATIENTS

This prospective study was approved by our institutional ethics board and performed at a tertiary pediatric hospital from September 2012 to October 2013. With written informed parental consent, 3D ultrasound scanning was added to the first routine clinical 2D hip ultrasound scan of each patient. Scan indications were clinical suspicion of DDH due to hip laxity, asymmetric skin creases, and/or risk factors such as positive family history. We considered only one hip per patient except in cases of bilateral hip dysplasia (to increase our number of total scans of dysplastic hips). A total of 51 hips in 42 patients, 27 (64%) females, were analyzed. Patients were scanned at age 4-111 days, average 44.2 days. They received routine care at a pediatric orthopedic clinic from surgeons blinded to 3D ultrasound images and findings. We observed clinical care for at least 6 months (of the infants age) to classify each imaged hip as: normal at first orthopedic assessment/ultrasound (category 0, n = 20 patients, 20 hips); questionably abnormal initially but resolved spontaneously on follow-up imaging/clinical examination (category 1, n = 10 patients, 10 hips); dysplastic requiring treatment by Pavlik harness and/or surgery (category 2, n =12 patients, 21 hips).

#### 4.2.2 IMAGING

Ultrasound used Philips iU22 (Philips Healthcare, Andover, MA, USA) platforms. Conventional 2D ultrasound of both hips was performed using a 12 MHz linear transducer (12L5, Philips Healthcare, Andover, MA, USA), including static imaging in Graf standard coronal plane and axial plane, by usual clinical protocol. 2D ultrasound was interpreted by a pediatric radiologist, with results and images made available to referring clinicians. In addition, a 13 MHz 3D linear (13VL5, Philips Healthcare, Andover, MA, USA) transducer was used in coronal orientation to obtain a 3D ultrasound data set at each hip. The transducer swept through a range of +/- 15° in 3.2 seconds to generate a 3D data set of 256 ultrasound slices 0.13 mm thick, each containing 411 x 192 pixels measuring 0.11 x 0.20 mm. Each user scanned each hip twice by 2D and twice by 3D ultrasound.

#### 4.2.3 IMAGE PROCESSING

Images were analyzed off-line using custom software (MATLAB R2010a-2012; The MathWorks, Natick, USA). For each hip, two observers performed trials on both 3D and 2D data sets of each hip (MM, author of this Thesis, and JJ, the lead radiologist with fellowship training in pediatric and musculoskeletal radiology and 11 years experience). In each 3D data set each user selected two points based on predefined landmarks which would be used to define a rotation of the data-set into an orientation that contains a standardized 2D central plane. The landmarks were located on opposite ends of the lateral margin of the acetabular ridge, with Point A located at its most posterior limit and Point B at its most anterior limit (Figure 4.2)

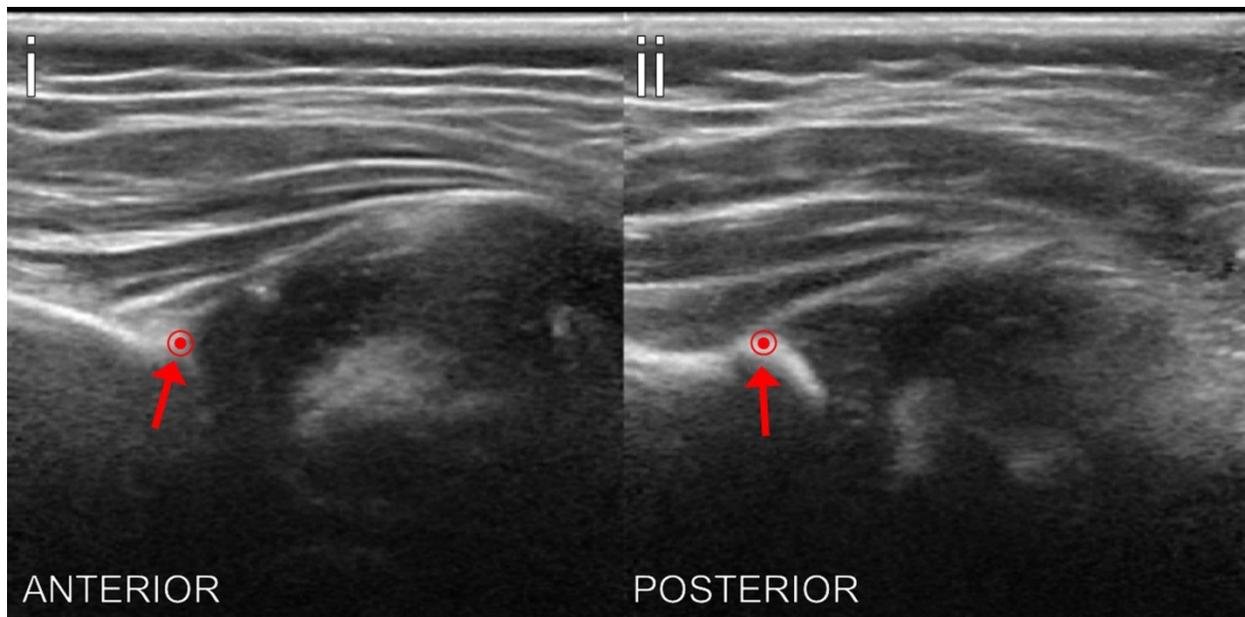


Figure 4.2 Images from 3D ultrasound of the right hip in a 101 day old boy diagnosed clinically with normal hips, showing clearly defined landmark points (i) anterior, (ii) posterior. Slices more anterior to (i) and more posterior to (ii) no longer showed acetabular edges.

The data was rotated based on the angle between the line created from these points and the original orientation of the data as well as the orientation of the Iliac wall, realigning the images into a ‘central plane’ in which points A and B were directly antero-posterior to one another (Figure 4.3).

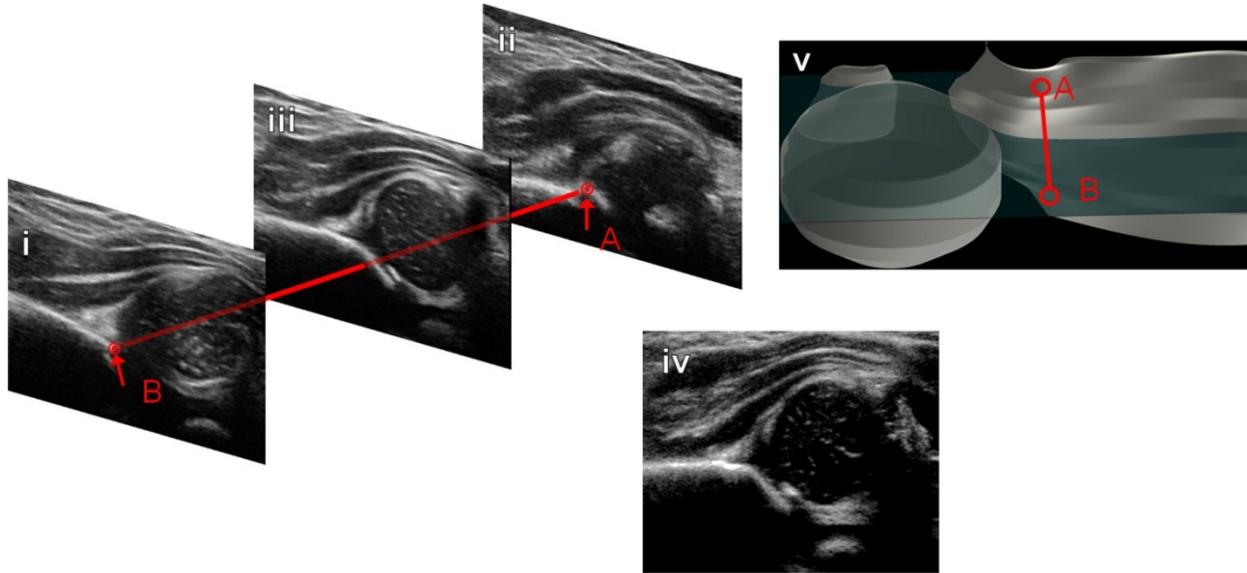


Figure 4.3 Ultrasound images of a girl diagnosed with DDH, aged 6 days at time of scan. (i) anterior landmark selection, (ii) posterior landmark selection, (iii) 3D reconstructed image at central plane, (iv) 2D clinical scan image for comparison with (iii). (v) Location of landmarks A and B on a 3D model of the acetabulum generated from this 3D ultrasound scan. In this scan the central plane was rotated only slightly from the image acquisition plane ( $2^\circ$ ).

Finally, the user would select the coronal slice in the reoriented data (Figure 4.3(iii)) which most accurately resembled a Graf standard plane image; this was considered the “central slice”. The user would then measure the alpha angle in this extracted 2D image with the exact same process used to measure the alpha angle from the conventional 2D ultrasound. For 2D ultrasound, each user measured the alpha angle in all of the clinical Graf standard plane images.

#### 4.2.4 STATISTICAL ANALYSIS

Statistics were calculated on SPSS (Chicago, IL, USA, v. 20). Descriptive statistics were recorded as mean  $\pm$  standard deviation (SD). For repeated measurements we calculated the mean difference in values and to characterize the limits of agreement, we reported the repeatability coefficient (RC), i.e., the range of values within which there is 95% confidence that a second observation will fall from the first [1]. The intraclass correlation coefficients (ICC)

between individual components of landmark locations were also recorded to characterize consistency between observers.

For each scan, inter-observer variability of several parameters was calculated: 3D location of landmarks A and B; 3D angular orientation of the line joining A-B (indicating the orientation of the central plane); the location of the selected central slice along this line, the alpha angle measured on this reoriented slice of 3D ultrasound data; and the 2D ultrasound alpha angle. For the alpha angle we also assessed inter-scan variability at 3D ultrasound and 2D ultrasound, based on a user performing the same analysis on two separate scans of the same hip. We compared the alpha angles measured on the 3D ultrasound central plane images and the routine 2D ultrasound images. We measured intra-observer variability of each parameter on 3D and 2D ultrasound by having each observer perform all studies again one week later.

### 4.3 RESULTS

The results of inter- and intra-scan variability studies are presented for the 3D ultrasound landmarks and central plane in Table 4.1, and for alpha angles in Table 4.2.

**Table 4.1 Reliability of 3D hip ultrasound landmark localization and central plane orientation.**

	Mean	SD*	RC*
<b>Central Plane Landmarks</b>			
<i>Inter Observer (mm)</i>	1.6	1.0	2.0
<i>Intra Observer (mm)</i>	0.9 / 0.9	0.9 / 0.9	1.8 / 1.8
	(User 1 / User2)		
<b>Central Plane Orientation</b>			

<b>Inter Observer (°)</b>	5.4	4.9	9.6
<b>Intra Observer (°)</b>	4.2 / 4.3	3.9 / 4.8	7.6 / 9.4
	(User 1 / User 2)		

\*SD = Standard Deviation, RC = Repeatability Coefficient

**Table 4.2 Reliability of alpha angle -- 2D vs. 3D Ultrasound**

	<b>Our 2D</b>	<b>Our 3D</b>	<b>2D Others</b>
<b>Alpha Angle Inter-Observer</b>			
<b>mean (°)</b>	0.8	2.6	6.5 <sup>(4)</sup>
<b>SD (°)</b>	3.9	5.4	2 <sup>(4)</sup> , 3.5 <sup>(6)</sup>
<b>Repeatability coefficient (°)</b>	7.8	10.8	17 <sup>(5)</sup>
<b>Alpha Angle Intra Observer</b>			
<b>mean (°)</b>	0.2	0.8 (user 1) /0.5 (user 2)	-
<b>SD (°)</b>	3.1	3.1 / 4.7	3 <sup>(2)</sup> , 3.2 <sup>(3)</sup> , 2.8 <sup>(6)</sup>
<b>Repeatability Coefficient (°)</b>	6.2	6.2 / 9.3	
<b>Alpha Angle Inter-Scan</b>			
<b>mean (°)</b>	0.1	0.6	4.0 <sup>(4)</sup>
<b>SD (°)</b>	5.5	4.1	2.0 <sup>(4)</sup>
<b>Repeatability coefficient (°)</b>	10.9	8.2	

#### 4.3.1 VARIABILITY OF LANDMARKS

The intra- and inter- observer variation in the position of the identified landmark points A and B was small (Figure 4.4), averaging 0.9 (MM) and 1.6 mm (MM vs JJ) respectively, with coefficients of repeatability 1.8-2.0 mm (Table 4.1). ICC values were recorded for the comparison of the two observer's (MM and JJ) 3D point locations in X, Y, and Z coordinates with values 0.93, 0.895, and 0.938 respectively. Paired t-tests showed that variation in location of point A was not significantly different than for point B ( $p = 0.36-0.87$ ).

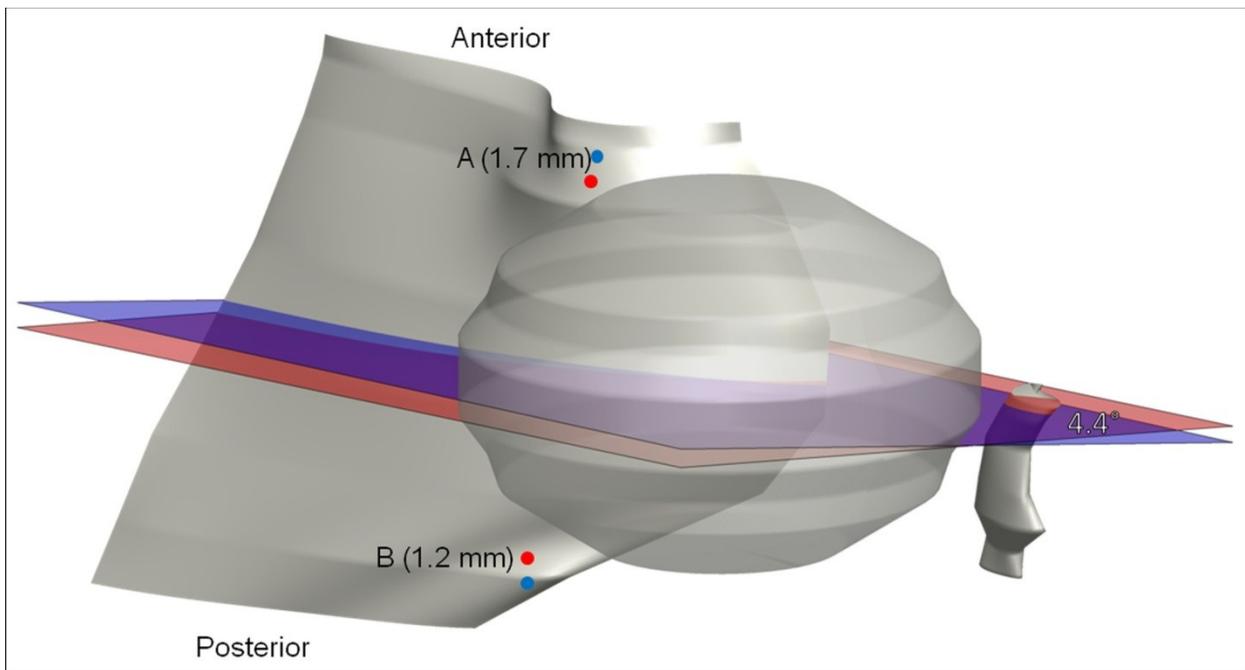


Figure 4.4 Inter-observer variation in selection of landmark locations and central plane orientations. In this scan of the left hip of a 13 day old girl, the two observers selected landmark A (top two dots) 1.7 mm apart, and landmark B 1.2 mm apart. The angle between the two central plane orientations, equal to the angle between lines AB drawn by observer 1 and by observer 2, was 4.4°. These levels of variation were typical in our data set, near the mean values seen in Table 4.1.

#### 4.3.2 VARIABILITY OF CENTRAL PLANE ORIENTATION

The angle between the line A-B drawn by observer 1 (MM) and the same line drawn by observer 2 (JJ) in the same 3D data set represents the difference in 3D orientation of the associated central planes (Figure 4.4). This varied by a mean of  $\sim 5^\circ$  and had a coefficient of

repeatability  $<10^\circ$  between observers. The largest intra-observer variation was  $15.7^\circ$  for observer 1 (MM) and  $13.7^\circ$  for observer 2 (JJ), while the largest inter-observer variation was  $22.8^\circ$ . These extreme values were in markedly dysplastic hips. Central plane orientation had significantly higher variability in hips clinically diagnosed as dysplastic (repeatability coefficients  $8.6, 11.8, 12.0^\circ$  for observer 1, observer 2 and inter-observer data, respectively) than in normal hips ( $6.2, 6.8, \text{ and } 7.6^\circ$ ,  $p < 0.05$ ). ICC values were calculated comparing the two observer's X and Y line components (Z components had no meaningful variation); the ICCs were  $0.867$  and  $0.901$  respectively. In additional analysis, once data was reoriented along the central plane one observer re-identified landmarks in there-sliced data; the amount of further rotation was small ( $1.8 \pm 1.2^\circ$ , coefficient of reliability  $2.4^\circ$ ).

#### 4.3.3 VARIABILITY OF SLICE SELECTION

To determine the variability of the selection of the 'central slice' most representative of the Graf standard plane, after data had been reoriented into the central plane using acetabular landmarks and the 'central slice' had been selected by each user, one user (MM) reselected the central slice on his own trial data and upon the other user's (JJ) data. Inter-observer variability of slice selection was within  $1.3 \pm 6.8$  slices (mean  $\pm$  SD), repeatability coefficient  $13.3$  slices (i.e., 95% of the time, the observers would select the best slice less than 14 slices away from each other, a distance of  $1.7$  mm, or 5% of the width of the 3D ultrasound scan performed). ICC for slice selection was  $0.975$ . The intra-observer variability was lower, at  $0.8 \pm 5.2$  slices, repeatability coefficient  $10$  slices, representing  $1.3$  mm distance or 4% of scan width. The visual impact of this was small (Figure 4.5).

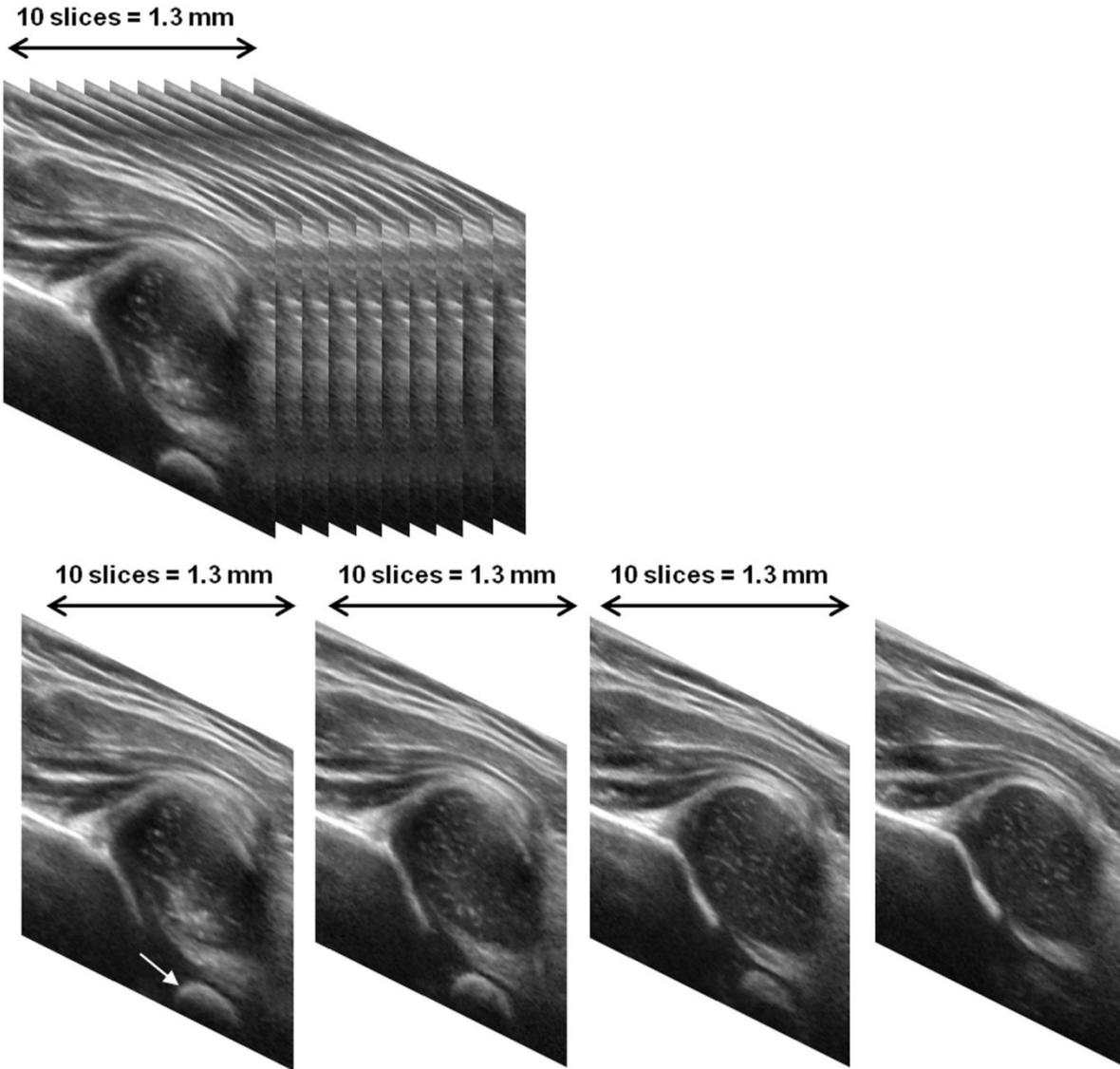


Figure 4.5 Selection of best central slice. 3D ultrasound scans are composed of hundreds of very closely spaced slices (spacing 0.13 mm in our data), with a schematic showing a stack of 10 slices at top. The four images below were taken from 3D ultrasound of a 48 day old girl's left hip, each 10 slices (1.3 mm) away from the next. Across the 30 slice (3.9 mm) distance shown above from left to right, the acetabular shape and femoral head slowly change in shape, and the os ischium (arrow) disappears. In this study, two observers selected the best slice within 14 slices of each other in 95% of cases, a distance of 1.7 mm.

#### 4.3.4 INTER-OBSERVER VARIABILITY OF ALPHA ANGLE

The inter-observer variation of the 3D-ultrasound alpha angle (Table 4.2) is the total variation based on two users completely reprocessing 3D scan data from landmark selection to slice selection and measuring the alpha angle. Inter-observer variation on 2D ultrasound is the variation only of the step involving measurement of the angle on the saved image. The repeatability coefficients indicate that 95% of the alpha angles lay within  $10.8^\circ$  of each other on 3D-ultrasound, vs.  $7.8^\circ$  on 2D-ultrasound ( $p < 0.05$ ). ICC values for the 2D alpha angle were 0.936 and 0.879 for the 3D alpha angle. Intra-observer variability showed a similar pattern with slightly lower values; repeatability coefficients were  $6.2^\circ$  2D vs.  $9.3^\circ$  3D,  $p < 0.05$ . The numeric differences in alpha angle were present despite subjectively similar appearance of the images processed by the two observers in nearly all cases (Figure 4.6).

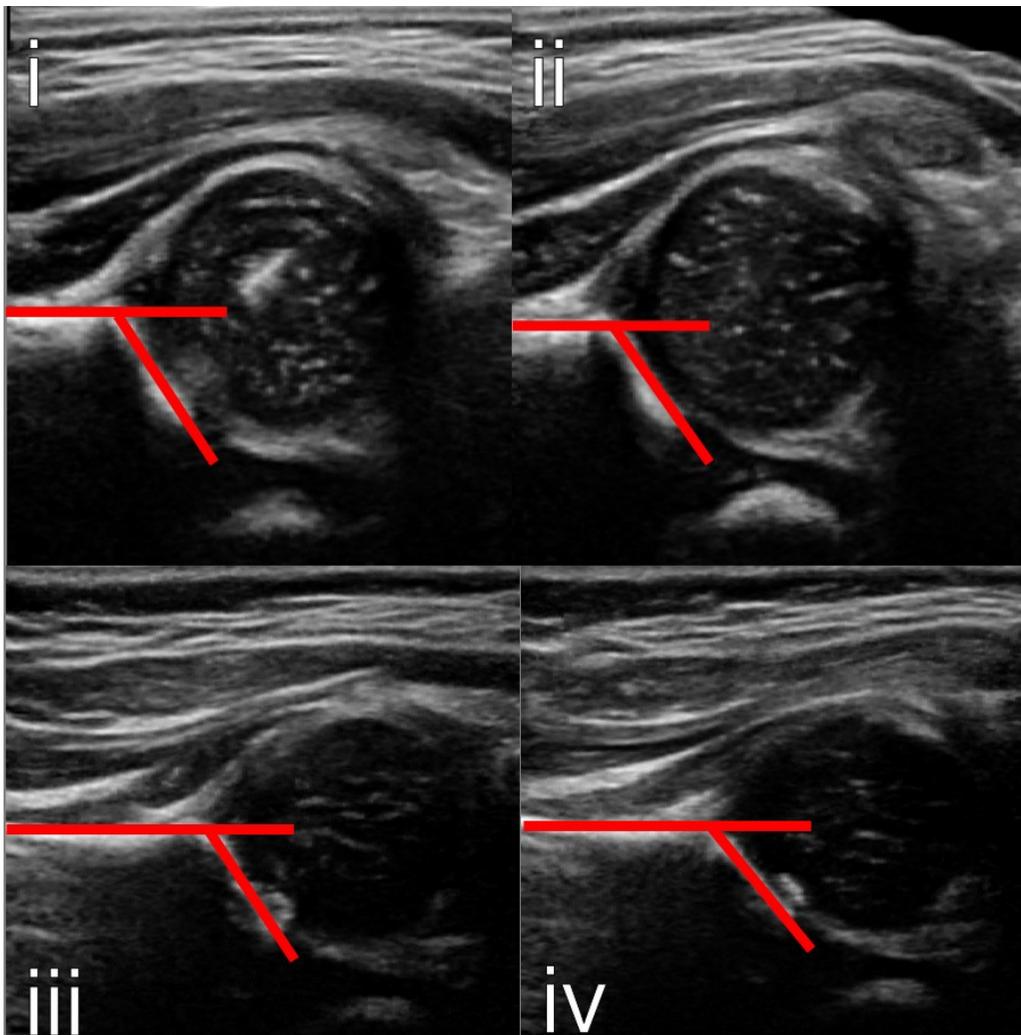


Figure 4.6 Effect of inter-observer variability in processing 3D ultrasound (i,ii) Comparison of 'best slice' reconstructed 3D images processed by two users, with alpha angles  $0.6^\circ$  from each other. Scans were of a 14 day old boy's left hip, which was initially classified as borderline for DDH but resolved at follow-up without intervention. The acetabular margin is nearly identical in the scans, although image (i) captures more of a femoral head ossification centre and less of the os ischium than (ii). (iii,iv) 'Best slice' reconstructed 3D ultrasound images from two users in a 40 day old girl's left hip, which was also initially classified as borderline for DDH but resolved at follow-up without intervention. This patient had one of the largest inter-observer variations in alpha angle we saw, from  $56.6^\circ$  to  $48.8^\circ$ . Note that both patients had slightly curved acetabular rims in all ultrasound images obtained, and the os ischium was only faintly visible for the second patient.

#### 4.3.5 INTER-SCAN VARIABILITY OF ALPHA ANGLE

This variability was computed by having each observer perform the entire process from 2D and 3D data acquisition to alpha angle measurement on separate ultrasound scans performed by different sonographers at the same visit. The inter-observer inter-scan variability was statistically significantly less for 3D ultrasound than 2D ( $8.2^\circ$  vs  $10.9^\circ$ ,  $p < 0.05$ ). Examples of inter-scan variability are given in Figure 4.7. A summary flowchart showing the processing steps and reliability of each step is shown in Figure 4.8.

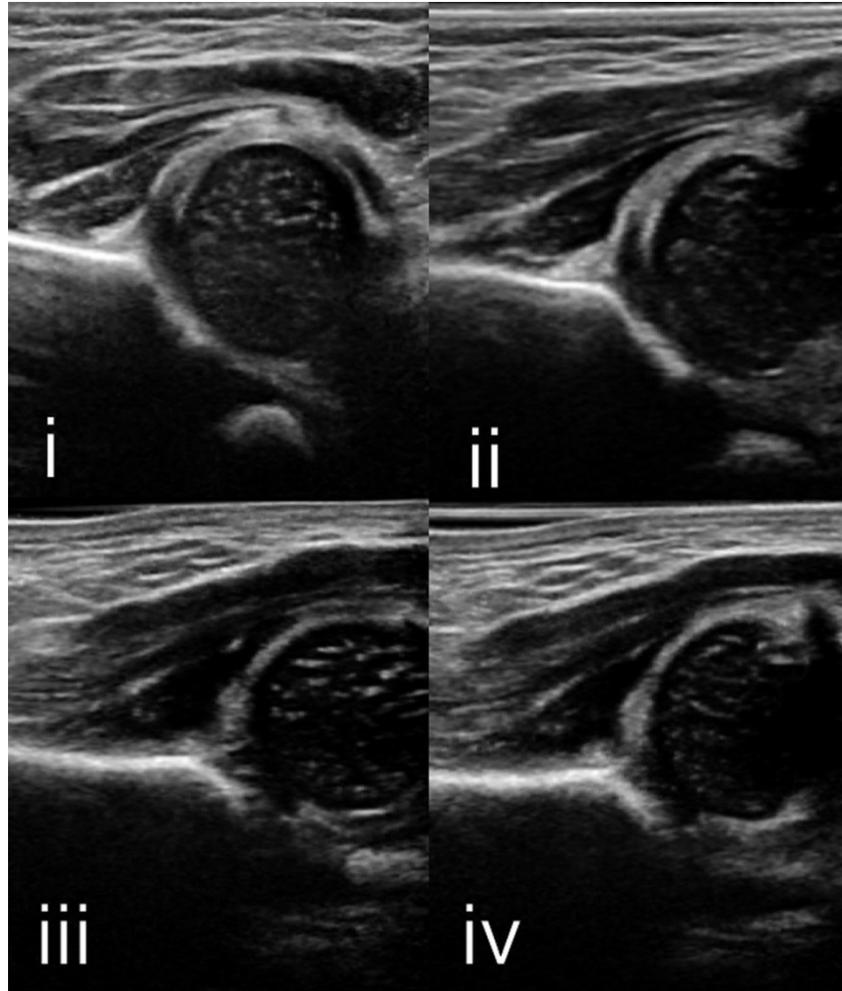


Figure 4.7 Inter-scan variability. Top row: images (i) and (ii) are at the central plane after reconstruction of two different 3D ultrasound scans, in a 6 day old girl with a dysplastic left hip. The scans are very similar visually despite being obtained by two different observers and processed separately. The alpha angles measured on these images varied by  $2^\circ$ .

Bottom row: images (iii) and (iv) were obtained in the same way by processing two different 3D ultrasound scans, in a 4 day old boy with a dysplastic right hip. These two images had the largest difference ( $11^\circ$ ) in measured alpha angles in our study, but are visually more similar than this would suggest. Note the limited visualization of the os ischium, which occurred on both 2D and 3D ultrasound of this patient.

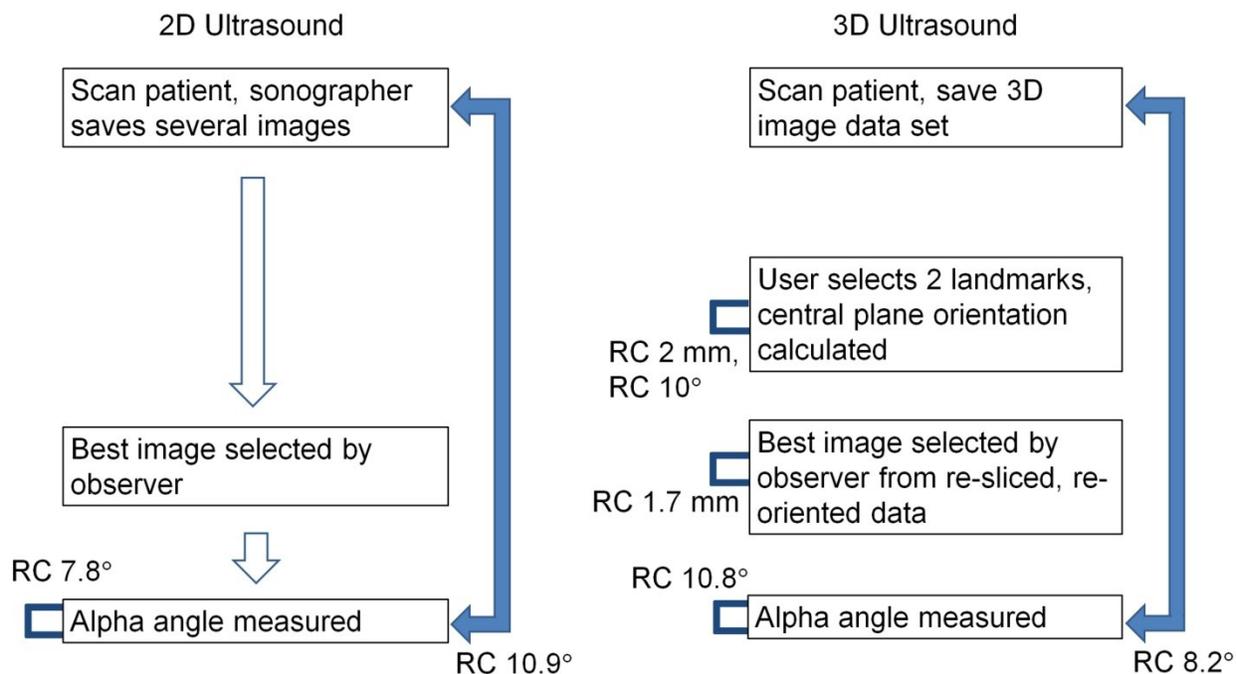


Figure 4.8 Data processing workflow and reliability for 2D and 3D ultrasound. RC = reliability coefficient of inter-observer variability for the relevant measurement, distance or angle. Note that the overall inter-scan reliability of 3D ultrasound alpha angle measurement, combining all processing steps, was significantly improved over 2D ultrasound, and that most of the variability in both steps occurred in actual measurement of the alpha angle.

#### 4.3.6 COMPARISON OF 3D AND 2D ALPHA ANGLES

The alpha angles measured on conventional 2D hip ultrasound and on the 3D ultrasound performed at the same visit were compared for each hip. The alpha angles differed by  $0.5 \pm 6.7^\circ$  for Observer 1 and  $3.8 \pm 7.3^\circ$  for Observer 2, with most of the variability occurring in dysplastic hips (SD of difference between 2D and 3D =  $9.4^\circ$  for dysplastic hips vs.  $3.3^\circ$  in normal hips). No patients with DDH had a 3D ultrasound alpha angle  $>60^\circ$  for either observer, while 2-4 clinically normal hips had a 3D ultrasound alpha angle  $<55^\circ$ .

#### 4.3.7 ALPHA ANGLES IN NORMAL VS. DYSPLASTIC HIP

On both 2D and 3D ultrasound the categories of dysplasia (0=normal, 1=borderline initially but ultimately normal, 2=dysplastic and given treatment) were separated into distinct

groups by their alpha angles. For 2D ultrasound, the mean alpha angle was  $67.8^{\circ}$  for category 0 (95% CI of mean  $65.1-70.5^{\circ}$ ),  $57.1^{\circ}$  ( $52.1-62.1^{\circ}$ ) for category 1, and  $45.7^{\circ}$  ( $42.4-49.0^{\circ}$ ) for category 2. For 3D ultrasound, values were  $65.3^{\circ}$  ( $61.5-69.1^{\circ}$ ) for category 0,  $57.2^{\circ}$  ( $52.9-61.6^{\circ}$ ) for category 1, and  $45^{\circ}$  ( $41.7-48.3^{\circ}$ ) for category 2.

#### 4.4 DISCUSSION

The purpose of this study was to evaluate the reliability of 3D hip ultrasound landmark localization, and whether this led to improved reliability of alpha angle measurement characterizing hip dysplasia. We found that landmarks at the anterior and posterior acetabular edges could be reliably identified by two users within 2 mm of each other, representing a high level of precision. The orientation of a standard plane generated using these landmarks was reproducible within  $10^{\circ}$ , representing a narrow range of variation compared to the  $24-45^{\circ}$  range that gave acceptable Graf standard-plane images in Chapter 2 of this thesis. Although these processing steps and central slice selection were highly reliable, the variability of the alpha angle was only slightly improved using 3D ultrasound compared to 2D ultrasound (RC  $8.2^{\circ}$  vs  $10.9^{\circ}$ ), with variation attributable mainly to the final processing step common to both 2D and 3D ultrasound – actual measurement of the alpha angle between two lines drawn on the ‘best’ 2D image from a 2D scan or a 3D data set.

The most important finding, that acetabular landmarks can be reliably identified and used to generate a reproducible central plane on 3D ultrasound of the hip, is a crucial prerequisite to effective use of 3D ultrasound in management of DDH. With this confirmation, we can justify fusing images from serial follow-up 3D ultrasound scans, each oriented along its central plane, allowing consistent visual and numeric tracking of DDH progression and response to therapy, and potentially leading to additional insights into 3D deformity in DDH.

The other main finding, that there is only slightly (although significantly) lower inter-scan variation between alpha angle measurements using 3D ultrasound compared to 2D

ultrasound, is expected in light of our sub-analyses showing that the error is dominated by the final processing step common to 2D and 3D ultrasound, i.e., drawing 2 lines and measuring the angle between them on the final 'best' ultrasound image. If a clinician's only purpose is to measure the alpha angle, 3D ultrasound adds little benefit. However, the alpha angle is an index designed for 2D ultrasound. Knowing that the orientation of a hip in space can be consistently determined using just two acetabular landmarks, we can begin investigating new indices of 3D hip dysplasia using geometric models derived from 3D ultrasound, which may improve on the limitations of the alpha angle.

The 2D alpha angle is known to be highly variable (Table 4.2). Our values of inter-observer, intra-observer, and inter-scan variability are comparable to the wide range of variability reported by others [2-6]. This range of variability is similar in magnitude to the 7-10° width of each Graf diagnostic category [7], suggesting that DDH may very easily be misdiagnosed if a clinician relies heavily on the alpha angle. This also confirms the previous findings using 3D ultrasound, as in chapter 2 of this thesis.

Once acetabular landmarks were defined and data rotated, a user was able to re-select these landmarks on the rotated data with high precision (RC 2.4°), implying that after one rotation the orientation of 3D ultrasound image data has 'snapped' into a stable final resting plane. Whether there is benefit from routinely adding a 2<sup>nd</sup> or 3<sup>rd</sup> iteration of landmark selection is uncertain and merits further study.

Once a central plane was identified, selection of the 'best' central slice meeting Graf criteria was highly reproducible between observers, varying by just 1.7 mm. The variability of the alpha angle on the 3D 'best' slice selected by users was comparable to the variability of the 2D alpha angle, as well as to inter-scan variability of 2D and 3D ultrasound (all with reliability coefficients ~8-11°). This implies that the majority of the 'error' in our 3D method is, as with traditional 2D ultrasound, contained within the actual drawing of the alpha angle measurement.

Users were able to recreate similar 2D images of each hip, even between totally separate 3D scans.

The two observers, MM and JJ, had different levels of experience. User MM (author of this thesis) has been working with 3D data sets throughout his degree, while JJ is a practicing radiologist with more experience in 2D ultrasound but less experience with 3D ultrasound, which may explain why user MM's intra-observer variability in 3D ultrasound was less than JJ's. In later studies more users and more training may be beneficial. However, it is actually a strength of this study that our two users had quite different backgrounds and experience; the levels of inter-observer reliability we report are therefore conservative and should only improve with further consensus training.

The variation in comparison of 3D and 2D parameters such as central plane orientation and alpha angle was higher in dysplastic than normal hips. This reflects the difficulty in identifying the acetabular margin which is smoother and less crisply defined in a dysplastic hip (e.g., Figure 4.3(iii)). The orientation of the central plane is more precise in normal patients, just as the alpha angle is measured with less variability in normal patients.

This study had limitations, related to its exploratory nature. To our knowledge we are the first clinic to systematically add 3D ultrasound for infant hip dysplasia, as was the case in chapter 2 of this thesis, this gave us a limited sample size to draw from; however, we were still able to include a full range of hips in each diagnostic category from normal to severely dysplastic. The clinical diagnoses and treatment decisions were made by clinic orthopedic surgeons as per usual practice, and were not validated by any external gold standard. Our results showing differences in reliability of ultrasound parameters based on clinical diagnosis should be interpreted cautiously as the ultimate diagnosis may vary in individual patients after longer term follow-up. Also, although we have confirmed that acetabular landmarks can be reliably identified and allow 3D image data to be rotated into a reproducible orientation, we have not demonstrated whether the landmarks or orientation have any physical meaning. To

study this would require fusion (ie. overlaying of specific landmarks or of entire bone tracings between two modalities) of 3D ultrasound data sets to other imaging modalities and/or a physical hip model, studies we intend to perform in the future. Finally, the clinical utility of measuring the alpha angle on 3D ultrasound is uncertain, as this is not what the alpha angle was designed for. Further testing of the 3D ultrasound alpha angle against long term clinical outcomes would be needed to confirm its validity in this context. There is opportunity to use the complete acetabular shape data available in 3D ultrasound to develop new indices of DDH which may improve on the obvious limitations of the alpha angle, which are discussed in the next chapter of this thesis.

#### 4.5 CONCLUSION

Acetabular landmarks can be reliably identified within 2 mm in 3D hip ultrasound, allowing rotation of data into a standardized central plane reproducible within  $10^\circ$  between observers. Use of this standard central plane only slightly improves inter-scan variability in alpha angle measurement vs. traditional 2D ultrasound, likely because most variation in the alpha angle occurs in the final step of its measurement. Applications of the standardized 3D ultrasound central plane will be to fuse serial ultrasounds for follow-up assessment, and to underpin development of new indices of 3D deformity in DDH which may improve on the diagnostic accuracy of the alpha angle.

#### 4.6 REFERENCES

- [1] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476), pp. 307-310. 1986.
- [2] E. A. Simon, F. Saur, M. Buerge, R. Glaab, M. Roos and G. Kohler. Inter-observer agreement of ultrasonographic measurement of alpha and beta angles and the final type classification based on the graf method. *Swiss Med. Wkly.* 134(45-46), pp. 671-677. 2004.

- [3] E. A. Roovers, M. M. Boere-Boonekamp, T. S. Geertsma, G. A. Zielhuis and A. H. Kerkhoff. Ultrasonographic screening for developmental dysplasia of the hip in infants. reproducibility of assessments made by radiographers. *J. Bone Joint Surg. Br.* 85(5), pp. 726-730. 2003.
- [4] M. Zieger. Ultrasound of the infant hip. part 2. validity of the method. *Pediatr. Radiol.* 16(6), pp. 488-492. 1986.
- [5] D. P. Gwynne Jones, A. G. Vane, G. Coulter, P. Herbison and J. D. Dunbar. Ultrasound measurements in the management of unstable hips treated with the pavlik harness: Reliability and correlation with outcome. *J. Pediatr. Orthop.* 26(6), pp. 818-822. 2006.
- [6] C. Morin, S. Zouaoui, A. Delvalle-Fayada, P. M. Delforge and H. Leclet. Ultrasound assessment of the acetabulum in the infant hip. *Acta Orthop. Belg.* 65(3), pp. 261-265. 1999.
- [7] R. Graf. Fundamentals of sonographic diagnosis of infant hip dysplasia. *J. Pediatr. Orthop.* 4(6), pp. 735-740. 1984.

# CHAPTER 5

## Chapter 5 ..

### A NOVEL INDEX IN THE DIAGNOSIS OF INFANT HIP DYSPLASIA USING 3D ULTRASOUND: THE INSTABILITY INDEX

#### 5.1 INTRODUCTION

This chapter expands on the potential uses of 3D ultrasound in the diagnosis of infant hip dysplasia. We have illustrated that we can reliably reorient 3D data sets of infant hips into a standardized view of the relevant bones (acetabulum and femoral head). Once in this view we can begin to analyze different patients in the same manner for a meaningful diagnosis of dysplasia. As mentioned earlier (section 1.22), we know that the acetabulum is remodeled based on forces and loading factors imposed by the femoral head in the ball and socket joint [1]; consequently we can develop indices which measure this ‘stability’ within the socket. This chapter introduces a new index for measuring infant hip dysplasia, the ‘inStability’ index, which may differentiate between normal and dysplastic hips. The ‘inStability’ index is based on the acetabulum resisting loads perpendicular to its surface (normal forces). The chapter focuses on illustrating the diagnostic utility, as well as inter/intra observer reliability, and inter scan reliability of the new index.

#### 5.2 METHODS

##### 5.2.1 PATIENTS

This prospective study was approved by the University of Alberta Health Research Ethics Board. Imaging was performed at a tertiary pediatric hospital from October 2012 to April 2014. At the first routine clinical 2D US of the hip in each patient, written informed consent was obtained from a parent to add 3D US of the hip at the same visit. The imaging indication was clinical suspicion of DDH because of laxity at examination, asymmetric hip creases, or other risk factors such as a family history of a positive finding for DDH. Because dysplasia can be unilateral or bilateral, we included each hip separately. We had 100 hips in 83 patients, and 60 (72%) of them were female. Patients underwent US at the mean age of 45.4 days (range, 4-173 days; for female infants, mean age was 46.1 days [range, 5 – 173 days]; for male infants, mean age was 42 days [range, 4 -111 days]). They received routine care at a pediatric orthopedic clinic from one or more of five clinic surgeons who were blinded to 3D US images and findings. We observed clinical care for at least 6 months to classify each imaged hip as normal at the first orthopedic assessment (category 0; n = 40), questionably abnormal initially but with findings that resolved spontaneously at follow-up imaging and clinical examination (category 1; n = 21), or dysplastic and requiring treatment by using a Pavlik harness and/or surgery (category 2; n = 39).

### 5.2.2 IMAGING

All imaging was performed by using platforms (Philips iU22; Philips Healthcare, Andover, Mass). We performed conventional 2D US in both hips by using a 12-MHz linear transducer (Philips L 12-5; Philips Healthcare, Andover, Mass), including static coronal imaging in the Graf standard plane and axial dynamic imaging to assess for hip stability by using our usual protocol. Two-dimensional US images were interpreted by a pediatric radiologist, with results and images made available to referring clinicians. In addition, two study team members, including a radiologist, technologist, or medical or graduate student trained by the study radiologists, used a high-resolution 13-MHz 3D linear (13VL5; Philips Healthcare, Andover, Mass) transducer in the coronal orientation to obtain a 3D US data set at each hip. With this

transducer, we performed a 3.2 second automated sweep through a range of  $\pm 15$  degrees to generate a 3D data set of 256 US sections that were 0.13 mm thick, each containing 411 X 192 pixels measuring 0.11 X 0.20 mm. These 3D scans were not released for use in clinical treatment.

### 5.2.3 IMAGE PROCESSING

Images were analyzed off-line using custom software (MATLAB R2010a-2012; The MathWorks, Natick USA). For each hip, 3 observers (MM, a graduate student in biomedical engineering/radiology and author of this thesis, JJ, the lead radiologist with fellowship training in pediatric and musculoskeletal radiology and 11 years experience, and EC, a medical student with experience in scanning infant hips and working with 3D data sets) performed trials on the 3D data sets of each hip. In each 3D data set each user selected two points based on predefined landmarks which would be used to define a rotation of the data-set into an orientation that would standardize the viewing plane of the acetabulum and femoral head, as per the descriptions laid out in chapter 4 of this thesis (also illustrated in Fig 5.1 (i) and Fig 5.1 (v)).

For each hip, the acetabulum was traced on all re-oriented slices between the landmark points on each end (Fig 5.1 (ii) and Fig 5. (iv) illustrate this), and the femoral head was also traced on all slices between the acetabulum landmarks which provided a clear view of the femoral head (Fig 5.1 (ii) and Fig 5. (iv)). The tracings were done using a customized semi-automated interactive interface that allowed contours traced on selected sections to be interpolated to intervening sections, then reviewed and corrected by using a nudging tool.

These tracings were used to create a surface model of both the acetabulum (Fig 5.1 (vi)) and femoral head. Surface normals were calculated based on tangent planes at each point on the acetabulum model. After performing a bicubic fit of the data in the x, y, and z directions, diagonal vectors and their cross products were computed to form the normal at each point (Fig 5.2 (i)). Each point on the acetabulum surface was then checked for the closest point on the femoral head on the same slice. Once these two calculations had been performed, all surface

normals were weighted according to a  $d^{-2}$  (distance) proportionality, and then averaged to create a single cumulated surface normal (Fig 5.2 (ii)). The user would then draw a line representing the iliac wall on the first 1-5 slices after the posterior landmark 'bump' had leveled off into a flat iliac wall (Fig 5.1 iii, and Fig 5.2 ii). The inStability angle was calculated through the dot product of the iliac wall line and the weighted surface normal (ie. the angle between the iliac wall and the vector intended to represent the line of action of the net force that can act from the acetabulum on the femoral head).

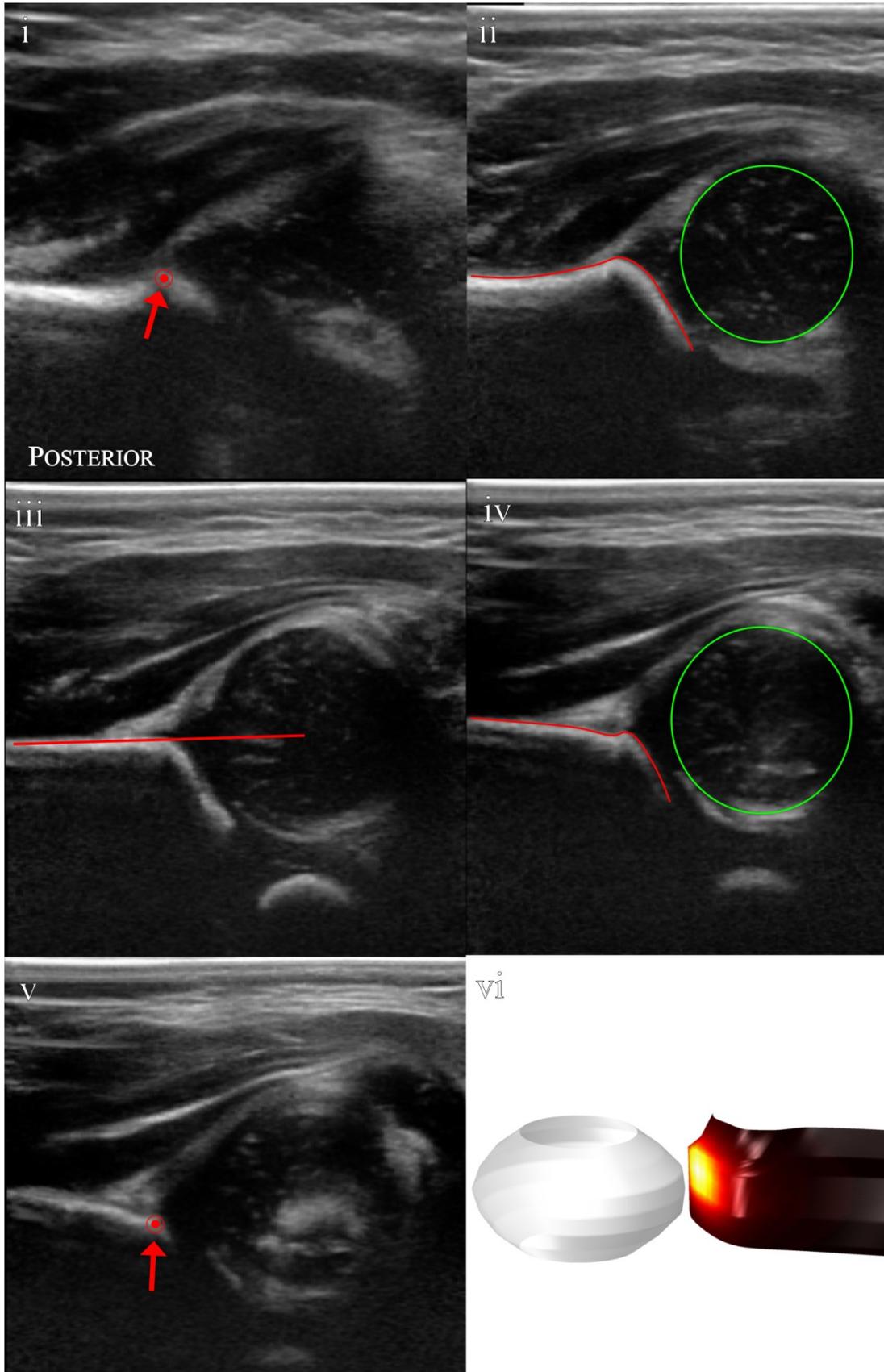


Figure 5.1 Acetabulum landmarks, tracing techniques, and surface model. Figures (i) and (v) show the acetabulum landmarks at the posterior and anterior edges of the acetabulum, slices more anterior or posterior no longer showed acetabulum edges. Figures (ii) and (iv) show acetabulum (red) and femoral head (green) tracings on slices between the landmarks. Figure (iii) shows a slice with a representative Iliac Wall line, which would be used in the calculation of the inStability angle. Figure (vi) shows the acetabulum surface model which was generated from the tracings on this patient, surface model is also color coded for distance to femoral head (white = least, black = largest distance), femoral head is shown semi transparently in white. Figure (vi) is shown from an orientation reversed from figures (i) through (v)

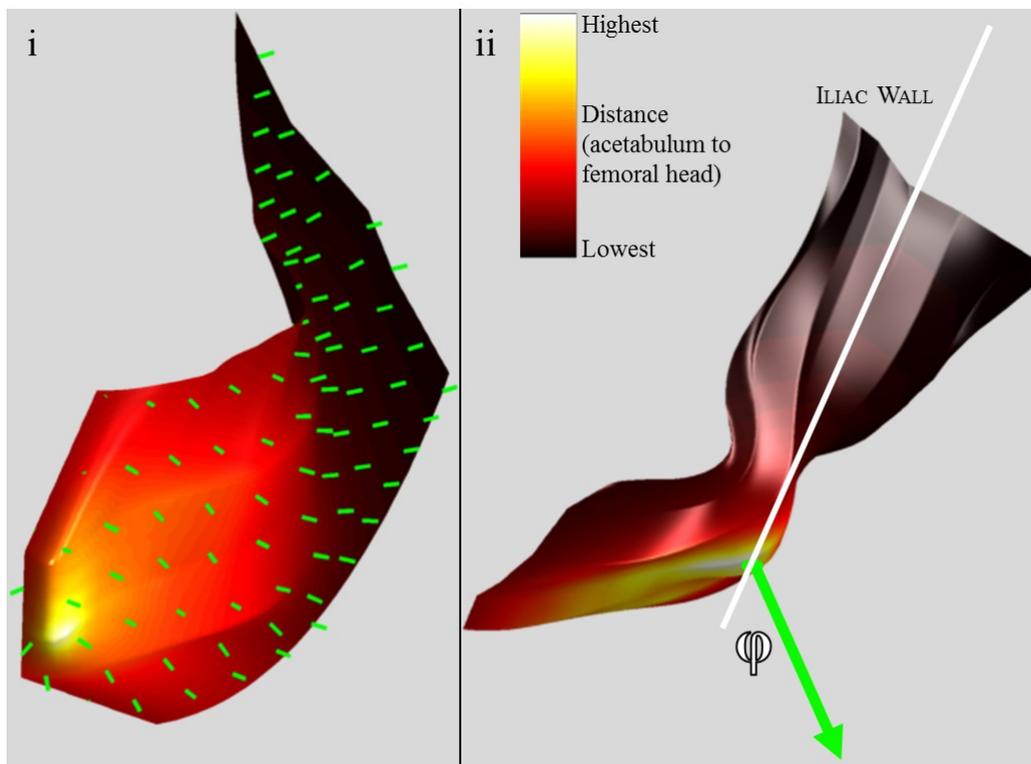


Figure 5.2 Surface normals and inStability Angle. Figure (i) shows a limited number of the surface normals used to calculate the net surface normal. Figure (ii) illustrates a representative Iliac Wall (white line) and net surface normal (green arrow) as well as the inStability Angle ( $\phi$ ).

User 1 (MM) did tracings on the total set of each hip scans ( $n = 100$  scan 1,  $n = 98$  scan 2), User 2 (JJ) did tracings on a random selection of  $n = 20$  hips in each set, and User 3 (EC) did one full set of scan 1 ( $n = 100$ ).

Lastly, the Acetabular surface models of 15 hips (two observers and two separate scans with 5 hips in each of the diagnostic categories) were matched and aligned with Amira software (FEI Visualization Sciences Group, Düsseldorf, Germany), and root mean square error (RMS) values were calculated, the RMS error represents the smallest average distance from a point on one model to the nearest point on another (Fig 5.3). Additionally, one user created detailed (high fidelity) models on all patients twice to compare the difference between simplified (representative of tracings done in the study) and detailed models. The detailed models consisted of drawing contours on 8-10 slices with approximately double the number of points as the simplified models which were drawn on only 5-6 slices with 5-6 points each and relied heavily on interpolation.



Figure 5.3 Illustration of the RMS between Surface Models. Three views of the same patient as compared by user 1, red is a detailed tracing and white is a simple tracing, this hip's RMS value was 0.21 mm indicating that the distance from a given point on one model to the closest similar point on the other was very small, typically an approximate size of 1 voxel.

The training process involved preliminary trials by four users (MM, JJ, EC, and MN; MN a pediatric radiologist) testing how best to trace the acetabulum and femoral head through an iterative process of random 3D data sets.

The alpha angle was also calculated for each hip by one user (MM) for comparison against the inStability angle. We have previously validated the reliability of alpha angle assessment by this user against the lead radiologist ( $0.8^\circ \pm 3.9^\circ$   $n = 51$ , which is similar to literature values for inter observer variability of the alpha angle).

#### 5.2.4 STATISTICAL ANALYSIS

Statistics were calculated on SPSS (Chicago, IL., USA, v. 20). Descriptive statistics were recorded as mean  $\pm$  standard deviation (SD). For repeated measurements we calculated the mean difference in values and to characterize the limits of agreement, reported the repeatability coefficient (RC), i.e., the range of values within which there is 95% confidence that a second observation will fall from the first [2]. The intraclass correlation coefficient (ICC) was also calculated to assess conformity between users.

For each scan, inter-observer variability of the inStability angle was calculated; intra-observer variability was also calculated (with the user leaving one week between consecutive tracings of the same hip) as well as inter-scan variability of the hips. The inStability angle was also compared against diagnostic category. Root mean square error values of surface models were calculated between models of high and low fidelity, between users, and between different scans.

### 5.3 RESULTS

#### 5.3.1 VARIABILITY OF INSTABILITY INDEX

The intra-observer variation of the inStability index was  $-0.86^\circ \pm 3.5^\circ$  for user 1 (MM on  $n = 98$  cases), and  $-0.05^\circ \pm 3.5^\circ$  for user 2 (JJ on  $n = 20$  cases) with the same repeatability coefficient for both users,  $6.86^\circ$ . The ICC for user 1 was 0.98 and 0.97 for user 2.

The inter-observer reliability of the inStability index was  $0.23^{\circ} \pm 4.25^{\circ}$  between users 1 and 2 (MM/JJ on  $n = 20$  cases) and  $-2.6^{\circ} \pm 4.0^{\circ}$  between users 2 and 3 (JJ/EC on  $n = 20$  cases), with similar repeatability coefficient between users,  $8.33^{\circ}$  and  $7.84^{\circ}$  respectively. The ICC value between user 1 and user 2 was 0.97, while between users 2 and 3 it was also 0.97.

The inter-scan reliability of the instability index for user 1 was  $-0.44^{\circ} \pm 5.82^{\circ}$  (MM on  $n = 100$ ) and  $0.9^{\circ} \pm 4.2^{\circ}$  (JJ on  $n = 38$  cases), with repeatability coefficients  $11.4^{\circ}$  and  $8.2^{\circ}$  respectively. The ICC values for user 1 was 0.924, and 0.965 for user 2.

### 5.3.2 NORMAL VS. DYSPLASTIC HIPS

Categories of dysplasia (0 = normal, 1 = borderline initially but ultimate normal, 2 = dysplastic and given treatment) were separated into distinct groups by both their 2D alpha angle, and their inStability index. For 2D ultrasound (standard Alpha Angle), the mean alpha angle was  $66.9^{\circ} \pm 5.5^{\circ}$  (95% CI of  $65.1^{\circ}$ ,  $68.6^{\circ}$ ) for category 0,  $54.4^{\circ} \pm 8.82^{\circ}$  ( $50.4^{\circ}$ ,  $58.4^{\circ}$ ) for category 1, and  $44.9^{\circ} \pm 8.09^{\circ}$  ( $42.3^{\circ}$ ,  $47.6^{\circ}$ ) for category 2, see Fig 5.4. All means were significantly different with p-values falling below 0.0001 between all categories. For 3D ultrasound, the mean inStability angle was  $30.98^{\circ} \pm 4.7^{\circ}$  ( $29.5^{\circ}$ ,  $32.5^{\circ}$ ) for category 0,  $36.8^{\circ} \pm 8.0^{\circ}$  ( $33.3^{\circ}$ ,  $40.6^{\circ}$ ) for category 1, and  $52.3^{\circ} \pm 8.4^{\circ}$  ( $49.6^{\circ}$ ,  $55.0^{\circ}$ ) for category 2, see Fig 5.4. Means were significantly different ( $p < 0.0001$ ) for categories 0 vs 2 and 1 vs 2, and significantly different for categories 0 vs 1 ( $p = 0.003$ ). Included are also pyramid graphs showing a case by case frequency distribution with lines showing the cut off threshold of  $60.0^{\circ}$  separated by diagnostic category (normal above  $60^{\circ}$ , dysplastic below  $60^{\circ}$ ) for Alpha Angle (Fig 5.5), and a cut off threshold of  $40.0^{\circ}$  for inStability angle (normal below  $40^{\circ}$ , dysplastic above  $40^{\circ}$ , Fig 5.5).

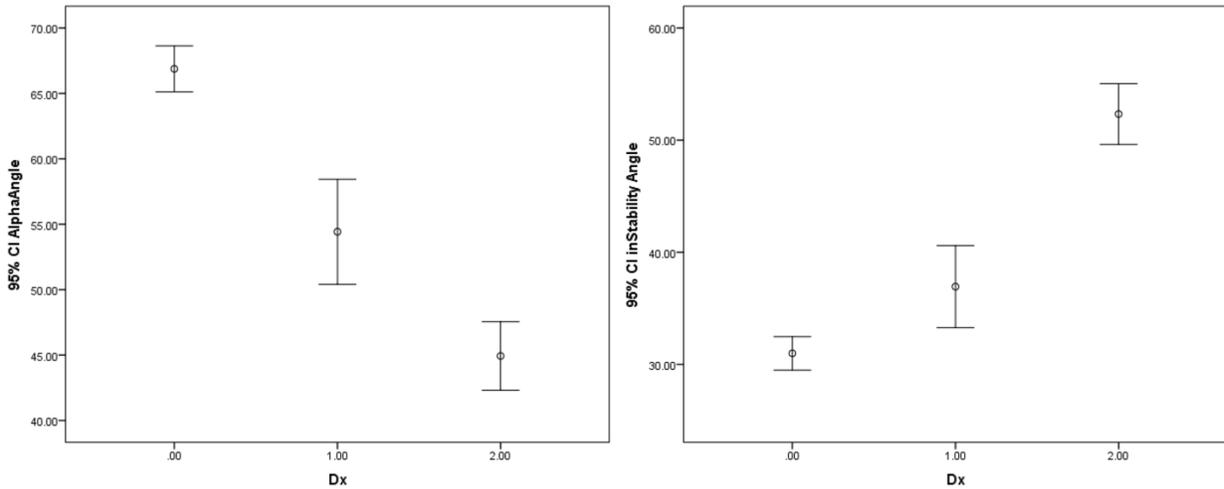


Figure 5.4 Alpha and inStability Angle by Diagnostic Category. At left, The 95% confidence intervals for each mean Alpha Angle value by each diagnostic category. Note the large gap between the two ‘normal’ categories (0 and 1). At right, The 95% confidence interval for each mean inStability angle by diagnostic category, note the large gap between normal (0 and 1) and dysplastic hips (2).

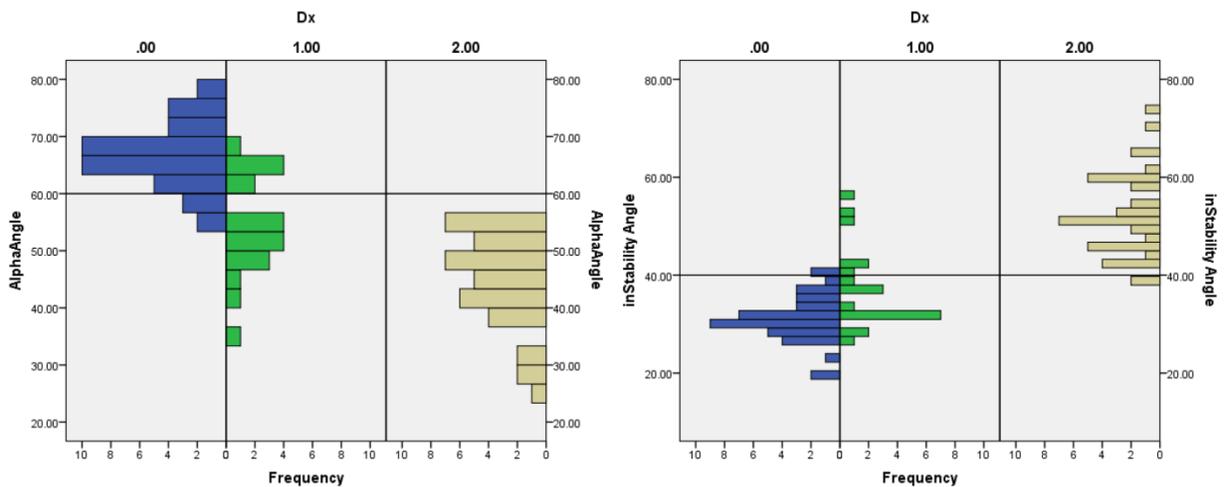


Figure 5.5 Case by Case Frequency Distribution Pyramid Graphs for Alpha and inStability Angle by Diagnostic Category. At left, Alpha Angle, Category 0 and 2 hips have largely different location of distributions, while category 1 hips have a distribution which stretches into both categories. Black horizontal line at 60° represents the cut-off defined by Graf for normal hips. At Right, inStability Angle. Categories 0 and 2 have largely different distributions, while category 1 is similar to category 0 with a few outliers.

### 5.3.3 RECEIVER OPERATING CHARACTERISTIC CURVES

The receiver operating characteristic curves are also included for both the Alpha Angle, and the inStability index, as shown in Fig 5.6, for a diagnostic test using these imaging indices individually to detect DDH compared to the clinical gold standard diagnosis by an orthopedic surgeon using all available clinical and imaging information. The area under the curve for Alpha Angle was 0.928, and for inStability index was 0.965.

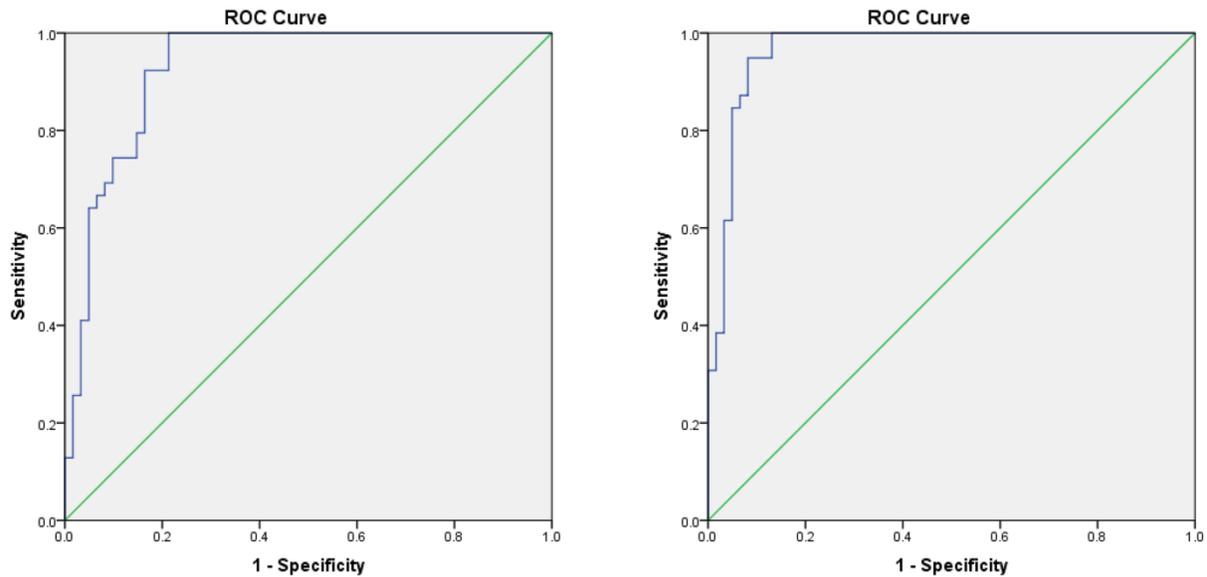


Figure 5.6 ROC Curve for Alpha and inStability Angles. At left, Alpha Angle, The area under the ROC curve was 0.928. At right, inStability Angle, the area under the ROC curve was 0.965.

### 5.3.4 CORRELATION OF INSTABILITY AND ALPHA ANGLES

The pearson correlation coefficient between user 1's (MM) instability angle and alpha angle values (n = 100) was -0.753 at a significance of  $p < 0.001$  (Fig 5.7).

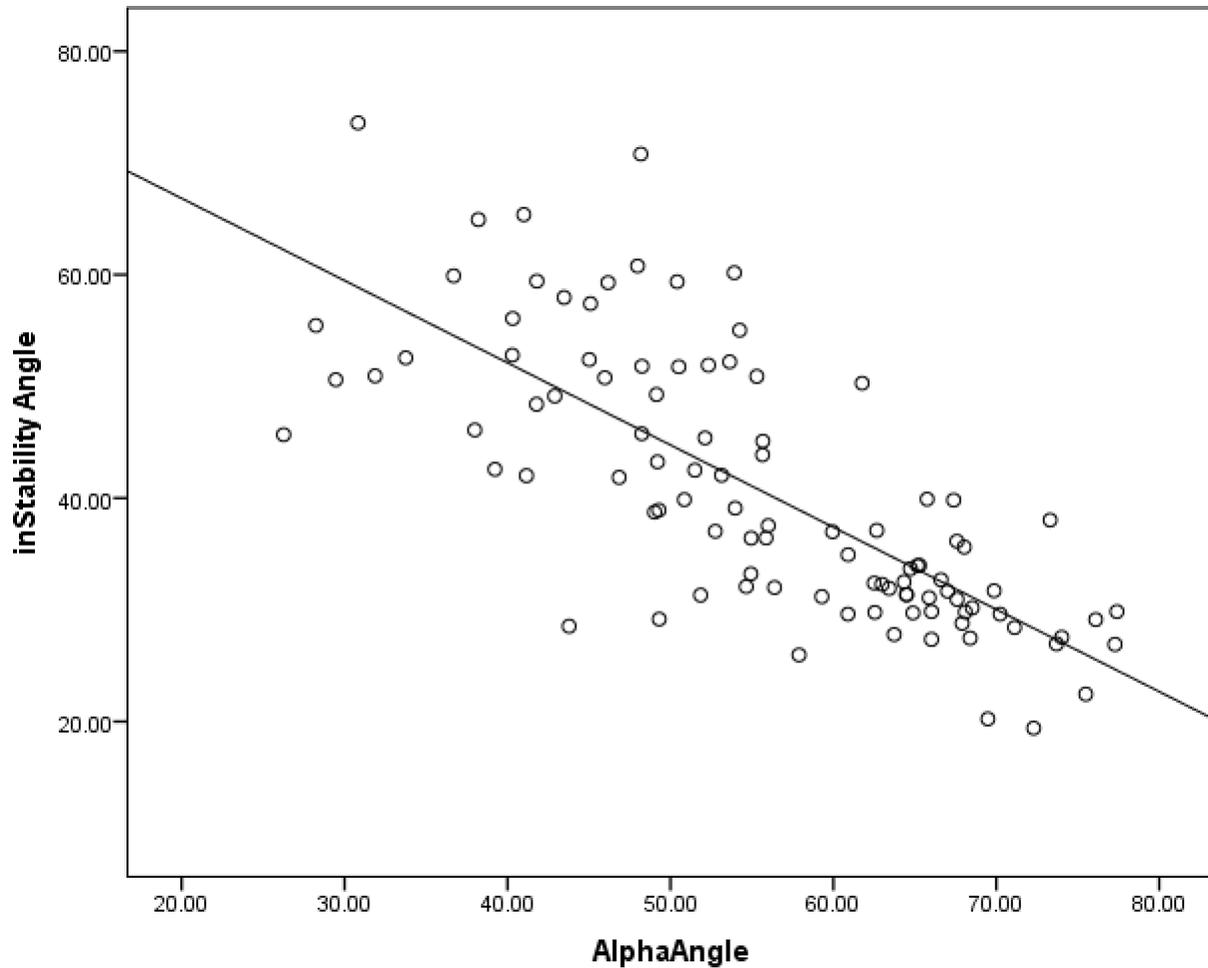


Figure 5.7 Correlation of Alpha Angle to inStability Angle. Alpha angle and inStability angle was moderately correlated, with a Pearson correlation coefficient of -0.753. This implies that inStability angle may measure different aspects of dysplasia than alpha angle.

### 5.3.5 ACETABULUM FUSION

The root mean square values for the surface models are summarized in the table below (table 5.1). The table shows the effect of changing fidelity, observer, and scan between surface models on the average distance between corresponding points on the models. The largest error was between different scans of the same hip for user MM, a value of  $0.436 \pm 0.19$  mm (mean  $\pm$  standard deviation).

**Table 5.1 RMS Values for Acetabulum Surface Models**

	Detailed vs Detailed	Simple vs Detailed	JJ v MM	MM Scan 1 v Scan 2	JJ Scan 1 v JJ scan 2
RMS in mm (mean +/- SD)	0.178 +/- 0.04	0.198 +/- 0.05	0.238 +/- 0.06	0.436 +/- 0.19	0.402 +/- 0.13
Change	Nothing	Fidelity	Observer	Scan	Scan

## 5.4 DISCUSSION

The purpose of this study was to introduce a new index which can be generated from 3D ultrasound to quantify and categorize infant hip dysplasia. Chapter 4 of this thesis dealt with the reliability of the landmarks required to standardize the orientation of the 3D data into viewpoints which can be compared between patients, while this chapter focuses upon the reliability and diagnostic utility of the inStability index derived using these landmarks. With intra-observer variation of the inStability index measuring in at  $-0.86^{\circ} \pm 3.5^{\circ}$  and  $-0.05^{\circ} \pm 3.5^{\circ}$  (user 1 (MM) and user 2 (JJ), respectively), the numbers are very similar to reported values of alpha angle intra-observer variation in literature and in our own studies (standard deviations of  $3^{\circ}$  [3],  $3.2^{\circ}$  [4], and  $2.8^{\circ}$  [5] literature;  $3.1^{\circ}$  in chapter 4). Similar trends were also observed for inter-observer and inter-scan variation between the 3D inStability index and 2D alpha angle. Inter-observer reliability of inStability index was  $0.23 \pm 4.25$  (repeatability coefficient  $7.8^{\circ}$ ) and alpha angle literature values were:  $2^{\circ}$  [6] and  $3.5^{\circ}$  [5] standard deviations with one paper reporting a high repeatability coefficient of  $17^{\circ}$  [7], while our own standard deviation of the alpha angle was  $3.9^{\circ}$  (chapter 4). Inter-scan variability of the inStability index was  $-0.44 \pm 5.82^{\circ}$  and  $0.9 \pm 4.2^{\circ}$  (user 1 and user 2, respectively) and was similar to our reported value in chapter 4 of a  $5.5^{\circ}$  standard deviation for the alpha angle. Comparison to the alpha angle is a

useful metric as it is the index most widely used and has the lowest variability of any 2D ultrasound measured index, especially in comparison to the beta angle [3, 4, 6, 7] or femoral head coverage [7, 8]. The inStability angle has inter-observer variability comparable to the alpha angle.

The receiver operating characteristic curves of the inStability index and alpha angle provide a useful insight into the diagnostic utility of this new metric, as shown in Fig 5.7. While the two curves are very similar and both appear to effectively diagnose hip dysplasia, it is actually the inStability index which has the greater area under its curve, 0.965 vs 0.928 for alpha angle, suggesting a trend toward the inStability index representing a better metric to measure the condition. These values appear very high, admittedly areas under an ROC curve which are greater than 0.90 are exceptional, but this is likely because the individuals who scanned the patients in this study are highly practiced at scanning for infant hip dysplasia (hundreds of scans in total as well as performing scans every week). Therefore, alpha angle's area under this ROC curve is probably better fit to the value of 0.836 as it was reported in chapter 3 of the thesis, this value dealt with scans retrieved from our tertiary hospital imaging records and reflects a more varied expertise of sonographers. Consequentially the inStability ROC value may also be slightly inflated.

Although all dysplastic hips had alpha angle less than  $60^\circ$  (which is generally accepted as the cut off between normal vs dysplastic [9]), the alpha angle did not as distinctly separate normal (categories 0 and 1) and dysplastic (category 2) hips as well as the inStability angle did. Category 1 hips appeared with a high variability in the alpha angle, frequently with alpha angles falling into the normal and dysplastic categories. Figure 5.4, helps illustrate the prevalent difference in alpha angle between category 0 and 1 hips (both of which are normal and did not undergo treatment), and have significantly different means ( $p < 0.0001$ ). The inStability index had a much less prominent difference between category 1 and 0 hips, with only a small gap of  $0.8^\circ$  between the two categories' 95% confidence intervals (Figure 5.4). Although categories 0

and 1 were still significantly different in terms of means for inStability angle, it was trending towards insignificance, with the removal of only two outliers from category 1 the p value grew to an insignificant value ( $p = 0.012$  from  $p = 0.003$ ). The inStability index pyramid graph, shown in Fig 5.5, also shows this trend, with category 0 and 1 hips appearing very similar and only 5 patients in category 1 resting in the defined 'dysplastic' region (above  $40^\circ$ ). Also to note is that, out of the total 5 patients who were misclassified in the category 1 diagnosis, 3 of them were scanned at too early of an age (4 days old, 11 days old, and 22 days old).

It is generally recommended to scan infants at around 6 weeks as before this the hip is immature [10-12]. Although delayed scanning is typically an accepted practice for 2D ultrasound, we have no reason to believe it is not applicable to 3D ultrasound as well, as this is a general statement on the physiological development of the hip. The two dysplastic patients whose hips had inStability angles appearing normal ( $<40^\circ$ ) were both very close to crossing the threshold. A  $40^\circ$  threshold for the inStability angle visually best distinguished normal from dysplastic hips in this data set, but should not be taken as an absolute tolerance for dysplasia or normal hip. To do this requires extensive further validation with a larger study and more precise knowledge of what exactly is being measured; however, it is still useful as a general tool for classifying the data collected in this study.

Much of the variability in the inStability index came from the user's selection of the best iliac wall representation. We carefully derived simple rules to optimize the reliability of this selection, but, like anything relying on user input, there will be disagreement and personal bias introduced. An automated technique for characterizing the iliac wall angle from a 3D surface model may be more reliable, but will require substantial effort outside of the scope of this proof of concept.

The moderate correlation of  $r=-0.753$  between inStability angle and alpha angle, with an  $r^2 = 0.5$  suggests that only about half the variation in one of these angles accounts for variation in the other; this implies that the inStability angle measures different aspects of dysplasia than

the alpha angle. If the inStability angle were simply a 3D version of the alpha angle, then the correlation between the two would be expected to be much higher. Also, the moniker ‘inStability angle’ may be more aptly renamed as something similar to ‘contact angle’ (as it is based on the area on the acetabulum most likely to be contacted on the femoral head).

As shown in Table 5.1, the fusion between surface models was extremely accurate, with the largest values (Scan 1 vs Scan 2) still below half a millimeter. Drawing a higher fidelity model of the acetabulum only slightly decreases variability when compared to a simple model. This result aids in the decision of choosing to trace the more time efficient simple models. The RMS values between users were also very similar to intra-observer results, confirming our qualitative impression that drawing the surface model contours is generally straightforward and unlikely to be heavily biased. Changing observer does not appear to be a great source of variability for tracing models. Performing completely separate ultrasound scans by different observers and fusing surface models from these has slightly higher error, essentially double the RMS values of using the same scan; however, even at double the variability, the variability is still very small at only half a millimeter.

This study had limitations similar to the rest of this thesis. We are the first clinic to systematically add 3D ultrasound for infant hip dysplasia, as with chapters 2 and 4 of this thesis, this gave us a limited sample size to draw from; however, we were still able to include a full range of hips in each diagnostic category ranging from normal to severely dysplastic. The clinical diagnoses and treatment decisions were made by clinic orthopedic surgeons as per usual practice, and were not validated by any external gold standard. Our results showing differences in reliability of ultrasound parameters based on clinical diagnosis should be interpreted cautiously as the ultimate diagnosis may vary in individual patients after longer term follow-up. Finally, the clinical utility of the inStability angle is uncertain, further testing against long term clinical outcomes would be needed to confirm its validity as a diagnostic tool.

The use of 3D ultrasound gives a more full view of the acetabulum than available from current 2D ultrasound, and is likely to represent a useful diagnostic tool. Even without the use of a quantifying index such as the inStability index, review of the 3D ultrasound images and surface models may provide subjective and overall impressions of a patient's physiological development of value beyond a numeric index dysplasia.

## 5.5 CONCLUSION

The inStability index generated from 3D ultrasound data of infant hips can be reliably reproduced within and between users (intra-observer  $-0.05^\circ \pm 3.5^\circ$ , inter-observer  $0.23 \pm 4.25$ ), and between scans (inter-scan  $0.9 \pm 4.2^\circ$ ). The area under the curve of receiver operating characteristics of the inStability index was high and trended to be greater than that of the alpha angle for our data set (0.965 vs 0.928). These findings imply there may be high, clinical utility of 3D ultrasound scans for infant hip dysplasia. The inStability index shows promising initial results and warrants further study, and provides a good case for the use of 3D ultrasound in infant hip dysplasia.

## 5.6 REFERENCES

- [1] M. B. Ozonoff. *Pediatric Orthopedic Radiology / M.B. Ozonoff* 1992[University of Alberta JW Scott Health Sciences RD 732.3 C48 O99 1992].
- [2] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476), pp. 307-310. 1986.
- [3] E. A. Simon, F. Saur, M. Buerge, R. Glaab, M. Roos and G. Kohler. Inter-observer agreement of ultrasonographic measurement of alpha and beta angles and the final type classification based on the graf method. *Swiss Med. Wkly.* 134(45-46), pp. 671-677. 2004.
- [4] E. A. Roovers, M. M. Boere-Boonekamp, T. S. Geertsma, G. A. Zielhuis and A. H. Kerkhoff. Ultrasonographic screening for developmental dysplasia of the hip in infants. reproducibility of assessments made by radiographers. *J. Bone Joint Surg. Br.* 85(5), pp. 726-730. 2003.
- [5] C. Morin, S. Zouaoui, A. Delvalle-Fayada, P. M. Delforge and H. Lecllet. Ultrasound assessment of the acetabulum in the infant hip. *Acta Orthop. Belg.* 65(3), pp. 261-265. 1999.

- [6] M. Zieger. Ultrasound of the infant hip. part 2. validity of the method. *Pediatr. Radiol.* 16(6), pp. 488-492. 1986.
- [7] D. P. Gwynne Jones, A. G. Vane, G. Coulter, P. Herbison and J. D. Dunbar. Ultrasound measurements in the management of unstable hips treated with the pavlik harness: Reliability and correlation with outcome. *J. Pediatr. Orthop.* 26(6), pp. 818-822. 2006.
- [8] A. Falliner, D. Schwinzer, H. J. Hahne, J. Hedderich and J. Hassenpflug. Comparing ultrasound measurements of neonatal hips using the methods of graf and terjesen. *J. Bone Joint Surg. Br.* 88(1), pp. 104-106. 2006.
- [9] R. Graf. Fundamentals of sonographic diagnosis of infant hip dysplasia. *J. Pediatr. Orthop.* 4(6), pp. 735-740. 1984.
- [10] S. K. Storer and D. L. Skaggs. Developmental dysplasia of the hip. *Am. Fam. Physician* 74(8), pp. 1310-1316. 2006.
- [11] V. Bialik, G. M. Bialik, S. Blazer, P. Sujov, F. Wiener and M. Berant. Developmental dysplasia of the hip: A new approach to incidence. *Pediatrics* 103(1), pp. 93-99. 1999.
- [12] A. Roposch, L. Q. Liu, F. Hefti, N. M. P. Clarke and J. H. Wedge. Standardized diagnostic criteria for developmental dysplasia of the hip in early infancy. *Clinical Orthopaedics and Related Research* (12), pp. 3451. 2011.

# CHAPTER 6

## Chapter 6 ..

### DISCUSSION AND CONCLUSIONS

#### 6.1 DISCUSSIONS

##### 6.1.1 THESIS OVERVIEW

This thesis illustrated the flaws of the current 2D ultrasound imaging techniques used to diagnose DDH and introduced new techniques using both 2D ultrasound methods and 3D ultrasound. Chapter 2 showed the inter-scan error of 2D ultrasound can be substantial, and confirmed the in-vivo results of Falliner, et al [1], which showed that inter-scan variability is non-negligible and that the ‘Graf standard plane’ can be recreated with multiple probe orientations. Chapter 3 focused on adding a meaningful index to quantify acetabular rounding on traditional 2D ultrasound images obtained to measure the alpha angle by Graf’s method. The fourth chapter outlined the reliability and techniques we used to standardize our 3D data, and demonstrated that 3D scans can be used to obtain a Graf standard image of the hip. Finally, in chapter 5 I generated a new index from 3D ultrasound scanning, the inStability angle, which can reliably discern between dysplastic and normal hips.

##### 6.1.2 PROBLEMS WITH 2D ULTRASOUND AND POTENTIAL BENEFITS OF 3D

As mentioned, the inter-scan error of 2D ultrasound measurements can be substantial, but in addition to that problem is that the 2D Graf ultrasound method is a difficult scan to perform. Even in our tertiary hospital where these types of scans are performed regularly, errors commonly occur. Sometimes sonographers are unable to obtain the proper orientation. While

searching for images for use in the acetabular rounding study in chapter 3, we found multiple 2D ultrasound hip images saved to clinical records which were not adequate. Some of the common errors can be seen in Fig 6.1.

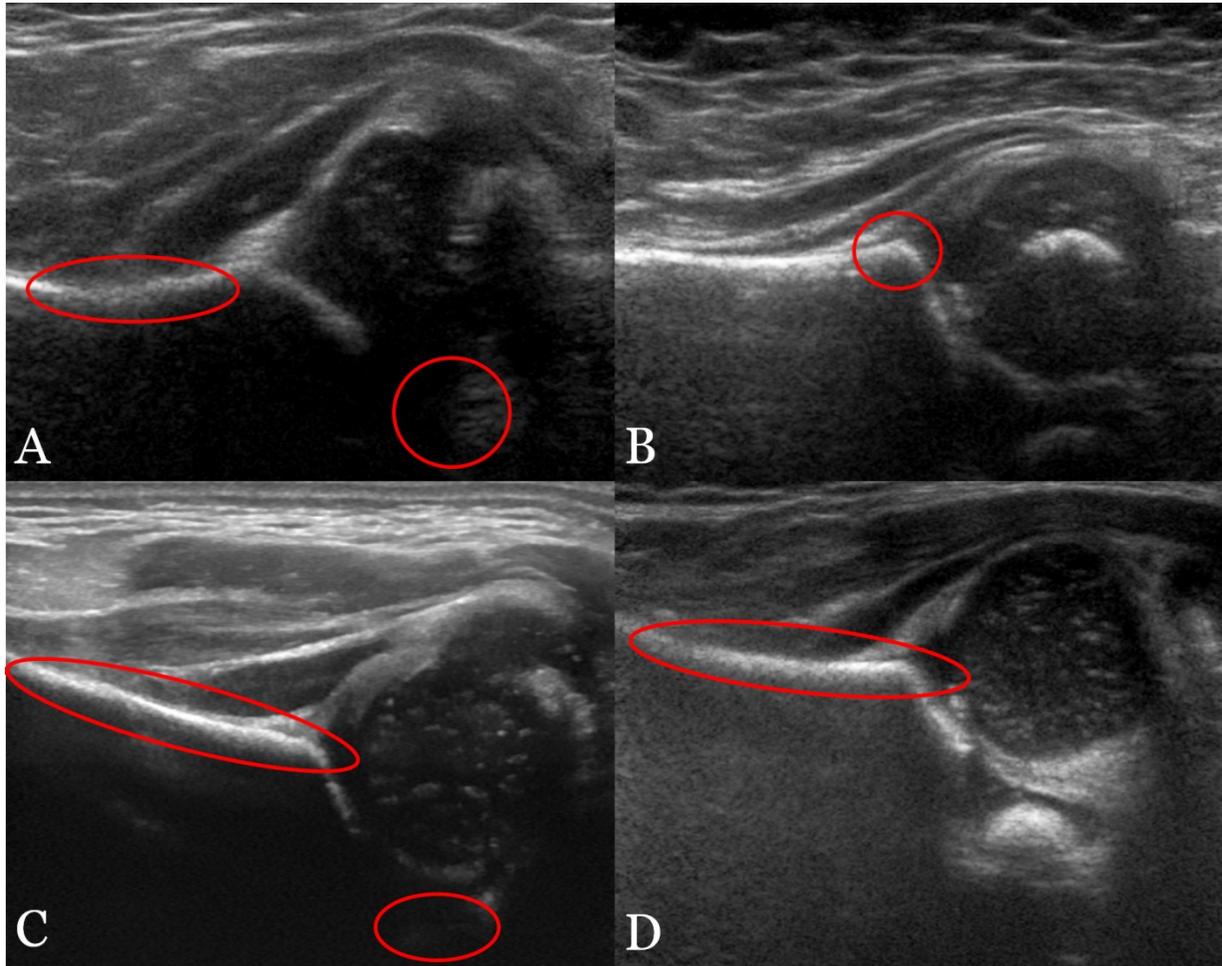


Figure 6.1 Common errors of 2D ultrasound scanning. A) has a curved iliac wing, and missing ischium. B) is too far posterior and includes a lip on the iliac wing and acetabulum corner. C) again has a curved iliac wing and very faint ischium, D) again has a curved acetabulum.

In contrast, by scanning the acetabulum in three dimensions, the entire bone can be observed and abnormalities should be visible. If the technician can obtain a scan in which the acetabulum is visible and little or no movement artifacts are present, then the scan will be technically successful. Consequently it is likely that less training will be required to produce 3D

ultrasound images than 2D, and the scan may be much easier to perform for centers which do not receive large numbers of infant hip scans. This also has great implications for rural areas (such as northern Canada) where a subspecialty trained radiologist may not be present to dictate the exam; an out of town radiologist can report the results correctly because they are provided a full sweep of the acetabulum and not just a 'snapshot' which they must assume is oriented correctly.

The added benefit of providing a radiologist an entire sweep of the hip is already a strong reason for the inclusion of this method; the radiologist can observe the entire 3D shape of the hip without actually entering the exam room and scanning the infant themselves. This benefit shaves substantial time off both the radiologist's and the technician's workload, as the technician no longer has to spend large amounts of time sifting and searching for the best 'snapshots' of the infants hip, but rather sweeps the whole hip and sends it straight to the radiologist. The radiologist is also able to reduce the number of 'suspect' hips that must be checked in person. It is likely that the quantification of hip dysplasia can also be much more accurate and precise than the current 2D method as it is based upon the entire hip.

### 6.1.3 REVIEW OF QUANTITATIVE RESULTS

Chapter 2 provided convincing quantitative results on the limitations of current 2D ultrasound methodology, specifically the inter scan variability. Adequate clinical quality images could be created in probe orientations varying by an average of  $24^\circ$ , or by as much as  $45^\circ$  in each of two perpendicular axes. These varying probe orientations caused significant changes in the alpha angle, with an average of  $13.9^\circ$  and standard deviation of  $7.1^\circ$  for each hip. Observed variation in alpha angle would cause crossing of clinical dysplasia categories, corresponding to over half of our data set traversing into at least two clinical categories. This variation was also even more significant in younger patients, specifically those under 1 month of age, where almost three quarters of the imaged hips crossed categories.

The indices which we introduced in chapter 3 to improve upon the current 2D ultrasound diagnosis of DDH, arc length and AROC, were shown to be quantifiably useful. My recommendation to incorporate only one of these indices, with the pick being AROC, came about as a result of the high correlation (Pearson correlation coefficient of 0.94) between the two. AROC displayed a high ICC value, comparable to the alpha angle (0.835 vs. 0.895, respectively) and a moderate correlation was displayed between the two ( $r = -0.588$ ) leading us to infer that the two measure different aspects of hip dysplasia and are both useful. The area under the ROC curves of alpha angle and AROC showed that similar diagnostic utility exists between the two (0.836 vs. 0.758, respectively). The AROC was able to detect between normal and dysplastic hips with significantly different average values between the two categories, 1.6 mm normal vs. 3.1 mm dysplastic.

Landmarks introduced in chapter 4 for the standardization of a 3D coordinate system were reproducible within 2 mm between observers and 1 mm within observers. The central plane created from these landmarks was also highly reliable, with the angle between user's planes being on average  $5.4^\circ$  away from each other. The inter-observer variability between the 3D and 2D alpha angles were comparable with repeatability coefficients of  $10.8^\circ$  and  $7.8^\circ$ , respectively. The inter-scan variability of the 3D alpha angle was slightly smaller than that of the 2D alpha angle with repeatability coefficients  $8.2^\circ$  vs.  $10.9^\circ$ .

The inStability index had inter and intra observer variability similar to alpha angle with values of  $-0.05^\circ \pm 3.5^\circ$  and  $0.23^\circ \pm 4.25^\circ$  (mean  $\pm$  SD) and good inter scan variability (coefficient of variability of  $8.2^\circ$ ). The root mean square error values of the surface models showed high reproducibility of models with the largest error (comparing models between two different scans of the same hip) falling below 0.5 mm and the smallest error (comparing two high fidelity models) resting at an average of 0.18 mm. The correlation between alpha angle and inStability angle had an  $R^2$  value of 0.5 explaining that only half of the variability in inStability

angle is accounted for by alpha angle and that they do measure different aspects of dysplasia. While the area under ROC curves trended towards higher clinical utility for inStability than alpha angle (0.965 vs. 0.928).

#### 6.1.4 EVALUATION OF CLINICAL UTILITY OF 3D ULTRASOUND

Due to the novel nature of 3D ultrasound imaging of infant hips, this thesis worked with small sample sizes. Although, the preliminary results were highly promising, it is still imperative that larger trials be conducted in the routine course of clinical care. We are the first center to begin to use 3D ultrasound routinely on infant hips (to our knowledge). In the coming years our center will have amassed several hundred (potentially in the thousands) infant hip scans for use in studies, as of May 31st 2014 we have acquired over 400 patients worth of scans, additionally each week approximately 8 normal and 2 abnormal hips are added to the database. As the sample size grows and clinical applicability becomes more realistic, there will be an increasing need to follow a formal and recognized approach to validate the clinical utility of the novel imaging techniques developed. We have been following and will continue to follow the OMERACT [2] guidelines to standardize our results. With our preliminary work we have already covered many of the original OMERACT filter guidelines in the three main categories (Truth, Discrimination, and Feasibility) for determining applicability of a measurement instrument in a clinical setting. In the truth section, we have already displayed face validity with our initial results, as well as content validity (3D ultrasound is conceptually better, as it can intrinsically acquire more information about the acetabulum's 3D geometry), with construct validity (how well does the measurement correlate to variables that are known to be related to the condition) remaining to be fully exhibited. For discrimination, we have already illustrated preliminary reliability (but would benefit from larger trials and the correspondingly more convincing statistics which would come with larger samples) and sensitivity to change (our inStability index clearly separated normal and dysplastic hips, but could again benefit from larger trials and

longer follow up). Lastly, we have also illustrated preliminary feasibility; our 3D scans are done in the same time frame as 2D scans, and in most cases are easier to perform. The post processing which we perform on the 3D data sets is minimal, taking about 2-3 minutes of user interaction per scan, which fits well within the 30 minute standard examination and processing time frame at the tertiary hospital from which we collected data from. Our feasibility is also increased by the prospect of remote reporting by radiologists (as mentioned earlier). To further increase the real-world feasibility of this analysis, work with a commercial ultrasound manufacturer could be done to incorporate our indices and measurement capabilities; once this is done we could begin scanning in multiple centers with relative ease.

#### 6.1.5 LIMITATIONS

As we are the first group to attempt to use 3D ultrasound on infant hips, we did not know in advance the best technique for scanning. This has resulted in some of our initial 3D scans being of poorer quality for our ultimate analysis than others, despite this, new sonographers have generally been able to acquire scans of adequate quality on their first attempt with minimal coaching, which leads us to believe that this is indeed easier than the 2D Graf method. To date, 7 individuals have been trained to perform routine scans for our studies.

Our available (13VL5; Philips Healthcare, Andover, Mass) probe, although high resolution, was sometimes lacking in penetration (section 1.1.4). Our most common penetration issue was a lack of a visible ischium. This generally occurred in older and larger patients; consequently it may be of use to use a lower frequency probe for some patients. The ischium is generally a landmark used for the standard 2D coronal Graf plane, but we have no reason to question the utility of still using it to obtain images of the hip in this general reference plane. Additionally, it was sometimes difficult to image very non-compliant infants and motion artifacts hindered the utility of some of our 3D scans. Acquisition of full sweeps of the hip took slightly less than 3 seconds; and while this is fast, it is still susceptible to motion artifacts.

Motion artifacts will soon be overcome with the increasing sophistication and usability of 4D ultrasound probes; these probes (section 1.1.5) will also increase the ease of use for technicians.

#### 6.1.5 FUTURE DIRECTIONS

This thesis introduces several innovations to imaging diagnosis of hip dysplasia, each of which could be incorporated clinically as follows:

The acetabular radius of curvature (AROC) and the arc length of the acetabulum which were introduced in chapter 3 would be easy to incorporate into clinical exams as much of the functionality is easy to program and build into existing PC programs for radiologist use. With others [3] also attempting to improve on Graf's current technique, we can be certain that new indices and knowledge of the abnormal development of the infant hip is needed. Potential benefits of working on the AROC method to quantify the amount of acetabular rounding, other than the obvious quantification on Graf standard images, is that it can be incorporated into three dimensions.

A three dimensional version of the AROC may potentially capture and quantify an index which reflects rounding. This 3D AROC naturally lends itself to comparison with the simplified surface models like those in this thesis, which would be incredibly useful in the creation of an infant hip atlas. This atlas would create the 'typical' normal hip (based on hundreds of surface models) and subsequent scans could be compared against this normal to find aspects of dysplasia. It is also unlikely that there would be one single normal hip for all ages; the atlas would be generated based upon trends observed from the models which are acquired and would likely have an age group dependent typical normal hip (as infant hips change quickly over time). Not only would this atlas be beneficial in diagnosing dysplastic hips quickly and accurately, but it could be beneficial in terms of surgical planning. Surgeons could see the exact locations of where a specific infant's hip differs from the normal and plan their treatments accordingly.

The inStability index discussed in chapter 5 is only one technique based on 3D ultrasound; there is room for many other quantification techniques to be explored. The inStability index is similar to alpha angle, and although we believe it is stronger and incorporates aspects of rounding, there is the same potential to use secondary indices analogous to indices like coverage on 2D scans. Additionally, this is an easy index to incorporate secondary indices onto as there is already a surface model created in the methodology to calculate the inStability index.

Another beneficial avenue to explore to further understand the stability of the hip joint is finite element modeling. Finite element modeling of the hip is already a thriving area of research with many groups publishing papers [4-10]; however, there are fewer studies on the pediatric hip which is structurally much different, especially in young infants. Using 3D ultrasound to generate the finite elements would create high fidelity models with sub millimeter precision that represent the hip very accurately, and could be used for increasing the knowledge of infant hip dysplasia.

## 6.2 CONCLUSIONS

The studies in this thesis show that 3D ultrasound is of promising utility in the diagnosis of DDH. There are significant flaws in the current Graf standard techniques which are widely used currently (high inter/intra- observer variability and high inter-scan variability – that can cause a change in diagnostic category in half of all scans), that are alleviated with the use of 3D ultrasound. The acetabular radius of curvature introduced in this thesis attempts to increase the reliability of the current 2D ultrasound method by quantifying the rounding which clinicians already incorporate into their impression of the hip. Standardization of 3D ultrasound scans of the hip may be done through the use of our landmarks which were proven reliable (identifiable within 2 mm), and the inStability angle shows great initial utility with more accurate ROC curves than the alpha angle. Work is still needed to be done in the form of larger trials,

exploration of different 3D indices, and finite element modeling, but this thesis provides the rough framework for 3D ultrasound to be incorporated into the diagnosis of infant hip dysplasia.

### 6.3 REFERENCES

- [1] A. Falliner, H. J. Hahne, J. Hedderich, J. Brossmann and J. Hassenpflug. Comparable ultrasound measurements of ten anatomical specimens of infant hip joints by the methods of graf and terjesen. *Acta Radiol.* (2), pp. 227. 2004.
- [2] M. Boers, J. R. Kirwan, G. Wells, D. Beaton, L. Gossec, M. A. d'Agostino, P. G. Conaghan, C. O. Bingham 3rd, P. Brooks, R. Landewe, L. March, L. S. Simon, J. A. Singh, V. Strand and P. Tugwell. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J. Clin. Epidemiol.* 67(7), pp. 745-753. 2014.
- [3] O. Y. Yavuz, I. Uras, B. A. Tasbas, M. H. Ozdemir, M. Kaya and M. Komurcu. A new measurement method in graf technique: Prediction of future acetabular development is possible in physiologically immature hips. *J. Pediatr. Orthop.* 2014.
- [4] C. L. Abraham, S. A. Maas, J. A. Weiss, B. J. Ellis, C. L. Peters and A. E. Anderson. A new discrete element analysis method for predicting hip joint contact stresses. *J. Biomech.* 46(6), pp. 1121-1127. 2013.
- [5] A. A. Ali, L. Cristofolini, E. Schileo, H. Hu, F. Taddei, R. H. Kim, P. J. Rullkoetter and P. J. Laz. Specimen-specific modeling of hip fracture pattern and repair. *J. Biomech.* 47(2), pp. 536-543. 2014.
- [6] G. Chen, L. Yang, K. Li, R. He, B. Yang, Y. Zhan, Z. Wang, B. Yu and Z. Jian. A three-dimensional finite element model for biomechanical analysis of the hip. *Cell Biochem. Biophys.* 67(2), pp. 803-808. 2013.
- [7] S. P. George and G. Saravana Kumar. Patient specific parametric geometric modelling and finite element analysis of cementless hip prosthesis. *Virtual & Physical Prototyping* 8(1), pp. 65-83. 2013.
- [8] C. R. Henak, G. A. Ateshian and J. A. Weiss. Finite element prediction of transchondral stress and strain in the human hip. *J. Biomech. Eng.* 136(2), pp. 1-11. 2014.
- [9] C. Henak, A. Kapron, A. Anderson, B. Ellis, S. Maas and J. Weiss. Specimen-specific predictions of contact stress under physiological loading in the human hip: Validation and sensitivity studies. *Biomechanics & Modeling in Mechanobiology* 13(2), pp. 387-400. 2014.
- [10] Y. Luo, Z. Ferdous and W. D. Leslie. Precision study of DXA-based patient-specific finite element modeling for assessing hip fracture risk. *International Journal for Numerical Methods in Biomedical Engineering* 29(5), pp. 615-629. 2013.

# CHAPTER 7

## Chapter 7

### COMPLETE BIBLIOGRAPHY

- [1] C. L. Abraham, S. A. Maas, J. A. Weiss, B. J. Ellis, C. L. Peters and A. E. Anderson. A new discrete element analysis method for predicting hip joint contact stresses. *J. Biomech.* 46(6), pp. 1121-1127. 2013.
- [2] A. Abuhamad. Automated multiplanar imaging: A novel approach to ultrasonography. *J. Ultrasound Med.* 23(5), pp. 573-576. 2004.
- [3] M. Akiyama, Y. Nakashima, M. Oishi, T. Sato, M. Hirata, D. Hara and Y. Iwamoto. Risk factors for acetabular retroversion in developmental dysplasia of the hip: Does the pemberton osteotomy contribute? *Journal of Orthopaedic Science* 19(1), pp. 90-96. 2014.
- [4] A. A. Ali, L. Cristofolini, E. Schileo, H. Hu, F. Taddei, R. H. Kim, P. J. Rullkoetter and P. J. Laz. Specimen-specific modeling of hip fracture pattern and repair. *J. Biomech.* 47(2), pp. 536-543. 2014.
- [5] American Institute of Ultrasound in Medicine. AIUM practice guideline for the performance of an ultrasound examination for detection and assessment of developmental dysplasia of the hip. *J. Ultrasound Med.* 32(7), pp. 1307-1317. 2013.
- [6] C. E. Bache, J. Clegg and M. Herron. Risk factors for developmental dysplasia of the hip: Ultrasonographic findings in the neonatal period. *J. Pediatr. Orthop. B* 11(3), pp. 212-218. 2002.
- [7] E. Bar-On, S. Meyer, G. Harari and S. Porat. Ultrasonography of the hip in developmental hip dysplasia. *J. Bone Joint Surg. Br.* 80(2), pp. 321-324. 1998.
- [8] V. Bialik, G. M. Bialik, S. Blazer, P. Sujov, F. Wiener and M. Berant. Developmental dysplasia of the hip: A new approach to incidence. *Pediatrics* 103(1), pp. 93-99. 1999.
- [9] K. Bin, J. M. Laville and F. Salmeron. Developmental dysplasia of the hip in neonates: Evolution of acetabular dysplasia after hip stabilization by brief pavlik harness treatment. *Orthop. Traumatol. Surg. Res.* 100(4), pp. 357-361. 2014.
- [10] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476), pp. 307-310. 1986.

- [11] M. Boers, J. R. Kirwan, G. Wells, D. Beaton, L. Gossec, M. A. d'Agostino, P. G. Conaghan, C. O. Bingham 3rd, P. Brooks, R. Landewe, L. March, L. S. Simon, J. A. Singh, V. Strand and P. Tugwell. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J. Clin. Epidemiol.* 67(7), pp. 745-753. 2014.
- [12] J. T. Bushberg. *The Essential Physics of Medical Imaging / Jerrold T. Bushberg ... [Et Al.]* 2012[University of Alberta JW Scott Health Sciences RC 78.7 D53 E78 2012].
- [13] S. Carley, S. Dosman, S. R. Jones and M. Harrison. Simple nomograms to calculate sample size in diagnostic studies. *Emerg. Med. J.* 22(3), pp. 180-181. 2005.
- [14] J. P. Cashman, J. Round, G. Taylor and N. M. Clarke. The natural history of developmental dysplasia of the hip after early supervised treatment in the pavlik harness. A prospective, longitudinal follow-up. *J. Bone Joint Surg. Br.* 84(3), pp. 418-425. 2002.
- [15] G. Chen, L. Yang, K. Li, R. He, B. Yang, Y. Zhan, Z. Wang, B. Yu and Z. Jian. A three-dimensional finite element model for biomechanical analysis of the hip. *Cell Biochem. Biophys.* 67(2), pp. 803-808. 2013.
- [16] C. Dezateux and K. Rosendahl. Developmental dysplasia of the hip. *Lancet* 369(9572), pp. 1541-1552. 2007.
- [17] C. Erturk, M. A. Altay and U. E. Isikan. A radiological comparison of salter and pemberton osteotomies to improve acetabular deformations in developmental dysplasia of the hip. *Journal of Pediatric Orthopaedics-Part b* 22(6), pp. 527-532. 2013.
- [18] A. Falliner, H. J. Hahne, J. Hedderich, J. Brossmann and J. Hassenpflug. Comparable ultrasound measurements of ten anatomical specimens of infant hip joints by the methods of graf and terjesen. *Acta Radiol.* (2), pp. 227. 2004.
- [19] A. Falliner, D. Schwinzer, H. J. Hahne, J. Hedderich and J. Hassenpflug. Comparing ultrasound measurements of neonatal hips using the methods of graf and terjesen. *J. Bone Joint Surg. Br.* 88(1), pp. 104-106. 2006.
- [20] A. Fenster and D. B. Downey. 3-D ultrasound imaging: A review. *Ieee Engineering in Medicine and Biology Magazine* 15(6), pp. 41-51. 1996.
- [21] P. R. Fleissner Jr, C. J. Ciccarelli, R. E. Eilert, F. M. Chang and G. L. Glancy. The success of closed reduction in the treatment of complex developmental dislocation of the hip. *J. Pediatr. Orthop.* 14(5), pp. 631-635. 1994.
- [22] M. Fu, S. Xiang, Z. Zhang, G. Huang, J. Liu, X. Duan, Z. Yang, P. Wu and W. Liao. The biomechanical differences of rotational acetabular osteotomy, chiari osteotomy and shelf procedure in developmental dysplasia of hip. *BMC Musculoskelet. Disord.* 15pp. 47-2474-15-47. 2014.
- [23] O. Furnes, S. A. Lie, B. Espehaug, S. E. Vollset, L. B. Engesaeter and L. I. Havelin. Hip disease and the prognosis of total hip replacements: A review of 53 698 primary total hip replacements reported to the norwegian arthroplasty register 1987-99. *J Bone Joint Surg (Br)* 83B(4), pp. 579-586. 2001.

- [24] S. P. George and G. Saravana Kumar. Patient specific parametric geometric modelling and finite element analysis of cementless hip prosthesis. *Virtual & Physical Prototyping* 8(1), pp. 65-83. 2013.
- [25] E. O. Gerscovich, A. Greenspan, M. S. Cronan, L. A. Karol and J. P. McGahan. Three-dimensional sonographic evaluation of developmental dysplasia of the hip: Preliminary findings. *Radiology* 190(2), pp. 407-410. 1994.
- [26] R. Graf. Fundamentals of sonographic diagnosis of infant hip dysplasia. *J. Pediatr. Orthop.* 4(6), pp. 735-740. 1984.
- [27] R. Graf and K. Lercher. [Experiences with a 3-D ultrasound system in infant hip joints]. *Ultraschall Med.* 17(5), pp. 218-224. 1996.
- [28] R. Graf, M. Mohajer and F. Plattner. Hip sonography update. quality-management, catastrophes - tips and tricks. *Med. Ultrason.* 15(4), pp. 299-303. 2013.
- [29] D. P. Gwynne Jones, A. G. Vane, G. Coulter, P. Herbison and J. D. Dunbar. Ultrasound measurements in the management of unstable hips treated with the pavlik harness: Reliability and correlation with outcome. *J. Pediatr. Orthop.* 26(6), pp. 818-822. 2006.
- [30] H. T. Harcke and L. E. Grissom. Performing dynamic sonography of the infant hip. *AJR Am. J. Roentgenol.* 155(4), pp. 837-844. 1990.
- [31] C. R. Henak, G. A. Ateshian and J. A. Weiss. Finite element prediction of transchondral stress and strain in the human hip. *J. Biomech. Eng.* 136(2), pp. 1-11. 2014.
- [32] C. Henak, A. Kapron, A. Anderson, B. Ellis, S. Maas and J. Weiss. Specimen-specific predictions of contact stress under physiological loading in the human hip: Validation and sensitivity studies. *Biomechanics & Modeling in Mechanobiology* 13(2), pp. 387-400. 2014.
- [33] P. Hoskins, K. Martin and A. Thrush. *Diagnostic Ultrasound : Physics and Equipment* 2010[University of Alberta JW Scott Health Sciences RC 78.7 U4 D516 2010].
- [34] D. L. Hykes, W. R. Hedrick and D. E. Starchman. *Ultrasound Physics and Instrumentation / David L. Hykes, Wayne R. Hedrick, Dale E. Starchman* 1992[University of Alberta JW Scott Health Sciences RC 78.7 U4 H99 1992].
- [35] T. Jackson, J. Watson, J. LaReau and B. Domb. Periacetabular osteotomy and arthroscopic labral repair after failed hip arthroscopy due to iatrogenic aggravation of hip dysplasia. *Knee Surgery, Sports Traumatology, Arthroscopy* 22(4), pp. 911-914. 2014.
- [36] D. Kane, W. Grassi, R. Sturrock and P. V. Balint. A brief history of musculoskeletal ultrasound: 'from bats and ships to babies and hips'. *Rheumatology (Oxford)* 43(7), pp. 931-933. 2004.
- [37] K. Karadayi, R. Managuli and Y. Kim. Three-dimensional ultrasound: From acquisition to visualization and from algorithms to systems. *Ieee Reviews in Biomedical Engineering* 2pp. 23-39. 2009.

- [38] Y. Lefevre, J. M. Laville and F. Salmeron. Early short-term treatment of neonatal hip instability with the pavlik harness. *Rev. Chir. Orthop. Reparatrice Appar. Mot.* 93(2), pp. 150-156. 2007.
- [39] Y. Luo, Z. Ferdous and W. D. Leslie. Precision study of DXA-based patient-specific finite element modeling for assessing hip fracture risk. *International Journal for Numerical Methods in Biomedical Engineering* 29(5), pp. 615-629. 2013.
- [40] M. Mafalda Santos and G. Filipe. Treatment of congenital hip dislocation using pavlik's harness. long term results. *Rev. Chir. Orthop. Reparatrice Appar. Mot.* 83(1), pp. 41-50. 1997.
- [41] L. Mercier, T. Lango, F. Lindseth and D. L. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine and Biology* 31(4), pp. 449-471. 2005.
- [42] P. D. Mitchell and R. C. Redfern. The prevalence of dislocation in developmental dysplasia of the hip in britain over the past thousand years. *J. Pediatr. Orthop.* 27(8), pp. 890-892. 2007.
- [43] C. Morin, S. Zouaoui, A. Delvalle-Fayada, P. M. Delforge and H. Lecllet. Ultrasound assessment of the acetabulum in the infant hip. *Acta Orthop. Belg.* 65(3), pp. 261-265. 1999.
- [44] A. K. Mostert, N. J. Tulp and R. M. Castelein. Results of pavlik harness treatment for neonatal hip dislocation as related to graf's sonographic classification. *J. Pediatr. Orthop.* 20(3), pp. 306-310. 2000.
- [45] D. Olszewski C. and L. Karol A. The medial ludloff open reduction in developmental dysplasia of the hip before the age of walking. *Oper Techniq Orthop* 23(3), pp. 109-114. 2013.
- [46] M. B. Ozonoff. *Pediatric Orthopedic Radiology / M.B. Ozonoff* 1992[University of Alberta JW Scott Health Sciences RD 732.3 C48 O99 1992].
- [47] A. Pavlik. Stirrups as an aid in the treatment of congenital dysplasias of the hip in children. by arnold pavlik, 1950. *J. Pediatr. Orthop.* 9(2), pp. 157-159. 1989.
- [48] K. Perry I., R. Trousdale T. and R. Sierra J. Hip dysplasia in the young adult: An osteotomy solution. *Bone Joint j (Br)* 95-B(11), pp. 21-25. 2013.
- [49] D. H. Pretorius, N. Borok, M. S. Coffler and T. R. Nelson. Three-dimensional ultrasound in obstetrics and gynecology. *Radiol. Clin. North Am.* 39(3), pp. 499-+. 2001.
- [50] E. A. Roovers, M. M. Boere-Boonekamp, T. S. Geertsma, G. A. Zielhuis and A. H. Kerkhoff. Ultrasonographic screening for developmental dysplasia of the hip in infants. reproducibility of assessments made by radiographers. *J. Bone Joint Surg. Br.* 85(5), pp. 726-730. 2003.
- [51] A. Roposch, L. Q. Liu, F. Hefti, N. M. P. Clarke and J. H. Wedge. Standardized diagnostic criteria for developmental dysplasia of the hip in early infancy. *Clinical Orthopaedics and Related Research*] (12), pp. 3451. 2011.

- [52] R. B. Salter, G. Hansson and G. H. Thompson. Innominate osteotomy in the management of residual congenital subluxation of the hip in young adults. *Clin Orthop Relat Res* (182), pp. 53-68. 1984.
- [53] W. N. Sankar, C. R. Young, A. G. Lin, S. A. Crow, K. D. Baldwin and C. F. Moseley. Risk factors for failure after open reduction for DDH: A matched cohort analysis. *J. Pediatr. Orthop.* 31(3), pp. 232-239. 2011.
- [54] E. SEVERIN. Congenital dislocation of the hip; development of the joint after closed reduction. *J. Bone Joint Surg. Am.* 32-A(3), pp. 507-518. 1950.
- [55] D. Shorter, T. Hong and D. A. Osborn. Cochrane review: Screening programmes for developmental dysplasia of the hip in newborn infants. *Evid Based. Child. Health.* 8(1), pp. 11-54. 2013.
- [56] E. A. Simon, F. Saur, M. Buerge, R. Glaab, M. Roos and G. Kohler. Inter-observer agreement of ultrasonographic measurement of alpha and beta angles and the final type classification based on the graf method. *Swiss Med. Wkly.* 134(45-46), pp. 671-677. 2004.
- [57] C. Sohn, G. P. Lenz and M. Thies. 3-dimensional ultrasound image of the infant hip. *Ultraschall Med.* 11(6), pp. 302-305. 1990.
- [58] S. K. Storer and D. L. Skaggs. Developmental dysplasia of the hip. *Am. Fam. Physician* 74(8), pp. 1310-1316. 2006.
- [59] U. von Jan, H. M. Overhoff and D. Lazovic. 3-D visualization of the newborn's hip joint using ultrasound and automatic image segmentation. *Stud. Health Technol. Inform.* 77pp. 1170-1174. 2000.
- [60] K. K. White, D. J. Sucato, S. Agrawal and R. Browne. Ultrasonographic findings in hips with a positive ortolani sign and their relationship to pavlik harness failure. *J. Bone Joint Surg. Am.* 92(1), pp. 113-120. 2010.
- [61] A. B. Wolbarst. *Physics of Radiology / Anthony Brinton Wolbarst ; with Illustrations by Gordon Cook* 2005[Cross Cancer Institute RC 78.7 D53 W65 2005].
- [62] O. Y. Yavuz, I. Uras, B. A. Tasbas, M. H. Ozdemir, M. Kaya and M. Komurcu. A new measurement method in graf technique: Prediction of future acetabular development is possible in physiologically immature hips. *J. Pediatr. Orthop.* 2014.
- [63] M. Zieger. Ultrasound of the infant hip. part 2. validity of the method. *Pediatr. Radiol.* 16(6), pp. 488-492. 1986.