Evaluating the Reliability of Four-Dimensional Computed Tomography Scans of the Wrist

by

Nicholas Bruce Chang MBBS

A thesis submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree of

Master of Science in Medicine

Clinical Epidemiology
Faculty of Medicine
Memorial University of Newfoundland

May 2018

St. John's
Newfoundland and Labrador
Canada

Table of Contents

ABSTRACT ..... **III**

ACKNOWLEDGMENTS ..... **IV**

DEDICATION ..... **V**

LIST OF TABLES ..... **VI**

LIST OF FIGURES ..... **VII**

LIST OF ABBREVIATIONS ..... **VIII**

CHAPTER 1: INTRODUCTION **1**

1.1 Background 1

1.2 Objective 2

1.3 History of Computed Tomography 3

1.4 Development of Four-Dimensional Computed Tomography 6

CHAPTER 2: DYNAMIC MUSCULOSKELETAL IMAGING **9**

2.1 Dynamic Imaging Modalities 9

2.2 Literature Search 12

2.3 Literature Review – Four-Dimensional CT 15

2.3.1 Hip *15*

2.3.2 Patellofemoral Joint *18*

2.3.3 Foot/Ankle *20*

2.3.4 Shoulder *23*

2.3.5 Hand *26*

2.3.6 Wrist *28*

2.3.7 Multiple Joints *35*

2.3.8 Excluded Studies *35*

CHAPTER 3: THE SCAPHOLUNATE JOINT **37**

3.1 Scapholunate Anatomy 37

3.2 Scapholunate Dissociation 40

3.3 Diagnostic Imaging of the Scapholunate Joint 42

CHAPTER 4: VALIDITY OF DIAGNOSTIC TESTS **49**

4.1 Accuracy and Reliability 49

4.2 Inter- and Intra-Rater Reliability 49

4.2.1 Percent Agreement *50*

4.2.2 Cohen's Kappa and Scott's π coefficient *50*

4.2.3 Fleiss' Kappa *52*

4.2.4 Gwet's AC_1 Coefficient *53*

4.2.5 Conger's Kappa *54*

4.2.6 Krippendorff's Alpha *54*

4.2.7 Brennan-Prediger Coefficient *55*

CHAPTER 5: METHODOLOGY 57
5.1 Research Proposal / Ethics Approval 57
5.2 Study Design 58
5.3 Statistical Analysis 61
CHAPTER 6: RESULTS 63
6.1 Inter-Rater Reliability Results 63
6.2 Subgroup Analysis of Upper Extremity Surgeons 65
6.3 Preference of View for Assessment of Scapholunate Instability 65
6.4 Intra-Rater Reliability Results 67
6.5 Repeat Inter-Rater Reliability Results 68
CHAPTER 7: DISCUSSION 69
7.1 Inter-rater coefficients 69
7.3 Standard Errors and 95% Confidence Intervals 72
7.3 Upper Extremity Subgroup 74
7.4 Surgeons' Preferred View 75
7.5 Study Limitations 75
CHAPTER 8: CONCLUSION 80
BIBLIOGRAPHY 82
APPENDIX A: SUMMARY OF LITERATURE REVIEW 92
APPENDIX B: HREB ETHICS APPROVAL LETTER 99
APPENDIX C: RPAC APPROVAL LETTER 101
APPENDIX D: LINK TO NORMAL 4D CT SCAN OF WRIST 102
APPENDIX E: LINK TO 4D CT SCAN OF SCAPHOLUNATE INSTABILITY 103
APPENDIX F: RAW DATA FOR INITIAL RATER RESPONSES OF SCAPHOLUNATE STABILITY 104
APPENDIX G: RAW DATA OF SURGEONS' PREFERRED VIEWS OF 4D CT SCANS 105
APPENDIX H: RAW DATA FOR RESPONSES OF SCAPHOLUNATE STABILITY AT 3 MONTHS 106

Abstract

Introduction: Four-dimensional CT (or 4D CT) scans are a novel approach to diagnosing musculoskeletal pathology. Although still in its infancy, there has been a surge of interest in identifying clinical applications for musculoskeletal 4D CT. The scapholunate joint has received the most attention thus far due to the complex articulations and challenges faced with prompt diagnosis of scapholunate injuries. The objective of this thesis is to review current literature on musculoskeletal 4D CT and to evaluate the inter- and intra-rater reliability of the assessment of scapholunate stability in 4D CT wrist scans.

Methodology: 4D CT scans of thirteen healthy volunteers and four patients were prepared. Seven orthopaedic and plastic surgeons were recruited to qualitatively assess the stability of the scapholunate joint in the 4D CT scans. Statistical analysis included percent agreement, Fleiss' kappa, and Gwet's AC1 coefficient.

Results: The percent agreement amongst all raters was 0.80392 (95% CI: 0.675 - 0.932). Fleiss' Kappa was 0.54895 (95% CI: 0.252 - 0.846) and Gwet's AC_1 was 0.54895 (95% CI: 0.391 - 0.915). The intraclass correlation coefficient (ICC) for intra-rater reliability was 0.71631 (95% CI: 0.5567 – 0.8423).

Conclusion: Our pilot study suggests good inter- and intra-rater reliability for the qualitative assessment of scapholunate instability in 4D CT scans. Although further studies are required, this thesis highlights the vast potential of 4D CT as a non-invasive diagnostic technique of dynamic musculoskeletal injuries.

Acknowledgments

This thesis would not have been possible without the patience, guidance, and support of several individuals.

To Dr. Frank O'Dea and the Division of Orthopaedic Surgery at Memorial University. Thank you for your unconditional support in enabling and encouraging me to pursue a Master's degree alongside my residency training program.

To the participating orthopaedic and plastic surgeons. Thank you for taking the time out of your day to partake in this project. I am grateful for your patience and tolerance.

To Dr. Paul Kelly. Thank you for sharing your contributions to the project and for your help navigating the computer software.

To Dr. John Hopkins. Thank you for your assistance with the technical aspects of four-dimensional computed tomography. Your enthusiasm for the subject was inspiring.

To my supervisory committee: Dr. Andrew Furey, Dr. John Harnett, and Dr. Craig Stone. Thank you all for your selfless time and valued advice. I could not have asked for a more dedicated and devoted group.

To my supervisor and mentor, Dr. Andrew Furey. Thank you for always having a solution for every setback we encountered. This undertaking would not have been possible without your guidance and leadership.

Thank you all, words will never truly express my profound gratitude and appreciation.

Dedication

This project is dedicated to my family. To my mother and father, Cindy and Bruce, who instilled in me the true value of perseverance and work ethic. To my brothers, Ryan, Justin, and Alex, for our unwavering bond. And to my partner, Hannah, for her unconditional love and support despite all the obstacles we have faced. Thank you for always being there for me.

List of Tables

Table 1 - Inclusion / Exclusion Criteria 14

Table 2 - Included Articles Classified by Joint 14

Table 3 - Geissler Arthroscopic Classification 44

Table 4 - Staging Algorithm for Scapholunate Dissociation 45

Table 5 – Management of Scapholunate Dissociation 47

Table 6 - Altman Interpretation of Cohen's Kappa 51

Table 7 - Distribution of Scapholunate Stability Responses by Rater 64

Table 8 - Inter-rater Variability of Scapholunate Instability 64

Table 9 - Inter-rater Variability of Scapholunate Instability (Upper Extremity Surgeon Subgroup) 65

Table 10 - Distribution of Preferential View of Each Rater 66

Table 11 - Inter-rater agreement on preferred views for assessing scapholunate instability 66

Table 12 - Intra-Rater Reliability Results 67

Table 13 - Repeat Inter-Rater Reliability Results 68

Table 14 - Sample Size as per the Desired Error Margin 76

Table 15 - Number of Raters as per Desired Variation Coefficient 77

List of Figures

FIGURE 1 - REPRESENTATION OF A SINGLE-SLICE SCANNER VS. A MULTI-SLICE SCANNER.....7

FIGURE 2 - LITERATURE SEARCH FLOW DIAGRAM 13

List of Abbreviations

4D CT – Four-Dimensional Computed Tomography

AC – Acromioclavicular

ACL – Anterior Cruciate Ligament

CIC – Carpal Instability Complex

CID – Carpal Instability Dissociative

CIND – Carpal Instability Nondissociative

DISI – Dorsal Intercalated Segmental Instability

FAI – Femoroacetabular impingement

ICC – Intra-class Correlation Coefficient

MRI – Magnetic Resonance Imaging

MSCT – Multi-Slice Computed Tomography

SC – Sternoclavicular

TG – Trochlear Groove

TT – Tibial Tuberosity

Chapter 1: Introduction

1.1 Background

One of the most profound technological advancements witnessed in the field of orthopaedics has been the development of diagnostic imaging. Diagnostic imaging is ubiquitous across all facets of orthopaedic surgery. It is an essential tool for diagnosis, prognosis, pre-operative planning, intra-operative visualization, and clinical surveillance.

The first documented plain x-ray was performed in Germany in 1895 (Waters, 2011). Since then, its form and function has continuously evolved. Modern diagnostic imaging techniques in orthopaedics include plain x-rays, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and dynamic fluoroscopy. Three-dimensional digital reconstructions of axial scans can be produced within seconds. The digitalization of diagnostic imaging has enabled instantaneous access to local and remote imaging, and has encouraged collaboration within and between specialties.

Four-dimensional CT (4D CT), also known as dynamic CT, is a recent technological advancement that has witnessed a surge of interest over the last couple of decades. Prior to the introduction of 4D CT, 'dynamic' CT scans referred to static CT images of joints held in dynamic positions (e.g. knee flexion or extension). As implied, 4D CT adds the dimension of time to axial and three-dimensional imaging. Initial clinical utility was focused on cardiac imaging and targeted radiotherapy. However, advanced imaging

techniques have enabled dynamic imaging of musculoskeletal joints in motion. Although still in its infancy, the potential clinical applications of four-dimensional musculoskeletal imaging are vast. Early research has evaluated anatomic structures such as the human scapholunate joint, patellofemoral joint, acromioclavicular joint, and femoroacetabular joint.

1.2 Objective

Dynamic imaging of the scapholunate joint has been highlighted as a topic of interest due to the challenges faced with prompt diagnosis of scapholunate ligamentous injuries. Late recognition of a scapholunate ligament injury may already demonstrate irreversible osteoarthritic changes. Current diagnostic techniques for investigating the scapholunate joint include plain x-rays, 'stress' views, CT, and MRI. However, these modalities are far from perfect. 4D CT has shown promise as a diagnostic approach to scapholunate injuries. A limited number of studies have described various methods to measure scapholunate widening in both normal and abnormal wrists in 4D CT scans. However, there is no literature to date on the inter- or intra-observer variability of 4D CT scapholunate scans of in-vivo subjects. The objective of this thesis is to investigate inter- and intra-observer reliability in the assessment of scapholunate four-dimensional CT scans.

1.3 History of Computed Tomography

In order to evaluate potential clinical applications for 4D CT, it is helpful to have a broad understanding of how it works. This paper does not intend to delve into the highly-technical physical properties of 4D CT image acquisition. However, a brief history of the evolution of CT scanners sheds light on modern 4D CT imaging capabilities.

Prior to axial CT imaging, conventional plain film x-ray was confined to two-dimensional representations of 3D anatomical structures. The first iteration of the CT scanner was developed by Sir Godfrey Hounsfield (Hounsfield, 1973). In order to acquire a head CT scan, the head was conceptually divided into axial 'slices'. An x-ray tube would transmit a narrow beam of x-rays through the head, which was received by a linked scintillation detector positioned directly opposite to it. 160 images were taken as the x-ray tube and detector scanned across the head in a linear fashion. This motion is referred to as translation (Goldman, 2007). At the end of each transmission, the entire gantry would be rotated 1°, and another set of 160 images would be taken. After a total of 180 rotations, the images were processed by a mini computer. The absorption coefficient per corresponding volumetric unit was calculated and corresponded with absolute values of absorption coefficients for various tissues. The processed image could subsequently be visualized on a cathode-ray tube, and printed using a line printer.

In 1974, the emergence of second generation CT scanners substantially reduced the scan time (Goldman, 2007). Like the first-generation CT scanner, it also utilized a translation-rotation motion. However, multiple narrow-beams and detectors were incorporated, which decreased the scan time for each translation. For example, CT scanners with twenty narrow beams and detectors would take 1/20 of the time as the first-generation CT scan. This was clinically important, as slices could now be performed in less than 20 seconds, allowing a patient to hold their breath for the entire duration (ibid).

Third generation scanners eliminated the translation motion (Goldman, 2007). To achieve this, the single x-ray beam was expanded into a fanbeam, which could incorporate the entire width of the body. Multiple detectors would receive the x-ray beams. These detectors were linked to the x-ray tube, and would rotate together around the body. This rotation-rotation motion was significantly smoother than the previous translation-rotation designs, and scans of each slice could be completed in mere seconds.

By 1976, fourth generation CT scanners were introduced. In this 'rotate-fixed' design, only the x-ray tube would rotate around a patient (Goldman, 2007). The x-ray beams were received by a fixed ring of detectors around the patient. Similar to the 3rd generation CT scanners, this also produced a fanbeam view. Although introduced following the 3rd generation, there were no obvious benefits of the 4th generation CT

scanner over the previous generation. One drawback was the large ring diameter required for the multi-detectors. The expanding number of detectors grew too costly for further development of the 4th generation scanners. The design was eventually abandoned, and modern conventional CT scanners are still based on the 3rd generation design.

However, one of the issues with both 3rd and 4th generation CT scanners at the time was the significant interscan delay. After every 360° slice was performed, the gantry would need to be rotated back before a new slice was scanned. The interscan delay would take 8-10 seconds, as opposed to the 1-2 seconds for each slice acquisition (Goldman, 2007). Advances in CT imaging techniques remained stagnant, until slip-ring technology was developed in 1987. Slip rings enabled the conductive transfer of electrical power to the rotating gantry, while the image data was wirelessly transmitted (ibid). This allowed for continuous rotation of the gantry, without the need for braking and reversal. However, there was still a delay from the stepwise translation of the table through following every slice acquisition. This was known as axial step-and-shoot CT.

To address this issue, the concept of helical CT (also known as spiral CT) was developed (Kalender, et al., 1990). This strategy involves continuous rotation of the x-ray beam while the table and patient are simultaneously slid through the gantry in a smooth fashion. As a result, path of the x-ray beam travels in a spiral or helical pattern along the

z-axis (Goldman, 2007). Images would be processed using interpolation and were reconstructed into axial slices at any given point along the z-axis. Quality and spatial resolution of helical CT scans varied depending on the slice thickness and helical pitch (distance the table moves per complete rotation of the gantry). As the quality of CT scans increased and scan times decreased, helical CT scanning became the standard of care by the mid-1990s (Goldman, 2007).

1.4 Development of Four-Dimensional Computed Tomography

The concept of musculoskeletal four-dimensional CT was not technologically feasible until the development of multi-slice CT in 1998. Multi-slice CT (MSCT) utilizes a multi-row detector array whereas previous CT scanners used only a single-row detector array (Hu, 1999) (see Figure 1). This expanded the field-of-view in the z-axis, by acquiring multiple slices at different z-points simultaneously. This enabled the entire joint to be scanned without having to translate the CT table. This also significantly improved scanning time and temporal resolution. The first multi-slice CT scanner was capable of 4-slice image acquisition. Subsequent advancements led to 16-slice (Ropers, et al., 2003), 64-slice (Nikolaou, et al., 2004), 256-slice, and 320-slice CT scanners (Hsiao et al., 2010). The 320-slice scanners are capable of imaging 16cm in the z-axis per rotation (Pan, 2013).

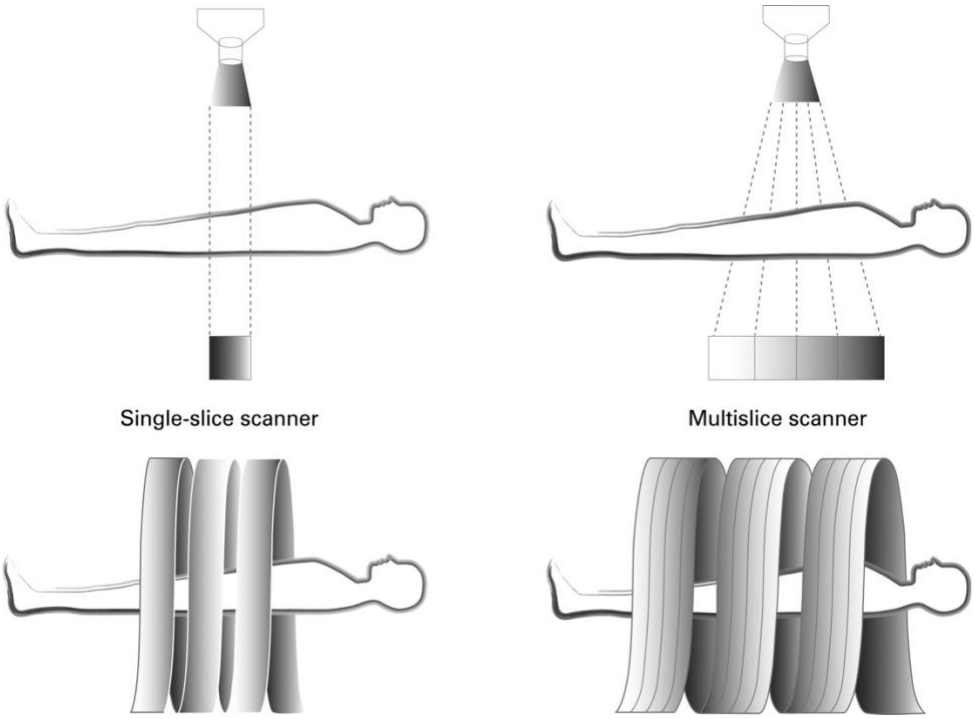

Figure 1 - Representation of a single-slice scanner (left) vs. a Multi-slice scanner (right) (Roberts, et al., 2008)

Initial clinical applications for multi-slice CT focused on cardiac CT imaging. Prior to 4D CT, axial imaging of the heart was challenging because of its dynamic motion. However, the wide field-of-view capabilities of multi-slice CT enabled simultaneous imaging of the entire heart. Images could now be acquired at any point in the cardiac cycle. Contrast administration with CT angiography enabled digital reconstructions of the pulmonary arteries, aorta, and coronary arteries (Nikolaou, et al., 2004).

The clinical utility of 4D CT quickly spread to other specialties. Four-dimensional CT was found to be advantageous in radiotherapy. This is because thoracic tumours are prone to intrafraction variability, which is the variation of a tumour's location and size due

to breathing, heart motion, and gastrointestinal motion (Kipritidis, et al., 2015). For radiotherapy planning, the clinical tumour volume (CTV) included both the gross tumour volume (GTV) and any potential microscopic disease. This area was known as the clinical tumour volume (CTV). To compensate for the dynamic positioning of the tumour, a wider zone of radiation was required. This was referred to as the planning target volume (PTV).

Advances in four-dimensional CT permitted four-dimensional tracking of tumour position with respiratory motion. When synchronized with the radiation beam, this significantly reduced the amount of collateral irradiation of healthy tissue, whilst ensuring the desired radiation coverage of the tumour (Wang, et al., 2009). Similar findings have also been documented for breast cancer, liver tumours, and unresectable pancreatic tumours (Wang, et al., 2012; Yeo, et al., 2014; Cattaneo, et al., 2010). Four-dimensional CT has also been utilized to accurately localize parathyroid adenomas for pre-operative planning (Hunter, et al., 2012). The volume of research pertaining to four-dimensional CT has exploded over the last decade. One of the more promising fields is musculoskeletal 4D CT.

Chapter 2: Dynamic Musculoskeletal Imaging

2.1 Dynamic Imaging Modalities

One of the drawbacks with conventional musculoskeletal diagnostic imaging techniques is that static images are taken at a single snapshot in time. This can prove challenging to radiographically diagnose musculoskeletal pathologies that have a predominant dynamic component, such as ligamentous injuries. Radiologists look for clues that imply injury to a ligament. For example, edema at a ligamentous insertion site on T2-weighted MRI slices indicates inflammation. Abnormal contour, discontinuity, or absence of the ACL ligament suggests a rupture. Widened diastasis between two bones on x-ray may imply a ligamentous injury. However, these clues are not always apparent, and there are several limitations with these diagnostic techniques. Subacute or chronic ligamentous injuries may no longer show inflammation or edema on MRI. Abnormal diastasis between the scaphoid and lunate may be a late radiographic finding, which may be only found after irreversible changes to wrist dynamics.

Imaging techniques have been developed to assist with diagnosis of dynamic musculoskeletal injuries. As the name suggests, stress view radiographs are x-rays taken when a joint is 'stressed' to accentuate joint instability. Numerous radiographic stress views have been described, including nine exclusively for evaluating the scapholunate joint (Lee et al., 2011). The stress force can be provided by an assistant (e.g. external

rotation stress of the ankle syndesmosis joint), by the patient (e.g. clenched-fist stress view), or by gravity (gravity-stress view of the ankle). There are several limitations with radiographic stress views. First, it can be difficult to define the difference between a normal and abnormal stress view. Reference values are available in the literature, but these are based on case series with relatively small sample sizes (Ozçelik et al., 2005). There is also variability in ligamentous laxity amongst 'normal' individuals (Park, 2002). Second, reproducibility of stress views is difficult. X-rays are two-dimensional representations of three-dimensional structures. Any change in distance, angulation, or rotation can introduce variation in measurements. Third, standardization of the amount of force in Newtons required for each 'stress' has not been translated into clinical practice.

Dynamic fluoroscopy, or videofluoroscopy, is another modality that evaluates dynamic motion. Like standard radiographs and CT-scans, x-rays are utilized to produce images. While early iterations of radiographs produced images on film, fluoroscopes displayed the x-ray images on a fluorescent screen. This enabled dynamic, real-time videos. Technological advancements have modernized fluoroscopes by replacing the fluorescent screen with digital intensifiers. In addition, mobile C-arm units have eased their intra-operative utility. Real-time imaging enables confirmation of implant position in multiple planes.

Stress maneuvers may also be performed intra-operatively in real-time. The syndesmosis ligament of the ankle is commonly assessed intra-operatively to determine its stability. Limitations concerning the accuracy and reproducibility of real-time fluoroscopy have been reported. One study determined that syndesmosis malreduction up to 30 degrees of external rotation may go undetected using intra-operative fluoroscopy (Marmor, et al., 2011).

Dynamic ultrasound has shown potential as a dynamic imaging modality. Advantages of ultrasound are its low cost, lack of radiation, and capability of incorporating time as a dimension. Musculoskeletal applications of ultrasound include the dynamic assessment of the rotator cuff, acromioclavicular joint, extrinsic and intrinsic wrist ligaments, syndesmosis, peroneal tendon, and ulnar collateral ligament of the elbow (Petscavage-Thomas, 2014). Unfortunately, dynamic ultrasonography is limited by its poor penetration of bone, susceptibility to artefact, and the requirement for a skilled technician (Finnoff, 2017).

Protocols for evaluating joint mechanics using kinematic and dynamic MRI have been introduced over the last couple of decades. Kinematic MRI scans require images to be taken repeatedly at different joint positions, including their 'stressed' positions. Dynamic MR images are produced in real-time throughout a joint's range of motion (Gold, 2003). Biomechanical studies have evaluated patellofemoral, glenohumeral, and wrist

joints (Draper, et al., 2009) (Hodge, et al., 2001) (Boutin RD, et al., 2013). Whilst extremely promising, dynamic MRI is currently limited by its technological constraints. Image resolution is accurate to within 2 mm, and requires slow motion of the joint (maximum speed 217 mm per second in a 1.5T scanner and 38 mm per second in a 0.5T scanner) (Draper, et al., 2008).

Four-dimensional CT scans has the potential to avoid some of the limitations of other dynamic imaging modalities. The modality is less operator dependent. Image resolution is superior with the capacity for future improvement. Dynamic CT scans can be performed quickly, with relatively low doses of radiation. Dynamic CT protocols have been developed to evaluate a wide range of joints.

2.2 Literature Search

A comprehensive literature search was conducted to review musculoskeletal applications of 4D CT scans. This was performed in April 2017, using PubMed MEDLINE, Embase, and Cochrane databases (see Figure 2). The MESH/Emtree term 'Four-Dimensional Computed Tomography' was utilized for the MEDLINE and Embase searches respectively. Other search terms included 'dynamic CT', 'dynamic computed tomography', and 'joint'. Boolean operators were utilized. A publication filter was applied to include articles from 2007 onwards. This is the year where 256- and 320- multi-slice CT

scanners were introduced. These scanners had the wide field-of-view capabilities ideal for performing musculoskeletal four-dimensional CT scans (Kwong, et al., 2015).

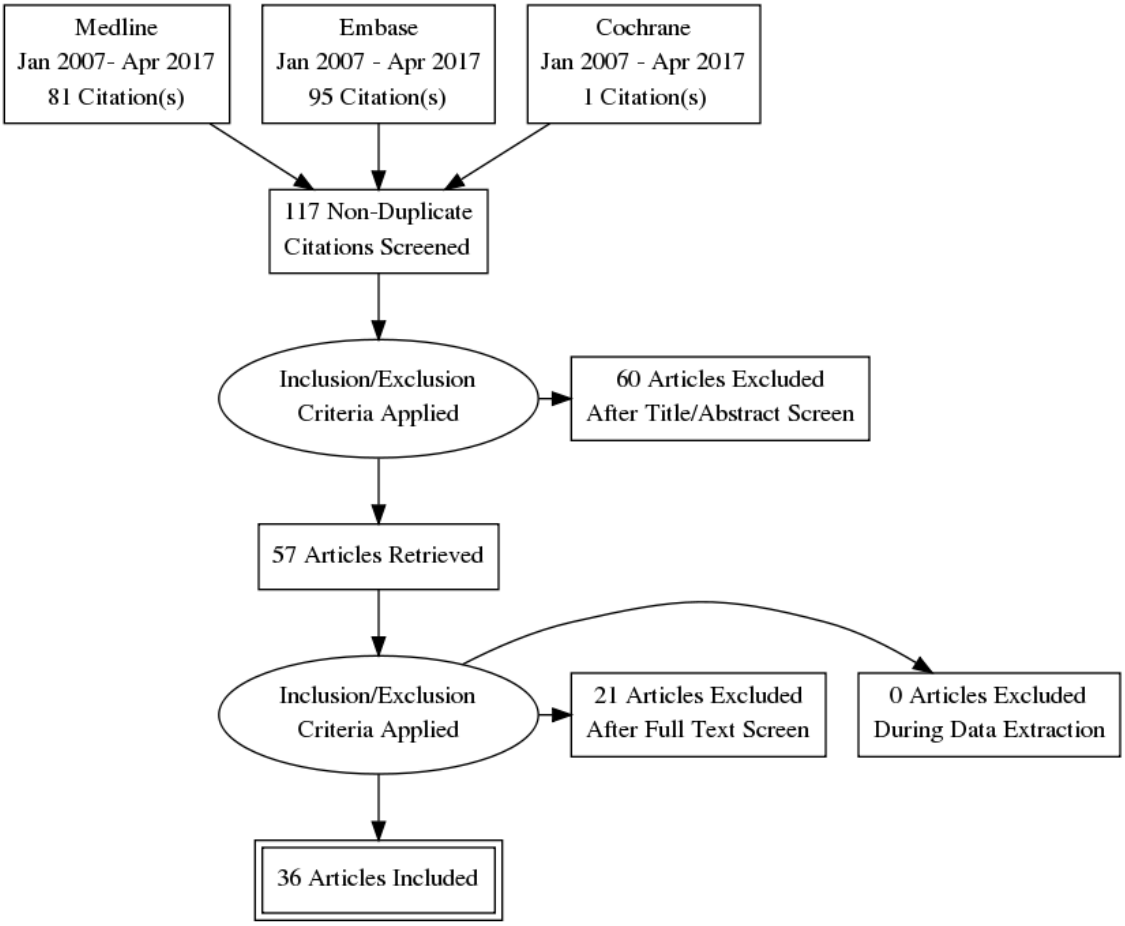

Figure 2 - Literature Search Flow Diagram

A total of 177 articles were found. The articles were uploaded into the RefWorks citation manager. Sixty duplicate articles were identified and removed, leaving 117 articles. A title and abstract screen was completed, using the pre-determined inclusion

and exclusion criteria (see Table 1). The remaining 57 articles were retrieved and the full texts were reviewed (see Table 2). Of these, 36 articles were included in the final literature review. A summary chart of the literature review is listed in Appendix A.

Table 1 - Inclusion / Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Study using Four-Dimensional CT	Static CT of dynamic joint positions
Study involving musculoskeletal joint(s)	Four-dimensional computer simulation studies
Human and/or cadaver subjects	Studies involving face (i.e. temporomandibular joint)
	Animal studies, review articles, technical/engineering/physics-based studies
	Non-English articles

Table 2 - Included Articles Classified by Joint

Joint	Number of Articles
Hip	2
Patellofemoral	6
Foot/Ankle	2
Shoulder	4
Hand	5
Wrist	16
Multiple Joints	1
TOTAL:	**36**

2.3 Literature Review – Four-Dimensional CT

2.3.1 Hip

Two studies investigated the utility of pre-operative assessment of femoroacetabular impingement (FAI). In one study, thirty patients with a clinical and radiographic diagnosis of FAI were prospectively recruited (Wassilew, et al., 2013). Both cam-type FAI lesions of the femur and pincer-type lesions of the acetabulum were included. The patients were consented for four-dimensional CT scanning of their affected hip, followed by surgical treatment. The dynamic CT scans were performed in a 320-slice CT scanner (Aquilion One, Toshiba, Nasu, Japan), capable of a 16-cm wide field-of-view. Patients were positioned supine and placed head-first into the CT scanner. The patients were asked to perform two active, dynamic movements of their hip. First, patients were asked to flex their affected hip, with their leg kept in full extension. Second, patients were asked to perform a simultaneous abduction and external rotation maneuver. A belted contraption was utilized to help guide the patient's leg through a standardized, consistent motion. The CT image data were uploaded to a specialized software called Vitrea® (Vital Images, MN, USA). Dynamic cine-images were viewed in 2D and 3D planes by 3 reviewers, consisting of 2 musculoskeletal radiologists and 1 orthopaedic surgeon. The reviewers assessed for subluxation of the hip joint secondary to mechanical impingement. They defined this as either lateralization of the femoral head relative to the deepest portion of the acetabulum or a vacuum phenomenon noted during the dynamic movements.

Surgical management for all patients consisted of an open hip dislocation with an acetabuloplasty and/or femoral head-neck osteoplasty. The pre-operative dynamic CT scan were compared with the intra-operative findings. The operating surgeon was blinded from the pre-operative assessment.

Results revealed that pre-operative dynamic CT was highly accurate for predicting FAI intraoperative findings. Dynamic CT was also able to determine between anterior and posterior subluxation of the hip with high inter-observer agreement. The kappa (κ) values ranged from 0.87 to 0.92 and 0.72 to 0.94 respectively. Anterior subluxation was detected with a sensitivity of 96% and specificity of 100%, whereas posterior subluxation was detected with a sensitivity of 100% and specificity of 78%.

Whilst the results are promising, the article noted some limitations. Femoroacetabular impingement is associated with labral tears. Unlike magnetic resonance imaging (MRI) or magnetic resonance arthrograms (MRA), dynamic CTs are unable to delineate any cartilaginous pathology. The average effective radiation dose (9.8 mSv) of the dynamic CT scan was higher than the mean effect radiation dose for a standard CT pelvis (7.5 mSv) (Sodickson, et al., 2009). In addition, the patients recruited to the study initially presented with early osteoarthritic symptoms and radiographic findings. It was difficult to discern whether the early osteoarthritis occurred secondary to FAI, or if impingement symptoms developed secondary to osteoarthritis. Finally, no post-

operative clinical correlation was performed to see if pre-operative assessment of dynamic CT was associated with better patient outcomes.

A second study prospectively performed 4D CT scans of the hip in patients prior to hip arthroscopy for femoroacetabular impingement (Fernquest, et al., 2017). Fifty patients underwent passive flexion, adduction, and internal rotation within the CT scanner. Measurements of the alpha angle between the femoral neck and head were determined at every 30° increment. Using a clockface concept, measurements were taken in the 11, 12, 1, 2, and 3 o'clock positions with the 1, 2, and 3 o'clock positions corresponding to the anterior aspect of the femoral head.

Results demonstrated that the mean point of impingement in this cohort was at 41.36° of flexion. Alpha angles measured at the 2 and 3 o'clock positions were significantly associated with reduced hip flexion. In addition, the value of alpha angles measured in the 2 o'clock position was inversely correlated with hip flexion and pelvic tilt. In other words, cam lesions located more anterosuperiorly were more likely to cause impingement. The larger the cam lesion, the greater reduction in hip flexion and the greater degree of pelvic retroversion.

While the study demonstrated the feasibility of evaluating femoroacetabular impingement dynamically, it did highlight potential issues with 4D CT. Range of motion

of the patients' hips were restricted secondary to pain in 25.3% of patients and secondary to the physical constraints of the CT scanner in 20.5%. Current CT dimensions and designs may limit hip 4D CT scans by patient size and body morphology. Nevertheless, these studies suggest that dynamic hip CT scans are accurate and feasible, and future research can help determine whether these scans can assist with the diagnosis, management, pre-operative planning, or prognosis of FAI.

2.3.2 Patellofemoral Joint

Stability of the patellofemoral joint is important for normal knee function and range of motion. The patellofemoral joint is a complex articulation influenced by several anatomic factors. These factors can be static or dynamic. Stability is influenced by the shape of the patella, shape of the trochlear groove, integrity of the medial patellofemoral ligament, and q-angle. Dynamic stability is afforded by contraction of the vastus medialis and lateralis muscle. Diagnostic imaging modalities have been used to evaluate for indicators of patellar instability. Plain radiographs of the patella may show evidence of patellar subluxation and/or dislocation, excessive patellar tilt, patella alta, or an abnormal congruence angle. Superimposition of axial CT slices enables measurement of the distance between the tibial tuberosity and trochlear groove (TT-TG distance) (Dejour, et al., 1994). An increase in the TT-TG distance is associated with patellar instability. One potential issue with the TT-TG distance is that the measurements are derived from a static CT scan, with complete relaxation of the extensor mechanism. Static MRI has been used

to evaluate the integrity of the medial patellofemoral ligament and assess for osteochondral injuries.

Early 'dynamic' CT scans of the patellofemoral joint involved static CT scans of patients actively contracting their quadriceps (Guzzanti, et al., 1994). Guzzanti et al. reported 52% of their cases revealed radiographic alteration of patellofemoral stability with active quadriceps contraction.

4D CT enables direct visualization of the patellofemoral articulation during active patient flexion and extension. Two studies have been performed utilizing 4D CT in patients with patellar instability. Symptomatic knees with patellar instability were compared to the patient's contralateral asymptomatic knee. In one study, the 4D CT scans revealed a statistically significant difference in the TT-TG distance, trochlear groove depth, patellar height ratio, and patellar lateral displacement between the symptomatic and asymptomatic knee (Demehri, et al., 2014a). Williams et al. noted a significant difference between the symptomatic and asymptomatic sides in patellar tilt and bisect offset when the knee was at 5° and 15° of flexion (Williams, et al., 2016). A statistically significant difference for the TT-TG distance was found at 5°, 15°, 25°, and 35° of flexion.

One study compared manual and semi-automatic measurement of patellar displacement and patellar tilt in dynamic CT scans (Forsberg, et al., 2016). Their semi-

automatic method utilized a specialized computer software program. It required manual selection of pre-defined anatomic points on the 3D-reconstructed CT scans. The software automatically calculated the distances and angles between the points throughout the duration of the cine-clip. When compared to manual measurements, the authors determined the semi-automatic method had better inter- and intra-observer reliability. In addition, the semi-automatic method was anecdotally found to be more advantageous. It allowed continuous imaging and automatic measurements throughout the joint's range of motion, rather than selecting static slices for each manual measurement.

Four-dimensional CT scans have also been utilized for research purposes to evaluate patellar-stabilizing procedures (Cosgarea, et al., 2013) (Elias, et al., 2014) (Gobbi, et al., 2016). Qualitative and quantitative analyses of patellar tracking were compared between pre- and post-operative dynamic CT scans. These studies have demonstrated the vast potential for research and clinical applications of 4D CT patellofemoral scans.

2.3.3 Foot/Ankle

Two studies have evaluated 4D CT of the foot and ankle. A proof of concept study evaluated subtalar joint motion and stability (Gondim Teixeira P. , et al., 2017a). The study was composed of two arms: one evaluating cadaveric specimens, and one evaluating healthy volunteers. Three cadaver legs were scanned three times using 4D CT. For the first scan, the leg was placed in the anatomic position on the scanner table. An evaluator

controlled an elastic band wrapped around the hindfoot to guide the subtalar joint through a pronation-supination motion. Prior to the second scan, one of the primary stabilizers of the subtalar joint (cervical or interosseous ligament) was dissected and resected. The third and final scan would be performed after the remaining ligament was released.

In the other study arm, eighteen healthy volunteers were recruited. They were positioned supine on the scanner table. The volunteers were given prior training on how to guide their foot through pronation and supination using the elastic band. All CT data sets were processed using a specialized imaging software. Reference points were manually placed on pre-determined anatomic landmarks to enable semi-automatic processing of various distance and angular measurements. The authors measured the coronal talocalcaneal angle, axial talocalcaneal angle, medial facet gap, and posterior calcaneal facet uncovering.

The results showed that resection of the stabilizing ligaments of the subtalar joint increased the joint amplitude (19% partial resection vs 22% complete resection). The medial facet gap measurement was identified as a potential clinically important value, as it showed the greatest change after ligament resection. Overall inter-observer variability was rated between moderate and excellent. Limitations to the study include a small sample size, absence of pathologic specimens, and the non-physiologic stress upon the

subtalar joint (non-weight-bearing). In addition, the subtalar joint is a complex joint with multiple secondary stabilizers, such as the inferior extensor retinaculum, calcaneofibular ligament, and calcaneotibial ligament. Injury to these secondary stabilizers depends on the mechanism of injury and position of the foot at the time of injury. The study did not individually isolate and resect the secondary stabilizers separately. However, the article successfully demonstrated the technical feasibility of 4D CT scans of the subtalar joint. Further research should be directed towards the clinical applicability of these scans.

Gondim Teixeira et al. also conducted a pilot study utilizing a wide field-of-view CT scanner on a phantom device and a cadaver ankle (Gondim Teixeira P. , et al., 2017b). Their goal was to identify which variables could be controlled to improve image quality. They determined that the most important factor for high image quality was volume acquisition speed. This determines the maximum velocity of joint motion that can be adequately imaged. In addition, image quality was influenced by the motion characteristics of the joint. Image quality was greatest at the highest amplitudes of joint range of motion (e.g. terminal flexion and extension), whereas the resolution was reduced in the middle of the joint's range. Finally, the orientation of joint motion affected the final image quality. Motions parallel to the CT gantry rotation axis (e.g. wrist pronation/supination or hip internal/external rotation) had more intense imaging artefacts compared to perpendicular motions (e.g. flexion/extension).

2.3.4 Shoulder

The role of four-dimensional CT for the shoulder is still in its infancy. Of the four articles published, three are case reports or case series. Curiously, no studies have evaluated the glenohumeral joint. One reason may be because the physical dimensions of the gantry limit the dynamic movements possible within the CT scanner. This may impede evaluation of the full range of motion of the shoulder. However, articles have described 4D CT scans of the acromioclavicular, stenoclavicular, and scapulothoracic joints of the shoulder.

Two papers have investigated the acromioclavicular joint using dynamic CT. Alta et al. studied acromioclavicular joint motion in healthy volunteers (Alta, et al., 2012). They were initially positioned supine on the CT scanner table. Over a 7-second period, the volunteers were asked to forward elevate their shoulder to 90° and place their hand on the outer ring of the gantry. Then, the volunteers would adduct their arm whilst actively pushing against the gantry. The resistance of the gantry would effectively place a loading force on the acromioclavicular joint, simulating the Bell-van Riet (BvR) clinical examination test (van Riet & Bell, 2011). Their results revealed that the clavicle consistently translates posteriorly and superiorly (mean 1.1±0.9 mm and 0.6±0.5mm respectively) when loaded in an adducted position.

Dyer et al. authored a case report regarding the use of 4D CT for a young lady with a previous acromioclavicular dislocation (Dyer, et al., 2015). X-rays initially revealed a low-grade AC injury. However, clinical examination and failure of non-operative management suggested a higher-grade dislocation. For the scan, the patient was also placed supine on the CT table. Three movements were conducted in succession. The first maneuver was forward flexion of the shoulder to 45°. The second maneuver was internal rotation to 90°. The third maneuver was adduction of the shoulder to 45°. The dynamic CT scan revealed significant superior and posterior translation, exceeding than the normal values reported by Alta et al. The joint space widening was 5.3mm/9.4mm/7.8mm after each respective maneuver (ibid).

Despite their low level of evidence, these two papers on the acromioclavicular joint do suggest that it is technically possible to visualize four-dimensional loading of the acromioclavicular joint within a CT scanner. Limitations were noted in the studies. Due to the CT scanner physical constraints, patients were placed supine. Physiologic loading of the acromioclavicular joint due to gravity is normally in the superior-inferior plane. Within a CT scanner, this is now a posterior-directed force. Attempts were made to elevate the scapula on the CT scanner to alter this vector (Alta, et al., 2012). In addition, the effective radiation dose was relatively high in both studies (2.5-3.5mSv and 5.0301mSv). This was due to the relatively long duration of the scans and the high organ-specific conversion factor for the neck (Alta, et al., 2012). Despite these concerns, these

studies have demonstrated clinical potential for dynamic CT scans of the acromioclavicular joint.

A case report described the use of dynamic CT for evaluation of a 69 year old female with atraumatic posterior instability of the stenoclavicular joint (Hislop-Jambrich, et al., 2016). Clinical examination of the patient was limited by pain. Dynamic CT scans were performed on her bilateral sternoclavicular joints. The dynamic movements visualized were forward flexion to 45° and cross-arm adduction to 45°. This revealed significant asymmetry in the sternoclavicular joint motion and pathologic narrowing between the tracheal lumen and medial end of the clavicle. Analysis of the dynamic CT scan subsequently led to a sternoclavicular joint-stabilizing procedure.

Bell et al. conducted a retrospective case series of 12 patients with snapping scapula syndrome. In this uncommon disorder, the superomedial corner of the scapula impinges upon the posterior thoracic cavity. In some cases, soft tissue creates a tethering effect on the scapula (Bell, et al., 2015). Dynamic CT scans were utilized to pre-operatively determine the amount of bone for resection. Patients were positioned on the CT table and asked to recreate their 'snapping' maneuver. The results did not reveal any significant difference in measurements between the symptomatic or asymptomatic sides. However, only the symptomatic side was dynamically visualized.

Posterior sternoclavicular instability and snapping scapular syndrome are uncommon disorders. High-quality research studies are not realistically feasible given their low prevalence. Nevertheless, these case series and case reports illustrate how dynamic CT scans can provide valuable clinical information that directs surgical management.

2.3.5 Hand

Five studies were identified where 4D CT was utilized to evaluate thumb or finger motion. Two were kinematic studies to investigate normal thumb range of motion. Goto et al. compared one healthy volunteer to one cadaver specimen. Thumb carpometacarpal flexion-extension, abduction-adduction, and pronation-supination were scanned using a wide field-of-view dynamic CT scanner (Goto, et al., 2014). In addition to determining normal range of motion values, Goto also evaluated the change in contact area at the 1st carpometacarpal joint. The results reported that the maximum calculated contact area highest during thumb palmar abduction, and lowest during thumb adduction. The authors suggested their results may contribute to the understanding of normal thumb biomechanics.

Kerkhof et al. evaluated thumb opposition in six cadaver specimens (Kerkhof, et al., 2016). A motion simulator was designed to passively guide the thumb to maximum opposition at a slow, constant pace. A 4D CT scanner tracked the movement of the 1st

metacarpal, trapezium, trapezoid, scaphoid, and distal radius. Digital measurements of joint motion from the 4D CT scans were recorded. These measurements were validated against the dynamic distances between multiple beads embedded in the bone or attached to the skin. The authors noted that the dynamic CT scan measurements of thumb motion were precise when compared to the bead validation tool. In addition, the authors observed significant motion of the 1st metacarpal and trapezium during thumb opposition, with little movement of the scaphoid. Kerkhof suggested that 4D CT is a valid tool for investigating in-vitro thumb kinematics, and may have a clinical role in diagnosing instability patterns.

Two studies reported on the use of 4D CT scans for post-operative evaluation of hand function. In a case series of 14 paediatric patients who had undergone pollicization of their index finger for various congenital deformities, 4D CT scans were utilized to demonstrate radiographic evidence of finger motion (Strugarek-Lecoanet, et al., 2016). A total of 23 pollicizations were reviewed. The authors noted 20 of the 23 thumbs had good motion at the reconstructed joint. Sixteen thumbs showed evidence of remodelling, with flattening of the metacarpal heads. Motion was observed in all three axes. Although the authors found the 4D CT scans helpful to observe the reconstructed joint mechanics, they suggest future long-term studies to investigate the natural history of osteoarthritis in the altered basilar joint mechanics. George et al. reported two cases of upper extremity allograft transplantation (George, et al., 2013). Although not the focus of the case series,

they commented on the use of 4D CT scan to record finger and thumb motion post-transplantation.

Chae et al. conducted a proof-of-concept study where 3D-printed thumb models were derived from 4D CT scans. (Chae, et al., 2015) A healthy volunteer's thumb was scanned during thumb abduction, opposition, and 'thumb pinch'. A 3D-printed model of the thumb in each representative position was produced, a process the authors referred to as "4D-printing." The angles between the 1st and 2nd metacarpals were measured in the coronal and sagittal planes using a goniometer. This was compared to digital measurements of these angles from the dynamic CT scans. The authors noted a small discrepancy between the values, with the goniometer over-estimating the angles. The authors believe that their study may set the ground work for 4D-printing to be utilized for pre-operative surgical planning in the future.

4D CT hand scans have a limited clinical role at present. Studies have shown that finger and thumb motion can be observed in real-time. Further studies may wish to evaluate pathological hand conditions to expand their clinical utility.

2.3.6 Wrist

Motion of the wrist is complex. Subtle injuries may compromise dynamic motion, which may lead to irreversible pathology. There is little surprise that it has been a major

focus of early research in musculoskeletal 4D CT. Studies have focused primarily on midcarpal articulation or the scapholunate joint. Case reports have also utilized 4D CT to assist with the diagnosis of capitate subluxation, pisotriquetral instability, and trigger lunate syndrome.

Two studies evaluated midcarpal stability using cadaveric specimens. In one study, a custom motion simulator took a cadaveric forearm and hand specimen through flexion-extension and radioulnar range of motion within a 4D CT scanner (Zhao, et al., 2015). They sought to measure the accuracy of measuring joint motion using proximity maps derived from specialized software. This was compared to measurements obtained from fiducial beads that were manually inserted on the carpal bones. Their results reported their measurements were accurate within 0.00–0.68° for rotation and within 0.02-0.30 mm for translation. The accuracy was comparable to other studies which utilized the similar image-based techniques. Jais et al. also utilized a motion simulator which brought a cadaveric, non-pathologic wrist through a range of motion (Jais, et al., 2014). Carpal hysteresis was evaluated. This is the lag of effect exhibited when the forces upon the wrist are changed. Carpal motion is not only dependent on the forces acting on it, but it is also influenced by its current and previous positions, direction, and velocity. Their study determined that the carpal hysteresis effect was highest in the lunate, followed by the triquetrum and scaphoid. Inter-observer reliability was greatest when quantifying hysteresis for the scaphoid.

Choi et al. reported a proof-of-concept study which demonstrated the feasibility of using dual-source CT to produce kinematic four-dimensional images of wrist motion (Choi, et al., 2013). Two healthy volunteers were recruited. Choi et al. found this method decreased scan time, increased temporal resolution, and reduced radiation exposure compared with single-source CT.

One study endeavored to visualize the dart throwing motion of the wrist and determine its rotational axis (Edirisinghe, et al., 2014). Health volunteers were recruited and trained to perform the dart throwing motion (extension-radial deviation to flexion-ulnar deviation). The 4D CT scans showed substantial motion at both the radiocarpal and midcarpal joints. The rotational axis of the dart throwing motion was calculated to be 27° of anteversion in the coronal plane 44° of varus in the transverse plane.

Repse et al. published two papers on his clinical experience with four-dimensional CT. A case report described how 4D CT was invaluable for diagnosing dynamic capitate subluxation (Repse, et al., 2015a). Static plain film radiographs and MRI did not suggest any abnormalities, but the 4D CT scan revealed dynamic instability of the capitate. The asymptomatic contralateral wrist was also imaged for a baseline comparison. Repse et al. also documented a case series of five patients with carpal instability nondissociative (CIND) (Repse, et al., 2015b). These patients demonstrated instability between the

proximal and distal rows of the carpal bones using 4D CT. Multi-planar reconstructions of the carpus were produced in the radio-lunate-capitate axis. Repse et al. described three independent pathologic findings on 4D CT for CIND: a vacuum phenomenon within the lunate-capitate joint, triggering of the lunate, and subluxation of the capitate. None of these findings were found in the asymptomatic wrists.

Two papers have evaluated the pisotriquetral joint. A small case series documented two cases of pisotriquetral instability (Demehri, et al., 2014b). In both cases, static imaging did not reveal any pathology. Diagnosis was assisted by the dynamic four-dimensional CT images and multi-planar reconstructions. Demehri et al. subsequently established a reference for normal range of motion of the pisotriquetral joint (Demehri, et al., 2015). Ten asymptomatic wrists were scanned using 4D CT while performing wrist flexion and extension motions. The anteroposterior (AP) and craniocaudal (CC) intervals were measured. Two MSK radiologists demonstrated high inter-observer reliability when measuring AP translation of the pisotriquetral joint (ICC = 0.80). The inter-observer reliability was not significant when measuring the craniocaudal interval.

One case report described a case of trigger lunate syndrome diagnosed using 4D CT (Troupis & Amis, 2013). The patient's clinical examination suggested mid-carpal instability. Abnormal lunate motion was assessed both qualitatively through visual assessment, and quantitatively through the arc of the radiolunate angle. Interestingly,

there was no evidence of lunate subluxation. Instead, an abnormal cessation of lunate motion was detected within its arc of motion. This unique finding had never previously been described in the literature. The authors have encouraged further 4D CT studies to help improve our knowledge of functional and pathological carpal motion.

Six studies have investigated scapholunate motion using 4D CT. Leng et al. evaluated radioulnar deviation in a cadaveric hand guided with a custom motion simulator (Leng, et al., 2011). 4D CT scans were performed prior to, and following resection of the scapholunate ligament. Comparison between the scans revealed increased motion between the scaphoid and lunate during radioulnar deviation. In addition, three orthopaedic surgeons all correctly identified the 4D clips with scapholunate instability.

Two case series/case reports have demonstrated early clinical utility of 4D CT for scapholunate instability. Halpenny et al. described a case of a patient with dynamic scapholunate instability diagnosed with 4D CT (Halpenny, et al., 2012). Initial static plain radiographs were normal. However, flexion-extension movement of the wrist under the 4D CT scanner revealed 6mm diastasis between the scaphoid and lunate. Kakar et al. reported their clinical utility of 4D CT in diagnosing two cases of scapholunate instability (Kakar, et al., 2016). They also produced coloured proximity maps to depict diastasis of the scapholunate joint in flexion-extension, radioulnar deviation, and the dart throwing motion. Two surgeons evaluated the 4D CT scans for scapholunate instability. They both

found the flexion-extension and radioulnar deviation motions were most helpful in assessing scapholunate instability.

Shores et al. compared 4D CT scans of asymptomatic wrists to 4D CT scans of wrists with previous scapholunate reconstructions (Shores, et al., 2013). Post-operative x-rays suggested subtle widening of the scapholunate interval. However, the 4D CT scans did not reveal any dynamic widening and the scapholunate joint resembled normal motion. Interestingly, wrist motion was abnormal compared to normal wrist biomechanics. Initiation of flexion-extension of the post-operative wrists was initiated by the radiocarpal joint, as opposed to the mid-carpal joint in normal wrists. In addition to evaluating the scapholunate joint, the authors also performed 4D CT scans on patients with ulnocarpal impaction for pre-operative assessment.

Garcia-Ellis et al. conducted a study examining the dart throwing motion under 4D CT (Garcia-Elias, et al., 2014). They compared healthy volunteers with patients with known scapholunate instability. In the healthy controls, the scaphoid and lunate translated together along the coronal plane, with little change in the scapholunate interval. Patients with scapholunate instability had increased translation of the scaphoid relative to the radial styloid, whereas the lunate translated slightly less than normal. This resulted in a significantly increased scapholunate interval, which dynamically changed during the dart throwing motion. The scapholunate interval was widest just before

terminal ulnar-flexion of the wrist. They concluded that the dart throwing motion is not recommended immediately following scapholunate ligament repair.

Demehri et al. completed a similar study investigating the scapholunate interval during dynamic wrist range of motion (Demehri, et al., 2016). Asymptomatic wrists and wrists with suspected scapholunate instability were compared. Three active movements were imaged using 4D CT: active fist clench, flexion-extension, and radioulnar deviation. The scapholunate interval consistently remained < 1 mm in all movements in the asymptomatic controls. In the symptomatic patients, the scapholunate interval was widest in extension (2.54±1.48mm), followed by clenched fist (2.53±1.19mm), then ulnar deviation (2.06±1.12mm). In addition, mild-moderate correlation was found between the patient's symptoms and the change in the scapholunate interval.

Foumani et al. evaluated the dynamic changes in cartilage thickness of the radiocarpal joint between static and dynamic CT scans (Foumani, et al., 2012). Ten healthy volunteers and 3 patients with radio-carpal osteoarthritis were included. A mechanical hand-shaking device guided the subjects through flexion–extension, radioulnar deviation and the dart throwing motion during a 4D CT scan. Static and dynamic distance maps were constructed of the radioscaphoid joint distance. There was a significantly reduced distance map in the dynamic versus the static CT scans. This difference was larger in the healthy controls compared to the patients with osteoarthritis.

2.3.7 Multiple Joints

Kalia et al. completed a feasibility study using 256 multi-slice CT scanners (Kalia, et al., 2009). Healthy volunteers were recruited. 4DCT scans were performed of their wrists and patellofemoral joints. The acquired dynamic CT scans were compared favourably to those from 64 multi-slice CT scanners.

2.3.8 Excluded Studies

A number of studies were excluded according to the inclusion and exclusion criteria. Gervaise et al. performed a review highlighting the differences in CT dose exposure between static and dynamic scans (Gervaise, et al., 2013). In their paper, they noted that new dose reduction parameters in the multi-slice CT scanners enable dynamic CT scans to be conducted with less radiation exposure compared to conventional CT. This was especially apparent in dynamic scans of peripheral joints. They concluded that it was very feasible to conduct dynamic acquisitions with total effective dose less than 1 mSv, which is the equivalent to a chest x-ray.

Twenty four articles from the literature search featured 'dynamic CT' scans of the spine. However, none of the articles incorporated four-dimensional CT using a wide field-of-view scanner and were subsequently excluded. These 'dynamic CT scans' acquired static datasets taken with the spine in different positions (e.g. flexion, extension,

maximum lateral rotation). Pathologies investigated included atlantoaxial rotatory displacement (AARD), rheumatoid arthritis, and nerve root impingement (McGuire, et al., 2002) (Soderman, et al., 2015) (Penning & Wilmink, 1987).

Chapter 3: The Scapholunate Joint

The literature review of musculoskeletal 4D CT demonstrates that it is at an early, evolving stage. After initial consideration of investigating other joints, it was decided that this thesis would focus on the scapholunate joint. The reasons for such was five-fold: 1) the clinical entity of scapholunate instability has challenges with early diagnosis using current imaging techniques, 2) the early success of studies proving the feasibility of 4D CT scans of the scapholunate joint, 3) the reduced relative radiation exposure with scanning peripheral extremities, 4) the availability of a small database of 4D CT scapholunate scans in healthy volunteers, and 5) anecdotal evidence from local hand and wrist surgeons utilizing scapholunate 4D CT scans to assist with the clinical management of patients.

3.1 Scapholunate Anatomy

The scapholunate joint is essential for normal carpal stability and motion. The word carpus is derived from the Greek word *karpos*, which means wrist. The main function of the wrist is to position the hand in space so that it can function in an effective and efficient manner (Garcia-Elias & Lluch, 2016).

The carpus is composed of eight bones arranged in two rows. There are 20 articulations within and between the carpus, radius, ulna, and bases of the five metacarpals. The proximal row consists of, in order from radial to ulnar direction, the scaphoid, lunate, and triquetrum. Although the pisiform is occasionally included as a bone

within the proximal row, it is a separate sesamoid bone within the tendon of flexor carpi ulnaris. The two joints within the proximal row are the scapholunate joint and the lunotriquetral joint. The distal row contains the trapezium, trapezoid, capitate, and hamate.

Stability of the wrist requires intact extrinsic and intrinsic ligaments, muscles, and tendons working in concert. Extrinsic ligaments connect the wrist to the forearm, whereas intrinsic ligaments link bones within the carpus. Important extrinsic ligaments on the volar surface include the radioscaphoid, radioscaphocapitate, long radiolunate, short radiolunate, ulnocapitate, ulnotriquetral, and ulnolunate ligaments. The only dorsal extrinsic ligament is the dorsal radio triquetral ligament (also known as the dorsal radiocarpal ligament). Intrinsic ligaments can either be transverse intercarpal (between bones in the same row), or midcarpal (between the proximal and distal row).

The scapholunate joint has two transverse carpal ligaments: the dorsal and volar (or palmar) scapholunate ligaments. The dorsal scapholunate ligament is biomechanically stronger and can withstand forces up to 260 N. The volar scapholunate ligament comparatively can yield up to 118 N (Garcia-Elias & Lluch, 2016). The lunotriquetral joint also has two transverse carpal ligaments (dorsal and volar). The volar lunotriquetral ligament is stouter than its dorsal counterpart. There is little motion within the distal row, as they are linked by strong dorsal, intra-articular, and volar intercarpal ligaments. The

midcarpal ligaments include the triquetrohamate, triquetrocapitate, scaphocapitate, dorsolateral scaphotrapeziotrapezoid, and dorsal intercarpal ligaments.

Kinematic motion of the wrist is complex. Planes of motion of the wrist include flexion/extension in the sagittal plane, radial/ulnar inclination in the coronal plane, and pronation/supination for wrist rotation. However, most functional motion of the wrist is in an oblique plane from extension and radial inclination to flexion and ulnar inclination. This is known as the dart throwing motion. In this plane, the proximal row is stable, and wrist motion primarily occurs through the mid-carpal joint between the proximal and distal row. In radial inclination, the trapezium and trapezoid are brought closer to the radial styloid. To accommodate this decreased in space, the scaphoid flexes palmarly. In ulnar inclination, the scaphoid extends due to its attachments to the trapezium and trapezoid (scapho-trapezio-trapezoid ligament). The other bones in the carpal row (lunate and triquetrum) flex and extend in conjunction with the scaphoid, but to a lesser degree (Moojen, et al., 2003).

The proximal row of the carpus is described as the intercalated segment of the wrist because there are no tendinous insertions on the proximal row. Thus, motion of the proximal row is governed by the motion of the distal row. When tendons of the forearm are contracted, they act directly on their attachments at the distal row. When the midcarpal ligaments become tense, the proximal row follows in motion. Due to its shape

and orientation, the scaphoid contributes more to wrist flexion and extension than the lunate. The scaphoid has an inherent tendency to flex. The shape of the lunate is wider on the volar aspect and narrower on the dorsal aspect; this causes it to preferentially extend. This has implications for carpal biomechanics when scapholunate ligament insufficiency is present. This will be discussed in more detail in the scapholunate dissociation section.

Stability of the wrist occurs when the carpal bones maintain normal alignment under physiologic stresses in all functional positions (Garcia-Elias & Lluch, 2016). This relies on intact ligaments, muscles, and wrist capsule, because the osseous anatomy of the carpus is inherently unstable. Mechanoreceptors in the carpal ligaments provide proprioceptive feedback to the brain, so that muscles and tendons can dynamically stabilize the wrist. Important dynamic stabilizers include extensor carpi radialis longus, abductor pollicis longus, and extensor carpi ulnaris (Garcia-Elias & Lluch, 2016).

3.2 Scapholunate Dissociation

Carpal instability has been classified into three main groups: carpal instability dissociative (CID), carpal instability nondissociative (CIND), and carpal instability complex (CIC). The CID pattern of instability includes conditions where there is an abnormal relationship within the same carpal row. The CIND group describes patterns of instability between the proximal and distal row. The CIC group incorporates instability between and

within the carpal rows. Scapholunate dissociation is the most common type of carpal instability, and is a recognized pattern of the carpal instability dissociative group. The abnormal relationship between the scaphoid and lunate can be predicted by the stage of the condition and the anatomical structures that have been disrupted.

Scapholunate instability may present after an acute injury or as a chronic degenerative tear. Acute injuries typically occur following a fall onto an outstretched hand. The position of the hand at the time of injury is normally wrist extension and ulnar inclination. Symptoms include pain over the dorsal and/or radial aspect of the wrist, weakness, and instability. A "clunk" or "click" may be described by patients with abnormal kinematics of the scaphoid and lunate. If left unrecognized or untreated, the malalignment of the carpus will follow a recognized pattern of progression until the carpus has undergone irreversible, degenerative changes. This is known as scapholunate advanced collapse (SLAC). Prompt diagnosis may lead to treatment options to correct and maintain stability of the scapholunate joint.

Diagnosis of scapholunate instability can be challenging, and is commonly missed on initial presentation. It may be associated with other injuries or fractures, masking its presence. Physical examination signs include tenderness over the scapholunate joint, pain with range of motion, and moderate swelling (if acute). Special examination tests

include the Watson shift test, resisted finger extension test, and scapholunate ballottement test.

3.3 Diagnostic Imaging of the Scapholunate Joint

Plain x-rays are often the first line of diagnostic tests arranged. Pathologic wrists may show radiographic signs, such as the "Terry Thomas sign" (widened scapholunate interval), scaphoid ring sign (flexed and foreshortened scaphoid), an increased radiolunate angle (> 15°), or an increased scapholunate angle (> 60°). Stress views are indicated when there is potential for dynamic scapholunate instability. In some stages of scapholunate dissociation, the abnormal carpal alignment may only be revealed when the joint is loaded or viewed in certain positions. Advanced disease states will display instability in both static and dynamic positions. The clenched fist view serves to axially load the scapholunate joint from distal to proximal. With scapholunate instability, the scapholunate joint may widen as the capitate is driven proximally between the scaphoid and the lunate.

Arthrography (radiographs with contrast injection) has lost favour as a diagnostic approach, as it cannot distinguish between asymptomatic degenerative tears and tears resulting in instability. Cineradiography (or continuous fluoroscopy) has been described as a highly sensitive and specific imaging modality for diagnosing scapholunate dissociation. One retrospective study reported a 90% sensitivity and 97% specificity for

coronal and sagittal cineradiography of wrists undergoing radial/ulnar deviation and flexion/extension (Sulkers, et al., 2014). Ultrasound has also been shown to be a valid diagnostic tool. Sensitivity for scapholunate dissociation ranges from 46 to 100% and specificity from 92 to 100% (Renoux, et al., 2009).

Magnetic resonance imaging (MRI) enables enhanced visualization of the soft tissues of the wrist. It has shown reasonable good sensitivity (63%) and specificity (86%) for diagnosing scapholunate ligament injuries (Garcia-Elias & Lluch, 2016). MR arthrography with contrast improves diagnostic accuracy.

Wrist arthroscopy is considered by many to be the gold standard for diagnosis of scapholunate tears. High volume centres may scope wrists before considering reconstructive procedures. However, wrist arthroscopy is a highly technical procedure that is not routinely performed everywhere. The Geissler classification is an arthroscopic grading system of the integrity of intercarpal ligaments (see Table 3) (Geissler, et al., 1996).

Table 3 - Geissler Arthroscopic Classification

Grade	Appearance	Probe
Grade I	Attenuation or hemorrhage of interosseous ligament as seen from the radiocarpal space. No incongruency of carpal alignment in midcarpal space.	
Grade II	Attenuation or hemorrhage of interosseous ligament as seen from the radiocarpal space. Incongruency or step-off seen in midcarpal space.	There may be a slight gap (less than width of probe) between carpal bones
Grade III	Incongruency or step-off of carpal alignment as seen from both radiocarpal and midcarpal space.	Probe may be passed through gap between carpal bones.
Grade IV	Incongruency or step-off of carpal alignment as seen from both radiocarpal and midcarpal space. There is gross instability with manipulation	A 2.7 mm arthroscope may be passed through gap between carpal bones.

Garcia-Elias described seven stages of progressive scapholunate instability (see Table 4). The staging system is based on six clinical questions: 1) Is it a partial ligamentous injury of the volar side only, with the dorsal ligament intact? 2) Is the ligament biologically capable of repair? 3) Is there normal alignment of the scaphoid? 4) Is there normal alignment between the radius and lunate? 5) Is the carpal misalignment reducible? and 6) Is the articular cartilage normal?

Table 4 - Staging Algorithm for Scapholunate Dissociation (adapted from Garcia-Elias & Lluch, 2016)

Stage	Partial injury?	Repairable?	Normal radioscaphoid angle?	Normal lunate alignment?	Reducible?	Normal Cartilage?
Stage I	Yes	Yes	Yes	Yes	Yes	Yes
Stage II	No	Yes	Yes	Yes	Yes	Yes
Stage III	No	No	Yes	Yes	Yes	Yes
Stage IV	No	No	No	Yes	Yes	Yes
Stage V	No	No	No	Yes	Yes	Yes
Stage VI	No	No	No	No	Yes	Yes
Stage VII	No	No	No	No	No	Yes

Stage I injuries are often referred to as a 'pre-dynamic' instability. Plain radiographs do not show evidence of scapholunate dissociation, even during stress views. Stage II and III injuries are 'dynamic' instabilities. Secondary stabilizers, such as the scapho-trapezio-trapezoid (STT) ligament, volar capsule, and dorsal intercarpal ligaments, help maintain carpal alignment in static positions. However, when the scapholunate joint is loaded, this may widen the scapholunate joint. Stages IV through VI may exhibit findings of 'static instability,' particularly dorsal intercalated segment instability, or DISI. As stated previously, the scaphoid has a tendency to flex, and the lunate to extend. Without an intact scapholunate ligament or secondary stabilizers, the scaphoid and lunate tend to

dissociate fall into their preferential positions. This is apparent with an increased scapholunate angle on lateral plain films of the wrist.

The stage of the scapholunate instability influences clinical treatment (see Table 5). Acute injuries diagnosed quickly may enable direct repair of the scapholunate ligament. After approximately 6 weeks, the torn ends of the ligament degenerate and lose its potential to heal. Reconstructive procedures aim to restore the function of the scapholunate ligament. Once the articular cartilage displays evidence of arthritic changes, salvage procedures will need to be considered.

Table 5 – Management of Scapholunate Dissociation (Garcia-Elias & Lluch, 2016)

Stage	Description	Treatment Options
Stage I	- Partial scapholunate ligament injury	- percutaneous or arthroscopic-assisted K-wire - physiotherapy - arthroscopic debridement + electrothermal shrinkage
Stage II	- Complete SL ligament injury, repairable	- ORIF + dorsal SL ligament repair
Stage III	- Complete SL ligament injury, Irreparable - Normal alignment of scaphoid/lunate	- Dorsal capsulodesis - Dorsal SL ligament soft tissue reconstruction - Bone-ligament-bone grafts - Dynadesis
Stage IV	- Complete SL ligament injury, Irreparable - Reducible flexion deformity of scaphoid	- SL ligamentoplasty with tendon graft - Reduction-association of SL joint (RASL)
Stage V	- Complete SL ligament injury, Irreparable - Reducible carpal collapse - Instability of radiolunate joint	- Spiral tenodesis
Stage VI	- Complete SL ligament injury, Irreparable - Irreducible carpal collapse - Normal cartilage	- Scaphoid-trapezium-trapezoid fusion - Scapholunate fusion - Scaphoid-capitate fusion - Scaphoid-lunate-capitate fusion - Radioscaphoid-lunate fusion + scaphoidectomy
Stage VII	- Complete SL ligament injury, Irreparable - Irreducible carpal collapse - Abnormal cartilage	- Radial styloidectomy - Four-corner fusion + scaphoidectomy - Proximal row carpectomy - Total wrist arthroplasty - Total wrist fusion

It is not uncommon for scapholunate injuries to be missed on initial presentation. They can be masked by concomitant injuries. They may also be brushed off as simple wrist sprains and ignored. There may be absent, or very subtle, radiographic findings in 'pre-dynamic' or 'dynamic' scapholunate injuries. MRI has shown reasonable effectiveness in diagnosing scapholunate injuries, but MRI studies of the wrist are lengthy and resource-dependent.

4D CT wrist scans may have the potential to assist with diagnosing 'dynamic' scapholunate injuries. These scans may provide additional information of carpal alignment in loaded states and during range of motion. This thesis proposes to examine the potential role of 4D CT scans in diagnosing scapholunate injuries. In particular, it will focus on analyzing the inter- and intra-observer reliability of assessing these scans in normal and abnormal wrists.

Chapter 4: Validity of Diagnostic Tests

4.1 Accuracy and Reliability

When new diagnostic techniques are introduced, they are ideally subjected to rigorous testing to ensure their validity. This requires both accuracy and reliability. The accuracy of a diagnostic test refers to its ability to discern between a target condition and health (Simundić, 2009). Parameters of accuracy include sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio. These are calculated when a diagnostic test is directly compared to an accepted gold standard.

However, an accurate test cannot be valid in the absence of reliability. This refers to how reproducible a given test result is, including between different observers or between different time points. There are numerous statistical methods to assess the reliability, or extent of agreement, between observers. These are expressed through various correlation coefficients.

4.2 Inter- and Intra-Rater Reliability

The most common statistic utilized for inter- and intra-rater reliability is intra-class correlation coefficient (ICC). There are multiple variations of ICC, and depend on the desired statistical 'model,' 'type' of study, and 'definition' of agreement (McGraw & Wong, 1996). In addition, intra-class correlation coefficients are typically calculated for

measuring agreement of ordinal, interval, or ratio variables (Hallgren, 2012). In our study, a nominal scale is used. This requires different statistical methods to examine inter-rater reliability.

4.2.1 Percent Agreement

The most rudimentary form of inter-observer variability is percent agreement. However, the downside of percent agreement as a statistical method is it does not consider the possibility of agreement based on chance (Glen, 2017). It also tends to overestimate inter-rater correlation. One of the earliest methods of assessing inter-rater correlation was described by Goodman and Kruskal in 1954 (Goodman & Kruskal, 1954). Their lambda coefficient measured the observed proportion of agreement of classifications. However, like percent agreement, it does not take chance into consideration.

4.2.2 Cohen's Kappa and Scott's **π** coefficient

Cohen's kappa coefficient, κ, was proposed in 1960 (Cohen, 1960). It is a chance-corrected coefficient that assumes that some instances of agreement are due to chance alone. The kappa coefficient is utilized when two raters classify subjects into two or more mutually exclusive nominal categories (Banarjee, et al., 1999). Cohen's kappa coefficient, κ, is defined as:

$$\kappa = \frac{p_o - p_c}{1 - p_c}$$

where p_o = the observed proportion of agreement
and p_c = the proportion of agreement expected by chance

Cohen's kappa is an advancement on the work conducted by Scott in 1955, who described the **π** coefficient. (Banarjee, et al., 1999). However, Scott's definition of p_c assumes that the proportion of subjects is a known quantity, and is equal between the two raters. On the other hand, Cohen's kappa equation considers the potential different rating patterns between the two raters (Gwet, 2002). Interpretation of Cohen's kappa has been incorporated into a classification as proposed by Altman (see Table 6) (Altman, 1991).

Table 6 - Altman Interpretation of Cohen's Kappa

Value of κ	Strength of Agreement
< 0.20	Poor
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Good
0.81 – 1.00	Very Good

Cohen's kappa remains a popular statistical method for measuring inter-rater variability in modern literature. However, there are flaws with Cohen's kappa when utilized or interpreted inappropriately. It is prone to rater bias, which is the difference of marginal distributions between raters (Banarjee, et al., 1999). Marginal distributions describe how each rater individually rates each subject into a category. In other words, Cohen's kappa is reliant on trained raters.

Cohen's kappa is affected by the true prevalence of subjects within each category. This is known as the first kappa paradox (Kuppens, et al., 2011). Effectively, Cohen's kappa will vary depending on the proportion of cases within each population. The second kappa paradox relates to scenarios where a symmetrical unbalanced marginal will produce a lower kappa value compared to an asymmetrical unbalanced marginal (Kuppens, et al., 2011). Finally, Cohen's kappa does not differentiate between rater disagreements, no matter how similar or dissimilar they are. Cohen addressed this by proposing a weighted kappa analysis, κ_w, which incorporates a ratio-scaled degree of agreement (Banarjee, et al., 1999).

4.2.3 Fleiss' Kappa

Both Cohen's kappa and Scott's pi coefficients are restricted to measuring agreement between two raters. Fleiss's kappa was proposed in 1971 to accommodate multiple raters (Fleiss, 1971). Fleiss' kappa considers the generalization of unweighted

kappa to the agreement between a constant number of independent raters. Despite Fleiss' coefficient nomenclature as κ, it is a mathematical extrapolation of Scott's **π** coefficient rather than Cohen's κ coefficient (Gwet, 2008). It is still closely associated with Cohen's kappa. Fleiss' kappa is defined as:

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

where $\bar{P} - \bar{P}_e$ is equivalent to the degree of agreement achieved above chance
$1 - \bar{P}_e$ is equivalent to the degree of agreement obtainable by chance

4.2.4 Gwet's AC_1 Coefficient

As described previously, high levels of agreement may paradoxically cause low values of κ when the prevalence of a trait in a population is very high or low. To adjust for this paradox, Gwet proposed a new coefficient to assess inter-rater reliability that was not affected by the two kappa paradoxes. He called it Gwet's AC_1 coefficient. (Gwet, 2008). This stands for agreement coefficient. His statistical method is based on the following two principles (Gwet, 2008):

"1. Chance agreement occurs when at least one rater rates an individual randomly

2. Only an unknown portion of the observed ratings is subject to randomness"

Based on his statistical models, Gwet noted that the AC_1 coefficient surpassed the **π** and κ coefficients when a trait's prevalence was high or low within a population.

However, all the coefficients performed similarly when the trait prevalence approached 50% (Gwet, 2008).

4.2.5 Conger's Kappa

Fleiss' original intention was for his kappa coefficient to be an extension of Cohen's kappa. However, it was Conger who established that when Fleiss' kappa was reduced to two raters, it corresponded with Scott's **π** coefficient (Conger, 1980). Conger subsequently described the true extension of Cohen's kappa for multiple raters. Although the equation was computationally more complex, Conger noted that there was a small but distinct difference between his and Fleiss' coefficient. As Cohen's kappa considers the different marginal distributions between raters, this can be extrapolated to Conger's kappa.

4.2.6 Krippendorff's Alpha

Krippendorff's alpha is another statistical method of quantifying agreement amongst raters (Gwet, 2011). It is popular amongst coders for content analysis. It is defined by the following formula:

$$a = \frac{p_a - p_e}{1 - p_e}$$

where p_a = the weighted percent agreement

and p_e = the weighted percent chance agreement

One of the advantages of Krippendorff's alpha coefficient is its ability to compensate for missing data. It can also be weighted according to the different response categories. However, like Cohen's kappa and Fleiss' kappa coefficients, it is susceptible to the paradox where high or low trait prevalence will result in a relatively lower kappa value.

4.2.7 Brennan-Prediger Coefficient

The Brennan-Prediger coefficient is an alternative measurement of inter-rater agreement and was described in 1981 (Brennan & Prediger, 1981). It is a multi-rater extrapolation of the G-index, which was initially described for two raters (Holley & Guilford, 1964). The Brennan-Prediger coefficient is still commonly referred to as the G-index. Its advantage is its resistance to the two kappa paradoxes. Brennan and Prediger argued that Cohen's kappa is appropriate when the marginals are fixed a priori. However, when the marginal distributions are free to vary, they proposed that the "chance" term in the formula be replaced by $1/n$, with *n* equal to the number of categories. The Brennan-Prediger coefficient is defined as:

$$k_n = \frac{P_a - 1/n}{1 - 1/n}$$

where P_a = the weighted percent agreement
and n = the number of response categories

Statistical models have shown that the G-index and Gwet's AC_1 are more appropriate statistical methods than Cohen's kappa and Scott's **π** coefficient when the trait prevalence is at extremes (very high or low) (Gwet, 2008). However, Gwet argues that the AC_1 coefficient still outperforms the G-index, particularly when a rater's propensity for a random rating is high. In this scenario, his Monte-Carlo statistical model showed an increase of the G-index's relative bias, compared to a decrease of the AC_1's relative bias (Gwet, 2008).

There are multiple valid methods to assessing inter-rater reliability. The various correlation coefficients tend to follow the same generic formula:

$$\kappa = \frac{p_o - p_c}{1 - p_c}$$

The calculation to determine the observed percent agreement (p_o) is constant. The differentiating factor between the correlation coefficients is how the percent chance agreement (p_c) is calculated. These can be complex calculations, and are significantly aided by statistical software. Numerous papers have critiqued the limitations of each method. However, knowledge of the different coefficients of inter-rater reliability enables the reader to critically analyze which coefficients are most applicable for a given study.

Chapter 5: Methodology

A number of study designs were entertained when devising the study protocol. The main objective of the thesis was to evaluate 4D CT wrist scans as a novel approach to diagnosing scapholunate injuries. As mentioned in the literature review section, a limited number of studies had been published on four-dimensional CT scans of the scapholunate joint. Clinical studies are restricted to cross-sectional studies, case reports, or case series. Other articles have focused on healthy controls or cadaveric specimens.

Our institution has the privilege of access to a 4D CT-capable scanner since 2014. Pilot studies had been conducted at Memorial University to explore its potential uses. Orthopaedic and plastic surgeons have begun to explore its clinical utility as an adjunct to current diagnostic imaging. As a result, a small number of patients with pathologic scapholunate injuries have had 4D CT scans of their wrists. In each scan, patients had performed a slow active radial-ulnar deviation motion in a clenched fist position. In addition, a cohort of healthy wrist 4D CT studies had been collected and anonymously saved into a database. A decision was made to take advantage of those opportunities.

5.1 Research Proposal / Ethics Approval

A study proposal was devised and reviewed by the MSc selection committee. An ethics application was approved by the Newfoundland and Labrador Health Research Ethics Board (NL HREB) (see Appendix B). Potential ethical issues were addressed and

reflected in the final study design. Given that the study involves data from both patients and volunteers, multiple steps were taken to avoid recording any personal identifying information. CT scans retrieved from the Picture Archiving and Communication System (PACS) were de-identified prior to exportation. Video editing software was utilized to crop out any extraneous information irrelevant to the study. A computerized random number generator assigned every scan a random letter. Verbal consent was obtained from the recruited surgeons. The surgeons were also assigned a number to maintain anonymity. All data was kept on encrypted devices. Local health authority approval was granted by the Eastern Health Research Proposals Approval Committee (RPAC) (see Appendix C).

5.2 Study Design

Pathologic and normal 4D CT wrists scans were collected for this study. Wrists were deemed "pathologic" if the diagnosis of scapholunate instability was confirmed in subsequent intra-operative findings, MRI scans, or future progression into a SLAC wrist. The "normal wrists" were scans of healthy volunteers who were asymptomatic and had no history of significant injuries to their wrist.

All studies had been performed on a 320-row multi-detector Toshiba Aquilion ONE CT scanner (Tokyo, Japan). Approval had been granted to access the Picture Archiving and Communication System (PACS) for 4D CT scans of real patients. Recent 4D CT scans were identified by searching for CT studies under the heading of "EH UPPER EXTREMITY

DYNAMIC." However, the original studies did not differentiate between 'static' or 'dynamic' CT scans. Thus, all upper extremity CT scans from 2013 were filtered. The radiological reports were screened to identify the 4D CT scans. All wrist 4D CT scans were viewed, but only those with dynamic scapholunate injuries were included. Unfortunately, the raw CT data was unavailable, so the 'views' of each 4D clip were limited to what the x-ray technologist and musculoskeletal radiologist had previously saved on to PACS. This limited the scope of our study, as we would be unable to perform quantitative analysis on the pathologic CT scans.

Each pathologic 4D CT scan had a minimum of three 'views': 1) a 4-dimentional clip from the volar side, 2) a 4-dimentional clip from the dorsal side, and 3) a clip of a coronal slice, taken approximately at the mid-coronal plane of the scapholunate joint. When 4D CT scans were uploaded to the PACS system, each video clip was imported as separate, individual JPEG images. The PACS software enables the user to cycle through the images in a cine-loop, thus enabling them to be viewed as a dynamic clip. For our study, the images were converted back into a video file so they could be viewed without the PACS system.

For each pathologic CT scan, the JPEG images of the dynamic sequences were de-identified and subsequently exported from the PACS system on to an encrypted drive. The JPEG images were imported into a video processing software, iMovie® (version

10.106, Cupertino, CA, USA. The video clips were uniformly cropped and adjusted to maintain a standard frame rate. Each subject's 4D volar, 4D dorsal, and coronal sequences were duplicated and spliced together. This resulted in a single video file containing two clips of each view that could be viewed sequentially (see Appendix D).

Normal scans were retrieved from a saved database. In this case, the raw CT scan data was available. This data was uploaded to an advanced radiological imaging software called Vitrea® VITAL (version 7, Minnesota, MN, USA). The CT data was manipulated to replicate the same views obtained from the pathological scan: a four-dimensional clip from the volar surface, a four-dimensional clip from the dorsal surface, and a two-dimensional clip at the mid-coronal plane of the scapholunate joint. This involved cropping each video clip, digitally removing extraneous artefact (e.g. CT table, soft tissues, vascular system), and orientating each view. Each sequence was separately downloaded as a .avi file. The .avi files were uploaded into iMovie® and converted into a single .mp4 video file (see Appendix E).

Overall, fifteen 4D CT scans were identified through a PACS search. Sixteen normal wrists were obtained from the database. However, only four of the pathologic wrists could be included in our study. Reasons for exclusion included: inadequate views on PACS, no scapholunate pathology, or advanced arthritic changes. Thirteen normal wrists were

included from the database. All scans that met the inclusion criteria were assigned a random letter of the alphabet via online random generator software.

Local orthopaedic and plastic surgeons were recruited to participate in the study. A total of 4 orthopaedic surgeons and 3 plastic surgeons agreed to participate. After receiving instructions, verbal consent was obtained. Questions were answered, and the surgeons were given a quick demonstration of a sample 4D CT clip that was not utilized in the study. The surgeons subsequently viewed each movie clip on a laptop. They were given the opportunity to repeat each clip as necessary. They were asked to determine if a) the wrist was stable or unstable, and b) which view was most beneficial to making that decision. Results were recorded by the principle investigator.

The same group of surgeons was invited back three months later for a repeat assessment of the 4D CT scans. The same scans were utilized, but rearranged in a different order. The surgeons were asked to answer the same questions as before. No training or feedback was provided between the index and repeat assessments.

5.3 Statistical Analysis

Statistical analysis of the data was performed using the statistical software AgreeStat 2015.6 for Windows (Gaithersburg, MD, USA). Inter-observer variability was calculated using percent agreement, Fleiss' Kappa, Gwet's AC_1, Conger's Kappa,

Krippendorff's Alpha, and the Brennan-Prediger coefficient. Intraclass Correlation Coefficients (ICCs) were utilized for analyzing intra-observer reliability.

Chapter 6: Results

The tabulated results were recorded on an Excel spreadsheet. A total of 17 wrist studies were evaluated. Thirteen of these scans were taken from healthy volunteers and four were from wrists with known scapholunate pathology. Seven surgeons agreed to participate. Four of the surgeons were from the department of orthopaedic surgery, and three were plastic surgeons. Five of the surgeons had a clinical special interest in upper extremity surgery.

6.1 Inter-Rater Reliability Results

The distribution of the raters' responses for the assessment of scapholunate instability are listed in Table 7. The raw data for the responses is listed in Appendix F. The percent agreement amongst all raters was 0.80392 (95% CI: 0.675 – 0.932) (see Table 8). Fleiss' Kappa was 0.54895 (95% CI: 0.252 - 0.846) and Gwet's AC_1 was 0.65313 (95% CI: 0.391 – 0.915). Conger's Kappa was 0.55229 (95% CI: 0.261 – 0.843), Krippendorff's Alpha was 0.55274 (95% CI: 0.256 – 0.846), and Brennan-Prediger coefficient was 0.60784 (95% CI: 0.351 – 0.865).

Table 7 - Distribution of Scapholunate Stability Responses by Rater

	Stable	Unstable	Total
Rater 1	11	6	17
Rater 2	13	4	17
Rater 3	10	7	17
Rater 4	11	6	17
Rater 5	9	8	17
Rater 6	14	3	17
Rater 67	13	4	17
Average	11.6	5.4	17.00

Table 8 - Inter-rater Variability of Scapholunate Instability

		Inference/Subjects			Inference/Subjects & Raters		
METHOD	Coefficient	StdErr	95% C.I.	p-Value	StdErr	95% C.I.	p-Value
Conger's Kappa	0.55229	0.12344	0.291 to 0.814	0.0003835	0.14667	0.261 to 0.843	0.0002863
Gwet's AC_1	0.65313	0.10542	0.43 to 0.877	0.00001284	0.13187	0.391 to 0.915	0.000003131
Fleiss' Kappa	0.54895	0.12607	0.282 to 0.816	0.0004919	0.14973	0.252 to 0.846	0.0004037
Krippendorff's Alpha	0.55274	0.12607	0.285 to 0.82	0.0004621	0.14931	0.256 to 0.849	0.0003569
Brennan-Prediger	0.60784	0.10261	0.39 to 0.825	0.00002141	0.12943	0.351 to 0.865	0.000008.814
Percent Agreement	0.80392	0.05131	0.695 to 0.913	3.965×10^{-11}	0.06471	0.675 to 0.932	0.000E+00

6.2 Subgroup Analysis of Upper Extremity Surgeons

A subgroup analysis was performed for the inter-rater variability of the upper extremity surgeons (see Table 9). The percent agreement was 0.77647 (95% CI: 0.604 – 0.949). Fleiss' Kappa was 0.47360 (95% CI: 0.106 – 0.841) and Gwet's AC_1 was 0.61105 (95% CI: 0.254 – 0.969). Conger's Kappa was 0.48154 (95% CI: 0.129 – 0.834), Krippendorff's Alpha was 0.47979 (95% CI: 0.115 – 0.845), and Brennan-Prediger coefficient was 0.55294 (95% CI: 0.209 – 0.897).

Table 9 - Inter-rater Variability of Scapholunate Instability (Upper Extremity Surgeon Subgroup)

		Inference/Subjects			Inference/Subjects & Raters		
	Coefficient	**StdErr**	**95% C.I.**	**p-Value**	**StdErr**	**95% C.I.**	**p-Value**
Conger's Kappa	0.48154	0.13523	0.195 to 0.768	0.002605	0.17655	0.129 to 0.834	0.008225
Gwet's AC_1	0.61150	0.11053	0.377 to 0.846	0.00004546	0.17913	0.254 to 0.969	0.001116
Fleiss' Kappa	0.47360	0.14122	0.174 to 0.773	0.004036	0.18379	0.106 to 0.841	0.01228
Krippendorff's Alpha	0.47979	0.14122	0.18 to 0.779	0.003679	0.18268	0.115 to 0.845	0.01079
Brennan-Prediger	0.55294	0.10783	0.324 to 0.782	0.0001012	0.17237	0.209 to 0.897	0.002090
Percent Agreement	0.77647	0.05391	0.662 to 0.891	1.404×10^{-10}	0.08619	0.604 to 0.949	5.449×10^{13}

6.3 Preference of View for Assessment of Scapholunate Instability

The distribution of each surgeon's preferred view that was most helpful in their assessment of wrist stability is listed in Table 10. The raw data can be found in Appendix G. The percent agreement was 0.51821 (95% CI: 0.341 – 0.695) (see Table 11). Fleiss' Kappa was -0.04290 (95% CI: -0.177 – 0.091) and Gwet's AC_1 was 0.37349 (95% CI: 0.074

– 0.673). Conger's Kappa was 0.01016 (95% CI: -0.089 – 0.11), Krippendorff's Alpha was -0.03414 (95% CI: -0.167 – 0.099), and Brennan-Prediger coefficient was 0.27731 (95% CI: 0.012 – 0.543).

Table 10 - Distribution of Preferential View of Each Rater

	4D Volar	4D Dorsal	2D Coronal	Total
Rater 1	0	10	7	17
Rater 2	0	0	17	17
Rater 3	0	9	8	17
Rater 4	1	7	9	17
Rater 5	2	10	5	17
Rater 6	0	1	16	17
Rater 7	0	0	17	17
Average	0.4	5.3	11.3	17.00

Table 11 - Inter-rater agreement on preferred views for assessing scapholunate instability

		Inference/Subjects			Inference/Subjects & Raters		
	Coefficient	StdErr	95% C.I.	p-Value	StdErr	95% C.I.	p-Value
Conger's Kappa	0.01016	0.03127	-0.056 to 0.076	0.7496	0.05008	-0.089 to 0.11	0.8397
Gwet's AC_1	0.37349	0.06719	0.231 to 0.516	0.00004324	0.15074	0.074 to 0.673	0.01497
Fleiss' Kappa	-0.04290	0.03576	-0.119 to 0.033	0.2477	0.06737	-0.177 to 0.091	0.5258
Krippendorff's Alpha	-0.03414	0.03576	-0.11 to 0.042	0.3539	0.06690	-0.167 to 0.099	0.6110
Brennan-Prediger	0.27731	0.06391	0.142 to 0.413	0.0005077	0.13369	0.012 to 0.543	0.04072
Percent Agreement	0.51821	0.04261	0.428 to 0.609	1.693×10^{-9}	0.08912	0.341 to 0.695	7.954×10^{-8}

6.4 Intra-Rater Reliability Results

The results of intra-observer reliability are listed in Table 12. All surgeons repeated their assessments of all scans three months following their index assessment. The intraclass correlation coefficient (ICC) for intra-rater reliability was 0.71631 (95% CI: 0.5567 – 0.8423).

Table 12 - Intra-Rater Reliability Results

Intraclass Correlation Coefficient - ICCa:		**0.71631**
95% Confidence Interval:		**0.5567 to 0.8423**
p-Value for Various Levels of ICCa:		
	ICCa	P-Value
	0	1.975×10^{-12}
	0.1	4.992×10^{-10}
	0.3	0.000005194
	0.5	0.005461
	0.7	0.3893
	0.9	0.9995

6.5 Repeat Inter-Rater Reliability Results

A second round of inter-rater reliability tests were conducted for the surgeons' repeat assessments (see Table 13). The raw data for the repeat responses are listed in Appendix H. Percent agreement was 0.85994 (95% CI: 0.725 – 0.995). Fleiss' kappa was 0.66270 (95% CI: 0.337 – 0.988) and Gwet's AC1 coefficient was 0.76050 (95% CI: 0.507 – 1). Conger's kappa was 0.66483 (95% CI: 0.344 – 0.986), Krippendorff's alpha was 0.66553 (95% CI: 0.341 – 0.99), and the Brennan-Prediger coefficient was 0.71989 (95% CI: 0.45 – 0.99).

Table 13 - Repeat Inter-Rater Reliability Results

		Inference/Subjects			Inference/Subjects & Raters		
METHOD	**Coefficient**	**StdErr**	**95% C.I.**	**p-Value**	**StdErr**	**95% C.I.**	**p-Value**
Conger's Kappa	0.66483	0.12290	0.404 to 0.925	5.785×10^{-05}	0.16156	0.344 to 0.986	8.185×10^{-05}
Gwet's AC_1	0.76050	0.08151	0.588 to 0.933	7.141×10^{-08}	0.12760	0.507 to 1	4.169×10^{-08}
Fleiss' Kappa	0.66270	0.12501	0.398 to 0.928	7.170×10^{-05}	0.16402	0.337 to 0.988	1.076×10^{-04}
Krippendorff's Alpha	0.66553	0.12501	0.401 to 0.931	6.855×10^{-05}	0.16335	0.341 to 0.99	9.504×10^{-05}
Brennan-Prediger	0.71989	0.08814	0.533 to 0.907	4.230×10^{-07}	0.13599	0.45 to 0.99	7.580×10^{-07}
Percent Agreement	0.85994	0.04407	0.767 to 0.953	1.398×10^{-12}	0.06800	0.725 to 0.995	0

Chapter 7: Discussion

Overall, our study suggests good inter- and intra-observer reliability for the assessment of scapholunate instability in 4D CT wrist scans. The fact that the intra-observer reliability results were higher than our index inter-observer results supports our interpretation of the data. Additionally, our repeat inter-reliability results obtained three months following our initial data is noticeably higher. The difference between the index and repeat results is not statistically significant. We are therefore unable to conclude whether the improvement in reliability was related to a 'training effect' or simply due to chance. It is important to note that no feedback or training was provided to the surgeons after they completed their first assessment. Perhaps the surgeons became more comfortable interpreting the 4D CT scans with more experience. However, the improved results are a positive indication that assessing scapholunate instability in 4D CT scans may have good reliability.

7.1 Inter-rater coefficients

As documented previously, there are multiple ways to measure inter-observer reliability. The measurement of inter-observer variability does not reflect the accuracy of the diagnostic technique. Rather, it measures the precision or reliability of raters to reproduce the same results.

Percent agreement amongst the raters was 80.39%. This implies very good agreement. However, percent agreement is a poor method for analyzing reliability because it does not take chance into account. As our study design incorporated multiple raters, we were unable to utilize Cohen's kappa or Scott's **π** coefficient. Our calculated Fleiss' kappa coefficient for inter-rater variability of the 4D CT wrist scans was 0.54895. According to Altman's kappa scale, this correlates with moderate inter-observer variability. However, the distribution of normal and pathologic wrists was not equal in our study (13 vs. 4). As such, this altered the probability of agreement occurring due to chance alone. According to the first kappa paradox, this may result in the underestimation the true extent of agreement. Scott's **π** coefficient, Cohen's κ coefficient, Fleiss' kappa, and Conger's kappa are all susceptible to this paradox.

Since the prevalence of scapholunate pathology within our population sample was 23.5%, it can be argued that Gwet's AC_1 coefficient is a more appropriate statistical method for our study than Fleiss' kappa coefficient. As mentioned earlier, Gwet's AC_1 is resistant to the kappa paradoxes. Our Gwet's AC_1 coefficient was 0.65313. This correlates with good inter-rater reliability. This compares favourably to the moderate rating of our Fleiss' kappa coefficient.

In our data, Conger's kappa was 0.55229 compared to Fleiss' kappa coefficient of 0.54895. This also correlates to a moderate rating of inter-observer reliability. Although

the difference in value is small, Conger's kappa coefficient is arguably more robust than Fleiss'. This is because Fleiss' kappa is an extension of Scott's **π** coefficient, whereas Conger's kappa is an extension of Cohen's kappa for multiple raters.

Our results revealed a Krippendorff's alpha coefficient was 0.55274. This is remarkably close to our Conger's kappa of 0.55229. However, Krippendorff's alpha coefficient is primarily utilized for inter-coder reliability for content analysis. It is not directly applicable for our study design.

Given the Brennan-Prediger coefficient, or G-index, is another paradox-resistant coefficient, one may expect that the value would be higher than Fleiss' kappa, Conger's kappa, and Krippendorff's alpha. Our calculations revealed that the Brennan-Prediger coefficient, or G-index, was 0.60784. This is similar to Gwet's coefficient of 0.65313, and higher than the coefficients susceptible to the kappa paradox.

When considering the low trait prevalence of scapholunate instability in our study, the most appropriate statistical method is Gwet's AC_1 coefficient, followed by the Brennan-Prediger coefficient. These two coefficients are resistant to the kappa paradoxes witnessed when the trait prevalence is high or low. This is reflected in our results. Gwet's AC_1 and Brennan-Prediger coefficients both correlated with 'good' inter-rater reliability.

Our values for Fleiss' kappa and Conger's kappa exhibit 'moderate' inter-rater variability, and suggests that these coefficients underestimated the true inter-observer reliability.

7.3 Standard Errors and 95% Confidence Intervals

Our calculated Gwet's AC_1 coefficient suggests that there is good inter-rater reliability when assessing stability of scapholunate joints. However, one must examine the confidence intervals when interpreting inter-rater reliability. In inter-rater reliability studies, the confidence intervals are more important than p-values for data interpretation. P-values measure the statistical significance of the correlation coefficients (e.g. Cohen's kappa) against the null hypothesis (e.g. no correlation; kappa = 0). For instance, a study may demonstrate a p-value less than 0.05, but with a low correlation coefficient value. Although the correlation coefficient may be statistically significant, the study would reveal low inter-rater reliability. After analysis of our results for the assessment of scapholunate stability, the p-values for all the correlation coefficients were 0.002605, 0.00004546, 0.004036, 0.003679, and 0.0001012 (see Table 8). These were all well below the arbitrary 0.05 p-value threshold. We can therefore conclude that there was a statistically significant degree of correlation. However, we are unable to ascertain the degree of agreement from the p-values alone.

On the other hand, the width of the confidence intervals reflects how close the measured correlation coefficient is to its 'true' value. The tighter the confidence interval,

the more confident we can be that the correlation coefficient is accurate. Unfortunately, our results revealed relatively large standard errors and corresponding wide confidence intervals. This affects our ability to confidently interpret our results. For Fleiss' Kappa, Conger's Kappa, Gwet's AC_1 and the Brennan-Prediger coefficient, the wide confidence intervals suggest the inter-rater reliability for assessing scapholunate instability may be anywhere from fair, moderate, good, or very good. It is interesting to note that the standard errors for the two kappa paradox-resistant correlation coefficients (Gwet's AC1 and Brennan-Prediger coefficient) were lower than the standard errors of the other correlation coefficients (see Table 8).

Examination of our statistical output of correlation coefficients reveals two sets of values for the standard error, 95% confidence intervals, and p-values. The two separate headings are: "Inference/Subjects" and "Inference/Subjects & Raters." For the first set of values, the standard errors and confidence intervals are determined according to the sampling population only (Gwet, 2011). It assumes that the variability amongst the raters is fixed. This is more applicable to situations such as content analysis, where the statistical inference of rater variability is not typically required or desired. On the other hand, statistical inference for both subjects and raters accounts for possible variability amongst subjects and raters. This set of values would be more applicable when the raters are representative of a broader population.

One of the main objectives of our study was to assess inter-rater reliability of assessing stability of the scapholunate joint in 4D CT scans. Although local surgeons were recruited, we hope our results will be generalizable to other surgeons. Thus, it is more appropriate to use the standard errors and confidence intervals that are unconditional on our rater sample. Examination of the data under the "Inference/Subjects & Raters" heading demonstrates that the standard errors are higher and that 95% confidence intervals are wider compared to the "Inference/Subjects" data. If we were only concerned with the extent of agreement amongst our local surgeons, the confidence intervals for Gwet's AC1 coefficient are 0.43 – 0.877. According to Altman's Scale, this correlates with anywhere from moderate to very good agreement.

7.3 Upper Extremity Subgroup

The Gwet's AC_1 coefficient for the subgroup of surgeons with a special interest in upper extremity surgery was 0.61150 ($95\%_{CI}$: 0.254 - 0.969). This was somewhat unexpected; the subgroup AC_1 value was lower than the AC_1 coefficient for the entire group. The wider confidence intervals could potentially be related to the smaller rater pool (5 vs. 7 raters). However, one should not interpret these results too closely because there is no statistically significant difference between the groups.

7.4 Surgeons' Preferred View

One does not require correlation coefficients to determine there is little agreement amongst surgeons for their preferred views for assessing scapholunate stability on the 4D CT scans (see Appendix G). The purpose of assessing surgeons' preferred views was not to test the statistical significance of inter-rater agreement. There is no right or wrong answer, and it is a completely subjective opinion. It is also neither realistic nor recommended for surgeons to rely on only one view when making diagnostic assessments. It does, however, provide insight into how the surgeons perceive relative stability or instability using four-dimensional CT scans.

Nevertheless, inter-rater agreement was poor amongst surgeons for the preferred 4D CT view. This was the case for all the correlation coefficients discussed previously. Interestingly, only Gwet's AC1 and the Brennan-Prediger coefficient were statistically significant. However, as discussed previously, this does not affect our statistical interpretation of the correlation coefficients.

7.5 Study Limitations

While the data suggests that there is good agreement when assessing 4D CT wrist scans, the wide confidence intervals limit the strength of our conclusions. It is not by any means a perfectly designed study. If unbounded by resources and reality, the study could be repeated under the following ideal conditions.

The most straightforward way to tighten the confidence intervals would be to increase the sample size. We were limited by the number of 4D CT scans available to us at the time of our data collection. A number of different methods are available to estimate sample size for inter-rater reliability studies. Some require knowledge of the percent agreement and percent chance agreement, which was known during the study design stage. Gwet proposed two different sample size charts that list the required subject or rater sample size depending on the desired error margin or coefficient of variation, respectively (Gwet, 2013) (see Table 14 and Table 15). An arbitrary value of 20% is commonly used. This would require a minimum of 25 subjects and 10 raters. Increasing the sample size or number of raters would subsequently decrease the error margin and coefficient of variation.

Table 14 - Sample Size as per the Desired Error Margin (Gwet, 2013)

Desired Error Margin	Required Subject Sample Size
5%	400
10%	100
15%	44
20%	25
25%	16
30%	11

Table 15 - Number of Raters as per Desired Variation Coefficient (Gwet, 2013)

Desired Coefficient of Variation	Required Number of Raters
5%	40
10%	20
15%	13
20%	10
25%	8
30%	7

Another aspect of the study that could be altered is the trait prevalence. If the number of pathologic 4D CT scans approached 50% of the total subjects, this would render Fleiss' kappa and Conger's kappa coefficients more applicable. It would be interesting to observe if all the correlation coefficients would approach equivalence. The designation of a "normal" wrist scan also relied on a healthy volunteer self-reporting any symptoms or past injuries.

In an ideal study, I would also attempt to standardize the wrists scan protocols further. The 4D CT scans utilized in this study were all conducted prior to this study design. The healthy wrists scans were collected in a previous study, where the principle investigator provided instructions and observed each of the scans taking place. The pathologic wrists utilized in this study were likely not provided the same instructions.

Although the patients are still performing the equivalent radial-ulnar deviation active motion, there are subtle differences in hand position and cadence compared to the healthy wrist studies. Some of the videos required digital cropping and alteration of the playing speed to remove any obvious visual differences between the scans. In addition, there is a degree of subjectivity when determining the true AP or lateral orientation of the wrist. The post-processing CT software enables manipulation of the raw data in any orientation. There was an automatic re-orientation function, but this may be influenced by the position of the hand within the CT gantry. It is uncertain whether this had any measurable effect on our data collection.

We also do not know which views of the wrist provide the best diagnostic information for assessing the scapholunate joint. The volar, dorsal, and mid-coronal views were previously selected by local musculoskeletal radiologists, and limited our ability to include different views (e.g. lateral). In addition, scapholunate joint does not lie in perfect accordance with the true anatomic planes. An oblique view looking along the joint axis would certainly influence our perception of joint stability Perhaps this could be an area of interest in future studies.

Finally, we were unable to investigate the inter- or intra-rater reliability of quantitative assessment of 4D CT scapholunate scans. This was not possible because the raw CT data was unavailable. A semi-automatic software is available that has the ability

to graph the dynamic distances between two manually set points throughout motion. It would be interesting to compare the qualitative and quantitative assessments of scapholunate stability in 4D CT scans.

Chapter 8: Conclusion

The validity of a new diagnostic tool must be fully assessed prior to being recommended for clinical application. For a diagnostic test to be valid, it must display accuracy and reliability. The accuracy of a test is essential so that the results truly reflect the specific question asked. Without accuracy, the diagnostic test is completely irrelevant. Reliability is another essential component for validity. An accurate test is clinically undependable if the true results cannot be reproduced. Reliability can relate to the extent of agreement amongst multiple raters, or the proportion of agreement at separate time points. Four-dimensional CT scans of the wrist are a relatively new diagnostic approach for assessing scapholunate pathology. Our objective was to shed light on the inter- and intra-rater reliability of the assessment of scapholunate stability in normal and abnormal wrists.

Our study suggests that utilizing 4D CT scans for assessing scapholunate stability has good inter- and intra-rater reliability. Although this study may be underpowered, it highlights the potential 4D CT has as a diagnostic tool. However, for 4D CT to be acceptable as a valid diagnostic test, studies must be performed to assess its accuracy. This may involve comparing 4D CT wrist scans to the gold standard diagnostic approach, wrist arthroscopy. This will yield important statistical parameters, such as sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio.

The future of musculoskeletal 4D CT scans is promising. Early literature has demonstrated creative applications of the four-dimensional capabilities of modern CT scanners. 4D CT scans have been performed on joints, including the hip, patellofemoral joint, acromioclavicular joint, scapulothoracic joint, subtalar joint, wrist, and thumb. Early literature has focused on technical feasibility of performing these musculoskeletal scans. Since the majority of clinical studies are descriptive in nature, rigorous testing of the validity of each diagnostic test must be performed prior to mainstream use. Reliability can be determined following a similar approach of this study protocol. We have discussed the merits of multiple statistical methods of measuring agreement amongst raters. These include percent agreement, Cohen's Kappa, Scott's Pi, Fleiss' Kappa, Conger's Kappa, Gwet's AC_1, Krippendorff's Alpha, and the Brennan-Prediger coefficient. Future inter-rater reliability study designs should consider the various correlation coefficients, and determine which ones are most applicable for that study.

This paper may contribute to the evidence base assessing the validity of assessing scapholunate stability in 4D CT scans. With future research, there is optimism that four-dimensional CT scans can prove to be a valuable, non-invasive diagnostic tool for investigating dynamic musculoskeletal conditions.

Bibliography

Alta, T., Bell, S., Troupis, J., Coghlan, J., & Miller, D. (2012). TitleThe new 4-dimensional computed tomographic scanner allows dynamic visualization and measurement of normal acromioclavicular joint motion in an unloaded and loaded condition. *J Comput Assist Tomogr, 36*(6), 749-754.

Altman, D. (1991). *Practical statistics for medical research.* London: Chapman and Hall.

Banarjee, M., Capazzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond Kappa: A Review of Interrater Agreement Measures. *The Canadian Journal of Statistics, 27*(1), 3-23.

Bell, S., Troupis, J., Miller, D., Alta, T., Coghlan, J., & Wijeratna, M. (2015). Four-dimensional computed tomography scans facilitate preoperative planning in snapping scapula syndrome. *J Shoulder Elbow Surg*(24), e83-90.

Boutin RD, B. M., Immerman, I., Ashwell, Z., Sonico, G., Szabo, R., & Chaudhari, A. (2013). Real-Time Magnetic Resonance Imaging (MRI) during Active Wrist Motion—Initial Observations . *PLoS ONE, 8*(12), e84004.

Brennan, R., & Prediger, D. (1981). Coefficient Kappa: Some Uses, Misuses, and Alternatives. *Educational and Psychological Measurement*(41), 687-699.

Cattaneo, G., Passoni, P., Sangalli, G., Slim, N., Longobardi, B., Mancosu, P., . . . Calandrino, R. (2010). Internal target volume defined by contrast-enhanced 4D-CT scan in unresectable pancreatic tumour: evaluation and reproducibility. *Radiother Oncol, 97*(3), 525–9.

Chae, M., Hunter-Smith, D., De-Silva, I., Tham, S., Spychal, R., & Rozen, W. (2015). Four-Dimensional (4D) Printing: A New Evolution in Computed Tomography-Guided Stereolithographic Modeling Principles and Application. *J Reconstr Microsurg, 31*(6), 458-463.

Choi, Y., Lee, Y., Kim, S., Cho, H., Song, H., & Suh, J. (2013). Four-dimensional real-time cine images of wrist joint kinematics using dual source CT with minimal time increment scanning. *Yonsei Med J, 54*(4), 1026 - 1032.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Conger, A. (1980). Integration and Generalization of Kappas for Multiple Raters. *Psychological Bulletin, 88*(2), 322-328.

Cosgarea, A., Carrino, J., Saranathan, A., Guseila, L., Tanaka, M., & Elias, J. (2013). Tibial tuberosity realignment alters in vivo patellar tracking. *Orthop J Sports Med, 1*(4).

Dejour, H., Walch, G., Nove-Josserand, L., & Guier, C. (1994). Factors of patellar instability: An anatomic radiographic study. *Knee Surg, Sports traumatol, Arthroscopy, 2*(19), 19-26.

Demehri, S., Hafezi-Nejad, N., Morelli, J., Thakur, U., Lifchez, S., Means, K., & Shores, J. (2016). Scapholunate kinematics of asymptomatic wrists in comparison with symptomatic contralateral wrists using four-dimensional CT examinations: initial clinical experience. *Skeletal Radiol*(45), 437-446.

Demehri, S., Hafezi-Nejad, N., Thakur, U., Morelli, J., Lifchez, S., Means, K., & Shores, J. (2015). Evaluation of pisotriquetral motion pattern using four-dimensional CT: initial clinical experience in asymptomatic wrists. *Clin Radiol, 70*(12), 1362-1369.

Demehri, S., Thawait, G., Williams, A., Kompel, A., Elias, J., Carrino, J., & Cosgarea, A. (2014a). Imaging characteristics of contralateral asymptomatic patellofemoral joints in patients with unilateral instability. *Radiology, 273*(3), 821-830.

Demehri, S., Wadhwa, V., Thawait, G., Fattahi, N., Means, K., Carrino, J., & Chhabra, A. (2014b). Dynamic evaluation of pisotriquetral instability using 4-dimensional computed tomography. *J Comput Assist Tomogr, 38*(4), 507-512.

Draper, C., Besier, T., Santos, J., Jennings, F., Fredericson, M., Gold, G., . . . Delp, S. (2009). Using Real-Time MRI to Quantify Altered Joint Kinematics in Subjects with Patellofemoral Pain and to Evaluate the Effects of a Patellar Brace or Sleeve on Joint Motion. *J Orthop Res, 27*(5), 571-577.

Draper, C., Santos, J., Kourtis, L., Besier, T., Fredericson, M., Beaupre, G., . . . Delp, S. (2008). Feasibility of using real-time MRI to measure joint kinematics in 1.5T and open-bore 0.5T systems. *J. Magn. Reson. Imaging, 28*, 158-166.

Dyer, D., Troupis, J., & Kamali Moaveni, A. (2015). Wide field of view CT and acromioclavicular joint instability: A technical innovation. *J Med Imaging Radiat Oncol , 59*(3), 326-330.

Edirisinghe, Y., Troupis, J., Patel, M., Smith, J., & Crossett, M. (2014). Dynamic motion analysis of dart throwers motion visualized through computerized tomography and calculation of the axis of rotation. *J Hand Surg Eur, 39*(4), 364-372.

Elias, J., Carrino, J., Saranathan, A., Guseila, L., Tanaka, M., & Cosgarea, A. (2014). TitleVariations in kinematics and function following patellar stabilization including tibial tuberosity realignment. *Knee Surg Sports Traumatol Arthrosc, 22*(10), 2350-2356.

Fernquest, S., Arnold, C., Palmer, A., Broomfield, J., Denton, J., Taylor, A., & Glyn-Jones, S. (2017). Osseous impingement occurs early in flexion in cam-type femoroacetabular impingement. *Bone Joint J, 4 Supple B*, 41-48.

Finnoff, J. (2017, March). *Musculoskeletal ultrasound of the shoulder*. Retrieved April 16, 2017, from UpToDate: https://www.uptodate.com/contents/musculoskeletal-ultrasound-of-the-shoulder

Fleiss, J. (1971). Measuring Nominal Scale Agreement Among Many Raters. *Psychological Bulletin, 76*(5), 378-382.

Forsberg, D., Lindblom, M., Quick, P., & Gauffin, H. (2016). Quantitative analysis of the patellofemoral motion pattern using semi-automatic processing of 4D CT data. *Int J Comput Assist Radiol Surg, 11*(9), 1731-1741.

Foumani, M., Strackee, S. v., Jonges, R., Blankevoort, L., & Streekstra, G. (2012). In-vivo dynamic and static three-dimensional joint space distance maps for assessment of cartilage thickness in the radiocarpal joint. *Clinical Biomechanics*(28), 151-156.

Garcia-Elias, M., & Lluch, A. (2016). Wrist Instabilities, Misalignments, and Dislocations. In S. Wolfe, W. Pederson, R. Hotchkiss, S. Kozin, & M. Cohen, *Green's Operative Hand Surgery* (7th Edition ed., Vol. 1, pp. 418-478). USA: Elsevier.

Garcia-Elias, M., Alomar Serrallach, X., & Monill Serra, J. (2014). Dart-throwing motion in patients with scapholunate instability: a dynamic four-dimensional computed tomography study. *J Hand Surg Eur Vol, 39*(4), 346-352.

Geissler, W., Freeland, A., Savoie, F., McIntyre, L., & Whipple, T. (1996). Intracarpal soft-tissue lesions associated with an intra-articular fracture of the distal end of the radius. *J Bone Joint Surg Am, 78*(3), 357-365.

George, E., Mitsouras, D., Kumamaru, K., Shah, N., Smith, S., Schultz, K., . . . Rybicki, F. (2013). Upper extremity composite tissue allotransplantation imaging. *Eplasty, 13*, e38.

Gervaise, A., Teixeira, P., Villani, N., Lecocq, S., Louis, M., & Blum, A. (2013). CT dose optimisation and reduction in osteoarticular disease. *Diagn Interv Imaging, 94*(4), 371-388.

Glen, S. (2017, 11 17). *Inter-rater Reliability IRR: Definition, Calculation*. Retrieved 12 03, 2017, from Statistics How To: http://www.statisticshowto.com/inter-rater-reliability/

Gobbi, R., Demange, M., de Avila, L., Araujo Filho, J., Moreno, R., Gutierrez, M., . . . Camanho, G. (2016). Patellar tracking after isolated medial patellofemoral ligament reconstruction: dynamic evaluation using computed tomography. *Knee Surg Sports Traumatol Arthrosc*, 1-9.

Gold, G. (2003). Dynamic and Functional Imaging of the Musculoskeletal System. *Seminars in Musculoskeletal Radiology, 7*, 245-8.

Goldman, L. (2007). Principles of CT and CT technology. *Journal of Nuclear Medicine Technology, 35*(3), 115-30.

Gondim Teixeira, P., Formery, A., Hossu, G., Winninger, D., Batch, T., Gervaise, A., & Blum, A. (2017b). Evidence-based recommendations for musculoskeletal kinematic 4D-CT studies using wide area-detector scanners: a phantom study with cadaveric correlation. *Eur Radiol, 27*(2), 437-446.

Gondim Teixeira, P., Formery, A.-S., Jacquot, A., Lux, G., Loiret, I., Perez, M., & Blum, A. (2017a). Quantitative analysis of subtalar joint motion with 4D CT: Proof of

concept with cadaveric and healthy subject evaluation. *Am J Roentgenol, 208*(1), 150-158.

Goodman, L., & Kruskal, W. (1954). Measures of Association for Cross Classifications. *Journal of the American Statistical Association, 49*(268), 732-764.

Goto, A., Leng, S., Sugamoto, K., Cooney, W., Kakar, S., & Zhao, K. (2014). In vivo pilot study evaluating the thumb carpometacarpal joint during circumduction. *Clin Orthop Relat Res, 472*(4), 1106-1113.

Guzzanti, V., Gigante, A., Di Lazzaro, A., & Fabbriciani, C. (1994). Patellofemoral Malalignment in Adolescents: Computerized Tomographic Assessment with or without Quadriceps Contraction. *Am J Sports Med, 22*(1), 55-60.

Gwet, K. (2002). Kappa Statistic is not Satisfactory for Assessing the Extent of Agreement Between Raters. *Statistical Methods for Inter-rater Reliability Assessment*(1), 1-5.

Gwet, K. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*(61), 29-48.

Gwet, K. (2011). *On Krippendorff's Alpha coefficient*. (Advanced Analytics LLC Publications) Retrieved 12 03, 2017, from AgreeStat: http://agreestat.com/research_papers/onkrippendorffalpha_rev10052015.pdf

Gwet, K. (2013). *Estimating the number of subjects and number raters when designing an inter-rater reliability study*. (Advanced Analytics LLC) Retrieved 12 05, 2017, from AgreeStat: http://agreestat.com/blog_irr/sample_size_calculation.html

Hallgren, K. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol, 8*(1), 23-34.

Halpenny, D., Courtney, K., & Torreggiani, W. (2012). Dynamic four-dimensional 320 section CT and carpal bone injury - a description of a novel technique to diagnose scapholunate instability. *Clin Radiol, 67*(2), 185-187.

Hislop-Jambrich, J., Troupis, J., & Moaveni, A. (2016). The Use of a Dynamic 4-Dimensional Computed Tomography Scan in the Diagnosis of Atraumatic

Posterior Sternoclavicular Joint Instability. *J Comput Assist Tomogr, 40*(4), 576-577.

Hodge, D., Beaulieu, C., Thabit, G., Gold, G., Bergman, A., Butts, R., . . . Herfkens, R. (2001). Dynamic MR imaging and stress testing in glenohumeral instability: Comparison with normal shoulders and clinical/surgical findings. *J. Magn. Reson. Imaging, 13*, 748–756.

Holley, W., & Guilford, J. (1964). A Note on the G-Index of Agreement. *Educational and Psychological Measurement*(24), 749-753.

Hounsfield, G. (1973). Computerized transverse axial scanning (tomography): part I. *British Journal of Radiology, 46*, 1016-22.

Hsiao, E., Rybicki, F., & Steigner, M. (2010). CT coronary angiography: 256-slice and 320-detector row scanners. *Curr Cardiol Rep, 12*, 68-75.

Hu, H. (1999). Multi-slice helical CT: Scan and reconstruction. *Medical Physics, 26*(1), 5-18.

Hunter, G., Schellingerhout, D., Vu, T., Perrier, N., & Hamberg, L. (2012). Accuracy of Four-dimensional CT for the Localization of Abnormal Parathyroid Glands in Patients with Primary Hyperparathyroidism. *Radiology, 264*(3), 789-95.

Jais, I., Liu, X., An, K., & Tay, S. (2014). A method for carpal motion hysteresis quantification in 4-dimensional imaging of the wrist . *Med Eng Phys, 36*(12), 1699-1703.

Kakar, S., Breighner, R., Leng, S., McCollough, C., Moran, S., Berger, R., & Zhao, K. (2016). The Role of Dynamic (4D) CT in the Detection of Scapholunate Ligament Injury. *J Wrist Surg, 5*(4), 306-310.

Kalender, W., Seissler, W., Klotz, E., & Vock, P. (1990). Spiral volumetric CT with single-breath-hold technique, continuous transport, and continuous scanner rotation. *Radiology, 176*, 181-83.

Kalia, V., Obray, R., Filice, R., Fayad, L., Murphy, K., & Carrino, J. (2009). Functional joint imaging using 256-MDCT: Technical feasibility. *Am J Roentgenol, 192*(6).

Kerkhof, F., Brugman, E., D'Agostino, P., Dourthe, B., van Lenthe, G., Stockmans, F., . . . Vereecke, E. (2016). Quantifying thumb opposition kinematics using dynamic computed tomography. *J Biomech, 49*(9), 1994-1999.

Kipritidis, J., Hugo, G., Weiss, E., Williamson, J., & Keall, P. (2015). Measuring interfraction and intrafraction lung function changes during radiation therapy using four-dimensional cone beam CT ventilation imaging. *Medical Physics, 42*(3), 1255-1267.

Kuppens, S., Holden, G., Barker, K., & Rosenberg, G. (2011). A Kappa-related Decision: κ, Y, G, or AC1. *Social Work Research, 35*(3), 185-189.

Kwong, Y., Mel, A., Wheeler, G., & Troupis, J. (2015). Four-dimensional computed tomography (4DCT): A review of the current status and applications. *J Med Imaging Radiat Oncol, 59*(5), 545-554.

Lee, S., Desai, H., Silver, B., Dhaliwal, G., & Paksima, N. (2011). Comparison of Radiographic Stress Views for Scapholunate Dynamic Instability in a Cadaver Model. *J Hand Surg Am, 36*(7), 1149-57.

Leng, S., Zhao, K., Qu, M., An, K., Berger, R., & McCollough, C. (2011). Dynamic CT technique for assessment of wrist joint instabilities. *Med Phys, 38*(Suppl 1), S50.

Marmor, M., Hansen, E., Han, H., Buckley, J., & Matityahu, A. (2011). Limitations of Standard Fluoroscopy in Detecting Rotational Malreduction of the Syndesmosis in an Ankle Fracture Model. *Foot & Ankle International, 32*(6), 617-622.

McGraw, K., & Wong, S. (1996). Forming inferences about some intraclass correlation coefficients. *Psychol Methods*(1), 30-46.

McGuire, K., Silber, J., Flynn, J., Levine, M., & Dormans, J. (2002). Torticollis in children: can dynamic computed tomography help determine severity and treatment. *J Paediatr Orthop, 22*(6), 766-770.

Moojen, T., Snel, J., Ritt, M., Venema, H., Kauer, J., & Bos, K. (2003). In vivo analysis of carpal kinematics and comparative review of the literature. *J Hand Surg, 28*(1), 81-87.

Nikolaou, K., Flohr, T., Knez, A., Rist, C., Wintersperger, B., T, J., . . . Becker, C. (2004). Advances in cardiac CT imaging: 64-slice scanner. *Int J Cardiovasc Imaging, 2*-(6), 535-40.

Ozçelik, A., Günal, I., & Köse, N. (2005). Stress views in the radiography of scapholunate instability. *Eur J Radiol, 56*, 258-61.

Pan, T. (2013). Helical 4D CT and Comparison with Cine 4D CT. In J. Ehrhardt, & C. Lorenz (Eds.), *4D Modelling and Estimation of Respiratory Motion for Radiation Therapy* (pp. 25-41). Springer-Verlag Berlin Heidelberg.

Park, M. (2002). Normal anteroposterior laxity of the radiocarpal and midcarpal joints. *JBJS Br, 84*, 73-76.

Penning, L., & Wilmink, J. (1987). Posture-dependent bilateral compression of L4 or L5 nerve roots in facet hypertrophy. A dynamic CT-myelographic study. *Spine, 12*(5), 488-500.

Petscavage-Thomas, J. (2014). Clinical applications of dynamic functional musculoskeletal ultrasound. *Reports in Medical Imaging, 7*, 27-39.

Renoux, J., Zeitoun-Eiss, D., & Brasseur, J. (2009). Ultrasonographic Study of Wrist Ligaments: Review and New Perspectives. *Semin Musculoskelet Radiol, 13*(1), 55-65.

Repse, S., Amis, B., & Troupis, J. (2015a). Four-dimensional computed tomography and detection of dynamic capitate subluxation. *J Med Imaging Radiat Oncol, 59*(3), 331-335.

Repse, S., Koulouris, G., & Troupis, J. (2015b). Wide field of view computed tomography and mid-carpal instability: the value of the sagittal radius-lunate-capitate axis--preliminary experience. *Eur J Radiol, 84*(5), 908-914.

Roberts, W., Bax, J., & Davies, L. (2008). Cardiac CT and CT Coronary Angiography: Technology and Application. *Heart, 94*(6), 781-792.

Ropers, D., Baum, U., Pohle, K., Anders, K., Ulzheimer, S., Ohnesorge, B., . . . Achenbach, S. (2003). Detection of Coronary Artery Stenoses With Thin-Slice Multi-Detector Row Spiral Computed Tomography and Multiplanar Reconstruction. *Circulation, 107*(5), 664-666.

Schädel-Höpfner, M., Böhringer, G., Gotzen, L., & Celik, I. (2005). Traction radiography for the diagnosis of scapholunate ligament tears. *J Hand Surg Br, 30*, 464–7.

Shores, J., Demehri, S., & Chhabra, A. (2013). Kinematic "4 Dimensional" CT Imaging in the Assessment of Wrist Biomechanics Before and After Surgical Repair. *Eplasty, 13*, e9.

Simundić, A. (2009). Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC, 19*(4), 203-211.

Soderman, T., Olerud, C., Shalabi, A., Alavi, K., & Sundin, A. (2015). Static and dynamic CT imaging of the cervical spine in patients with rheumatoid arthritis. *Skeletal Radiol, 44*(2), 241-248.

Sodickson, A., Baeyens, P., Andriole, K., Prevedello, L., Nawfel, R., Hanson, R., & Khorasani, R. (2009). Recurrent CT, Cumulative Radiation Exposure and Associated Radiation-induced Cancer Risks from CT of Adults. *Radiology, 251*(1), 175-184.

Strugarek-Lecoanet, C., Chevrollier, J., Pauchard, N., Blum, A., Dap, F., & Dautel, G. (2016). Morphology and Mobility of the Reconstructed Basilar Joint of the Pollicized Index Finger. *J Hand Surg (USA), 41*(9), e267-e272.

Sulkers, G., Schep, N., Maas, M., van der Horst, C., Goslings, J., & Strackee, S. (2014). The diagnostic accuracy of wrist cineradiography in diagnosing scapholunate dissociation. *J Hand Surg Eur Vol, 39*(3), 263-271.

Troupis, J., & Amis, B. (2013). Four-dimensional Computed Tomography and Trigger Lunate Syndrome . *J Comput Assist Tomogr, 37*(4), 639-643.

van Riet, R., & Bell, S. (2011). Clinical evaluation of acromioclavicular joint pathology: sensitivity of a new test. *J Shoulder Elbow Surg, 20*.

Wang, L., Hayes, S., Paskalev, K., Jin, L., Buyyounouski, M., Ma, C., & Feigenberg, S. (2009). Dosimetric comparison of stereotactic body radiotherapy using 4D CT and multiphase CT images for treatment planning of lung cancer: Evaluation of the impact on daily dose coverage. *Radiotherapy and Oncology, 91*(3), 314-24.

Wang, S., Li, J., Wang, W., Zhang, Y., Li, F., Fan, T., & Shang, D. (2012). A study on the displacements of the clips in surgical cavity for external-beam partial breast irradiation after breast-conserving surgery based on 4DCT. *J Radiat Res, 53*(3), 433-8.

Wassilew, G., Janz, V., Heller, M., Tohtz, S., Rogalla, P., Hein, P., & Perka, C. (2013). Real time visualization of femoroacetabular impingement and subluxation using 320-slice computed tomography. *J Orthop Res, 31*(2), 275-281.

Waters, H. (2011). The First X-ray, 1895. *The Scientist, 25*(7).

Williams, A., Elias, J., Tanaka, M., Thawait, G., Demehri, S., Carrino, J., & Cosgarea, A. (2016). The Relationship Between Tibial Tuberosity-Trochlear Groove Distance and Abnormal Patellar Tracking in Patients With Unilateral Patellar Instability. *Arthroscopy, 32*(1), 55-61.

Yeo, U., Taylor, M., Supple, J., Siva, S., Kron, T., Pham, D., & Franich, R. (2014). Evaluation of dosimetric misrepresentations from 3D conventional planning of liver SBRT using 4D deformable dose integration. *J Appl Clin Med Phys, 15*(6), 188-203.

Zhao, K., Breighner, R., Holmes, D., Leng, S., McCollough, C., & An, K. (2015). A technique for quantifying wrist motion using four-dimensional computed tomography: approach and validation. *J Biomech Eng, 137*(7).

Appendix A: Summary of Literature Review

Joint	Study	Design	Sample	Intervention	Summary
HIP	Wassilew et al., 2013	Cross-sectional comparative study	30 patients with FAI	4D CT scans of hips were compared with intra-operative findings during FAI surgery	There was 97% accuracy and a κ coefficient of 0.87-0.92 for anterior subluxation, and 93% and a κ coefficient of 0.72-0.94 for posterior subluxation
	Fernquest et al., 2017	Case series/Cross-sectional study	50 patients	4D CT scans were performed pre-operatively prior to hip arthroscopy	Mean point of FAI impingement was 41.36° of hip flexion
PATELLO-FEMORAL	Demehri et al., 2014a	Cross-sectional retrospective review	25 patients / 25 controls	Asymptomatic contralateral patella-femoral joints had 4D CT scans and were compared to controls	Trochlear groove depth (TGD) (median=3.0 mm), TT-TG (median= 15 mm), patellar height ratio (PHR) (median= 1.17), and patellar lateral displacement showed a statistically significant difference compared to healthy controls
	Williams et al., 2016	Cross-sectional comparative study	25 patients with PF instability	4D CT scans comparing unilateral patellofemoral instability with contralateral knee	Significant difference of lateral patellar tilt (LPT), bisect offset (BO), and TT-TG distance at 5° and 15° of knee flexion between symptomatic and asymptomatic knee. TT-TG distance was also significantly increased at 25° and 35°
	Forsberg et al., 2016	Case series / Cohort study	5 patients with PF instability	4D CT scans obtained pre- and post-MPFL reconstructions	A semi-automatic method of quantifying patellofemoral motion was preferable to a manual method. MPFL reconstruction reduced the patellar displacement by 10-15 mm and the patellar tilt by 15-20° in 3/5 patients

Joint	Study	Design	Sample	Intervention	Summary
	Cosgarea et al., 2013	Case series / Cohort study	6 patients with PF instability	4D CT scans obtained pre- and 1 year post- tibial tuberosity realignment	Patellar lateral shift and tilt decreased by an average of 3 mm and 3° respectively post-operatively. Tibial external rotation increased by an average of 2°.
	Elias et al., 2014	Case series / Cohort	5 patients (6 knees) with PF instability	4D CT scans obtained pre- and 1 year post- MPFL reconstructions	The patellar lateral shift and tilt decreased by 5.3mm and 4.2° at 5° of knee flexion. There was a significant difference at 5 and 10° of flexion
	Gobbi et al., 2016	Case series / Cohort	10 patients with PF instability	4D CT scans obtained pre- and 6 months post- MPFL reconstructions	There was no difference in patellar tracking following surgery. Clinical scores were statistically significantly improved and no relapses of instability were noted
FOOT/ANKLE *Subtalar joint*	Gondim Teixeira et al., 2017a	Cross-sectional comparative study	3 cadavers and 15 healthy volunteers	4D CT evaluated healthy subjects and cadavers. The scan of the cadaver was repeated after partial, then full resection of the subtalar ligaments (cervical and interosseous)	Joint amplitudes ranged from 6.4-22.8% in the healthy sample. Joint amplitudes increased 19% and 22% following partial and full subtalar ligament resection, respectively
Ankle joint	Gondim Teixeira et al., 2017b	Case report	1 cadaver	4D CT scan captured ankle F/E (flexion/extension) and external-internal rotation (ER/IR)	The authors determined the most important factor for high image quality was volume acquisition speed. Image quality was greatest at the highest amplitudes of joint motion. Motions parallel to the CT gantry had more significant artefacts.
SHOULDER *AC joint*	Alta et al., 2012	Cross-sectional study	16 healthy volunteers	Motion at the AC joint was evaluated with shoulder adduction, superior elevation, with and without resistance using 4D CT	Motion at the AC joint is predominantly posterior 1.1±0.9 mm, P = 0.001) in the coronal plane, followed by superior translation of the clavicle (0.6 ± 0.5 mm, P = 0.001)

Joint	Study	Design	Sample	Intervention	Summary
AC joint	Dyer et al., 2015	Case report	1 patient with a Grade II or III AC joint injury	4D CT captured motion at the AC joint, with shoulder forward flexion to 45°, internal rotation to 90°, and adduction to 45°	Motion was detected in 3 planes: anterior-posterior, superior-inferior, and narrowing of the width of the AC joint
SC joint	Hislop-Jambrich et al, 2016	Case report	1 patient with atraumatic posterior sternoclavicular instability	4D CT captured motion at the SC joint, with shoulder flexion, and cross-body adduction	Significant motion of the clavicle posteriorly and inferiorly was detected.
Scapulothoracic joint	Bell et al., 2015	Case series / Cross-sectional study	12 patients with snapping scapula syndrome	4D CT recorded patients reproducing their 'snapping' motion	The 4D CT scans showed 5/12 patients had direct bony contact between the scapula and the posterior ribs, 4/12 had close contact with the posterior ribs, and 3/12 had no contact but a likely soft tissue point of impingement. The authors reported the 4D CT scans assisted with their pre-operative evaluation
HAND *Thumb*	Goto et al., 2014	Cross-sectional study	1 volunteer, 1 cadaver	4D CT scan captured thumb circumduction in 1 healthy volunteer and 1 cadaver	Maximum joint contact area between the thumb and trapezium occurred during thumb palmar abduction; the least joint contact area was during thumb adduction
Thumb	Kerkhof et al., 2016	Cross-sectional study	6 cadavers	4D CT captured thumb opposition of cadaveric thumbs. This was validated against an attached calibration cube, and against inter-osseous beads	The precision of 4D CT scan was within 0.13 mm when compared to the inter-osseous beads validation method. Motion was noted at the trapezio-metacarpal and scapho-trapezio-trapezoidal joints during thumb opposition

Joint	Study	Design	Sample	Intervention	Summary
Thumb	Strugarek-Lecoanet et al., 2016	Case series	14 patients (23 thumbs)	4D CT scans were performed on patients with previous index finger pollicization procedures for reconstruction of congenital thumb hypoplasia	20 cases had good motion of the reconstructed thumb. 16 thumbs showed remodeling of the metacarpal heads.
Thumb	Chae et al., 2015	Case report	1 healthy volunteer	4D CT captured thumb abduction, opposition, and key pinch of a healthy volunteer	The 4D CT scan data was derived for the basis of constructing 3D printed models of the hand in the aforementioned positions
Hand	George et al., 2013	Case series	2 patients (3 hands)	4D CT scans evaluated hand motion 8 months post-hand transplantation	The authors noted the importance of advanced musculoskeletal imaging pre- and post- hand transplantation procedures
WRIST	Zhao et al., 2015	Anatomical study	1 cadaver	Comparison of motion capture (wrist F/E and R/U [radioulnar] deviation) between 4D CT scan and fiducial beads.	4D CT measurements were accurate within 0.00–0.68° for rotation and within 0.02-0.30 mm for translation compared fiducial beads
	Jais et al., 2014	Anatomical study	1 cadaver	Carpal hysteresis was measured following 4D CT of cadaveric wrist during radioulnar deviation	The carpal hysteresis effect was highest in the lunate, followed by the triquetrum and scaphoid. Inter-observer reliability was highest when quantifying hysteresis for the scaphoid
	Choi et al., 2013	Proof-of-concept study	2 volunteers	Wrist R/U deviation and P/S (pronation-supination) were captured using a dual-source 4D CT scanner	Dual-source CT scanner enabled a reduction in scan time and radiation exposure compared to previously reported studies
	Edirisinghe et al., 2014	Cross-sectional study	7 healthy volunteers	4D CT captured volunteers performing the dart thrower's motion (DTM)	The rotational axis of the dart throwing motion was measured to be 27° of anteversion in the coronal plane 44° of varus in the transverse plane
Capitate	Repse et al., 2015a	Case report	1 patient with dynamic capitate subluxation	4D CT of the abnormal and normal wrists were performed during R/U deviation and F/E	4D CT of the wrist displayed abnormal subluxation of the capitate relative to the lunate, compared to the contralateral normal side

Joint	Study	Design	Sample	Intervention	Summary
Mid-carpal joint	Repse et al., 2015b	Case series / Cross-sectional study	20 patients with mid-carpal instability	4D CT captured motion of every patient's bilateral wrists in R/U deviation, F/E, and the DTM	The abnormal wrists demonstrated instability between the proximal and distal rows of the carpal bones using 4D CT. Three independent pathologic findings on 4D CT were determined for CIND: vacuum phenomenon within the lunate-capitate joint, triggering of the lunate, and subluxation of the capitate
Piso-triquetral joint	Demehri et al., 2014b	Case series	2 patients with pisotriquetral instability	4D CT scans performed on wrists during P/S, F/E, and F/U deviation	One patient demonstrated abnormal volar subluxation of the pisiform relative to the triquetrum. The second patient demonstrated abnormal proximal subluxation of the pisiform.
Piso-triquetral joint	Demehri et al., 2015	Cross-sectional study	10 healthy volunteers	4D CT captured wrist flexion-extension for measurements of the anteroposterior (AP) interval and craniocaudal (CC) excursion of the pisotriquetral joint	There was high inter-rater reliability measuring the AP interval (ICC=0.80), and moderate inter-rater reliability for CC excursion of the pisotriquetral joint (ICC=0.40)
Lunate	Troupis & Amis, 2013	Case report	1 patient with trigger mid-carpal instability	4D CT recorded wrist motion during F/E and R/U deviation	A triggering phenomenon of the lunate was observed during the flexion-extension arc, with abrupt cessation and subsequent resumption of motion. The authors referred to this as trigger lunate syndrome
Scapholunate joint	Leng et al., 2011	Biomechanical study	1 cadaver	4D CT recorded R/U motion of a cadaveric wrist pre- and post-resection of all portions scapholunate ligament	Mean image quality was similar between static and dynamic CT images. All 3 orthopaedic surgeons correctly identified the 4D CT scans of the wrist with the simulated scapholunate joint instability

Joint	Study	Design	Sample	Intervention	Summary
Scapholunate joint	Halpenny et al., 2012	Case report	1 patient with a scapholunate injury	4D CT recorded F/E of the injured wrist. Initial static studies were normal	4D CT revealed the scapholunate distance increased to 6 mm during the dynamic study. This assisted the authors with their diagnosis of dynamic scapholunate instability
	Kakar et al., 2016	Case series	2 patients with scapholunate instability	4D CT capture wrist F/E, R/U deviation, and the DTM	The 4D CT scans were clinically useful in diagnosing scapholunate injury. The flexion-extension and radioulnar deviation motions were most helpful in assessing scapholunate instability
	Shores et al., 2013	Cross-sectional study	4 patients and 1 healthy volunteer (6 wrists)	4D CT recorded wrist F/E, R/U deviation, P/S, and DTM in patients with previous scapholunate surgery. Pre-op 4D CT scans of ulnocarpal impaction were also performed with F-E, R-U deviation, and clenched fist. 1 healthy volunteer was scanned	The 4D CT scans of the post-surgical patients showed altered biomechanics of the mid-carpal joint, with initiation of motion at the radio-carpal joint.
	Garcia-Elias et al., 2014	Cross-sectional study	6 patients with scapholunate instability and 6 healthy volunteers	4D CT captured the dart-throwing motion. 2 radiologists and 1 surgeons focused on the movement of the scaphoid and lunate in the coronal plane, and the scaphotrapezial angle in the dart-throwing plane of motion	In normal wrists, the scaphoid and lunate translate in the coronal plane, but do not flex or extend. Scapholunate injury induces a gap during the dart-throwing motion. The authors do no recommend immediate dart-throwing range of motion following scapholunate ligament repair
Scapholunate joint	Demehri et al., 2016	Cross-sectional comparative study	12 wrists with suspected SL pathology and 10 asymptomatic wrists	4D CT scans were performed with active wrist fist clench, F/E, R/U deviation	In the symptomatic patients, the scapholunate interval was widest in extension (2.54±1.48mm), followed by clenched fist (2.53±1.19mm), then ulnar deviation (2.06±1.12mm). The scapholunate interval consistently remained < 1 mm in all movements in the healthy controls

Joint	Study	Design	Sample	Intervention	Summary
	Foumani et al., 2012	Cross-sectional study	10 healthy volunteers and 3 patients with radiocarpal OA	Static and 4D CT scans were performed. A mechanical hand-shaker device guided patients through F/E, R/U deviation and DTM. Static and dynamic distance maps were produced of the radiocarpal joint	There was a significant reduction in the dynamic distance map compared to the static distance map. This difference was smaller in the patients with osteoarthritis
MULTIPLE JOINTS *Patella-femoral and wrist joints*	Kalia et al., 2009	Cross-sectional study	6 healthy volunteers	4D CT scans captured F/E of the knee (patellofemoral joint), and P/S and R/U deviation of the wrist	The authors highlighted the advantages of using 256-slice CT scanners over 64-slice CT scanners for 4D CT scans

Key:

4D CT = four-dimensional CT; AC = acromioclavicular, AP = anteroposterior; BO = bisect offset; CC = craniocaudal; CIND = carpal instability nondissociative; ER = external rotation; FAI = femoroacetabular impingement; F/E = flexion-extension; ICC = intraclass correlation coefficient; IR = internal rotation; LPT = lateral patellar tilt; MPFL = medial patellofemoral ligament; SC = sternoclavicular; TGD = trochlear groove depth; TT-TG = tibial tuberosity-trochlear groove

Appendix B: HREB Ethics Approval Letter

Ethics Office
Suite 200, Eastern Trust Building
95 Bonaventure Avenue
St. John's, NL
A1B 2X5

March 21, 2017

Dear Dr. Chang:

Researcher Portal File # 20171891
Reference # 2017.068

RE: "Evaluation of Inter-Observer Variability in Assessment of Dynamic CT Scans"

Your application received a delegated review by a sub-committee of the Health Research Ethics Board (HREB). ***Full approval*** of this research study is granted for one year effective **March 21, 2107**.

This is your ethics approval only. Organizational approval may also be required. It is your responsibility to seek the necessary organizational approval from the Regional Health Authority (RHA) or other organization as appropriate. You can refer to the HREA website for further guidance on organizational approvals.

This is to confirm that the HREB reviewed and approved or acknowledged the following documents (as indicated):

- Application, approved
- Study information sheet version 2 dated March 14, 2017, approved
- Revised version 2.0 - qualitative assessment form, dated March 13, 2017, approved
- Revised version 2.0 - quantitative assessment form, dated March 13, 2017, approved
- Letter for request, approved

MARK THE DATE

This approval will lapse on March 21, 2018. It is your responsibility to ensure that the Ethics Renewal form is submitted prior to the renewal date; you may not receive a reminder. The Ethics Renewal form can be found on the Researcher Portal as an Event form.

If you do not return the completed Ethics Renewal form prior to date of renewal:

- ***You will no longer have ethics approval***
- *You will be required to stop research activity immediately*
- *You may not be permitted to restart the study until you reapply for and receive approval to undertake the study again*
- *Lapse in ethics approval* ***may result in interruption or termination of funding***

You are solely responsible for providing a copy of this letter, along with your approved HREB application form; **to Research Grant and Contract Services** should your research depend on funding administered

through that office.

Modifications of the protocol/consent are not permitted without prior approval from the HREB. **Implementing changes without HREB approval may result in your ethics approval being revoked, meaning your research must stop**. Request for modification to the protocol/consent must be outlined on an amendment form (available on the Researcher Portal website as an Event form) and submitted to the HREB for review.

The HREB operates according to the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2), the Health Research Ethics Authority Act (HREA Act) and applicable laws and regulations.

You are responsible for the ethical conduct of this research, notwithstanding the approval of the HREB.

We wish you every success with your study.

Sincerely,

Ms. Patricia Grainger (Chair, Non-Clinical Trials Health Research Ethics Board)
Dr. Joy Maddigan (Vice-Chair, Non-Clinical Trials Health Research Ethics Board)

CC: Dr. Andrew Furey

Appendix C: RPAC Institutional Approval Letter

Eastern Health

Department of Research
5th Floor Janeway Hostel
Health Sciences Centre
300 Prince Philip Drive
St. John's, NL A1B 3V6
Tel: (709) 752-4636
Fax: (709) 752-3591

May 9, 2017

Dr. Nicholas Chang
Room H-1385, HSC
St. John's, NL
A1B 3V6

Dear Dr. Chang,

Your research proposal *HREB Reference #: 2017.068 "Evaluation of Inter-Observer Variability in Assessment of Dynamic CT Scans"* was reviewed by the Research Proposals Approval Committee (RPAC) of Eastern Health at a meeting dated May 9, 2017 and we are pleased to inform you that the proposal has been granted full approval.

The approval of this project is subject to the following conditions:

- The project is conducted as outlined in the HREB approved protocol;
- Adequate funding is secured to support the project;
- In the case of Health Records, efforts will be made to accommodate requests based upon available resources. If you require access to records that cannot be accommodated, then additional fees may be levied to cover the cost;
- A progress report being provided upon request.

If you have any questions or comments, please contact Sharon Newman, Manager of the Patient Research Centre at 777-7283 or by email at sharon.newman@easternhealth.ca.

Sincerely,

Mike Doyle, PhD
Director of Research
Chair, RPAC

MD/rg

Appendix D: Link to Normal 4D CT Scan of Wrist

LINK:

https://youtu.be/Gpno58ropyE

QR CODE:

Appendix E: Link to 4D CT Scan of Scapholunate Instability

LINK:

https://youtu.be/GCy97_zw5d4

QR CODE:

Appendix F: Raw Data for Initial Rater Responses of Scapholunate Stability

Wrist Study	**Rater 1**	**Rater 2**	**Rater 3**	**Rater 4**	**Rater 5**	**Rater 6**	**Rater 7**
A	Stable	Stable	Stable	Stable	Stable	Stable	Stable
B	Stable	Stable	Stable	Stable	Stable	Stable	Stable
C	Stable	Stable	Stable	Stable	Stable	Stable	Stable
E	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable
G	Stable	Stable	Stable	Stable	Unstable	Stable	Stable
I	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable	Stable
M	Unstable	Stable	Unstable	Stable	Stable	Stable	Stable
N	Stable	Stable	Stable	Stable	Stable	Stable	Unstable
O	Stable	Stable	Unstable	Stable	Stable	Stable	Stable
P	Stable	Stable	Unstable	Unstable	Unstable	Stable	Stable
Q	Unstable	Unstable	Unstable	Unstable	Unstable	Stable	Unstable
R	Stable	Stable	Stable	Stable	Stable	Stable	Stable
S	Stable	Stable	Stable	Stable	Unstable	Stable	Stable
T	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable
W	Stable	Stable	Stable	Stable	Stable	Stable	Stable
X	Stable	Stable	Stable	Stable	Stable	Stable	Stable
Y	Unstable	Stable	Stable	Unstable	Unstable	Stable	Stable

Appendix G: Raw Data of Surgeons' Preferred Views of 4D CT Scans

Wrist Study	**Rater 1**	**Rater 2**	**Rater 3**	**Rater 4**	**Rater 5**	**Rater 6**	**Rater 7**
A	4D Dorsal	2D Coronal	4D Dorsal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal
B	4D Dorsal	2D Coronal	2D Coronal	2D Coronal	2D Coronal	2D Coronal	2D Coronal
C	2D Coronal	2D Coronal	2D Coronal	2D Coronal	2D Coronal	2D Coronal	2D Coronal
E	4D Dorsal	2D Coronal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal	2D Coronal
G	4D Dorsal	2D Coronal	4D Dorsal	2D Coronal	4D Volar	2D Coronal	2D Coronal
I	4D Dorsal	2D Coronal	4D Dorsal	4D Volar	4D Dorsal	2D Coronal	2D Coronal
M	4D Dorsal	2D Coronal	4D Dorsal	2D Coronal	2D Coronal	2D Coronal	2D Coronal
N	2D Coronal	2D Coronal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal	2D Coronal
O	4D Dorsal	2D Coronal	4D Dorsal	2D Coronal	4D Dorsal	2D Coronal	2D Coronal
P	2D Coronal	2D Coronal	4D Dorsal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal
Q	2D Coronal	2D Coronal	2D Coronal	2D Coronal	4D Volar	4D Dorsal	2D Coronal
R	4D Dorsal	2D Coronal	2D Coronal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal
S	4D Dorsal	2D Coronal	2D Coronal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal
T	2D Coronal	2D Coronal	4D Dorsal	4D Dorsal	4D Dorsal	2D Coronal	2D Coronal
W	4D Dorsal	2D Coronal	2D Coronal	2D Coronal	4D Dorsal	2D Coronal	2D Coronal
X	2D Coronal	2D Coronal	2D Coronal	2D Coronal	4D Dorsal	2D Coronal	2D Coronal
Y	2D Coronal	2D Coronal	2D Coronal	2D Coronal	4D Dorsal	2D Coronal	2D Coronal

Appendix H: Raw Data for Responses of Scapholunate Stability at 3 Months

Wrist Study	**Rater 1**	**Rater 2**	**Rater 3**	**Rater 4**	**Rater 5**	**Rater 6**	**Rater 7**
A	Stable	Stable	Stable	Stable	Stable	Stable	Stable
B	Stable	Stable	Stable	Stable	Stable	Stable	Stable
C	Stable	Stable	Stable	Stable	Stable	Stable	Stable
E	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable
G	Stable	Stable	Stable	Stable	Unstable	Stable	Stable
I	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable
M	Stable	Stable	Stable	Stable	Unstable	Stable	Stable
N	Stable	Stable	Stable	Stable	Stable	Stable	Stable
O	Stable	Stable	Stable	Stable	Stable	Stable	Stable
P	Stable	Stable	Stable	Unstable	Unstable	Stable	Stable
Q	Unstable	Unstable	Unstable	Unstable	Unstable	Stable	Unstable
R	Stable	Stable	Unstable	Stable	Stable	Stable	Stable
S	Stable	Stable	Stable	Stable	Stable	Stable	Stable
T	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable	Unstable
W	Stable	Stable	Unstable	Unstable	Stable	Stable	Stable
X	Stable	Stable	Stable	Stable	Stable	Stable	Stable
Y	Stable	Stable	Unstable	Stable	Stable	Stable	Stable