4-2011

The Effect of Training on Novice Raters When Performing Radiographic Measurement of Humeral Retroversion: a Follow-up Study

Ryan Christensen
St. Catherine University

Danielle Grambo
St. Catherine University

Erin Ingram
St. Catherine University

Lyna Menezes
St. Catherine University

THE EFFECT OF TRAINING ON NOVICE RATERS WHEN PERFORMING RADIOGRAPHIC MEASUREMENT OF HUMERAL RETROVERSION: A FOLLOW-UP STUDY

by
Ryan Christensen
Danielle Grambo
Erin Ingram
Lyna Menezes

Doctor of Physical Therapy Program
St. Catherine University

April 28, 2011

Research Advisor: Cort J. Cieminski, PT, PhD, ATR

Abstract

Background and Purpose: Humeral retroversion angle (HRVA) is the angle between the proximal humeral axis and the distal humeral axis. HRVA is significantly higher in the dominant shoulder for individuals that partake in increased overhead activities. CT scans are recognized as the gold standard for measuring HRVA, although current literature is investigating the use of radiographs to measure HRVA as a more cost-effective method. The purpose of this study was to determine if there is a training effect in the ability to measure HRVA on radiographs as determined by physical therapy students.

Methods: Four second-year Doctor of Physical Therapy students measured HRVA on 35 shoulder radiographs on two separate occasions. Two of the students received training (trained group) from an experienced HRVA investigator which included written and verbal instruction on how to obtain HRVA measurements. Training also included HRVA measurement practice sessions with feedback from the experienced HRVA investigator prior to data collection. The other two students (untrained group) were provided only written instructions on how to obtain HRVA measurements. The experienced HRVA investigator served as the gold standard.

Results: There was minimal difference in the HRVA measurements between the trained and untrained groups with the individuals in both groups demonstrating excellent intra-rater reliability. The intra-rater reliability was 0.89 and 0.87 for the trained group members and 0.89 and 0.89 for the untrained group members, respectively. In addition, the inter-rater reliability between the trained and untrained groups in comparison to the experienced clinician was also minimal. The combined inter-rater reliability for the trained group was 0.82 and untrained group was 0.80.

Conclusion: Novice physical therapy students can accurately measure HRVA on radiographs with written instruction and no other formal training. This demonstrates the potential to further incorporate radiographic information into clinical practice. This supports the APTA's Vision 2020 for physical therapists to become autonomous practitioners and expand the physical therapy scope of practice.

Acknowledgements

We would like to extend a special thanks to our research advisor Cort Cieminski for his guidance and patience throughout this project. We appreciate your assistance with both the research process and our professional growth.

To the St. Catherine University Doctor of Physical Therapy faculty, we thank them for all the time and effort they have given wholeheartedly over the past three years. They have instilled a commitment to life long learning in all of us.

Thank you to our classmates for the support and collaboration over the course of our graduate studies.

Finally, we thank our friends and families for their encouragement and understanding during the past three years.

The undersigned certify that they have read, and recommend approval of the research project entitled…

THE EFFECT OF TRAINING ON NOVICE RATERS WHEN PERFORMING RADIOGRAPHIC MEASUREMENT OF HUMERAL RETROVERSION: A FOLLOW-UP STUDY

submitted by
Ryan Christensen
Danielle Grambo
Erin Ingram
Lyna Menezes

In partial fulfillment of the requirements for the Doctor of Physical Therapy Program

Primary Advisor___Cort J. Cieminski___________________________ Date___4/28/2011___

Table of Contents

Chapter I

Introduction

Humeral retroversion (HRV) is defined by Krahl[1] as "a difference in the relative position of the long axes of the proximal and distal ends of the bone".[1] The humeral retroversion angle (HRVA) is the angle between the proximal and distal humeral axes. Humeral retroversion is formed when rotation of the humerus occurs around the proximal epiphyseal plate. Krahl[1] determined that most retroversion formation occurs from the twentieth week until birth and the remainder continues to form until age 20. Humeral head retroversion angles also differ among various ethnic groups.[2]

A study conducted by Edelson[3] showed that although HRV is present in all ethnic groups the rate of derotation differs across ethnic groups. These findings are an indication that it is important for HRV to be assessed individually during surgical procedures in order to better serve the patient population. Clinically, the ability to accurately measure the HRVA is relevant to orthopedic cases.[3] For example, in a patient undergoing a total shoulder athroplasty, the HRVA allows the new joint components to mimic the original orientation of the humerus. Additionally, the process of developing a greater degree of HRV can decrease the potential risk of anterior dislocations in patients with a greater HRVA due to the increased time it would take for the humerus to glide anteriorly during external rotation (ER).

In overhead throwing athletes, HRV has been shown to be significantly increased in the dominant shoulder compared to the non-throwing shoulder. Crockett et al[4] looked at 25 male professional baseball pitchers who had started throwing overhead before the age of 10. The study found 40º of HRV present in the dominant shoulder of the throwing group compared to 18º of HRV in the dominant shoulder of the non-throwing group. In addition, Crockett et al[4] discovered 40º of HRV in the dominant shoulder of the throwing group compared to 23º of HRV in the non-dominant shoulder of the throwing group. These results confirmed that overhead throwing athletes have a higher HRVA angle in their dominant arm.

The HRVA gets smaller in the process of de-rotation up until the proximal epiphysis fuses. After the fusion of the proximal epiphysis, the HRVA remains static and is not further influenced by the exposure to overhead throwing. However, up until this closure of the growth plate, HRVA is strongly influenced by the amount of overhead throwing exposure as this repeated exposure slows down the rate of de-rotation. Therefore, overhead throwing athletes have a greater HRVA compared to individuals who do not have overhead throwing exposure.[1,5]

Osbahr et al[5] discussed possible benefits of throwing athletes having an increased HRVA. A greater HRVA allows for increased external rotation (ER) at the glenohumeral joint. Increased ER adds energy within the kinetic chain due to an increase in torque therefore allowing a greater throwing velocity to be generated. Secondly, with more HRV, the shoulder joint may be better equipped to withstand anterior forces. Soft tissue structures located anterior have less stress applied for a given amount of ER which allows the soft tissue structures to stay within their elastic range and better stabilize the glenohumeral joint. Osbahr et al[5] concluded that rotational changes in the throwing shoulder are due to bony as well as soft tissue adaptations in the shoulder.

Many studies have been completed in order to better understand the assessment of humeral head retroversion and the accuracy measuring humeral head retroversion angle using different imaging techniques such as radiography and computed tomography. Both Soderlund et al[6] and Oztuna et al[7] established the use of radiography to image the HRVA. Soderlund et al[6] compared the use of radiograph versus computed tomography when measuring the HRVA.

Upon comparison, it was found that 2° was the maximum angle measurement difference between the two different imaging procedures. The mean difference between the imaging measurements was 1.5°. In support, Oztuna et al[7] compared the radiographic imaging of the dry humeri HRVA with the anatomic HRVA of the dry humeri. The average difference between the radiograph and the anatomic measurement was 0.9°. Therefore, a single exposure radiograph is sufficient to measure the HRVA with high accuracy.

Evidence supports three identifiable sources of variability in radiograph measurements including patient, procedure, and observer. The study examined the use of both trained and untrained observers. Currently, physical therapists, along with other professionals not specifically trained in the field of radiology, are not allowed to order, read, interpret, or use radiographs as a method for measuring anatomical structures or angles. Studies examining the variations and reliability between trained and inexperienced observers, though limited, do exist. If a standard technique is implemented, clinical experience is not indicative of skill in rating radiographs as demonstrated by Owen et al.[8] Excellent intra-observer reproducibility was found for both novice and intermediate observes when measuring scoliosis as evaluated by Dang et al.[9] Additionally, Margolis et al[10] determined there was no advancement of radiographic interpretation skills following additional years of formal education.

Vision 2020[11] emerged from the American Physical Therapy Association as an effort to further the physical therapy profession through the practice of life-long learning, evidence-based practice, and integrity. Goals of Vision 2020[11] include becoming practitioners of choice, direct access for patients, and building an autonomous practice. Gaining the ability to order radiographs and interpret relevant information would support Vision 2020 and help provide patient services throughout the continuum of care.

The purpose, therefore, of this study is to establish the accuracy of HRVA measurements on a radiograph of untrained raters in comparison to trained raters and an expert rater.

The primary null hypothesis of the research study, given the information presented in the literature, was that no difference exists between untrained and trained raters when measuring HRV angles on radiographs. The secondary null hypothesis was that there is no difference in the intra-rater reliabilities between individual raters. The alternative hypotheses were that the trained group would achieve a greater level of inter-rater reliability compared to the untrained group and that the experienced rater will have a greater intra-rater reliability than the untrained raters.

Chapter II

Review of Related Literature

Etiology of Humeral Retroversion

A number of studies have been conducted in order to better understand various aspects of the etiology of humeral retroversion (HRV). Research studies differ in their findings about the cause and exact location on the humerus at which HRV occurs. A good knowledge base of the etiology of HRV allows for accurate measurements of the HRVA on CT scans and radiographs. Accurate HRVA measurements are clinically significant especially in shoulder joint replacement surgeries which require the new ball and socket to mimic the original orientation of the humerus. HRVA's typically decrease from adolescence to adulthood but the rate at which it decreases differs in certain populations and across ethnic groups. This section aims to review studies that look at different characteristics of HRV.

Krahl[1] conducted an early study in 1947 to determine the site, cause and duration of humeral retroversion. Krahl[1] defines torsion as "a difference in the relative position of the long axes of the proximal and distal ends of the bone." Humeral torsion was measured from the humeri of 42 cadavers using a torsiometer, an angle measurement apparatus. The torsional angle was measured by drawing a line A perpendicular to line B which bisected the long axis of the humeral head followed by line C which bisected the medial and lateral epicondyles. The angle formed when lines A and C met was measured as the torsional angle and the torsional angle value was complementary to 90^{o}.[1]

A common debate at the time was whether torsion occurred at the proximal or distal epiphysis and evidence from Krahl's[1] study proposed that it does not occur at the distal epiphysis. The spiral groove winds distally from dorsal to lateral around the bone making it appear as though the shaft has been twisted resulting in the movement of the distal end in a medial direction. From this, researchers assumed the path of the spiral groove provides evidence that torsion occurs at the distal epiphysis. This was disproved due to the fact that the spiral groove is not evident during the initiation of torsion as one would expect. Additional evidence shows that ossification of the diaphysis does not disrupt torsion which continues till maturity. Evidence shows that humeral torsion occurs at the proximal epiphysis. Torsion is completed at the age of 20 which is at approximately the same time the proximal epiphysis fuses indicating that torsion occurs at the proximal epiphysis.[1]

Furthermore, in order to measure the torsional angles from birth to 20 years the torsional angle was measured in 374 humeri, 21 were from fetuses and new born infants and 40 were from birth to 24 years and 313 ranged from 25 to 91 years of age. It was determined that from the twentieth week until birth torsional angle increased from 42^o to 60^o. Torsion then continues till the twentieth year and reaches an angle of 74^o. These results indicated that there is an increase in humeral torsion during development and childhood and that most of the torsion has taken place by the time of the birth and the rest of it completed by the twentieth year.[1]

Edelson[2] conducted a study in 1999 to determine the variations of humeral head retroversion in different ethnic groups. The study consisted of 363 bone specimens of humeral heads including different ethnic groups of white Americans, African American, Native Americans from New Mexico, northern Chinese, Alaskan Eskimos and Negev Desert Bedouins.[2]

In order to determine the HRV angle, a line was drawn by pencil to bisect the oblique surface of the humeral head in a superior inferior direction. The anterior surface of the condyles of the elbow was placed against a flat steel panel and the humeral head was then leaned back against the panel. Finally a standard goniometer attached to the panel was opened along the penciled line in order to measure the HRV angle.[2]

Results indicated that in general across the ethnic groups the right humeral head was more retroverted than the left humeral head and the retroversion angles were slightly greater in males than females. Edelson[2] stated that these differences were not statistically significant. Ethnic groups including the Chinese, Alaskan Eskimo and Native Americans had greater retroversion angles compared to the rest of the ethnic groups. For example, the mean retroversion angle for northern Chinese specimens was 46.4 °on the right and 42.7° degrees on the left. White American, African Americans and Bedouin specimens were similar in their retroversion parameters while Northern Chinese, Alaskan Eskimo and Native Americans had similar retroversion parameters.[2]

This study conducted by Edelson[2] indicates that although significant retroversion is present across all ethnic groups at birth the progression of de-rotation up until adulthood differs amongst various ethnic groups. It is generally accepted that retroversion angles are between 25° and 35° in adults. These degrees are expressed in terms of supplementary angles. Due to the difference of humeral retroversion among various ethnic groups the patient population will be better served during a surgical procedure if humeral retroversion is looked at separately for each individual.[2]

In addition to determining the variation in humeral retroversion in different ethnic groups Edelson[3] conducted a study in 2000 to determine the development of humeral head retroversion. It is clinically significant to know the degree of retroversion of the humeral head as it has to be anatomically reproduced during a shoulder replacement surgery. Edelson[3] studied 180 humeri of people aged 4 months to 19 years of age. The methods were the same as mentioned in the previous Edelson[2] study. Results indicated that humeral head retroversion was approximately 65 degrees bilaterally in children aged 4 months to 4 years. On average at eight years of age children tended to de-rotate into the adult parameters of retroversion but all children past 11 years of age had reached adult retroversion values. Additional retroversion continues to take place past age 11 and ceases between the ages of 16 to 20 which is in conjunction with the values of humeral retroversion determined in Krahls'[1] study.[3]

Edelson[3] believed that the radial groove represented retroversion cessation. The absence of a radial groove at age 11 and its presence at age 16 led Edelson[3] to conclude that humeral retroversion ceases by age 16. Krahl[1] in his study previously disproved this theory by stating that the absence of the formation of a radial groove during the initiation of retroversion as expected suggests that the radial groove is not an indication of the occurrence of humeral retroversion.[3]

A study conducted by Sabick et al[12] focused on better understanding the contribution of overhead throwing motion in youth baseball pitchers to the development of proximal humeral epiphysiolysis and HRV. Proximal humeral epiphysiolysis refers to a localized pain in the proximal humerus during overhead throwing motions. The two major causes for epiphysiolysis are suggested to be distraction and rotational stresses applied to the proximal humeral epiphysis during overhead throwing motions prior to epiphyseal fusion. It is also suggested that adult throwing athletes tend to have more HRV due to muscle forces acting on the humerus during over head throwing.[12]

The primary aims of these studies were to; 1) quantify the forces and torques acting on the humerus during pitching motions; 2) to further understand the cause of humeral epiphysiolysis with biomechanical evidence; 3) to further assess the mechanism of the development of HRV. There were 14 youth baseball pitchers around the age of 14 that were included in the study. They were filmed from the front and dominant side while throwing 10 fast balls with maximal effort in a simulated game. During the throwing motion the net force and torque acting on the humerus were calculated using standard biomedical techniques. Results showed that the external rotation torque reached its peak value of approximately 17.7 N·m right before maximum shoulder external rotation. This high torque during maximal external rotation in overhead throwing causes shear stress that leads to distortion of the proximal humerus that could lead to a development of greater HRV angles and humeral epiphysiolysis over time. It was also found that the distraction forces during over head throwing contribute less to the stresses applied to the proximal humerus in comparison to the stresses applied by the external rotation torque.[12]

Sabick et al[12] concluded that overhead throwing motions in baseball pitchers contribute to two clinical entities, HRV and humeral epiphysiolysis. HRV is considered to be a possible beneficial adaptation as it allows for more external rotation before the anterior capsule and ligaments are activated to restrict the motion. This increased ER allows more time during the arm- cocking phase of the throw (stride foot contact and maximum shoulder external rotation) that provides a greater angle to accelerate the arm and ball. On the other hand, humeral epiphysiolysis is not necessarily a beneficial adaptation but a pathological response to the shear forces placed on the proximal humerus during over head throwing.[12]

Cowgill[13] conducted a study where he proposed developing a model that assessed the populational differences of humeral torsional in adults from a combined ontogenetic and functional perspective. The specific research questions that Cowgill[13] addressed in his study included the existence of a pattern variation in humeral torsion during growth, the age at which humeral torsion is attained, the existence of population differences and gender differences of torsion patterns during growth and if so the time at which they emerge and finally the variation in bilateral asymmetry of torsion with age. Cowgill[13] assessed to see if the population differences in humeral torsion are due to differences in habitual patterns of upper limb use in the different populations due to research studies suggesting that sports participation for example causes variations in humeral torsion.

Six skeletal collections consisting of 407 individuals whose ages ranged from birth to 17.9 years were used in this study. Humeral torsion was also measured in a subgroup of 38 individuals ranging from 18 to 30 years to assess the level of torsion in each sample. HRVA was measured in this study by the measurement protocol set by Rhodes (2006).[14] Also, to assess habitual differences in upper limb use the six skeletal samples were chosen to represent a diverse presentation of life style and activity patterns.[12] Four of the six samples were from non-mechanized and non urban societies (Mistihalj, Indian Knoll, Point Hope and California Amerindian). In contrast the Dart Collection is an ethnically mixed urban and non urban sample where as the Louis Lopes collection consists of urban 20th century Portuguese. Gender information between the latter two samples was also present. Ages of 5 of the six skeletal samples were obtained from lateral mandibular radiographs and the ages from the Louis Lopes samples were determined from civil registrations of birth and death. Statistics including ANOVA and Kruskal-Wallis were used to answer the research questions proposed in the study. [13]

The study concluded that humeral torsion decreases by 25° from birth to adulthood but an exact age at which humeral torsion occurs could not be obtained from this study. As far as population differences and habitual patterns are concerned, certain populations did not meet the predictions of humeral torsion angles based on their activity levels. Levels of torsion in general were expected to be higher in the groups that took part in more strenuous activity and lower in the groups that were less active and more urban. The Mistihalj sample did not fit this prediction, as it is a non-mechanized medieval agricultural population yet the individuals in this group had torsional levels that were even lower than the more urban samples.[13]

While habitual differences in populations can contribute to the functional component of humeral torsion, it is suggested that there may be a genetic component that causes population variance of torsional angles. This is due to the fact that torsional differences between the populations can be seen at a very early age between birth and 2 years. There was also the presence of a connection of bilateral asymmetry in humeral torsion angles in adults and males in general were found to have higher torsional angles but the values were not statistically significant. Finally, Cowgill[13] also concluded that while the functional differences in torsional angles cannot necessarily confidently be attributed to repetitive throwing activities such as Sabick's[12] study suggested, it could still be attributed to a combination of activity patterns which result in a functional imbalance between medial and lateral rotators.[13]

Humeral Retroversion in Overhead Throwing Athletes

Early studies suggested that an increase in humeral retroversion (HRV) was due to soft tissue changes occurring around the proximal humerus and shoulder, but newer studies specifically examining overhead throwing athletes have suggested that HRV and range of motion (ROM) changes can be due to a combination of soft tissue changes and bony adaptations.

A study by Crockett et al[4] utilized 25 male professional baseball pitchers who all had started throwing overhead before the age of 10 to a group of 25 male control-matched subjects who were never involved in any overhead throwing sports. The goal of the study was to determine if there were significant osseous differences between the two groups and between bilateral shoulders and whether a potential difference was a factor in the motion seen in elite overhead throwing athletes.[4]

An independent *t*-test was used to compare ROM and retroversion between the elite overhead throwing group and the control group. Paired *t*-tests were used to compare the differences in glenohumeral ER and IR, capsular instability, and HRV for the dominant versus non-dominant shoulders of each subject.[4]

A computed tomography (CT) scan, which followed the protocol described by Hill et al[15], was utilized to determine the amount of humeral retroversion present for all subjects. Humeral retroversion was measured by determining the difference between a line drawn parallel to the distal articular surface and a line that bisected a spherical section of the head. External rotation and IR were measured using standard goniometry with the shoulder at 90° of abduction. Internal rotation was measured without stabilizing the scapula. Anterior and posterior glenohumeral laxity was determined according to the system proposed by Hawkins and Bokor.[16]

Results from the study showed 40° of HRV in the dominant shoulder of the throwing group compared to 18° of HRV in the dominant shoulder of the non-throwing group. The study also showed 40° of HRV in the dominant shoulder of the throwing group compared to 23° of HRV in the non-dominant shoulder of the throwing group.[4]

When comparing joint laxity between the dominant and non-dominant shoulder of throwing athletes, there was no significant difference in anterior glenohumeral laxity with the dominant shoulder of throwers having a grade of 1.36 versus a grade of 1.24 in the non-dominant shoulder of throwers. For posterior glenohumeral laxity, no significant difference was shown with the dominant shoulder of throwers having a grade 1.92 compared to 1.80 in the non-dominant shoulder of throwers.[4]

For ROM findings, IR at 90° of abduction of the dominant shoulder was 62° (±7.4°) compared to 71° (±9.3°) for the non-dominant throwing shoulder. For ER at 90° of abduction, the dominant shoulder measured 128° (±9.2°) compared to 119° (±7.2°) for the non-dominant shoulder. There was no significant difference in the total arc of rotational motion between the dominant and non-dominant shoulders of the throwing group.[4]

This study by Crockett et al[4] suggests that there was no significant difference in anterior and posterior capsular laxity between the dominant and non-dominant shoulders of the overhead-throwing group. A significant difference in HRV between the dominant and non-dominant shoulders of throwers was found, however. These findings suggest that the decrease in IR and increase in ER is due an osseous humeral change and is not due to increased anterior capsular laxity and posterior capsular tightness.[4]

A study performed by Osbahr et al[5] examined both shoulders of 19 male college baseball pitchers. The amount of HRV was calculated using the technique of Soderlund et al[6] that used a single radiograph to take a semiaxial view of the glenohumeral joint. Humeral retroversion was measured as the angle between the anatomic neck axis and the epicondylar axis. Measurements of passive glenohumeral ER at 0° to 90° of abduction and IR at 90° of abduction and were measured under a 3.5 kg load.[6]

An unpaired *t*-test was used to examine the differences between the dominant and non-dominant shoulder for HRV and ROM measurements. A Pearson's correlation coefficient was utilized to test the relationship between the ROM measurements, HRV angle, number of years pitched between the ages of 8 to 16 years of age, and the age of subjects ($P \leq 0.05$).[5]

The subjects had greater HRV in the dominant shoulder 33.2° (±11.4°) compared to the non-dominant shoulder 23.0° (±9.1°). The throwing athletes in this study also were found to have increased ER at 90° of abduction with 126.4° (±12.0°) in the dominant shoulder and 114.5° (±9.1°) in the non-dominant shoulder. They also found the throwing athletes to have decreased

IR with the dominant shoulder measuring an average of 79.3° (±13.3°) compared to 91.4° (±13.6°) in the non-dominant shoulder. The study demonstrated a relationship between HRV and ER in the dominant shoulder of pitchers in the study (P=0.0001, r=0.8639). However, no correlation was found between HRV and the number of years pitched between the ages of 8 to 16 years of age (P=0.2498, r=-0.2777).[5]

The study by Osbahr et al[5] concluded that rotational changes in the throwing shoulder are due to bony as well as soft tissue adaptations in the shoulder. Osbahr et al[5] suggested the two possible benefits of overhead throwing athletes having a greater of HRV angle in their throwing shoulder. First, the greater HRV allows for greater external rotation at the glenohumeral joint. This increased rotation potentially may add to the energy available within the kinetic chain due to an increase in torque. This increase will allow for a greater throwing velocity to be generated. Secondly, the glenohumeral joint may be more stable to anterior force when greater HRV is present. The anterior soft tissue structures have to stress less for a given amount of ER, which allows the soft tissue structures to stay within their elastic range and better stabilize the glenohumeral joint.[5]

A study performed by Kronberg et al[17] also suggested that overhead throwing athletes have greater ER and decreased IR in their throwing shoulder compared to their non-throwing shoulder. In the Kronberg et al[17] study, they found that in normal shoulders, greater HRV was typically seen with an increased range of ER at 90° of shoulder abduction but no difference was found between the subjects' dominant and non-dominant shoulders for each ROM tested.[17] The results of the Osbahr et al[5] study seemed to agree with these findings as it also showed greater HRV shows a statistical significance to increased ROM for ER at 90° of shoulder abduction. Both the Kronberg et al[17] and Osbahr et al[5] studies found a statistical difference in HRV when comparing the dominant and non-dominant shoulders, but the Osbahr et al[5] study also found a statistical difference between the subjects' dominant and on-dominant shoulders when tested for all ROM.[5,17]

The Osbahr et al[5] findings were also consistent with a study looking at HRV in handball players done by Pieper.[18] He noted that nearly every handball player he looked at had increased ER and decreased IR in their throwing shoulders. He looked specifically at 51 male professional handball players and 37 male subjects who had no history of being a throwing athlete. The handball players were found to have 9° more HRV on their dominant shoulder compared to their non-dominant shoulder. There was no significant difference found in HRV in the dominant and non-dominant shoulders of the non-throwing group.[18]

Thirteen of the 51 handball players had symptoms of chronic shoulder pain while the other 38 handball players had no shoulder problems. The 13 who had chronic shoulder pain had no significant difference in HRV between their dominant and non-dominant throwing shoulders. However, these subjects stated competitive throwing at a late age; therefore, they did not have enough exposure to lead to a greater HRVA. The 38 subjects who were asymptomatic for chronic shoulder pain were found to have a statistically significant increased HRV (14°) in their dominant shoulder compared to their non-dominant shoulder. This suggests that greater HRV from repetitive overhead throwing at an early age can serve as a protective mechanism against anterior instability.[18]

A study done by Reagan et al[19] looked at 54 asymptomatic college baseball players who had a mean age of 19.3 years. Twenty-five of the subjects were pitchers. The mean number of year of overhead throwing experience for the pitchers was 14 years (10 years to 17 years). Reagan et al[19] hypothesized that an increase in ER and a decrease in IR between the dominant and non-dominant shoulders of throwers would directly correlate with an increased HRV angle in the dominant shoulder.

External rotation and IR ROM were measured using standard goniometer measurements at 0° and 90° and with the scapula stabilized. Humeral retroversion was obtained using the methods described by Kronberg et al[17] and Soderlund et al[6] using radiographs. A subject number was utilized for identifying the radiographs to ensure proper blinding by two investigators.[6,17,19]

A paired *t*-test showed a statistically significant increase in HRV in the dominant shoulder compared to the non-dominant shoulder with the dominant shoulder measuring 36.6° (±9.8°) compared to 26.0° (±9.4°) for the non-dominant shoulder. A Pearson's product moment correlation was used to determine that the mean difference in HRV mean difference of 9.7° between the dominant and non-dominant shoulder in ER at 90° of abduction correlated significantly with the mean difference in HRV of 10.6° between the dominant and non-dominant shoulders ($P=0.001$). Similarly, the mean difference in IR at 90° of abduction between the dominant and non-dominant sides, 8.2°, correlated significantly with the difference in HRV as well ($P=0.003$).[19]

Paired *t*-tests also helped show significant differences between the dominant and non-dominant shoulder for IR at 90° of abduction and ER at 90° of abduction. For IR at 90° of abduction, the dominant shoulder measured an average of 43.0° (±7.4°) compared to an average measurement of 51.2° (±7.3°) for the non-dominant shoulder. For ER at 90° of abduction, the dominant shoulder measured an average of 116.3° (±11.4°) compared to an average measurement of 106.6° (±11.2°) in the non-dominant shoulder. There was no significant difference in the total arc of rotational motion between the dominant and non-dominant shoulders when measured at 90° of abduction with the dominant shoulder measuring 159.5° (±12.4°) and the non-dominant shoulder measuring 157.8° (±11.5°).[19]

The study by Reagan et al[19] suggests that an increase in HRV allows for an increase in ER without requiring a change in capsular laxity. Reagan et al[19] hypothesized that an osseous change that contributes to greater ER can serve as a protective mechanism against the injurious effects of a capsular laxity change such as anterior instability or chronic pain.

Measuring Humeral Head Retroversion

Numerous studies have been completed in order to better understand the assessment of HRV and the accuracy measuring HRVA using different imaging techniques. However, many studies differ on the anatomic landmarks from which to measure a HRVA. Also, several different imaging methods, such as radiography and computed tomography, have been used to quantify the HRVA. Each imaging procedure has had varying effects of accuracy and differing limitations. The following information is a review of studies that examined different factors that must be taken into account when measuring HRV.

First, in order to measure the HRVA during an imaging procedure, the correct anatomic landmarks must be determined. Soderlund et al[6] defined the HRVA with proximal and distal axes using radiographs and dry humeri examination. The proximal axis was defined as the humeral neck. The humeral neck, or proximal axis, was defined as being perpendicular to the line connecting the dorsal and ventral tangent points. The distal axis was defined as the line connecting the two epicondyles and the angle between the proximal and distal axes was the HRVA. Further studies that also defined the HRVA following Soderlund et al's[6] anatomic landmarks include Reagan et al[19] and Osbahr et al[5] who also utilized a radiographic imaging technique.

In another study by Oztuna et al[7], dry humeri were examined in order to correctly define the anatomic landmarks for HRVA measurement. The proximal axis was the line perpendicular to the humeral head surface diameter. The diameter of the humeral head surface was obtained by the two points where the humeral head articular surface became flat. The distal axis was the line tangent to the capitellotrochlear articular surface. The angle between the axes was the HRVA. Similarly, a study by Hernigou et al[20] also defined the proximal axis as a line perpendicular to the center of the humeral head; however the distal axis was established as a line directly through the epicondyles.

Conversely, Cassagnaud et al[21] ascertained that the landmarks used to measure the HRVA were the angle between the articular surface axis of the humeral head and bicondylar axis of the distal humeral epiphysis. In other words, the articular surface axis was obtained by placing a perpendicular line connecting the anterior and posterior extremities of the cephalic cartilage of the humerus. The bicondylar axis was classified as the straight line joining the most prominent part of the medial and lateral epicondyles. These two axes were overlaid in order to measure the HRVA between the articular surface axis and the bicondylar axis.

In a similar fashion, Doyle and Burks[22] used slightly different landmarks for the proximal humerus axis, but retained the same distal landmarks as the above mentioned study. For the proximal humeral axis, a line was drawn that connects the anterior and posterior margins of the articular cartilage of the humeral head. A line was then drawn perpendicular from the midpoint of line that connects that anterior and posterior margins. The HRVA was calculated by subtracting the distal axis angle from the proximal axis angle.

Frich and Moller[23] also utilized standard radiographs to measure the HRVA. However, in comparison to the previously mentioned studies, they attempted to simplify the equation that was used to measure the HRVA. Placing the subject's forearm in 35° of IR, a radiographic image was taken in an anterior-posterior direction. Line AB was established as a line drawn at a right angle to the long axis of the ellipse of the humeral head. Line CD was drawn perpendicular to the axial line of the prosthesis. Next, the angle between line AB and line CD was angle *a*. Then, instead of using a complicated equation to calculate the HRVA, they simply subtracted 40° from angle *a* in order to find a rough estimate of the HRVA. It was deduced in this study that this formula created a good correlation between the actual HRVA and the measured HRVA. However, the results in this study were not discussed in great detail. Further studies would have to investigate this method for reliability and validity before implementing this equation.

Using computed tomography, Hill et al[15] calculated the smallest concentric circle that coincided to the humeral head radius. Once this concentric circle was found, two points were placed at 90° to the anterior articular margin. The line that connected these two points was the humeral head axis. In order to determine the distal axis, a line was drawn parallel to the distal articular surface. The difference between the humeral head axis and the distal axis was the HRVA. A study by Crockett et al[4] also adapted this method to measure the HRVA.

Besides the above mentioned anatomic landmarks, several other studies utilized different anatomic landmarks for measuring the HRVA. A study by Yoshida et al[24] utilized the difference between two different baselines to establish the HRVA. The first baseline was referenced at the proximal humeral head. A line was drawn connecting the greater and lesser tuberosities. From this line of reference, the first baseline was established by drawing a line perpendicular to the reference line between the greater and lesser tuberosities. The distal baseline was established as the line parallel to the articular surface of the trochlea and capitellum.

Whitely et al[25] used an indirect method of ultrasound (US) visualization and palpation in order to quantify the HRVA. Ultrasound was utilized in order to visualize when the greater and lesser tuberosities were of equal height above the bicipital groove floor. This was completed by rotating the arm of the supine subject. Next, an inclinometer was placed on the ulnar styloid process and the medial border of the ulna with the subject's arm fully supinated. The HRVA was defined as the angle of the bent forearm subtracted from the vertical position which was referenced as 0° of shoulder rotation. This equation was also utilized in a study by Ito.[26] However, in this study, the US was utilized to locate when the bicipital groove faced upwards in comparison to the equal height of the greater and lesser tuberosities.

Once the anatomic landmarks had been defined from which to measure HRV, it must be remembered the humeral positioning may affect the ability to accurately measure the HRVA. Numerous studies have found that incorrect humeral positioning can generate an inaccurate measurement of HRVA. Soderlund et al[6] found that correct positioning of the arm during radiographic measurement is imperative for high accuracy of HRVA measurement. In this study, dry humeri were imaged using radiograph and computed tomography. Ninety degrees of shoulder flexion was operationally defined as 0°. Therefore, 10° of shoulder extension is equivalent to 80° of shoulder flexion and 10° of shoulder flexion would be equivalent to 100° of shoulder flexion. In the radiographic method, the dry humeri were imaged in a semi-axial view in 5°, 10°, 15°, 20°, 30°, 45° of abduction and in extension of 5° and 10°. These same positions of abduction were also imaged in 0°, 5° and 10° degrees of flexion. Humeral head retroversion angle radiographic measurements from these different positions were compared to computed tomography scans of the HRVA. It was found that when the dry humeri were in extension, measurement error between computed tomography and radiographs exceeded 2°. Overall, HRVA measurement differences did not go beyond 5° as long as the dry humeri were not abducted more than 20° and not extended.[6]

The HRVA measurement differences between radiographs and computed tomography due to positioning was taken into account when measuring the HRVA in five subjects without shoulder pathology.[6] The positioning was taken into account by creating a stand for the arm to be placed in during radiographic imaging. The subjects were supine with the cassette placed underneath the subject's shoulder. The shoulder was positioned in 90° of flexion and 10° of abduction. The stand helped to support the forearm parallel and horizontal to the long axis of the subject's body. Using this positional method during radiography, an average difference of 1.5° for the HRVA was found between radiography and computed tomography. Therefore, with correct positioning, a high degree of accuracy for measuring the HRVA can be completed by a single radiograph. Also, when imaging a patient population using radiograph, Soderlund et al[6] found that when the shoulder was abducted more than 20°, measurement error of humeral head retroversion increased.

Oztuna et al[7] established a different method for subject positioning for radiography. Like Soderlund et al[6], this study examined dry humeri in order to obtain anatomic HRVA measurements to compare to the radiographic HRVA measurement. In order to measure the dry humeri HRVA, two 360° goniometers were attached by a telescopic rod in parallel.[7] The distal axis was defined by the K-wire that was parallel to the capitellotrochlear joint surfaces. The proximal axis was defined where the K-wire passed through the longest diameter of the humeral head articular surface. Then, the goniometer was adjusted by moving the telescopic rod in order to measure the anatomic HRVA of the dry humeri.

In order to compare to effectiveness of the subject positioning, HRVA measurements of the dry humeri were positioned for radiographic imaging.[7] The dry humeri were placed into 20° of abduction in relation to the x-ray beam with a proximal axis of the line perpendicular to the humeral head diameter and a distal axis of the line tangent to the capitellotrochlear surface. These measurements were compared to the anatomic dry humeri HRVA measurements. It was determined that the average difference between anatomic and radiographic dry humeri measurements was 0.9°. As a result, reliability and correlation exists between these two methods of HRVA measurement. Due to this reliability and correlation, during radiographic imaging, subjects were positioned in 90° of shoulder elevation, 20° of shoulder abduction, 90 ° of elbow flexion and full supination while standing. The radiograph was taken posterior to anterior in a semiaxial view with the same proximal and distal axes listed for the dry humeri radiographic imaging.

A study by Debevoise et al also examined the importance of subject positioning.[27] In this study, in order to maintain a position of 15° shoulder abduction, neutral shoulder rotation and full supination, a positioning apparatus was utilized. In addition, a study by Farrokh et al[28] analyzed the importance of patient positioning while measuring HRV using computed tomography. Three different humeri positions were utilized. The first position was with the humerus parallel to the CT table. Second, the humerus was positioned in 20° of extension and 10° of abduction with the CT scan beam perpendicular to the diaphyseal axis. Lastly, the humerus was in the same position as the second position; however, the beam from the CT scan was oblique to the diaphyseal axis. It was found that when the humeri were positioned in the second position, there was no significant difference. Therefore, the HRVA measurement is dependent about positioning during computed tomography imaging.

Furthermore, in a study by Osbahr et al[5], patients were positioned in supine on the radiograph table with the arm in 90° of shoulder flexion and 10° of shoulder abduction; 90° of elbow flexion and a neutral position of the forearm. This position was maintained by having the patient hold onto a radiographic cassette table which was placed on the table with the patient. The benefit was that this positioning allowed the humeral head and the epicondyles to be projected visibly onto the radiographic film. In a similar manner, Reagan et al[19] positioned subjects in supine holding the imaged arm in 90° of shoulder flexion, 20° of shoulder abduction, 90° of elbow flexion, neutral shoulder rotation and neutral supination/pronation by using a positioning device.

Once the appropriate positioning of the humerus during imaging was determined, studies examined the reliability of HRVA measurement between different types of diagnostic imaging. Again, Soderlund et al[6] compared the difference between the measurements of the HRVA using radiography and computed tomography in subjects without current or past shoulder trauma. For the radiographs, patients were positioned as previously mentioned with the x-ray beam over the humeral head projecting through the biceps muscle. The HVRA was measured using the proximal and distal axes previously mentioned with a goniometer. The patient was supine with the humerus parallel to the body for computed tomography. It was found that 2° was the maximum angle measurement difference between the two imaging procedures and the mean difference between the imaging measurements was 1.5°. Thus, this study was able to conclude the using a single exposure radiograph to measure the HRVA is sufficient and it can be read with high accuracy.

While Soderlund et al[6] and Oztuna et al[7] established the use of radiography to image the HRVA, Hernigou et al used computed tomography to assess the HRVA.[20] First, the anatomic

HRVA was measured in order to compare this angle to the HRVA measured by computed tomography. It was found that the anatomic HRVA and computed tomography HRVA did not differ significantly. Also, a regression analysis displayed a strong linear relationship between the two methods of measurement. When compared to the direct anatomic measurement of the HRVA and the HRVA measured from computed tomography, it was found that there was no difference. Therefore, Hernigou et al[20] found that computed tomography should be used to quantify HRVA since it is a reliable and accurate measure. In addition, CT scans use lower levels of radiation and do not create distortion which is found in radiographic images.[28]

Like Hernigou et al[20], Nyffler et al[29] compared the reliability of radiography and computed tomography in a retrospective study by two orthopedists and one radiologist. This retrospective study examined glenoid version. For the radiographic images, the original films were taken with the patient in supine with the arm in 60° of abduction, neutral rotation and an axillary view. The x-ray tube was by the subject's hip and the cassette was held against the subject's neck. Lines were drawn through the middle of the scapular blade and along the anteroposterior diameter of the glenoid cavity. From these points of reference, the angle between a line perpendicular to the first line and the second line was the angle of glenoid version.

The computed tomography scans were taken with the patient in supine with scapula flat against the examination table and humerus parallel to the subject's body.[29] A line anterior to the glenoid posterior margin and a line at the glenoid fossa midpoint were drawn. Then, the angle of glenoid version was defined as the angle between the glenoid posterior margin and the line perpendicular to the glenoid fossa midpoint. When comparing the glenoid version angle using radiographs and computed tomography, it was found that radiographs overestimated the glenoid version angle in 86% of the images. For that reason, Nyffeler et al[29] concludes that radiographs cannot accurately assess glenoid version and computed tomography should be used since it is a more accurate and reproducible measurement for glenoid version.

Similar to the previously mentioned studies, a study by Boileu et al[30] supported the conclusion that computed tomography is more accurate in assessing HRVA than radiographs. In this study, humeri from cadavers, radiographs, computed tomography, computer-assistance and a direct method were employed to evaluate HRVA. The computer-assisted method was considered the gold standard for measuring HRVA. For radiography, the HRVA was defined as the angle between the elbow tangent axis and the humeral head axis using a frame to position the humerus. Computed tomography defined that the HRVA was the angle between the humeral head axis and epicondylar axis. The gold standard method, or computer-assisted method, used a custom frame to support the humerus. The humerus was supported in a stationary position at the elbow and head of the humerus which permitted for revolution around the diaphyseal axis. A three-dimensional model of the humerus was produced from the computer-assisted method using digitization of over one thousand points on each humerus. Lastly, the direct anatomic method used a device to measure the HRVA. This device held the humerus immobile and a 180 degree protractor was placed on the epicondyles. This was used to measure the HRVA by utilizing the angle between the humeral head axis and the epicondylar axis. After examination of these three different methods, it was found that radiography overestimated the retroversion angle when compared to the gold standard. No difference in HRVA measurement was seen when comparing the computer assisted method and computed tomography method. Therefore, computed tomography was superior to radiographic method when measuring humeral head retroversion.

Besides radiography and computed tomography, US has also been used as an imaging technique. In the previously mentioned studies by Whitley et al[25] and Yoshida et al[24], both employed US to visualize the bicipital groove and used either inclinometers or goniometers to measure the HRVA. Whitely et al[25] found that using the US method had an ICC value of 0.98 for the right arm and 0.94 for the left arm. In comparison, the direct palpation method of the bicipital groove had an ICC of 0.51 for the right arm and 0.49 for the left arm.

Similarly, Yoshida et al[24] found between trial reliability to be 0.95 for the dominant hand and 0.91 for the non-dominant hand. This inter-rater reliability is similar to the inter-rater reliability demonstrated by CT. This lends evidence to using US to image the HRVA since it is readily available in the clinic, has no radiation and little specialized training is needed since the physical therapists in the study only had 15 minutes of training prior to participating.

Many different factors must be taken into consideration when obtaining accurate HRVA measurements. It is vital to have correct anatomic landmarks from which to measure the HRVA in order to allow different raters to correctly measure the HRVA. Also, it important to maintain correct positioning of the arm during imaging; otherwise, the HRVA may be incorrectly measured. Lastly, different methods of imaging are used to assess the HRVA. It has been found in studies that computed tomography is the best method of measurement. Nonetheless, other studies have established that radiograph is as accurate as computed tomography when measuring the HRVA.

Variations between trained and inexperienced observers

Studies have identified three sources of variability in a radiograph measurement including patient, procedure, and clinician.[8,31] The amount of variability in clinician measurements can be quantified through intra-rater and inter-rater reliability. Studies have shown mixed reviews over the intra-rater reliability and the inter-rater reliability, given the amount of training of the raters, in reading radiographs.

Clinical decisions are often based largely on radiographic findings. Radiographic x-ray measurements need to be reliable to accurately support these clinical decisions. A study by Owen et al[8] examined the reliability between a pediatric orthopedic surgeon, a pediatric orthopedic fellow, and a research assistant in the measurement of angulation of femoral shaft fractures. Radiographs from 30 children aged 4 to 10, consisting of anterior, posterior, and lateral views were evaluated on two separate occasions, with 14 days between the measurements, by each observer. A standardized technique was provided to the observers but no formal training was conducted.[8]

The research assistant, who had no prior experience evaluating radiographic fracture positions, achieved the highest intra-rater reliability (ICC = 0.97). Inter-rater reliability as reported by the Pearson correlation coefficients were positive and extremely strong for all rater comparisons (r = 0.90). Owen et al[8] concluded that if a standard technique is implemented, skill in rating radiographs is not dependent on clinical experience.

Wright et al[31] attempted to reproduce a clinical setting by using unrehearsed observers to evaluate the accuracy of measuring lower limb alignment by means of radiographs. The purpose of the study was to investigate the contribution that inconsistent positioning plays in measurement variability. Additionally, the researchers reported on intra- and inter-rater reliability. Two orthopaedic surgeons served as observers measuring the tibiofemoral "anatomical" axis indicating lower limb alignment. Forty-three radiographs were measured by one observer once and then twice by the other observer, with one week separating those two measurements. The observers did not discuss measurement techniques or participate in joint practice sessions.[31]

The authors concluded that there were no significant differences between observers. Intra- and inter-rater reliability was determined with the interclass correlation coefficient. Reliability was reported for intra-observer measurements, inter-observer measurements, and an overall ICC with values of 0.96, 0.95, and 0.84, respectively. Further interpretation of the reliability among observers is limited due to it not being the primary aim of the study.[31]

Differences in the ability to measure radiographs, given specific parameters, between a "trained" novice and intermediate observer were examined in a study by Dang et al.[9] The intention of the study was to assess intra-observer reproducibility and inter-observer reliability through performance on measuring 14 of the 19 parameters outlined in the *AIS Radiographic Measurement Manual*, as identified by the Spinal Deformity Study Group (SDSG).[9] Ten radiographs were randomly selected from a sample of 49 nonoperative patients with moderate AIS (initial curves between 20° and 45°). The two observers independently evaluated the radiographs of the selected patients whose ages ranged between 11 and 15 years. The "trained" novice observer was an undergraduate student and the intermediate observer was a clinical engineer who regularly measured radiographs in a scoliosis clinic. Each observer reviewed the SDSG parameters prior to the study and measured the 14 designated parameters on all ten radiographs on 5 occasions, with 3-5 days between measurements.[9]

Excellent intra-observer reproducibility for both the novice and intermediate observers was found for the posteroanterior radiograph but only fair to good reproducibility for the sagittal radiographs. Inter-observer reliability was excellent for 7 of the 13 parameters observed, however one had a 95% confident interval that spanned zero. Overall, only one of the 14 parameters reported excellent intra-observer reproducibility for both observers and excellent inter-observer reliability. The authors theorized that discrepancies in measurement should not be attributed to examiner's skill level because the trained novice achieved similar intra-observer reproducibility compared to the intermediate examiner.[9]

Similar to the previous study, Margolis et al[10] looked at the progression of radiographic analysis skills with increased training in the realm of family practice. A total of 87 participants comprised of final-year medical students, family practice residents, general practitioners, and family practice educators were included in the study. Twelve sets containing 1- 3 radiographs of normal and abnormal chest and wrist x-rays were viewed by each participant in a darkened room. Participants were given as much time as needed to determine whether they thought the radiographs were normal or abnormal and to determine a diagnosis.[10]

The results of the study looked at a frequency analysis to describe the correct response rate. The family medicine educators had the highest mean scores for both identifying abnormal versus normal radiographs and correctly diagnosing the 12 sets of radiographs, but the researchers found no statistical difference in the scores between groups. The authors of the study concluded that there is no variation across four levels of family practice training indicating that additional years of formal education alone does not advance radiographic interpretation.[10]

Niemeyer et al[32] studied the reliability of Lenke's and King's classifications for adolescent idiopathic scoliosis when observers of varying levels of professional training were used. Three observers with 1, 5, and 10 years of orthopedic training examined 60 operative-case radiographs. Observer one had no experience with measuring radiographs with scoliosis or with either classification system. Observer two had experience measuring radiographs with scoliosis and knew only King's classification. Observer three had several years of experience measuring radiographs and surgical treatment of scoliosis and knew both classifications. Each observer measured and assessed both the nonmeasured and premeasured radiographs independently on 5 separate occasions with 1-2 weeks between sessions. Copies of the descriptions of both classification systems were available in each room.[32]

Results for the non-measured radiographs found the overall inter-observer agreement for Lenke's classification to be poor with a kappa coefficient of 0.23. King's classification agreement increased with each session improving from a kappa coefficient of 0.08 to 0.79 with an overall multi-rater agreement had a kappa coefficient of 0.42. Intra-observer agreement for the complete Lenke's classification was poor to fair with kappa coefficients of 0.30, 0.57, and 0.51 for Observers 1, 2, and 3, respectively. Intra-observer agreement for King's classification found kappa coefficients of 0.53 for Observer 1, 0.74 for Observer 2, and 0.64 for Observer 3. Results for the premeasured radiographs showed excellent agreement among all 3 observers for Lenke's complete classification and good agreement was found for King's classification with a kappa coefficient of 0.88 for all 3 observers. The results of the study did not show that the most experienced observer was capable of the measuring and classifying best and only marginal differences were found between the 3 observers. The authors did note several limitations of the study including the chance of recollection effect and that the observers became experienced during the evaluation process.[32]

Several studies have looked at the differences between varying health care professionals. In a study by Carman et al[33] researchers assessed intra-observer and inter-observer variation in measurement of the Cobb angle on radiographs of patients who had kyphosis or scoliosis between four orthopaedists and one physical therapist. Each observer measured 8 anteroposterior radiographs of patients with scoliosis twice, with a minimum of two weeks between measurements. The second phase involved the additional measurement of twenty lateral radiographs showing kyphosis.[33]

The results of the study showed a mean absolute value for scoliosis of intra-observer difference of 3.8° and 95 percent of the differences were by 8° or less. For kyphosis, the average difference between observers was 3.3° and 95 percent of the differences were 7° or less. The inter-observer differences were also small. The authors hypothesized the difference in readings was affected more by measurement error alone and not differences in the techniques or skills of different interpreters.[33]

Nelitz et al[34] looked at intra-observer and inter-observer measurement reliability during evaluation of the most common radiographic parameters in the assessment of hip dysplasia in adults. One-hundred randomly selected radiographs from patients aged 16 to 32 who had been diagnosed with unilateral hip dysplasia were used for assessment. Two senior orthopaedic surgery residents and one senior medical student served as the observers. The medical student had repeated training sessions and the observers agreed on precise definition of landmarks prior to the study. Measurements of the hip radiographs were independently carried out by all observers on one occasion and two of the three observers reviewed the radiographs a second time 3 months later.[34]

The results showed a wide span of measurement variability among radiological features. In general the inter-observer variability was lower than the intra-observer variability. Six of the 9 radiographic measurements showed high inter-observer reliability with interclass correlation coefficients between 0.76 and 0.87. The other 3 measurements had ICC values that ranged from 0.58 to 0.66. Similarly, the same 6 out of the 9 radiographic measurements showed sufficient intra-observer reproducibility with interclass correlation coefficients between 0.70 to 0.92. ICC values for the lowest 3 measurements ranged from 0.56 to 0.76. The researchers found the results to indicate that experience did not appear to influence intra-observer or inter-observer reliability because the less educated medical student's measurements did not significantly differ compared to the two more experienced observers. The authors attributed the variability to difficulty identifying certain landmarks and clarity of anatomical structures.[34]

In a study by Soderlund et al[6] inter- and intra-observer differences between two radiologists were examined. One radiologist had a great deal of experience and the other one had much less experience. Bilateral shoulder radiographs were obtained from 11 healthy subjects with no previous shoulder trauma or surgery. Each observer measured the humeral head retroversion angle twice in each radiograph, with two months between measurements.[6]

Intra- and inter-observer differences were determined by calculation of the mean difference. Assessment for mean intra-observer measurement variations yielded a mean difference of 1.7° for the experienced radiologist and 4.1° for the less experienced observer. The mean inter-observer difference was 1.0 degree. The suggestive low intra-observer reproducibility and inter-observer reliability were proposed by the authors to be largely contributed to by the positioning of the humerus during radiographic imaging.[6]

Computed tomography (CT) scans are the gold standard for assessing humeral retroversion. Nyffeler et al[29] conducted a study investigating the comparison of glenoid version measurement between radiographs and CT scans. Two orthopedic surgeons and 1 radiologist assessed 50 patient radiographs and CT scans for glenoid version. Twenty-five of the patients were assessed for anterior shoulder instability and the other 25 patients were assessed following total shoulder joint replacement surgery. Standard positioning and measurement parameters were predetermined and followed out as consistently as possible for minimal variability.[29] The results of this study revealed excellent inter-observer correlation in CT scan measurements for the instability group, ICC = 0.97. Differences between observers were never greater than 4° and averaged 1.5°. Computed tomography assessment for the total shoulder group was good as well (ICC = 0.93). Alternately, inter-observer reproducibility was lower for radiographs with an ICC of 0.77 and a range of examiner differences up to 35° were observed. The authors point out that the coefficient of correlation between CT scans and radiographs were 0.33 and 0.67 for the instability group and prosthesis group, respectively. Therefore, the authors conclude that to assess glenoid version, both pre- and post-operatively, CT should be used to ensure accuracy.[29] Determination for the necessity of training in reading or measuring radiographs remains inconsistent among studies. However, a central theme emerged from many of these studies examining the protocols and anatomical identification on radiographs as sources of measurement error more so than rater training or experience. Further studies need to be conducted to look at inter- and intra-rater reliability in specifically measuring HRVA using a radiograph between trained and inexperienced raters.

Chapter III

Methods

Subjects

Subjects were previously recruited on a voluntary basis to participate in a study conducted and described by Cieminski et al. Sixty-three male volunteers between the ages of 18-28 years with at least 5 years of competitive baseball experience had bilateral humeral radiographs taken. Thirty-eight of the sixty-three subjects were current pitchers and the twenty-five other subjects had little or no pitching experience. Radiographs obtained from these previous participants were used to measure HRVA. The present study included 35 randomly selected radiographs from a pool of 126 radiographs for assessment.

Raters

Trained, untrained, and experienced raters were involved in the study. Trained and untrained raters consisted of two second-year Doctor of Physical Therapy students for each group. All students had basic knowledge of radiographs, but no exposure to measuring HRVA prior to the study. Additionally, students had the same education on shoulder anatomy and physiology, arthrokinematics, osteokinematics, and kinesiology. Both groups were provided written and pictorial descriptions for obtaining the HRVA measurements. Trained raters received additional education and instruction in measuring HRVA during two sessions. The first session consisted of a one hour instructional meeting with the experienced clinician who explained the protocol for correctly establishing all points of reference on a radiograph and a demonstration of how to obtain the HRVA measurement. An additional follow-up session was completed the next day for thirty minutes in order to address any further necessary clarification. The untrained group did not receive any instruction or clarification beyond that explained in the written handout. The experienced clinician was a physical therapist with Bachelor's-trained entry-level physical therapy degree with subsequent Master's and PhD degrees earned. At the time of data collection the experienced clinician had 18 years of physical therapy experience and 50 hours of experience measuring HRVA using the semiaxial radiographic techniques utilized in this study.

Instrumentation

Measurement of HRVA was obtained from the selected radiographs using 1° increments from a standard goniometer. The standard goniometer was additionally used to ensure accurate positioning of the shoulder during radiographic measurement. Radiographs were taken using the following equipment: radiographic generator (Philips Optimus 65, Eindhoven, Holland), tube (Phillips SRO 33110, Eindhoven, Holland) and cassette. Hard copies were then produced.

Procedure

Two student groups, consisting of two Doctor of Physical Therapy students in each group, were randomly divided into trained and untrained groups. The written protocol and photograph were available for reference during HRV measurement sessions. The radiographs used in the instructional sessions were not included in the set of radiographs used in this research study. Thirty-five radiographs were analyzed on two separate occasions by both student groups, who were blinded to any indentifying information on the radiographs.

Measurement of the HRVA was obtained using two axes and is depicted in Appendix A – Figure 1. The humeral head axis was identified by placing two points, one anterior and one posterior, at the point where the articular surface of the humeral head flattens. A line was drawn connecting these two points of reference. The length of this line was measured in millimeters and crosshatches were placed at three equidistant points on this line. The perpendicular distances from these outermost crosshatches to the articular surface of the humeral head were measured in millimeters and compared to each other in order to ensure that the humeral head was truly bisected. If the humeral head was bisected correctly, these perpendicular distances should be equal lengths. If the measured distances were not found to be equal, the points on the articular surface were adjusted accordingly. Once these points on the humeral articular axis were confirmed a line was drawn between the points determining the proximal axis.

The trochlear, or distal, axis was established by drawing a line parallel to the anterior articular surface of the distal humerus by using the apex of both the medial and lateral surfaces. Through the proximal axis, a perpendicular line was drawn that intersected the distal axis. The angle of this intersection was the HRVA. The humeral head retroversion angle was defined as the acute angle between the proximal and distal axes and was measured directly on the radiographs using a standard goniometer. Between successive raters measuring HRVA, the lines on the radiographs were removed in order to protect the integrity of the research study results.

Data Analysis

Interclass correlation coefficients (ICC) were utilized to determine the differences between trained and untrained groups when measuring HRVA. The gold standard of HRVA measurements were the experienced clinician's HRVA measurements. Two measurements were taken for each of the 35 radiographs by each rater and were used to establish intra-rater reliability. As described in Portney and Wakins,[35] ICC Model (1, 1) was implemented to assess intra-rater reliability. ICC Model (1, 1) allows differences between measurements of HRV to be assessed, but does not assess raters. Therefore, ICC Model (1,1) was used versus other ICC

Models since it is a more conservative estimate of reliability, yielding lower values which allowed for generalization of the results to a larger population.

In order to evaluate inter-rater reliability, ICC Model (2, 1) was utilized. In ICC Model (2,1), each subject was measured by the same group of raters. Only the first measurement from each radiograph from each rater was used to calculate the inter-rater reliability. The average of the two measurements was not used in order to decrease the chance of masking rater errors. The experienced investigator's data is used in both the trained and untrained groups for analysis of inter-rater reliability.

For all ICC's, 95% confidence intervals (CI's) were also determined. A wider interval increases the confidence that the mean will fall within the interval, but increases the likelihood that our findings were influenced by outside factors.

Chapter IV

Results

Thirty-five radiographs were analyzed by two trained and two untrained raters on two separate occasions in an attempt to determine inter-rater and intra-rater reliability measuring HRVA. Intra-rater reliability was calculated using ICC model 1 for each trained and untrained rater comparing the separate measurement sessions. The trained and untrained raters all demonstrated excellent reliability coefficient measurements of 0.89 and 0.87 for the trained group and 0.89 and 0.89 for the untrained group, respectively. Each rater demonstrated statistical significance with a p value less than 0.01 within a 95% confidence interval. All intra-reliability coefficients are listed in Appendix B - Table 1.

Inter-rater reliability was calculated using ICC model 2, however only the first measurement was used against an experienced clinician's results for both the trained and untrained groups. The inter-rater reliability coefficient for the untrained group was 0.80 and 0.82 for the trained group, which demonstrated good reliability for both. Each group's results were statistically significant with a p value less than 0.01 within a 95% confidence interval. All inter-reliability coefficients are listed in Appendix B - Table 2.

Chapter V

Discussion & Conclusion

In the study, both inter- and intra-rater reliability for measuring HRVA between trained and untrained raters were examined. The reliability coefficients for inter- and intra-rater reliability for both the trained and untrained groups was good with values of 0.82 and 0.80, respectively, when compared against the experienced rater as the criterion or reference standard. Two measurements were obtained from each radiograph, but only the first measurement from each rater was used to determine ICC instead of using the average of the two to minimize errors being left undetected. The primary purpose of this study was to establish the accuracy of HRVA measurements on a radiograph of untrained raters in comparison to trained raters and an expert rater. The data gathered from this research found similar and accurate reliability results between all raters, therefore the theorized training effect was not found between Doctor of Physical Therapy students.

Intra-rater reliability was calculated using the data collected at the two measurement session. Each of the four students demonstrated good to excellent reliability with values falling between 0.87 to 0.89, showing good accuracy consistently between trained or untrained raters. Interclass correlation coefficient model 1,1 was used to measure intra-rater reliability to ensure the accuracy of the results. Model 1,1 allows incorporation of all sources of error, generally yielding lower numbers due to the conservative approach for the model, however the results have more credibility lending a higher level evidence.

The primary hypothesis of the study was that no difference exists between trained and untrained raters when measuring HRVA on radiographs. Results of this study support the hypothesis that written instructions are sufficient with no additional training for measuring

HRVA on plain radiographs. Previous studies have shown similar results. Differences in the ability to measure 14 spinal parameters on radiographs, given the specific parameters between a "trained" novice and intermediated observer were examined in a study by Dang et al.[9] Results from the study found similar intra-observer reproducibility for both the novice and intermediate observers, ranging from excellent on some parameters to fair to good on others, however no matter which parameter was being measured, both observers had similar values. In another study by Owen et al,[8] fracture angulation was determined on radiographs by a pediatric orthopedic surgeon, pediatric orthopedic fellow, and a research assistant that had no formal training with the specified technique. Findings of the study showed similar inter-rater results between all raters (r=0.90) regardless of profession and experience. Additionally, the research assistance demonstrated the highest intra-rater reliability (ICC=0.97). Owen et al[8] went on to conclude that if a standard technique is implemented, skill in rating radiographs is no dependent on clinical experience.

In contrast to these results, however, Soderlund et al[6] evaluated inter- and intra-observer differences between radiologists of varying experience levels and discovered minimal, yet noteworthy, discrepancies in observer reliability. Assessment for mean intra-observer measurement variations for the experienced radiologist and less experienced radiologist showed a difference of 1.7° and 4.1°, respectively. Soderlund et al[6] proposed the low levels of reproducibility had more contribution from humeral positioning during radiographic imaging than actual skill of the rater. Nelitz et al[34] supports this line of thought by attributing the variability of measurements encountered in their study to difficulty identifying certain landmarks and clarity of anatomical structures. Two senior orthopedic surgery residents and one senior medical student performed measurements of 9 common radiographic parameters in the assessment of hip dysplasia in adults, yielding reliability coefficients ranging from 0.56 to 0.76 (ICC). The results indicated as well that experience did not appear to have an influence because the less educated medical student's measurement did not significantly differ compared to the two more experienced observers.

Several limitations were identified in the study. One limitation is the lack of data comparison between CT scans and radiographs. Computed tomography scans are the accepted gold standard for measuring HRVA. More recently it has been proposed to utilize radiographs as comparable cost-effective alternative. The study, however, only included measurements from radiographs due to available resources, limited access to CT technology, time constraints, and limited funds. Therefore, this study cannot conclude that radiographs are more or less accurate in producing an image appropriate to correctly measure HRVA than CT scans. However, the study by Soderlund et al[6] examined the reliability of the HRVA measurement between radiographs and CT imagining. The study demonstrated that the maximum angle measurement difference between the two types of imaging was 2° and the mean difference was 1.5°. Soderlund concluded that using a single exposure radiograph to measure HRVA is sufficient to produce a comparable level of accuracy.

Another limitation is the difficulty in determining articular landmarks of the humeral head on the radiograph. As described in the methods, crosshatches were drawn and measured for humeral head bisection with a corrective plan established to make adjustments to ensure the accuracy of bisection. However, there is no procedure in place to help ensure the distal axis was determined accurately.

Finally, due to the smaller sample size of the study, generalizability and level of evidence are limited. Future studies should be performed both with a larger subject group and also with a larger rater group to support the findings of this research.

As discussed earlier, Vision 2020 emerged as an effort to further the physical therapy profession through the practice of life-long learning, evidence-based practice, and integrity. Goals of Vision 2020[11] include becoming practitioners of choice, direct access for patients, and building an autonomous practice. Based on the results of this study, Doctor of Physical Therapy students are able to read radiographs with good intra-rater and inter-rater reliability. Therefore, the radiographic imaging education provided in a doctoral level physical therapy program is sufficient to understand and identify bony anatomy on radiographic imaging. This proficiency in identification and understanding of radiographic imaging lends support to physical therapists becoming autonomous, practitioners of choice.

According to the *Guide to Physical Therapist Practice*,[36] states that physical therapists provide services to clients who have impairments, functional limitations, disabilities, or changes in physical function and health status resulting from injury, disease, or other causes. Therefore, based on this level of education, physical therapists may be able to identify dysfunction within these parameters autonomously. In order to provide patient services through the continuum of care, gaining the ability to order and evaluate radiographs would assist the physical therapy profession in its progress towards autonomous, practitioners of choice.

Conclusion

This research provides evidence to support the competencies of entry-level physical therapists at performing accurate measurements of HRVA on radiographs are comparable to experienced physical therapists with written instruction and no other formal training. This demonstrates the potential to further incorporate radiographic information into clinical practice. The study supports APTA's Vision 2020 for physical therapists to become autonomous practitioners and expand the physical therapy scope of practice.[11] Further research is needed to bridge the gap between Vision 2020 and physical therapy practice guidelines. We recommend continued research studies involving the examination of the effect of training on novice raters when performing radiographic measurements on a variety of joints and following other techniques or written protocols.

References

1. Krahl V. The Torsion of the Humerus its Localization, Cause and Duration in Man. *Am J Anat.* 1947; 80: 275-319.

2. Edelson G. Variations in the retroversion of the humeral head. *J Shoulder Elbow Surg.* 1999;8:142-5.

3. Edelson G. The development of humeral head retroversion. *J Shoulder Elbow Surg*. 2000; 9:316-8.

4. Crockett HC, Gross LB, Wilk KE et al. Osseous adaptation and range of motion at the glenohumeral joint in professional baseball pitchers. *Am J Sports Med.* 2002;30:20-26.

5. Osbahr DC, Cannon DL, Speer KP. Retroversion of the humerus in the throwing shoulder of college baseball pitchers. *Am J Sports Med.* 2002;30:347-363.

6. Soderlund V, Kronberg M, Brostrom LA. Radiologic assessment of humeral head retroversion: Description of a new method. *Acta Radiolgica.* 1989;30:501-5.

7. Oztuna V, Ozturk H, Eskandari MM, Kuyurtar F. Measurement of the humeral head retroversion angle: A new radiographic method. *Arch Orthop Trauma Surg.* 2002;122:406-9.

8. Owen JL, Stephens D, Wright JG. Reliability of radiographic measurement of fracture angulation in children with femoral shaft fractures. *Can J Surg*. 2007;50:115-118.

9. Dang NR, Moreau MJ, Hill DL, Mahood JK, Raso J. Intra-observer reproducibility and interobserver reliability of the radiographic parameters in the deformity study group's *AIS Radiographic Measurement Manual. Spine*. 2005;30:1064-1069.

10. Margolis SA, Nilsson KA, Reed RL. Performance in reading radiographs: does level of education predict skill? *J Contin Educ Health Prof*. 2003;23(1):48-53.

11. Vision 2020. American Physical Therapy Assciation website. 2011 Available at: www.apta.org. Accessed November 18, 2010.

12. Sabick M, Kim Y, Torry M, Keirns M and Hawkins J. Biomechanics of the Shoulder in Youth Baseball Pitchers: Implications for the Development of Proximal Humeral Epiphysiolysis and Humeral Retrotorsion. *Am J Sports Med*. 2005;33:1716-1722.

13. Cowgill L. Humeral torsion revisited: A Function and Ontogenetic Model for Populational Variation. *Am J Phys Anthropol.* 2007;134:472-480.

14. Rhodes J. Adaptations to humeral torsion in medieval Britain. *Am J Phys Anthropol*. 2006;130:160-166.

15. Hill JA, Tkach L, Hendrix RW. A study of glenohumeral orientation in patients with anterior recurrent shoulder dislocations using computerized axial tomography. *Orthopaedic Review*. 1989;18:84-91.

16. Hawkins RJ, Bokor DJ. Clinical Evaluation of Shoulder Problems. *The Shoulder.* Philadelphia: WB Saunders Co; 1998.

17. Kronberg M, Brostrom LA, Soderlund V. Retroversion of the humeral head in the normal shoulder and its relationship to the normal range of motion. *Clinical Orthop.* 1990;253:113-117.

18. Pieper HG. Humeral Torsion in the Throwing Arm of Handball Players. *Am J Sports Med.* 1998;26:247-253.

19. Reagan KM, Meister K, Horodyski MB, Werner DW, Carruthers C, Wilk K. Humeral retroversion and its relationship to glenohumeral rotation in the shoulder of college baseball players. *Am J Sports Med.* 2002;30:354-360.

20. Hernigou P, Duparc F, Hernigou A. Determining humeral retroversion with computed tomography. *Journal of Bone and Joint Surg.* 2002;84:1753-1862.

21. Cassagnaud X, Maynou C, Petroff E, Dujardin C, Mestdagh H. A study of reproducibility of an original method of CT measurement of the lateralization of the intertubercular groove and humeral retroversion. *Surg Radiol Anat.* 2003;25:145-151.

22. Doyle AJ, Burks RT. Comparison of humeral head retroversion with the humeral axis/biceps groove relationship: a study in live subjects and cadavers. *J Shoulder Elbow Surg.* 1998;7:453-7.

23. Frich LH, Moller BN. Retroversion of the humeral prosthesis in shoulder arthroplasty: measurements of angle from standard radiographs. *J Arthroplasty.* 1989;4:277-80.

24. Yoshida M, Saho Y, Katayose M. Reliability of measuring humeral retroversion using ultrasound imaging in a health nonthrowing population. *Journal of Sports Rehabilitation.* 2010;19:149-160.

25. Whiteley R, Ginn K, Nicholson L, Adams R. Indirect ultrasound measurement of humeral torsion in adolescent baseball players and non-athletic adults: reliability and significance. *J Science & Med in Sport.* 2006;9:310-8.

26. Ito N, Eto M, Maeda K, Rabbi ME, Iwasaki K. Ultrasonographic measurement of humeral torsion. *J Shoulder Elbow Surg.* 1995;4:156-61.

27. Debevoise NT, Hyatt GW, Townsend GB. Humeral torsion in recurrent shoulder dislocations: a technic of determination by x-ray. *Clinical Orthopaedics & Related Research.* 1971;76:87-93.

28. Farrokh D, Fabeck L, Descamps PY, Hardy D, Delince P. Computed tomography measurement of humeral head retroversion: influence of patient positioning. *J Shoulder Elbow Surg.* 2001;10:550-3.

29. Nyffeler RW, Jost B, Pfirrmann CW, Gerber C. Measurement of gleniod version: Conventional radiographs versus computed tomography scans. *Journal of Shoulder and Elbow Surgery.* 2003;12:493-496.

30. Boileau P, Bicknell RT, Mazzoleni N, Walch G, Urien JP. CT scan method accurately assesses humeral head retroversion. *Clin Orthop Relat Res.* 2008;466:661-669.

31. Wright JG, Treble N, Feinstein AR. Measurement of lower limb alignment using long radiographs. *J Bone Joint Surg Br*. 1991;73(5):721-723.

32. Niemeyer T, Wolf A, Kluba S, Halm H, Dietz K, Kluba T. Interobserver and intraobserver agreement of Lenke and King classifications for idiopathic scoliosis and the influence of level of professional training. *Spine*. 2006;18:2103--2107.

33. Carman DL, Browne RH, Birch JG. Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. *J Bone Joint Surg Am*. 1990;72(3):328-333.

34. Nelitz M, Guenther KP, Gunkel S, Puhl W. Reliability of radiological measurements in the assessment of hip dysplasia in adults. *Br J Radiol*. 1999;72(856):331-334.

35. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. New Jersey: Prentice-Hall; 2000.

36. American Physical Therapy Association (APTA). *Guide to Physical Therapist Practice*, 2nd edition. Alexandria, Va: APTA. 2003.

Appendix A

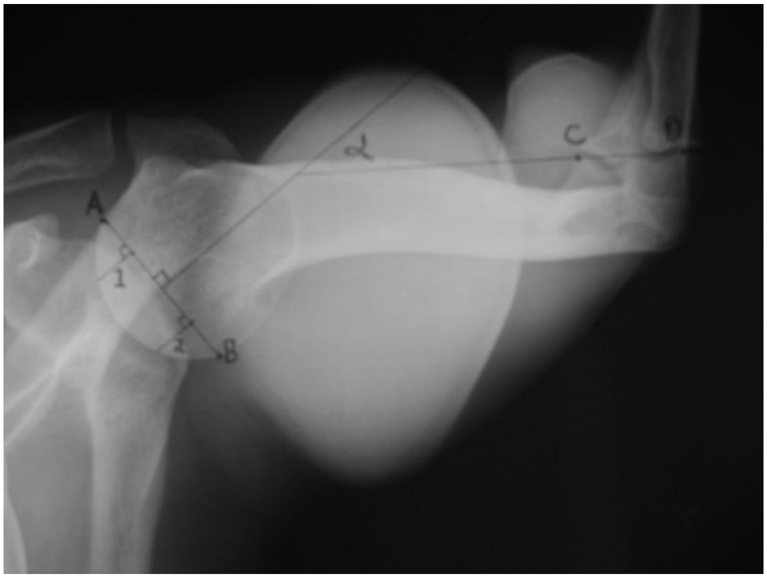

Figure 1. – Determination of humeral retroversion angle from radiograph.

Appendix B

Rater	ICC (95% CI)	Interpretation	Significance
Trained 1	0.89 (0.79-0.94)	Excellent	p<0.01
Trained 2	0.87 (0.77-0.94)	Good to Excellent	p<0.01
Untrained 1	0.89 (0.80-0.94)	Excellent	p<0.01
Untrained 2	0.89 (0.79-0.94)	Excellent	p<0.01

Table 1: Intra-rater reliability

Group	ICC (95% CI)	Interpretation	Significance
Trained Group	0.82 (0.72-0.90)	Good	p<0.01
Untrained Group	0.80 (0.56-0.89)	Good	p<0.01

Table 2: Inter-rater Reliability