EFFECTS OF BICYCLING ON THE KNEE: A COMPUTATIONAL STUDY

BIOMECHANICAL AND MEDICAL IMAGING INVESTIGATION OF THE EFFECTS OF BICYCLING ON THE KNEE USING COMPUTATIONAL METHODS

By ANTHONY A. GATTI, B.Sc (Hons), M.Sc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2020) McMaster University

(School of Rehabilitation Science) Hamilton, Ontario, Canada

TITLE: Biomechanical and Medical Imaging Investigation of the Effects Of Bicycling on the Knee Using Computational Methods

AUTHOR: Anthony A. Gatti, B.Sc (Hons), M.Sc.

SUPERVISOR: Dr. Monica R. Maly, PT, PhD.

NUMBER OF PAGES: xxi, 266

Abstract

Knee osteoarthritis (OA) is a global problem that causes joint pain and decreased mobility and quality of life. Knee OA costs the Canadian economy billions of dollars. Cartilage and bone are both implicated in knee OA pathogenesis. Obesity is a major risk factor for knee OA. Physical activity decreases pain and improves quality of life in those with knee OA. Nonetheless, we have limited biomechanical evidence to create concrete recommendations for prescription of aerobic exercise that improves clinical outcomes without exacerbating pain or worsening joint structures in knee OA. We have a limited understanding of how cartilage of the OA knee responds to physical activity, and the role of bone shape on the response.

This thesis fills four identified gaps in the literature. First, Chapter 2 used a fully-crossed random assignment study design where 40 healthy participants completed 18 bicycling positions to define novel equations for setting bicycle saddle position based on minimum or maximum knee flexion angle. This work is important because the current gold-standard of setting bicycle saddle position for mitigating injury focuses on a desired knee flexion angle; yet no easy methods exist. Second, Chapter 3 used the same dataset to identify how joint kinematics affect tibiofemoral and patellofemoral joint forces during bicycling. This work showed joint forces are least sensitive to the gold standard bicycle-fit parameter, minimum knee flexion angle; instead, minimum hip flexion angle was the most important. Third, Chapter 4 describes and validates a multi-stage convolutional neural network framework for efficiently segmenting cartilage and bone from magnetic

resonance imaging data. The algorithm produced state-of-the-art predictions on the commonly tested Osteoarthritis Initiative dataset in an average of 1.5 mins per knee. These methods will be crucial for improving experimental and epidemiologic studies of cartilage and bone. Fourth, Chapter 5 combines statistical shape models of the tibia and femur, joint forces estimated at the knee, and statistical parametric mapping to explore continuously over the cartilage surface how cartilage deforms after walking and bicycling. This study showed for the first time that the acute response of cartilage in women with symptomatic knee OA is dependent on bone shape and knee joint forces.

The bicycle-fit related studies provide the first comprehensive insights into how lower limb joint kinematics affect knee joint forces and provide novel equations to use this knowledge to easily set bicycle saddle position in the clinic, bicycle shop, or at home. The image analysis chapter describes an image segmentation framework that excels when applied to the knee. The final chapter integrates biomechanical measures with statistical shape models using custom data processing pipelines that yielded new insights and that hold great potential for evoking novel and specific findings about knee OA pathophysiology at the intersection of bone, cartilage, and mechanics.

Acknowledgements

This thesis would not have been possible without the support of numerous people. To my Supervisor Dr. Monica Maly. As any relationship that has lasted >7 years we've had our share of ups as well as downs. Thank you for sharing these experiences with me, for helping me grow through them, and for continuing to have mine and all of your students best-interests at heart. You put so much into your students at every level and you have been foundational in my growth across the breadth of skills and competencies necessary to succeed as a researcher. Your efforts have been instrumental in my (our) successes and I'm so happy to have been able to share them with you. I look forward to continuing to grow and work with you in varying capacities for many years to come.

To my committee members: Dr. Beauchamp, your clinical perspective has been critical in grounding the technical aspects, undoubtably this research was stronger as a result. Beyond your clinical perspective, your support has meant a lot to me; I always felt like you were truly there to help me succeed and I can't thank you enough for that. Dr. Keir, you and your lab group have been one of the biggest parts of my PhD experience. Thank you for being so generous and letting me takeover your lab space for weeks and months at a time. Participating in activities, meetings, and being part of your group has been a major part of my growth and learning throughout my PhD, thank you. Dr. Noseworthy, you have been a major influence on my path over the past 7-years. Thank you for sharing your passion for imaging, computer science, machine learning, and their intersection. These have since become a major part of my own passions; I am forever

grateful to you for that. Your enthusiasm for science and research of all kinds is unbounded and is infectious; thank you for sharing your passion with the rest of us.

I have had tremendous support from two lab groups. From the Mobilize group, no one has shared more of the journey and all of the trials and tribulations of it with me than Dr. Brenneman Wilson. Elora, thank you so much for being there to bounce ideas around or to listen to me rant/vent. You are one of the hardest working and most determined people that I know and truly are an exceptional researcher, I'm proud to say that I crossed the finish line a mere 2 months after you. To the rest of the Mobilize gang, you've all played unique roles in my journey and I can't thank you all enough. I hope that I get the opportunity to tell you all how I feel in person, when we all get through this trying time. To the Occupational Biomechanics Lab, you have been so welcoming to me and my antics, I could not have hoped for a better "step-group". Dan, you amaze me every time we interact. You are not only one of the most intelligent and creative people I know, but you are so kind, compassionate, and giving. You are an inspiration to me. Many of you, past and present, from both labs have become some of my closest friends, I truly hope that never changes.

Mom & Dad, I can't thank you enough for just being who you are, for passing a little bit of yourselves onto me, and for giving me Kathleen as a little sister. You are so different in many ways, but at the core you have much in common. You are both so generous and do everything for your families. Kathleen and I are lucky to have had, and to continue to have, you as parents, role-models, advisors, and friends to share the journey with. I would still be wandering around trying to find a place to live in Whistler

village if it wasn't for you both. Leslie & Paul, you've supported me in unique ways throughout the decades. Leslie, thank you for your unconditional love and support no matter what. Paul, thank you for engaging and sharing in our shared and unique passion for science and statistics. Kathleen, I couldn't have asked for a better sibling, team-mate, competitor, teacher, and confidant. Somehow, as we've grown (even if often thousands of kilometers apart), we've become closer and closer. I can't wait for the next part of our lives, now including Ben & Sarah, and the journeys that our families will share. Westons, you've always been on team-Gatti. Not only am I forever grateful for the amazing daughter that you raised into the woman that I love, but I am grateful for how you've welcomed me as part of your family over the past decade+ and how you've supported me on my journey.

Sarah, I could never have done any of this without you. Not only are you my wife, but you are my best-friend and my one and only partner on this journey. Many days when I didn't have energy to continue you quietly pushed me along. Your support has truly been selfless. Over the past 5 years, not only have you read every page of this thesis (except maybe this one ☺), but we've shared so many milestones from home renovations to setting the trend of "pop-up weddings", becoming doggie parents, and travelling to countless destinations. I am forever thankful for all that you've done and put up with to get me to this point. You are my everything, and I'm so happy that I have you to co-navigate the path ahead of us.

Table of Contents

List of Figures

List of Tables

List of all Abbreviations

3D	Three-Dimensional
AAM	Active Appearance Model
ACL	Anterior Cruciate Ligament
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
ASD	Average Surface Distance
BMI	Body Mass Index
CMC	Computed Muscle Control
CNN	Convolutional Neural Network
COMAK	Concurrent Optimization of Muscle Activations and Kinematics
CPD	Coherent Point Drift
CPU	Central Processing Unit
DESS	Dual Echo in the Steady State
dGEMRIC	delayed Gadolinium Enhanced Magnetic Resonance Imaging of Cartilage
DOF	Degrees of Freedom
DSC	Dice Similarity Coefficient

FEA	Finite Element Analysis
FOCUSR	Feature Oriented Correspondence Using Spectral Regularization
FSPGR	Fast Spoiled Gradient Recalled
GAQ	Get Active Questionnaire
gen-DSC	Generalized Dice similarity coefficient
GPT-3	Generative Pre-trained Transformer-3
GPU	Graphics Processing Unit
HR	Heart Rate
IJS	Integrated Joint System
ITK	Insight Toolkit
JAM	Joint Articular Mechanics
JRF	Joint Reaction Force
KAM	Knee Adduction Moment
KFM	Knee Flexion Moment
KL	Kellgren and Lawrence
KOPS	Knee Over Pedal Spindle
LASSO	Least Absolute Shrinkage and Selection Operator
MRI	Magnetic Resonance Imaging

MSE	Mean Squared Error
NBV	non background voxels
NIH	National Institutes of Health
OA	Osteoarthritis
OAI	Osteoarthritis Initiative
OARSI	Osteoarthritis Research Society International
pNBV	proportion of non background voxels
PReLU	Parametric Rectified Linear Unit
PT	Patellar Tendon
QT	Quadriceps Tendon
RCT	Randomized Controlled Trial
RMSE	Root Mean Squared Error
RPM	Revolutions Per Minute
SD	Standard Deviation
SPM	Statistical Parametric Mapping
SSM	Statistical Shape Model
$T_{1\rho}$	$T_{1\rho}$ relaxation time
T_2	T_2 Relaxation Time

TKR	Total Knee Replacement
TNF	Tumour Necrosis Factor
VD	Volume Difference
VTK	Visualization Toolkit
weighted-DSC	weighted generalized Dice similarity coefficient

Declaration of Academic Achievement

This thesis is the primary work of Doctor of Philosophy candidate Anthony A. Gatti. This thesis consists of six chapters. Chapter 1 is a review of pertinent background information relevant to osteoarthritis, and the relationships of cartilage, bone, obesity, physical activity, biomechanics, and medical imaging to osteoarthritis.

Chapter 2 contains a prepared manuscript entitled "Equations to Prescribe Bicycle-Fit based on Desired Joint Kinematics and Bicycle Geometry" by Anthony A. Gatti, Peter J. Keir, Michael D. Noseworthy, Marla K. Beauchamp, Monica R. Maly. This manuscript was submitted to the European Journal of Sport Science on September 14th 2020 (TEJS – 2020 – 1319). This study generated two equations that predict a bicycle saddle height that will elicit a desired amount of knee flexion based on the rider's anthropometrics and bicycle geometry. This study was conceived by Anthony Gatti and Monica Maly. Data was collected by Anthony Gatti with assistance from Emily Wiebenga. Data analysis, interpretation, and preparation of the manuscript were led by Anthony Gatti. All co-authors contributed substantially to the final submitted manuscript.

Chapter 3 contains a prepared manuscript entitled "Hip and ankle kinematics are the most important predictors of knee joint loading during bicycling" by Anthony A. Gatti, Peter J. Keir, Michael D. Noseworthy, Marla K. Beauchamp, Monica R. Maly. This manuscript is in-press at the Journal of Science and Medicine in Sport

(https://doi.org/10.1016/j.jsams.2020.07.001). This study explored the relationship between lower body joint kinematics during bicycling and the coinciding peak joint reaction forces at the knee. This study was conceived by Anthony Gatti and Monica Maly. Data was collected by Anthony Gatti with assistance from research assistants Emily Wiebenga, Elora Brenneman Wilson. Data analysis, interpretation, and preparation of the manuscript were led by Anthony Gatti. All co-authors contributed substantially to the final submitted manuscript.

Chapter 4 contains a prepared manuscript entitled "Automatic Cartilage and Bone Segmentation using Multi-Stage Convolutional Neural Networks: Data from the Osteoarthritis Initiative" by Anthony A. Gatti and Monica R. Maly. This manuscript was submitted to the journal Medical Image Analysis on October 8th 2020 (MEDIA-D-20-01172) and transferred to the journal Computerized Medical Imaging and Graphics on October 14th, 2020 (# TBD). This study describes and validates a novel multi-stage convolutional neural network framework for segmenting cartilage and bone from magnetic resonance imaging images. Study conception, data collection, data analysis, interpretation, and preparation of the manuscript were led by Anthony Gatti. Magnetic resonance imaging data collections were assisted by Elora Brenneman Wilson and Emily Wiebenga. All co-authors contributed substantially to the final submitted manuscript.

Chapter 5 contains a prepared manuscript entitled “Investigating Acute Changes in Osteoarthritic Cartilage by Integrating Joint Reaction Forces and Statistical Shape Models of Bone: Data from the Osteoarthritis Initiative” by Anthony A. Gatti, Peter J. Keir, Michael D. Noseworthy, Monica R. Maly. This manuscript is prepared for submission to the Journal of Biomechanics. This study explored the effect of an acute bout of walking and bicycling on the cartilage of women with symptomatic knee osteoarthritis. This study was conceived by Anthony Gatti and Monica Maly. Data was collected by Anthony Gatti with assistance from research assistants Emily Wiebenga, Elora Brenneman Wilson, and Andrew Wong. Data analysis, interpretation, and preparation of the manuscript were led by Anthony Gatti. All co-authors contributed substantially to the final submitted manuscript.

Chapter 6 includes a discussion of thesis Chapters 2 to 5 including integration of knowledge across chapters, important topics relevant to the work, future directions, and conclusions.

1 Background

1.1 Osteoarthritis

Osteoarthritis (OA) is a degenerative joint disease that can be described in terms of symptoms experienced by the individual and structural changes to the joint. For knee OA, symptoms include knee pain, joint stiffness, crepitus, and bony tenderness (Altman et al., 1986). Symptomatic knee OA is estimated to affect >15 million Americans (Deshpande et al., 2016), representing more than 7% of the American population; similarly, knee OA affects >8% of adults living in China (Tang et al., 2016). These statistics demonstrate that knee OA is a global problem (Safiri et al., 2020). Knee OA prevalence also increases with age (Tang et al., 2016), and OA in general is estimated to affect >14% of Canadians over 30 years of age (Birtwhistle et al., 2015) and >50% of adults over 65 years of age (Arden & Nevitt, 2006). Ultimately, painful knee OA leads to a 2.2 times greater risk of death compared to those without pain and osteoarthritis (Y. Wang et al., 2020). Knee OA has a greater prevalence in elderly adults than any other chronic disease (Guccione et al., 1994). OA is a global problem causing pain, and decreased mobility and quality of life.

Not only does OA, and knee OA in particular, affect numerous adults, it also poses significant financial burden. The Arthritis Alliance of Canada estimated that OA cost the Canadian economy > **$27 billion** ($10.2 billion direct costs, $17.3 billion indirect costs)(Bombardier et al., 2011). Another estimate of direct costs performed in 2015 indicated that direct healthcare costs related to OA (pharmacological and non-

pharmacological treatment, total joint replacement, hospital resource usage, treatment complications) were $2.9 billion in 2010 and are expected to rise to $7.6 billion by 2031 (Sharif et al., 2015). Regardless of discrepancies in the exact estimate, direct healthcare costs are in the billions and are expected to increase (Bombardier et al., 2011; Sharif et al., 2015). These increases were linked to increased age among Canadians over that time period, with 2010 including 26.3 million Canadians over 20 years of age, and 2031 expected to include 32.2 million Canadians over 20 years of age (>20% increase). It is also estimated that there will be a concurrent increase in OA prevalence over that same time period, increasing from 16% to 21% in women and from 11% to 16% in men (Sharif et al., 2015). These projections indicate that the economic burden of OA in Canada is expected to steadily increase for the foreseeable future (Bombardier et al., 2011; Sharif et al., 2015).

1.1.1 Pathophysiology

Damaged articular cartilage is hypothesized to be one of the earliest changes resulting from OA (Buckwalter et al., 2005). Knee OA is a degenerative joint disease that affects the whole joint. Knee OA includes deterioration of articular cartilage, changes to the bone that lays under cartilage (subchondral bone), and formation of bone cysts, bone marrow lesions, and osteophytes (bony-enlargements)(Creamer & Hochberg, 1997; Felson, 2006). Degenerative changes also extend to muscles, ligaments, and the synovium (Felson et al., 2000).

A recent theory of OA coined the "Integrated Joint System" (IJS) posits that healthy joints are at an equilibrium state between joint tissues (e.g., cartilage thickness and bone mineral density), as well as between joint mechanics and these tissues (e.g., peak joint reaction forces (JRF) during gait and cartilage thickness)(Edd et al., 2018). That is, the IJS theory states that in a healthy joint, tissues and mechanics are adapted to one another. For example, in a healthy joint, thicker cartilage is typically located in areas of higher subchondral bone mineral density (Babel et al., 2020). Also in healthy adults, higher joint loads are associated with both greater tibiofemoral cartilage thickness and greater subchondral bone mineral density (Edd et al., 2018; Van Rossom et al., 2017). The IJS theory states that initial disruption in homeostasis is considered initiation of OA and that once disrupted, with time, the healthy relationship turns into a degenerative one. For example, if subchondral bone mineral density increases, as it does with increased joint loads, but cartilage thickness does not concurrently increase – the healthy homeostasis is broken (Edd et al., 2018). Without this concurrent adaptation, joint loads will lead to even greater cartilage surface pressures (Li et al., 2001) and thus higher probabilities of cartilage damage. Cartilage damage further reduces cartilage thickness (or other characteristics relevant to load bearing capacity) and its ability to transmit load. A cycle ensues that progressively diverges from the original healthy homeostasis between joint tissues and mechanics. Lending support to the IJS theory, high bone mineral density adjacent to cartilage damage reflects Radin and Rose's theory that subchondral bone sclerosis is related to initiation and progression of cartilage degeneration (Radin & Rose, 1986).

1.1.2 Articular Cartilage

In a healthy joint, articular cartilage reduces friction and transmits forces between articulating bony surfaces in synovial joints. According to the triphasic theory of cartilage, (Lu & Mow, 2008) its load attenuation properties come from three primary phases: 1) a solid structure made up of collagen, proteoglycans, and chondrocytes, 2) a fluid phase, and 3) an ion phase. Collagen fibers within cartilage provide the scaffolding that prevent sheering and minimize swelling of cartilage. Fluid (water) within cartilage aids in attenuating deformation because under compression fluid will try to flow from areas of high pressure (compression) to areas of low pressure; frictional drag of fluid through the solid structure during this flow resists deformation and attenuates loads (Lu & Mow, 2008). Finally, the dissolved salts (typically Na^+ and Ca^{++} ions) are attracted to negatively charged chemical groups (sulfate: SO_3^-, carboxyl: COO^-) on glycosaminoglycans attached to the solid structure. This attraction creates an osmotic pressure gradient (Donnan osmotic pressure) that increases tissue hydration and further resists flow of fluid within cartilage (Lu & Mow, 2008). This osmotic pressure is thought to contribute between 30-50% of the equilibrium stiffness of cartilage (Lu & Mow, 2008). In addition to load bearing ability, cartilage provides a low friction surface between articulating bones with Lubricin playing a primary role (Y. Lee et al., 2018; Nordin & Frankel, 2012). Lubricin along with hyaluronic acid reduces friction between the adjacent cartilage surfaces by diminishing shear energy in joint gliding (Y. Lee et al., 2018). Lubricin also plays a role in preventing protein and synovial cell deposits on cartilage surfaces (Y. Lee et al., 2018; G Musumeci, 2013). Damage or disruption to any

component of this cartilage structure will hinder its ability to reduce friction and transmit forces to the underlying bone.

Cartilage undergoes remodeling based on mechanical and chemical changes in the joint induced by loading (Mazor et al., 2019; Ramage et al., 2009). For example, cyclic compression of cartilage under forces and frequencies considered physiologic produce anabolic responses that generate new cartilage extracellular structures, including but not limited to increased glycosaminoglycan synthesis, increased proteoglycan synthesis, and increased proliferation of chondrocytes (Ramage et al., 2009). Chondrocytes are central to this anabolic response (Ramage et al., 2009). In individuals with OA, 3-months of daily strengthening exercises increased hyaluronic acid and decreased markers of joint degeneration within joint fluid (Mazor et al., 2019). These changes are postulated to indicate increased anabolism and thus generation of new cartilage solid structure (Mazor et al., 2019). However, overloading is associated with increased proteases within the joint (e.g., matrix metalloproteinase: extracellular matrix degeneration; aggrecanase: aggrecan degeneration). Proteases are commonly responsible for joint tissue catabolism. By definition, OA occurs when catabolic pathways exceed anabolic ones leading to cartilage degeneration (Ramage et al., 2009). This body of research indicates that there is likely a yet to be identified ideal range of loading that promotes cartilage health without causing damage.

1.1.3 Bone in Relation to Osteoarthritis and Cartilage

While cartilage is a common focus in knee OA research, bone has increasingly been thought to play a central role in OA pathogenesis (Mansell et al., 2007). In particular, bone-remodeling is thought to lead to changes in bone material properties (Burr & Gallant, 2012; Goldring & Goldring, 2010) that may affect the biomechanical environment of the joint. Changes to bone shape are also a commonly noted feature of OA (Barr et al., 2015).

Animal models indicate that increased bone remodeling reduces thickness in subchondral cortical bone (subchondral plate) before other joint tissue changes (Burr & Gallant, 2012; Goldring & Goldring, 2010). Along with bone remodeling comes increased vascularization of the subchondral bone with potential for invasion of this vasculature into the deep layers of articular cartilage (Burr & Gallant, 2012; Glyn-Jones et al., 2015; Goldring & Goldring, 2010; Lories & Luyten, 2011). It is postulated that this vascular innervation may leak catabolic agents into cartilage, providing one mechanism of cartilage degeneration (Burr & Gallant, 2012; Lories & Luyten, 2011). As time progresses, bone resorption decreases but deposition rates are maintained, leading to increased thickness of the subchondral plate and calcified zone of cartilage (Burr & Gallant, 2012; Goldring & Goldring, 2010; Lories & Luyten, 2011). The increased calcified zone of cartilage reduces cartilage thickness on its own (Burr & Gallant, 2012; Goldring & Goldring, 2010; Lories & Luyten, 2011).

Greater subchondral plate thickness and bone mineral density is commonly believed to increase subchondral bone stiffness; this proposed phenomenon was the basis of Radin and Rose's theory on the role of subchondral bone in cartilage degeneration (Radin & Rose, 1986). However, recent belief is that the high rate of bone remodeling and deposition that occurs in OA in fact leaves behind bone with suboptimal mineralization and thus reduced stiffness (Burr & Gallant, 2012; Lories & Luyten, 2011). It is important to note that plain radiographs and dual-energy X-rays extract "apparent bone density" (bone mass divided by total tissue volume) whereas biopsy samples are needed to measure "material density" (bone mass divided by bone volume). Therefore, apparent bone density increases if the bone volume increases relative to the total tissue volume. When bone volume but not mineralization increases, there is an increase in the apparent bone density, but decrease in both material density and bone stiffness (Burr & Gallant, 2012). Nonetheless, current imaging modalities consistently observe subchondral sclerosis and increased apparent bone mineral density from underlying regions of cartilage degeneration and these changes are thought to be relevant in the disease course. The IJS theory relies entirely on the apparent bone mineral density reported from medical images and does not relate the theory to stiffness of the bone (or material density) (Babel et al., 2020; Edd et al., 2018). As Radin and Rose observed, there is in fact increased amounts of bone under regions of cartilage degeneration (Pugh et al., 1974; Radin & Rose, 1986). More work is necessary to examine the role of mineralization and stiffness of this bone in the early stages of OA to understand the validity of the IJS theory.

Bone shape has been of particular interest in relation to OA diagnosis and progression. It is postulated that increased rate of subchondral bone remodeling during early OA may lead to changes in joint shape and congruence (Burr & Gallant, 2012). These changes in bone shape and congruence are postulated to alter load transmission, potentially causing cartilage damage (Burr & Gallant, 2012; Goldring & Goldring, 2010; Lories & Luyten, 2011). Concurrent with changes to subchondral bone is osteophyte formation, one of the most noticeable changes in the OA knee[*](Goldring & Goldring, 2010). Osteophytes form at the joint margins, where cartilage and bone meet. Animal models indicate that osteophytes arise from proliferation of periosteal cells that differentiate into chondrocytes, hypertrophy and subsequently ossify to create skeletal outgrowths at the joint margin. Additional changes to the joint that occur in injured or diseased knees include increases in the absolute size of the tibial plateau, (Antony et al., 2011; Barr et al., 2015) and other changes to bone shape such as narrowing of the inter-condylar notch in those with ruptured anterior cruciate ligaments (Pedoia et al., 2015). It is postulated that many of these large changes to bone shape occur during periods of increased bone remodeling common in early OA (Burr & Gallant, 2012; Mansell et al., 2007).

[*] Osteophyte formation and changes to bone structure and organization precede detection in articular cartilage thickness changes and reductions in joint space narrowing (Goldring & Goldring, 2010).

Osteophytes, in particular, are one of three† main features that are used to diagnose and stage radiographic OA using the most common method, the Kellgren and Lawrence (KL) grading system (Kellgren & Lawrence, 1957).

1.1.4 Obesity

In addition to age, obesity is the major risk factor for knee OA (Blagojevic et al., 2010; Davis et al., 1990; Felson et al., 1988; Stürmer et al., 2000). The relationship between knee OA and obesity is hypothesized to exist due to the increased mechanical loads placed on the knee as a result of increasing body mass, (Stürmer et al., 2000) and systemic inflammation induced by fat mass (Kluzek et al., 2015). Increased body mass necessarily increases knee joint forces during weight-bearing activity. Also, white adipose tissue produces adipokines (e.g., leptin, resistin and chemerin) as well as inflammatory cytokines (e.g., tumour necrosis factor TNF, interleukin-1 IL-1, and interleukin-6 IL-6) which are both associated with and hypothesized to play a role in OA pathogenesis through altering the integrity of joint tissues (Kluzek et al., 2015; Giuseppe Musumeci et al., 2015). These inflammatory cytokines are linked to catabolic activity of cartilage, putting the anabolic/catabolic balance in favour of degeneration (catabolism)

---

† The Kellgren and Lawrence grading system utilizes three main features: joint space narrowing, osteophyte formation, and subchondral sclerosis to diagnose and stage knee OA. A further description of the KL system is provided in section 1.4.3.

(Ramage et al., 2009). Weight loss effectively reduces an individual's risk of developing knee OA (Felson et al., 1992).

1.2 Physical Activity and Osteoarthritis

Physical activity decreases an individuals' risk of cardiovascular disease, cancer, osteoporosis, hypertension, diabetes, depression, and it decreases all-cause mortality (I.-M. Lee & Skerrett, 2001; Warburton, 2006). In relation to knee OA, a Cochrane review of 54 randomized controlled trials (RCTs) found strong evidence that activity reduces pain and improves quality of life, as well as moderate evidence suggesting physical activity improves physical function in knee OA (Fransen et al., 2015). Another review concluded that physical activity may be beneficial to joint health (Urquhart et al., 2011). A study (n = 2,073) on the effect of physical activity on future incidence of symptomatic knee OA, and joint space narrowing progression, showed that individuals in the highest quartile of activity were less likely to worsen in terms of symptoms or joint structure (Felson et al., 2013). Even high-impact activity, such as running, may be beneficial, or at least cause no harm: participating in running decreased, (Williams, 2013) or had no effect (Chakravarty et al., 2008; Lo et al., 2018) on risk of developing OA. After accounting for obesity, running still decreases an individual's risk of OA, (Williams, 2013) highlighting that even high impact activities like running may maintain joint health, in the healthy joint. Furthermore, the Intensive Diet and Exercise for Arthritis (IDEA) trial showed that the combination of diet and exercise benefitted knee OA outcomes including systemic inflammation, pain, and function more than diet or exercise alone (Messier et al., 2013).

This body of work suggests that physical activity (1) reduces the risk of OA onset in healthy joints, (2) can ameliorate pain and improve quality of live in those that already have OA, and (3) that its benefits on knee OA are augmented by diet-based weight-loss.

In terms of joint structures, OA compromises the integrity of cartilage, therefore reducing the ability of this tissue to bear load. Knee OA is primarily characterized by pain at the joint that can be aggravated by simply walking (Farrokhi et al., 2017). In fact, walking >10,000 steps/day is associated with increased development of bone marrow lesions, meniscal pathology, and worsening cartilage defect scores in 405 community-dwelling adults (51-81 years) (Dore et al., 2013). It is important to note that this investigation found an interaction between step count and baseline cartilage volume, where those with higher baseline cartilage volume actually benefitted from higher step counts (Dore et al., 2013). These findings support the notion that activity may be beneficial and prevent OA in healthy adults, but may be detrimental to joint structure and pain in people with knee OA. This finding further supports the IJS theory that in OA a healthy homeostasis has been broken, initiating and propagating the disease process.

To allow individuals with knee OA to obtain the health benefits associated with physical activity, it is imperative that we identify activities that do not cause further damage, and ideally promote health in the knee affected by OA. Strides have been made at identifying (Longpré et al., 2015) and testing (Brenneman et al., 2015; Multanen et al., 2014)

strengthening exercise for knee OA. These investigations have utilized biomechanics to provide objective data about the forces being exerted on the body, and medical imaging to identify what is occurring inside of the joint. The next step is to produce aerobic activity recommendations for individuals with knee OA that do not aggravate symptoms or accelerate degeneration.

1.3 Biomechanical Indicators of Knee Osteoarthritis

Biomechanics assesses how mechanical principles relate to movement and biological function. In knee OA, it is postulated that joint loads beyond tissue tolerances lead to damage, initiating OA pathogenesis and deterioration of cartilage health. Therefore, biomechanics of knee OA primarily focuses on identification of measures of joint mechanics that represent cartilage loading. Measures of medial knee forces are of particular interest because medial knee OA is most prevalent (Metcalfe et al., 2012).

1.3.1 Joint Moments

A key outcome of interest for knee OA is the frontal plane joint measurement, the knee adduction moment (KAM). The KAM is an external moment of force (torque) that provides a surrogate measure of medial relative to lateral or relative to total knee compressive forces (Kutzner et al., 2013; Moyer et al., 2014). The KAM has also shown success at identifying those with and without knee OA, (Maly et al., 2013) and in predicting OA progression (Bennell et al., 2011; Brisson et al., 2017; Chang et al., 2015;

Chehab et al., 2014). Gait re-training to reduce the KAM also reduces self-reported knee OA severity (Cheung et al., 2018). A limitation of the KAM is that it does not actually measure the forces that are being translated between structures at the knee. For example, analysis of 6 male and 3 female participants with instrumented tibial prostheses showed R^2 of 0.56 between the KAM and medial contact force at the tibia, when measured over the whole stance phase of gait (Trepczynski et al., 2014). The high R^2 of 0.56 may be attributed to re-alignment of the joint after surgery or a statistical effect of analyzing the whole waveform instead of the peaks (Kutzner et al., 2013). Another study on one male participant with an instrumented tibial prosthesis showed that the knee adduction moment peak explained between 29-69% of the variance (R^2) in the peak medial contact force with coinciding root mean squared errors ranging from 41-129% of body weight depending on whether the first peak, second peak, or both peaks were analyzed‡ (Walter et al., 2010). Furthermore, the KAM on its own likely misses important muscular contributions across the knee. Accounting for muscular contributions to joint loading may be ameliorated by inclusion of the knee flexion moment (KFM) in quantifying knee joint loading. The KFM is indicative of the net muscle contraction across the knee

‡ The root mean squared errors and R^2 for these models were inversely related, with the higher R^2 models (peak 2, both peaks) producing poorer (higher) errors. These incongruent findings indicate that differences in variance of the dependent variable between models may be a larger driver of R^2 than prediction errors. In these cases, the root mean squared error is likely a better measure of fit and is relatively large ranging from 41 to 129% of body weight (Walter et al., 2010).

(Creaby, 2015). For example, during the early stance phase of gait the quadriceps counterbalance the KFM, therefore the KFM provides some indication of the quadriceps activation during this portion of stance. An editorial by Creaby argues that many interventions which aim to decrease the KAM may in fact increase the KFM, leading to no net change or even a possible increase in loading across the tibiofemoral joint, particularly in the medial compartment (Creaby, 2015). This rationale is supported by work showing that the combination of the KAM and KFM are better at predicting medial contact forces at the tibiofemoral joint (Manal et al., 2015; Trepczynski et al., 2014; Walter et al., 2010).

Including the KFM in modelling knee biomechanics helps account for muscular forces across the joint. The importance of these muscular contributions is highlighted by the fact that simple ground reaction forces during walking are less than half of the loads measured inside of the knee (Damm et al., 2017); this insinuates that muscular contributions may account for more than half of the forces across the knee. However, as stated in the editorial by Creaby, it is possible that opposite changes between different moments (e.g., KAM and KFM) may counteract one another. We therefore require methods of estimating forces acting inside of the joint.

1.3.2 Joint Forces

Use of advanced modeling software such as the open-source platform OpenSim created by the National Center for Simulation in Rehabilitation Research (Stanford University) allows for modeling of muscular forces, providing physiologic estimates of joint loading (Delp et al., 1990, 2007; Seth et al., 2018). OpenSim and other multi-body biomechanical modelling software packages sum conventional mechanical measures of joint reaction forces with muscular forces acting across the knee to estimate the total physiologic loading at the joint (Gerus et al., 2013; Hamner et al., 2010; Knarr & Higginson, 2015; Lerner et al., 2014; Saxby et al., 2016; Thelen & Anderson, 2006). After calculating the total knee JRF vector, it can be decomposed into components. Of particular interest in knee OA is compression between the tibia and femur. This compressive force is typically modeled as the component of the total knee JRF acting along the long-axis of the tibia (DeMers et al., 2014; Lerner et al., 2014; Steele et al., 2012).

To represent the total knee JRF and decompose it into its anatomic components, the muscular contributions must be estimated. One of the most common approaches of doing this is to perform traditional inverse kinematic and dynamic analyses of the acquired motion capture data, and then solve for the muscle activations and forces that would be necessary to observe these motions. These muscle activations can be calculated using static optimization or forward dynamic methods (Thelen et al., 2003; Thelen & Anderson, 2006).

Static optimization identifies a set of muscle activations that would produce the joint moments obtained from inverse dynamics while minimizing an objective function. When using this approach, most often the mathematical objective function is optimized to minimize the total muscular contribution required to produce the observed kinematics. It is important to note that researchers may choose to optimize other parameters. This approach typically analyzes each frame of recorded data independently. The default objective function for OpenSim minimizes the sum of muscle activations squared:

$$\text{Equation 1-1} \quad \min \sum_{i=1}^{N} a_i^2$$

Where a_i is the activation of muscle i and N is the number of muscles in the model. A function converting muscle activations into forces inherently models factors like the force length, force velocity, and muscle size (Delp et al., 2007). Squaring muscle activations helps distribute forces over multiple muscles by exponentially penalizing with every increase in activation of any one muscle; this results in greater distribution of activations across muscles. For example, if no exponent was used, then the solution that minimizes muscle activations would select the muscle that produces the largest moment (per unit of activation). By including the power of 2, the more activated that a muscle is the more costly it is to use (Equation 1-1). Therefore, as the activation of the most efficient muscle increases, it will become optimal to activate smaller muscles. A major limitation of static optimization is that this approach does not account for time dependence, such as the electromechanical delay between electrical muscle activations and force production, or the time-dependent characteristics of viscoelastic tissues like tendons and the muscles themselves (Zajac, 1989). Forward dynamic methods can include these time-dependent

features, often inputting a set of muscle activations and then relying on computer simulations to determine what the state of the system will be fractions of a second later (Delp et al., 2007; Thelen et al., 2003).

In walking and other non-ballistic tasks, static optimization produces comparable results to dynamic simulation methods (e.g., computed muscle control (CMC) in OpenSim)(Anderson & Pandy, 2001; Lin et al., 2012). Both methods report similar activation patterns as EMG, with neither outperforming the other (Roelker et al., 2020). Muscle weightings further improve prediction of physiologic muscle activations for static optimization by penalizing use of the gastrocnemius and hamstrings, making their use more costly in terms of the objective function, thus making them less likely to be used (Steele et al., 2012). Prediction of the compressive component of the knee JRF using static optimization and muscle weightings produce small reported errors (root mean squared error (RMSE) = 0.28 body weights) when compared to measured *in vivo* knee contact forces (DeMers et al., 2014; Gerus et al., 2013; Meyer et al., 2013; Steele et al., 2012; Walter et al., 2010). It is important to note that these JRF predictions, whether based on static or forward dynamics, can only be validated on individuals with instrumented joint replacements and thus generalization to natural knees is unknown. Nonetheless, given the comparable results by static and forward dynamics methods, the much lower computational load associated with static methods, and the accuracy of static optimization plus muscle weighting at predicting *in vivo* forces make static optimization

ideal, and recommended, for modelling activities such as walking and bicycling (Lin et al., 2012).

1.4 Medical Imaging Measures of Knee Osteoarthritis

Conventional magnetic resonance imaging (MRI), allows for quantification of joint characteristics including cartilage volume, cartilage thickness, bone shape, as well as joint tissue composition. Cartilage volume and thickness are traditional measures that characterize cartilage morphology; cartilage damage and loss are disease hallmarks. Knees affected by OA also have characteristic bone shapes, such as osteophytes at the joint margins, that are categorized using Kellgren and Lawrence grading and statistical shape models (SSMs) (Kellgren & Lawrence, 1957; Neogi et al., 2013). Of recent interest are MRI sequences such as transverse-relaxation time (T_2) that quantifies cartilage composition.

1.4.1 Cartilage Morphology

Measurement of cartilage morphology (thickness and volume) is important in the study of knee OA as loss of cartilage is a hallmark sign of the disease. Morphometric outcomes are recommended by the Osteoarthritis Research Society International (OARSI) to assess structural disease progression in clinical OA trials (Hunter et al., 2015). Finally, cartilage thickness is inversely related to cartilage surface pressures, von Misses stresses, and

hydrostatic pressure modeled using finite element analysis (FEA) of an axial load (Li et al., 2001). In theory, these higher loads are more likely to cause cartilage damage.

1.4.2 Cartilage Composition

Cartilage composition may be an early biomarker of knee OA (Razmjoo et al., 2020). Compositional measures are also used to test the acute response of cartilage to loading (Gatti et al., 2017). The MRI outcomes most commonly used to quantify cartilage composition are T_2 relaxation time, $T_{1\rho}$ relaxation time and delayed Gadolinium Enhanced Magnetic Resonance Imaging of Cartilage (dGEMRIC).

T_2 relaxation time has a positive relationship with free water, and is sensitive to collagen fiber alignment (Choi & Gold, 2011; Palmer et al., 2013). T_2 has been of particular interest as its relation to free water content is postulated as being indicative of cartilage degeneration. That is, as cartilage degenerates water content and mobility increases (Felix Eckstein et al., 1999; Xia, 2000), thus increasing free water content and T_2 relaxation time (Choi & Gold, 2011). Furthermore, T_2 relaxation is related to collagen alignment. Due to the depth dependence of collagen alignment (deep radial zone collagen fibres are aligned perpendicular to the bone, intermediate depth fibres are randomly oriented, and superficial fibres are oriented parallel to the articular surface), there is a depth dependence of T_2 relaxation times (Gatti et al., 2016; Timothy J. Mosher et al., 2005; Xia, 2000). Because T_2 reflects composition, it is ideal for *in vivo*

experiments designed to determine the response of cartilage to mechanical stimuli and to use baseline compositional measurements to determine the effect of intrinsic cartilage characteristics on said response. For example, under the triphasic theory of cartilage mechanics, fluid re-distribution occurs in response to mechanical loads. Therefore, changes in T_2 throughout cartilage may provide insight into not only where loads are being applied, but also how fluid is redistributed in response to these loads. Furthermore, T_2 is acquired at every voxel of an image and therefore we can characterize depth dependent responses of cartilage to mechanical loading (Gatti et al., 2016; Timothy J. Mosher et al., 2005). With carefully designed studies, these tools have the ability to confirm theoretical and computational models of cartilage mechanics, furthering our understanding of the role of mechanics on OA pathophysiology.

$T_{1\rho}$ relaxation time is another measure of cartilage composition. $T_{1\rho}$ is positively correlated with proteoglycan concentration (Akella et al., 2001; Choi & Gold, 2011; Wáng et al., 2015). $T_{1\rho}$ is thought to have less dependence on cartilage collagen content, and therefore may be a more sensitive measure of proteoglycan content than T_2 (Wáng et al., 2015). $T_{1\rho}$ relaxation time has been used to determine the *in vivo* response of cartilage to mechanical stimuli and how baseline composition affects deformation (Collins et al., 2018; Esculier et al., 2019; Souza et al., 2014). However, unlike T_2 which is directly proportional to free water content $T_{1\rho}$ is most strongly related to concentrations of proteoglycans. The absolute number of proteoglycans should not change from a single

bout of activity. Therefore, changes in $T_{1\rho}$ are most likely indirect measures of changes in water content caused by loading. Nonetheless, a combination of T_2 and $T_{1\rho}$ may provide a broader perspective on the acute changes of cartilage to loading.

Also of interest for cartilage imaging has been dGEMRIC (Choi & Gold, 2011). For dGEMRIC, gadolinium is injected (intravenously or intra-articular) and the participant is asked to walk around for a period of time to enable the gadolinium to diffuse throughout their joint and articular cartilage. Due to its negative charge, gadolinium will be repelled from negative chemical groups on glycosaminoglycans ultimately settling in areas of cartilage with low glycosaminoglycan concentrations. Because the presence of gadolinium affects T_1 relaxation a T_1 relaxation map is then acquired and can be used to infer gadolinium concentration, which is inversely related to glycosaminoglycan concentration. While dGEMRIC is used for assessing cross-sectional health of cartilage, (Van Ginckel et al., 2010) it is not appropriate for studies of the acute response to loading. Gadolinium has a half-life that will necessarily decrease the contrast and thus estimate of glycosaminoglycan concentration with time. Therefore, even in an unchanged individual there will be an apparent change in their acquired T_1 map and thus the perceived glycosaminoglycan concentration.

T_2, $T_{1\rho}$, and dGEMRIC all show promise as markers of cartilage composition. However, clinical usage of all these outcomes is hindered by difficulty in standardizing

measurements and dGEMRIC is limited by required injection of a contrast agent (Balamoody et al., 2013; Choi & Gold, 2011; Raya et al., 2009; Wáng et al., 2015).

1.4.3 Bone Shape

The KL system, the gold standard in classifying the severity of knee OA, identifies the presence of key bone features to stage OA severity. Stage 0 has no joint space narrowing (JSN; approximating femur and tibia) or other changes, Stage 1 has doubtful JSN and possible presence of small osteophytes (lipping), Stage 2 has definite osteophytes and possible JSN, Stage 3 has moderate osteophytes, definite JSN, some sclerosis, and possible deformity of the femur and tibia, Stage 4 includes large osteophytes, much JSN, severe sclerosis, and definite deformity of the tibia and femur[§](Braun & Gold, 2012; Kohn et al., 2016). While this classification system reflects elements of bone shape, the advent of 3D-meshes of knee bones calculate more precise measures of bone shape.

[§] Other systems of classifying radiographic knee OA have been proposed (Kohn et al., 2016). The primary reason for these alternate methods is to enable separate quantification of JSN and osteophyte formation in each compartment. Separate scores for JSN and osteophytes are important because under the KL system an individual with JSN but no osteophytes would be considered to not have OA as definite osteophytes is a requirement of KL grade 2, the threshold at which OA is diagnosed.

The 3D-meshes for bone shape are created from sets of MRI or CT images capturing three dimensions of a joint and provide detailed information about bone and joint morphology. Statistical shape models (SSM) built from these 3D-meshes provide a method to characterize features of bone shape, such as features that are characteristic of knee OA (Barr et al., 2015, 2016; Hunter et al., 2016; Neogi et al., 2013; Pedoia et al., 2015). The major features apparent on these OA models are broadening of the bone cartilage interface surface areas and the formation of osteophytes at the cartilage margins (Barr et al., 2016; Neogi et al., 2013). Knees that would be classified as "healthy" on the KL system (KL grade 0 or 1) but that showed features of OA based on these SSM models were 3-12.5 times more likely of developing KL grade 2 or worse knee OA 12 months later (Neogi et al., 2013). This predictive ability is likely owed, at least in part, to the fact that these 3D-meshes and SSM models capture more information than their 2D radiograph counterparts. Furthermore, KL grades from radiographs are shown to be highly sensitive to positioning error and inter-rater biases (Guermazi et al., 2011; Hayashi et al., 2018). The new information garnered from these models enables earlier identification and therefore greater treatment options to alter or divert the disease process.

1.4.4 Image Segmentation

While MRI is clearly useful for capturing cartilage morphology, cartilage composition and bone shape, we must first segment the MRIs to localize the anatomical tissues of interest and extract desired outcomes. For knee OA, the primary tissues of interest are cartilage and bone; other tissues that are relevant to knee OA disease include ligaments,

muscles, menisci, and bone abnormalities including subchondral cysts and bone marrow lesions. Segmentation of knee MRIs, particularly cartilage, requires fast, accurate, and reliable methods for basic science, clinical trials, and even clinical usage. The gold standard of cartilage segmentation from MRI is manual. Manual segmentation takes hours to perform, requires highly specialized knowledge, and shows within and between rater segmentation errors (Pedoia et al., 2016). Highlighting the lack of feasible methods of acquiring quantitative cartilage outcomes from MRIs is the small fraction of the Osteoarthritis Initiative (OAI) database for which quantitative cartilage outcomes are available. This dataset includes >4,000 participants with bi-lateral knee MRIs collected at 7 time-points over 8 years. This data source has been immensely valuable; however, the vast majority of this data has remained un-analyzed and thus its full potential is under-realized.

A number of automated methods of cartilage segmentation have been proposed. Historically, these methods used a variety of registration (SSMs, active appearance models, atlas-based algorithms) and traditional machine learning methods (support vector machines, k-nearest neighbours)(Ahn et al., 2016; Dam et al., 2015; Dodin et al., 2010; Fripp et al., 2010; Prasoon et al., 2013; Shan et al., 2014; Tamez-Pena et al., 2012; Q. Wang et al., 2014). However, these methods fell short of human level accuracies, (Shim et al., 2009) and many still took relatively long to analyze knees, with times ranging from tens of minutes to days. Currently, all of the most accurate methods of cartilage and bone

segmentation incorporate convolutional neural networks, a type of deep learning (Ambellan et al., 2019; Desai et al., 2020; Gaj et al., 2020; Panfilov et al., 2019).

1.4.4.1 Artificial Neural Networks and Deep Learning

Artificial neural networks (ANN) are at the basis of deep learning. ANNs were initially designed based on preliminary understanding of biological neural networks. That is, information or a signal is passed through a series of nodes (neurons) (Figure 1-1). Figure 1-1 describes the network parts and a potential example of predicting sex (male/female) from three inputs of height, weight, and eye colour. With the exception of the input nodes, the value at each node is calculated similar to a linear regression model (Goodfellow et al., 2016). For example, node B1 in Figure 1-1 would be calculated using Equation 1-2

$$\text{Equation 1-2} \quad \text{B1} = \beta_1^0 + \text{A1}(\beta_1^1) + \text{A2}(\beta_1^2) + \text{A3}(\beta_1^3)$$

where β_j^i is the beta coefficient for input node *i* connecting to hidden layer node *j* and *i* = 0 is equivalent to the intercept in a regression model. In the machine learning literature, the intercept is called the bias, and the beta coefficients for the individual nodes the weights;(Goodfellow et al., 2016) we use beta coefficients here due to familiarity in the health sciences domain.

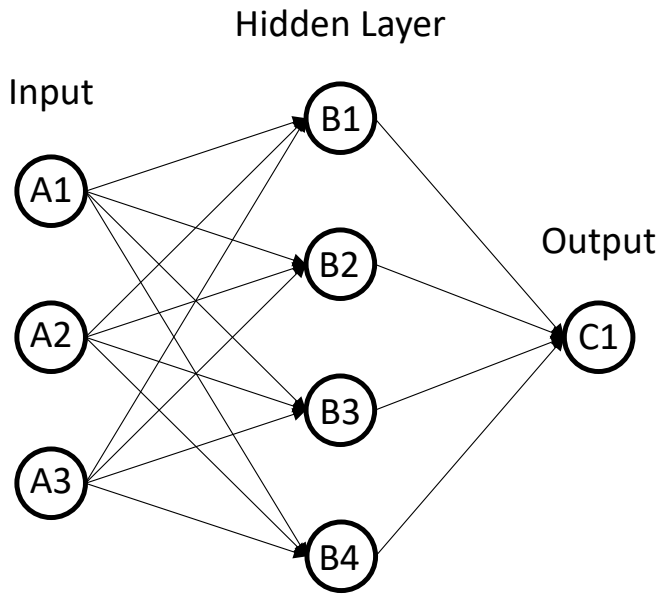


Figure 1-1. Example of an artificial neural network (ANN). The input layer includes three nodes (A1, A2, A3). Each input node is connected to every hidden layer node (B1, B2, B3, B4). Finally, the hidden layer nodes all connected to the output node (C1). In this example, A1, A2, and A3 could be any three input parameters, e.g., height, weight and eye colour. The output node can be anything we could model with a single output, e.g., a person's sex (male or female).

To enable the system to be non-linear, after calculating B1 (or the value for another hidden node), a non-linear activation function is applied to the output of Equation 1-2 (Goodfellow et al., 2016). The canonical activation function is the sigmoid or logistic function which turns all values to be in the range of 0 and 1 and is easily differentiated, which is important for learning the optimal parameters (Goodfellow et al., 2016). The output node (C1 in Figure 1-1) also uses Equation 1-2 but may differ in the activation function choice depending on the purpose of the ANN. For the example of predicting sex

in Figure 1-1 the logistic function would be used to classify the prediction to be binary; 0 (male) or 1 (female).

The optimal beta coefficients for each of the nodes is learned in a supervised fashion[**]. To learn the optimal coefficients, an error or loss function is defined, mean squared error (MSE) is commonly used (Goodfellow et al., 2016; Janocha & Czarnecki, 2017). The learning begins by starting with randomly assigned beta coefficients and inputting an example from the training dataset into the network (e.g., height, weight, and eye colour from Figure 1-1). The network outputs a prediction and the MSE of the prediction is calculated based on the known label. Based on the MSE, the beta coefficients are updated to improve the prediction. Specifically, 1) the partial derivative of the error function with respect to each individual beta coefficient is calculated[††]. 2) Each beta coefficient is then updated by taking a step (increase or decrease the beta coefficient) in

---

[**] Supervised learning includes example data for which the label is available and the network is trained to predict the right output label. Unsupervised learning infers new information from data or learns the labels, most commonly clustering. Reinforcement learning includes an agent (the ANN) and an environment (e.g., game). The agent learns to maximize some reward through trial and error and is often used for games; the popular AlphaGo or AlphaZero by Google's DeepMind are examples of reinforcement learning.

[††] The partial derivatives are the slope of the error function with respect to that particular beta coefficient. This method of calculating partial derivatives of the error function is known as back propagation (Goodfellow et al., 2016).

the direction which decreases MSE (has a negative slope)[‡‡]. This process of inputting a new example, calculating the MSE and performing 1) and 2) is conducted iteratively until the MSE converges.

Deep learning is an extension of ANNs that uses many layers of neurons in varying structures to predict an output. For example, the latest language network entitled generative pre-trained transformer-3 (GPT-3) contains 175 billion learned parameters (Brown et al., 2020). Convolutional neural networks for computer vision and image segmentation are purposely built for efficient analysis and perception of images. Instead of learning individual betas (weights and biases) for each input pixel, convolutional neural networks instead learn convolutional filters (Figure 1-2)(Goodfellow et al., 2016). Convolutional filters reduce memory usage, increase computational speed, and improve generalization (Goodfellow et al., 2016). A kernel is the primary element of a convolutional filter. A kernel is a data matrix, typically sized 3x3 or 5x5. This kernel slides over the entire image, multiplying the kernel elementwise by the pixels underneath it, summing all of the products and placing the total at the center of the kernel in the output image (Figure 1-2)(Goodfellow et al., 2016). Historically, hand-crafted convolutional filters such as the Sobel filter (Figure 1-2) which identifies edges have been

[‡‡] One of the most common methods of updating these coefficients is called stochastic gradient descent (Goodfellow et al., 2016).

used. Convolutional neural networks learn a sequence of filters necessary to achieve the task at hand (Goodfellow et al., 2016).

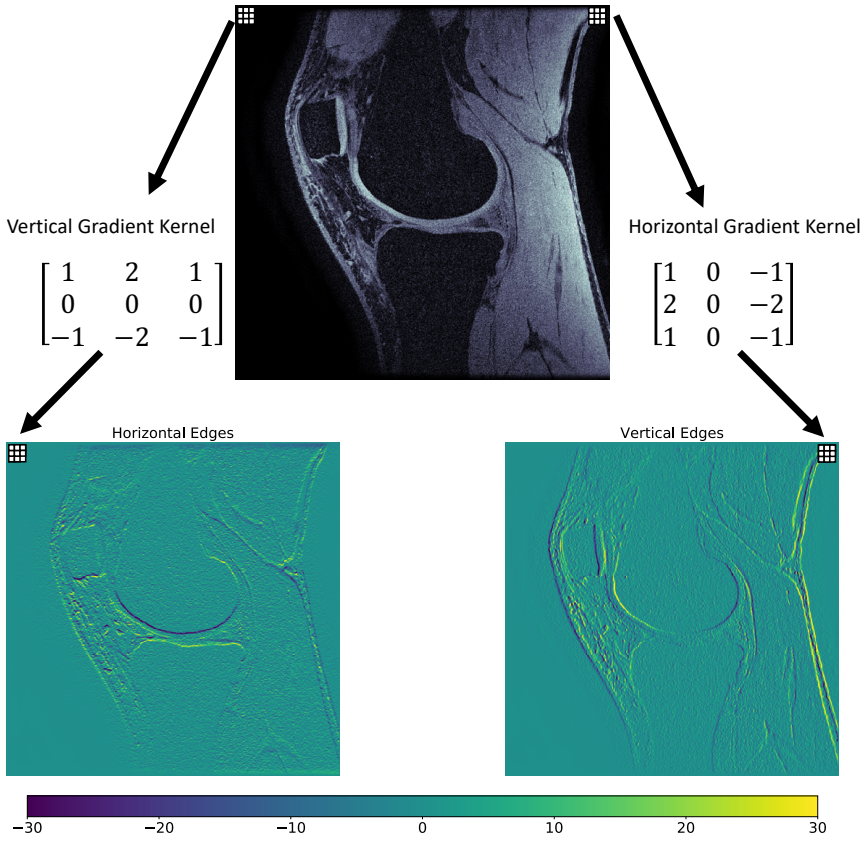


Figure 1-2. Sobel filter, an example of a convolutional filter. The Sobel filter calculates the gradient of the image in the horizontal or vertical directions. These gradients identify lines that run perpendicular to them because we expect there to be a gradient in the image when we cross an edge or boundary between two regions or colors in an image. If there is no gradient, then the output is zero (no edges). This can be seen by the mostly homogenously 0 (teal) coloring in the consistently dark regions of bone or background in the original image. To achieve this, the kernel (vertical gradient or horizontal gradient kernel shown in the figure) slides over the entire image. As it slides, the kernel is multiplied elementwise by the pixels that lay underneath it and then these products are summed together; the sum is placed at the center of the kernel in the output image. Examples for applying the vertical gradient kernel to get the horizonal edges and the horizontal gradient kernel to get the vertical edges are shown in the figure for a sagittal slice of a knee MRI. The absolute magnitude of the output images identifies how "prominent" the edge is, or how large the gradient change is. The sign of each pixel

identifies the direction of the gradient and therefore whether the image is going from a dark region to a light region or vice versa.

Deep learning methods have been successful at image tasks, however, a limitation when analyzing 3D medical image data is the high memory footprint of a 3D image, and the relatively low memory available on graphics processing units (GPU) used to train deep learning models.

1.5 Physical Activity, Osteoarthritis, Biomechanics and Medical Imaging – Current State of the Literature

Characterizing the response of cartilage to mechanical loading caused by physical activity can be categorized as acute changes/deformation or longitudinal adaptations. Acute changes (deformations) include changes to the joint that are measurable during or immediately post loading; acute changes are often transient. Longitudinal adaptations include changes to the actual anatomy or physiology of the joint induced by days, weeks, or months of an intervention and slowly occur/develop over time; longitudinal adaptations are not easily reversed.

Longitudinal investigations indicate that too much (Dore et al., 2013) or too little (Hinterwimmer et al., 2004; Souza et al., 2012) exposure to activity is detrimental to cartilage health. For example, unloading of the knee during recovery from an ankle sprain decreased cartilage thickness, (Hinterwimmer et al., 2004) and produced adverse

changes to cartilage composition, as measured using T_2 (reflecting water and collagen content (Choi & Gold, 2011; Palmer et al., 2013; Souza et al., 2012)) and $T_{1\rho}$ (proteoglycan concentration (Choi & Gold, 2011; Palmer et al., 2013; Souza et al., 2012))(Souza et al., 2012). Gradually increasing exercise participation benefits cartilage health. An exercise intervention in healthy young women showed that gradually increasing activity level increases glycosaminoglycan concentration, as measured using dGEMRIC (delayed gadolinium enhanced magnetic resonance imaging of cartilage)(Van Ginckel et al., 2010). Furthermore, a 12-month randomized controlled trial comparing a step-class to no exercise in women with mild knee OA (mean age 58 years) showed no change in cartilage composition as measured using T_2 and dGEMRIC (Multanen et al., 2014). These exercise classes consisted of aerobic and step-aerobic jumping exercises that were gradually increased over 12-months. The first 3-months had step heights of 5-10cm, progressing to 20cm after 3-months. This slow progression is likely important in minimizing adverse outcomes and allowing cartilage to adapt to these loads. Before continuing with longitudinal investigations, it is important that we first understand how the OA joint acutely responds to activity. A fundamental understanding of cartilage's acute response to activity and loading is crucial for understanding which exercises are safe for the OA knee.

Current literature on the acute response of cartilage (deformation) to activity focuses on healthy adults; we must include more studies that directly assess the OA knee. The current body of knowledge shows that cartilage volume and thickness is decreased after

as little as 30 knee bends, (F Eckstein, 2005) and that recovery is approximately linear (Felix Eckstein et al., 1999). Greater deformation and more rapid recovery is thought to occur due to lower resistance to fluid flow within cartilage, suggesting cartilage degeneration (Felix Eckstein et al., 1999; Lu & Mow, 2008). Assessment of acute changes to cartilage composition have mostly been performed in running, showing that superficial cartilage typically undergoes the greatest changes after running, (Cha et al., 2012; T.J. Mosher et al., 2010; Subburaj et al., 2012) and that older runners (>45 years of age) have greater changes in cartilage composition compared to younger runners (<20 years of age)(Cha et al., 2012). When comparing the acute effect of running on cartilage T_2 in healthy versus OA knees, it was found that healthy cartilage T_2 did not change (Esculier et al., 2019). However, OA cartilage had increased T_2, relative to pre-activity levels, at 60- and 90-minutes post-activity. This study had its earliest post-activity time-point at ~20 minutes and therefore may have missed the traditional decrease observed after activity. A more recent study also found that when measuring T_2 before and then within 10-hours post marathon that T_2 increased, (Z. Wang et al., 2020) and that this increase is more pronounced the higher an individuals' body mass index (BMI). These studies show that while the immediate acute response to mechanical load is decreased T_2, in the hours afterwards there appears to be an increase. Previously, we compared the acute changes in cartilage composition and morphology after running versus bicycling in healthy young men. We found that when normalized to total load just 15-minutes of running changed cartilage composition (reduced T_2) and reduced cartilage thickness and volume while 46-minutes of bicycling did not (Gatti, 2015; Gatti et al., 2017). The

finding that bicycling produced no changes in cartilage highlights its potential to enable exercise while minimizing the need for cartilage to dissipate loads by deforming for individuals with knee OA.

Most studies of cartilage change (acute or longitudinal) use similar rigid definitions of cartilage regions of interest, e.g., the entire medial or lateral tibial plateau, and the entire medial or lateral weight-bearing femur. It is possible that these standard definitions, which calculate mean outcomes over anatomically defined regions, may washout or miss subregional changes. In the worst case, it is possible that using these region definitions may actually produce erroneous results.

To combat problems encountered by broad region of interest definitions, some work in the knee OA literature has begun using statistical parametric mapping (SPM)(Brenneman et al., 2019a, 2019b; Gatti & Maly, 2018; MacKay et al., 2020). SPM has also gained popularity in the biomechanics field over the past decade, particularly owing to creation of the open-source package SPM-1D (Pataky, 2010; Pataky et al., 2008, 2016). SPM is a technique pioneered and made popular in brain MRI research (Forman et al., 1995; Friston et al., 1994). SPM enables hypothesis testing of spatially correlated data while accounting for the multiple comparisons problem inherent in performing statistical analyses in this type of data.

1.6 Bicycling & Knee OA

Bicycling may be the ideal activity to allow individuals with knee OA to exercise for weight management and general aerobic health. Supporting the use of bicycling for weight management, a study on young sedentary overweight women showed that indoor cycling training 3 times per week for 12 weeks resulted in 3.2% reductions in body weight and 5% reductions in fat mass (Bianco et al., 2010). Furthermore, an RCT in 100 adults between 67 and 76 years of age showed that participation in 3x1 hour cycling classes per week for 12-weeks resulted in significantly greater reductions in weight (0.83 kg; $p=0.007$), body mass index (0.26 kg/m^2; $p=0.032$), and waist circumference (1.7 cm; $p=0.023$) compared to the control group (Finucane et al., 2010).

Bicycling improves function, pain, and quality of life in adults with knee OA. In 2016, Alkatan and colleagues found that 12-weeks of bicycling (3 times/week; 45 mins/session) reduces joint pain, stiffness, and physical limitations while concurrently increasing quality of life and physical function (Alkatan et al., 2016). Another 12-week RCT found that bicycling classes lead to increased preferred gait speed, decreased joint pain after a 6-minute walk, and decreased joint pain and stiffness compared to a no-exercise control group (Salacinski et al., 2012). Furthermore, exercise intensity does not seem to matter; when randomized to a high-intensity or low-intensity stationary cycling program, regardless of group, participants with knee OA had significant improvements in functional tests, measures of overall pain, and aerobic capacity (Mangione et al., 1999).

These findings indicate that bicycling is a promising activity for minimizing joint contact forces at the knee, while allowing individuals with knee OA to improve body composition and physical function. Not only does bicycling produce small changes to cartilage, (Gatti et al., 2017) it has other health and economic benefits. Bicycling indoors, reduces the risk of injury from falls that may occur during walking, running, or other activities performed in an unpredictable environment. With appropriate positioning and cycling biomechanics, lower limb movements during bicycling are almost entirely in the sagittal plane and minimize movements in the frontal plane. Minimizing frontal plane movement may be important for maintaining stability for osteoarthritic and/or potentially unstable joints. Recent investigations also show that the peak KAM during bicycling is < 0.1 Nm/kg in adults with knee OA bicycling at a moderate 80 W (Gardner et al., 2015) and the KAM was <0.18 Nm/kg over a range of resistances and pedaling cadences in healthy adults between 40 and 79 years of age (Fang et al., 2016). These values are less than half of those reported during gait (0.4-0.5 Nm/kg)(Maly et al., 2013). Furthermore, the bicycle is highly adaptable; positioning of the bicycle seat, handlebars, and pedals can be optimized to minimize loading at the knee. Not only does the bicycle have mechanical and injury prevention benefits, bicycling can also be performed in a group format which augments exercise adherence, (Jordan et al., 2010) and minimizes costs (McCarthy et al., 2004).

1.6.1 Bicycle-fit

While bicycling does show promise as an intervention for knee OA, (Alkatan et al., 2016; Mangione et al., 1999; Salacinski et al., 2012) and many people (>40% of Canadians) bicycle, (Ramage-Morin, 2017) more than 51% of cyclists report overuse injuries (Silberman, 2013). Further cause for concern in knee OA, the joint that is most affected by overuse injuries is the knee (1/3 of overuse injuries)(De Bernardo et al., 2012). The primary method of mitigating musculoskeletal injuries during cycling is bicycle-fit (Dettori & Norvell, 2006; Silberman, 2013). Bicycle-fit primarily includes adjusting positioning of the saddle and handlebars and changing the length of the bicycle crank arms (Figure 1-3). For minimizing injuries at the knee, saddle positioning and crank arm length are of primary interest (Bini et al., 2011). However, a review of bicycle saddle positioning for knee injury risk and cycling performance concluded that:

"Methods for determining saddle height are varied and have not been comprehensively compared using experimental research studies. There is limited information on the effects of saddle height on lower limb injury risk, but more information on the effects of saddle height and cycling performance."(Bini et al., 2011)

There is limited new experimental research since this review (Bini, 2016; Bini et al., 2014; Menard et al., 2018). Given this state, the gold-standard recommendation for setting bicycle saddle height is based on knee flexion at the bottom of the revolution (6 o'clock), and is provided based on clinical expertise (Bini et al., 2011; Holmes et al., 1994). It is recommended that saddle height be such that it elicits a minimum knee flexion angle (knee flexion at the bottom of the revolution) of 25-40º, with 25-35º

recommended for static measurements and 30-40º recommended for dynamic measurements (Bini et al., 2011; Ferrer-Roca et al., 2012; Fonda et al., 2014; Priego Quesada et al., 2017). Within these ranges, individual recommendations are based on injury history and performance aspirations. Generally, saddle positions with less flexion (25-30º) are desired for performance and for those trying to prevent or alleviate conditions associated with the patellofemoral joint (Bini et al., 2011; Dettori & Norvell, 2006; Silberman, 2013). More knee flexion is commonly recommended for those with iliotibial band syndrome (Silberman, 2013). Additional research is needed to confirm the mechanical hypotheses linking bicycle fit (e.g., saddle height, knee flexion angle) and joint mechanics (e.g., joint reaction forces). With these data in hand, prospective studies are needed to determine whether these changes in bicycle-fit actually affect injury risk (Bini, 2016).

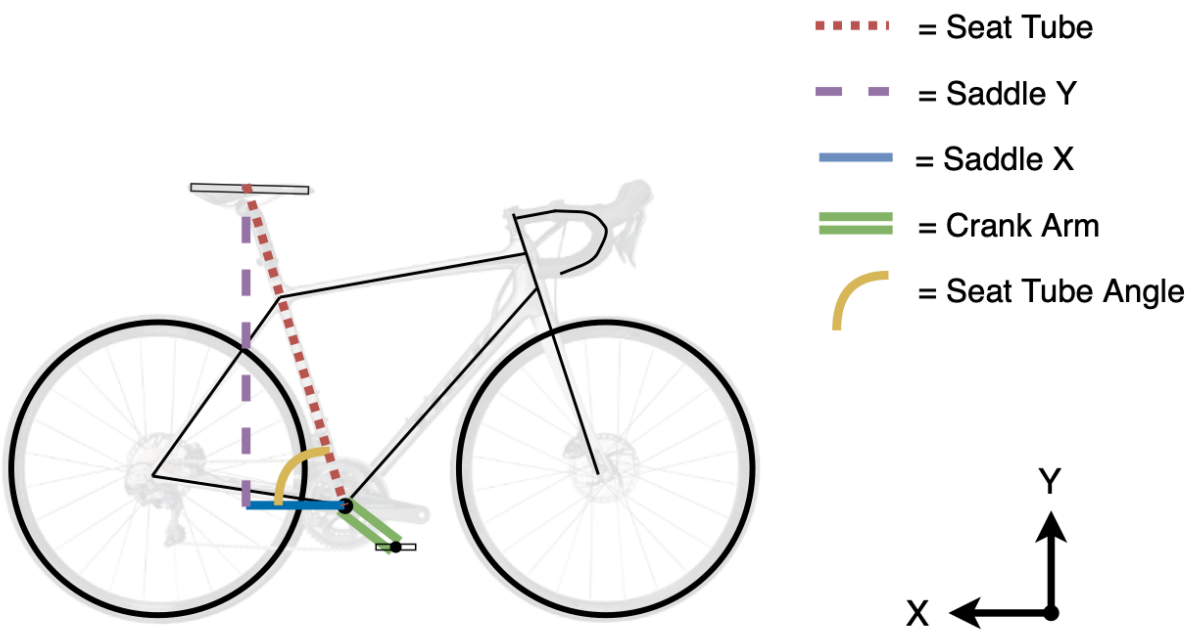


Figure 1-3. Bicycle-fit anatomy. Depiction of key bicycle components relevant for bicycle-fit of the lower extremity: saddle positioning and crank arm length. Saddle position can be described using two conventions. The first describes the position of the saddle using the straight-line distance from the bottom bracket to the top of the saddle (red dashed line, seat tube) and the angle of this line with respect to the horizontal (seat tube angle, yellow arch). This method effectively uses a polar coordinate system to describe saddle position relative to the bottom bracket. The second method that has become more popular, particularly for bicycle-fit, breaks this positioning into an x- and y-cartesian coordinate system, again centered at the bottom bracket. In this figure the cartesian coordinate system is noted in the bottom right corner; the x-axis always points toward the back of the bicycle. Using this method, the saddle position is described relative to the bottom bracket using an x-component (blue line) and a y-component (dashed purple line). In addition, crank arm length is the length of the arms to which the pedals attach (crank arm, green double line). The crank arm length is the radius of the circle which a rider's pedals rotate around when they are pedaling. The total distance from the pedal attachment to the top of the saddle is the saddle height. The saddle height can be calculated by adding the crank arm length (green double line) and the bottom bracket to top of saddle distance (red dashed line).

Even if specific recommendations backed by research existed, measuring knee flexion angles either statically or dynamically on the bicycle remains a challenge. This problem is highlighted by a reliability study of knee flexion angle measurements obtained at the bottom of a revolution (average of 15 revolutions) using high speed video recorded in laboratory conditions (Fonda et al., 2014). In the reliability study, the standard error of measurement of the knee flexion angle was 3º. This error is equivalent to a minimum detectable difference at 95% confidence of +/- 8.3º (i.e., 95% of unchanged individuals (same saddle height) will have random fluctuations of their average of 15 revolutions within +/- 8.3º). In the context of recommendations for dynamic knee flexion being within a 10º window (30-40º), +/- 8.3º is considerable. For this reason, we need either new technologies to measure joint kinematics during cycling that enable smaller measurement error, or we need alternative methods of prescribing saddle height that elicits a desired set of joint kinematics.

1.7 Gaps in the literature.

1. There is no easy way to prescribe or determine the optimal saddle position to elicit a desired amount of knee flexion.
2. There are no thorough experimental studies that determine how changing bicycle saddle position, and thus joint kinematics, affect forces acting inside of the knee.
3. Current methods of segmenting cartilage and bone from MRI for quantitative analyses are slow and prone to error. Faster and more accurate methods are required to enable efficient analysis.

4. We lack an understanding of how knee cartilage affected by OA acutely responds to physical activity *in vivo* and what biomechanical or joint health factors affect the response.

1.8 Thesis overview

The purpose of this thesis is to improve our understanding at the intersection of aerobic activity, knee biomechanics, quantitative MRI outcomes of the knee, and knee OA.

Chapters 2 and 3 of this thesis use data from one experimental study. That study had 41 healthy adult participants complete 18 different bicycling postures where 1 posture was derived from commercial guidelines and the remaining 17 were random deviations from commercial.

Chapter 2 uses this data to construct a set of equations that enable prediction of saddle position that elicits a user desired joint kinematic. The equations enable the operator to define a desired minimum or maximum knee flexion angle and it will return a saddle height predicted to achieve this kinematic outcome.

Chapter 3 determines how joint kinematics affect tibiofemoral and patellofemoral JRFs during cycling. Both the individual sensitivity of the JRFs to each lower-limb kinematic outcome (ankle flexion, knee flexion, hip flexion) as well as the best set of kinematic outcomes to predict each of the tibiofemoral and patellofemoral JRFs are reported.

Chapter 4 describes and validates a novel multi-stage convolutional neural network framework designed to efficiently segment cartilage and bone from high resolution MRI data acquired from both healthy and OA afflicted knees.

Chapter 5 explores how JRFs obtained during walking and bicycling, as well as bone morphology measured using SSMs affects the acute response of OA cartilage to walking and bicycling in a sample of older women with symptomatic knee OA. This study used results from chapters 2 and 3 to prescribe the bicycle-fits for women with knee OA and uses the convolutional neural network framework from chapter 4 to analyze the MRI data.

2 Equations to Prescribe Bicycle-Fit based on Desired Joint Kinematics and Bicycle Geometry

This paper was submitted to the European Journal of Sport Science on September 14th, 2020 (TEJS – 2020 – 1319).

Equations to Prescribe Bicycle-Fit based on Desired Joint Kinematics and Bicycle Geometry

European Journal of Sport Science

[1]Anthony A. Gatti, MSc, [2]Peter J. Keir, PhD, [3,4]Michael D. Noseworthy, PhD, [1,5]Marla K. Beauchamp, PT PhD, [1,2,6]Monica R. Maly, PT PhD

[1]School of Rehabilitation Sciences, [2]Department of Kinesiology, [3]School of Biomedical Engineering, [4]School of Electrical and Computer Engineering, and [5]Department of Medicine, McMaster University, Hamilton, Canada

[6]Department of Kinesiology, University of Waterloo, Waterloo, Canada

Corresponding Author:

Monica R. Maly
University of Waterloo
200 University Ave,
Waterloo, Ontario, Canada,
N2L 3G1
Fax: +1 519 885 4070
Email: mrmaly@uwaterloo.ca

Abstract word count: 224

Manuscript word count: 3,976

Running title: Bicycle-fit based on joint kinematics

Abstract

Background. Overuse knee injuries are common in bicycling and are often attributed to poor bicycle-fit. Bicycle-fit for knee health focuses on setting saddle height to elicit a minimum knee flexion angle of 25-40º . Equations to predict saddle height include a single input, resulting in a likely suboptimal bicycle-fit. The purpose of this work was to develop an equation to predict saddle height from anthropometrics, bicycle geometry, and user-defined joint kinematics.

Methods. Forty healthy adults (17 women, 23 men; mean (SD): 28.6 (7.2) years; 24.2 (2.6) kg/m^2) participated. Kinematic analyses were conducted for 18 three-minute bicycling bouts including all combinations of 3 horizontal and 3 vertical saddle positions, and 2 crank arm lengths. For both minimum and maximum knee flexion, predictors were identified using Least Absolute Shrinkage and Selection Operator (LASSO) regression, and final models were fit using linear regression.

Results. The equation to predict saddle position from minimum knee flexion angle (R^2=0.97; root mean squared error (RMSE)=1.15 cm) was: Saddle height (cm) = 7.41 + 0.82(inseam cm) – 0.1(minimum knee flexion º) + 0.003(inseam cm)(seat tube angle º). The maximum knee flexion equation (R^2=0.97; RMSE=1.15 cm) was: Saddle height (cm) = 41.63 + 0.78(inseam cm) – 0.25(maximum knee flexion º) + 0.002(inseam cm)(seat tube angle º).

Conclusions. These equations provide a novel, practical strategy for bicycle-fit that accounts for rider anthropometrics, bicycle geometry and user-defined kinematics.

Key words: Saddle; Biomechanical Phenomena; Injuries; Knee; Machine Learning

2.1 Introduction

Bicycling is associated with injuries caused by acute trauma (48.5%) and overuse (51.5%), with overuse injuries most commonly occurring at the knee (32.1%)[1,2]. Minimizing exposure to the biomechanics associated with overuse injury risk is a primary goal of "bicycle-fit"[2,3]. Bicycle-fit to prevent knee injuries involves adjusting saddle position[2], which alters lower extremity kinematics and joint reaction forces at the knee[4–10]. Yet, there is no consensus on the most effective method to prevent bicycling-related musculoskeletal injury[4]. It is likely that no single saddle height can be optimal for all riders. Instead, saddle height must be tailored based on anthropometrics, bicycle geometry and goals (i.e., maximize performance and/or minimize injury risk). There are multiple conventions that describe saddle position, with the most common requiring measurements of seat tube angle and saddle height or of the horizontal and vertical saddle positions (Figure 2-1).

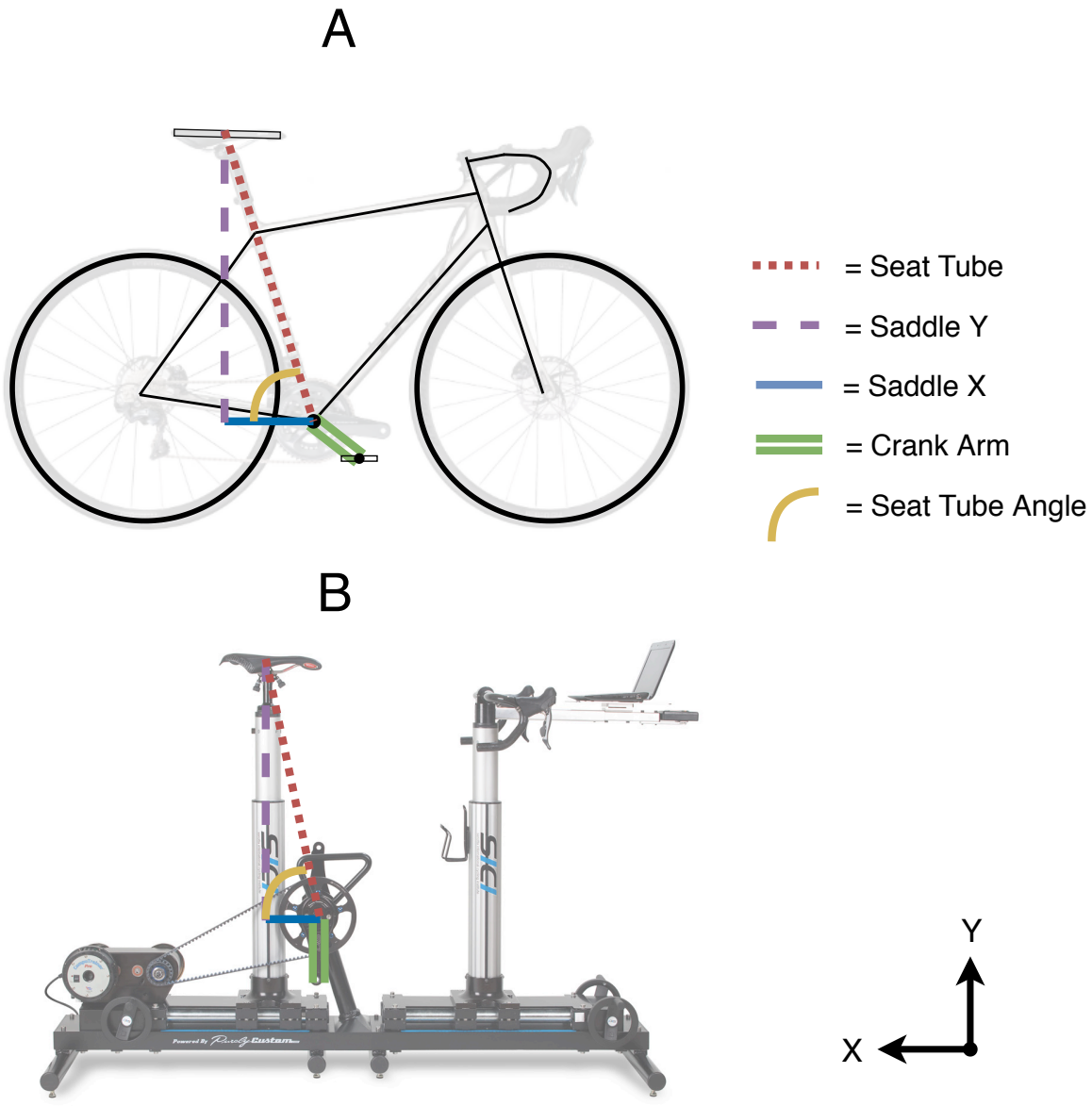


Figure 2-1. Bicycle fit parameters are shown on a conventional bicycle (A) and the bicycle-fit device (Purely Custom) used in this study (B). Both methods of describing saddle position have an origin at the bottom bracket (axis of rotation of the bicycle pedals), which is located at the intersection of the seat tube, saddle X, and saddle Y lines. Method 1 of describing saddle position uses the seat tube angle (gold) and the distance from the bottom bracket to the top of the saddle along the seat tube (red) to describe the location of the top of the saddle in space. Method 2 decomposes the seat tube into X and Y components on a Cartesian coordinate system. Methods 1 and 2 can be interchanged using trigonometry; e.g., to get conventions of Method 1 from Method 2, the seat tube angle can be calculated as $arctan(\frac{saddle\ Y}{saddle\ X})$ and the saddle height as

$crank\ arm\ length\ (green)\ +\ bottom\ bracket\ to\ saddle\ clamp\ distance\ +\ saddle\ clamp\ to\ top\ of\ saddle\ distance$, where bottom bracket to saddle clamp is calculated as $\sqrt{saddle\ X^2\ +\ saddle\ Y^2}$ and saddle clamp to the top of saddle along the seat tube is calculated as $\frac{saddle\ vertical\ depth}{sin(seat\ tube\ angle)}$.

"Ideal" saddle position is commonly achieved by assessing the minimum knee flexion angle, and the horizontal saddle alignment using knee-over-pedal-spindle (KOPS)[2], though both techniques can be challenging for clinicians to implement. Prior literature, based on clinical expertise, recommends minimum knee flexion angle is between 25-35º for static and 30-40º for dynamic measurements[4,11–13]. High knee flexion is hypothesized to instigate patellofemoral knee pain, while low knee flexion is hypothesized to aggravate the iliotibial-band and hamstrings[2–4]. The equipment or expertise to accurately measure knee flexion angles in the clinic or bicycle-shop is often unavailable and can be erroneous. For example, measurement of knee flexion angle at the bottom of a pedal revolution, acquired from video in a laboratory, showed a standard error of measurement of 3° [12] [†]. While this error seems small, the coinciding smallest detectable difference at 95% confidence was +/- 8.3° [14]. As reflected by the different recommended ranges, minimum knee flexion angle obtained from static measurements is systematically lower

[†] Standard error of measurement (SEM) was calculated as: $SEM\ =\ {}^{s_d}/_{\sqrt{2}}$ where s_d = standard deviation of the difference between repeated measurements. Fonda et al (2014) found that s_d of repeated measurements taken with dynamic 2D video was 4.3º. Therefore, $SEM\ =\ {}^{4.3}/_{\sqrt{2}}\ =\ 3.0°$

than dynamic measurements[12]. Horizontal saddle alignment using KOPS recommends the anterior knee be at or behind the pedal spindle with the pedal at 3 o'clock[2]. KOPS presumes that a more anterior knee alignment is associated with higher patellofemoral forces, however, recent work disputes this assertion[5,9,10]. Horizontal alignment is fine-tuned by sliding the saddle anterior/posterior at the seat clamp. Horizontal alignment affects the effective seat tube angle.

Minimum knee flexion angle and horizontal saddle alignment may not be ideal to mitigate patellofemoral conditions. During bicycling, peak quadriceps activations and forces occur near the top of the revolution[9,15], where the knee is most flexed. Greater knee flexion decreases the quadriceps moment arm, necessitating higher quadriceps forces, and therefore higher patellar compression, for a given knee extension moment[16]. In fact, patellofemoral joint reaction forces are three times more sensitive to maximum knee flexion angle as they are to minimum knee flexion angle[10]. In this scenario, bicycle-fit may be most effective if focused on identifying the maximum, rather than minimum, knee flexion angle.

A practical approach to setting saddle height is using an anthropometry-based equation, such as, Hamley and Thomas (109% of inseam)[17], trochanteric height (100% of trochanter height)[18], ischial tuberosity height (113% of ischial tuberosity height) [19], and the LeMond equation (distance from the bottom bracket to the top of saddle at 88% of

inseam height) [4]. While popular, these equations present key problems. All prescribe saddle height as a percentage of a single anthropometric measurement[4]. Their derivations are based on performance athletes, primarily men. These do not account for confounding effects of bicycle geometry (e.g., seat tube angle). These are based on data at a limited number of saddle heights. It is therefore not surprising these equations produce different saddle height recommendations from one another, with the inseam-based equations often yielding saddle heights outside the recommended range[20,21].

The purpose of this investigation was to develop an equation that enables accurate prediction of saddle height for a desired knee flexion angle based on simple anthropometrics and bicycle geometry (seat tube angle) that can be applied across a range of recreational bicyclists. Our secondary objectives were to determine (i) the best equations to predict bicycle saddle height when one of maximum knee flexion, minimum hip flexion, or maximum hip flexion were included in each equation; and (ii) whether saddle height prediction equations differed by sex. We hypothesized that all equations to predict saddle height would include inseam, foot length, and seat tube angle.

2.2 Methods

A fully-crossed random assignment experimental design, the same as described previously, was used[10]. Participants provided written, informed consent. The funders

played no role in the design, conduct, or reporting of this study. This study was approved by the Hamilton Integrated Research Ethics Board.

2.2.1 Participants

A convenience sample of healthy adults between 18-45 years of age who had cycling experience from recreational urban cycling to regional age-group level cyclists in mountain biking, cyclocross, and triathlon were recruited from the Hamilton, Ontario community. Participants were excluded if responses on the Get Active Questionnaire (GAQ) suggested exercise was unsafe[22], if they self-reported any injury to the lower extremity during the past 3-months, or self-reported arthritis, gout, unstable angina, neurologic conditions, or any previous surgery to the lower extremities. Descriptors included the Lower Extremity Functional Scale[23], sex, body height, body mass, body mass index, foot length, and inseam height.

2.2.2 Protocol

Participants completed one study visit (2.5-4 hour duration) and wore shorts, running shoes, and were either shirtless or wore a sports-bra. Participants could wear their own cycling shorts, though not required. The protocol included five steps: (i) measuring anthropometrics, (ii) defining commercial bicycle-fit, (iii) determining the experimental exercise intensity, (iv) instrumenting for kinematic analysis, (v) conducting the experimental bicycling protocol (18 x 3-min bicycling bouts).

2.2.2.1 Anthropometrics

Height, inseam, lateral malleolar height, greater trochanter height, and foot length were measured for both legs while standing barefoot. Inseam was measured from the floor to the top of a level which participants held between their legs while applying gentle pressure upward[4,20]. Malleolar, trochanter, and femoral condyle heights were measured from the floor. Foot length was measured from the posterior aspect of the calcaneus to the end of the longest toe using calipers. One researcher (AAG) palpated each anatomical structure and completed all measurements.

2.2.2.2 Commercial Bicycle-Fit

A fully-adjustable commercial fit-bike (Fit Bike Pro, Purely Custom, Twin Falls, ID, USA) was used. The commercial bicycle-fit included setting the saddle horizontal and vertical position using the Hamley and Thomas[4,17] equation, assuming a seat tube angle of 73.7° (mean of all 2018 Trek road and city bicycles), and a crank arm length of 172.5 mm. The vertical handlebar position was initially set level with the saddle, and the horizontal distance was set to achieve a 70-90° angle between the torso and upper arms. From the initial handlebar position, participants were encouraged to provide feedback to adjust handlebars to their preference. Flat pedals were used with the first metatarsal head aligned over the pedal spindle and the foot firmly secured using custom Velcro straps. Foot placement and fastening on the pedal were performed by the same researcher (AAG).

2.2.2.3 Determining Experimental Exercise Intensity

The exercise intensity used during the experimental bicycling bouts was determined from an incremental bicycling protocol that was a modified version of the YMCA submaximal test of maximal oxygen uptake[10,24,25]. Using the commercial bicycle-fit, participants started cycling with a resistance of 4 Newton-metres at a self-selected cadence (30-45W). Every three minutes resistance was increased by one gear. Heart rate (HR) was recorded every minute (Polar T31, Kempele, Finland). If HR differed by more than 5 beats between consecutive measurements, the stage was extended by 1 minute until steady-state was achieved. The protocol continued until participant HR was within 70-75% of their age-predicted maximum (208-0.7*Age)[10,26]. The average power output and cadence of the final stage were used for the experimental bicycling bouts. If the final stage HR was greater than 75% of age predicted maximum, power was set to the second last stage.

2.2.2.4 Instrumenting for Kinematic Analysis

For kinematic analysis with a passive motion capture system of 12 high-speed infrared cameras (Raptor-4, Motion Analysis Corporation, Santa Rosa, CA, USA), participants were outfitted with 40 retroreflective markers. Markers were affixed to both feet (1st metatarsal head, 5th metatarsal head, calcaneus, and lateral midfoot), lower legs (medial and lateral malleoli, tibial tuberosity, and 4 markers affixed to a rigid body placed on the lateral shank), thighs (medial and lateral femoral epicondyles, greater trochanter, and 4 markers affixed to a rigid body placed on the lateral midthigh), and the pelvis (right and left anterior superior and posterior iliac spines). Marker data were collected at 112.5 Hz.

2.2.2.5 Experimental Bicycling Bouts

Each participant completed 18 three-minute cycling bouts, each at the cadence and power derived from the incremental protocol, in randomized order. These 18 positions included all combinations of 3 vertical saddle positions (Y), 3 horizontal saddle positions (X), and 2 crank arm lengths. The Y component was the vertical distance from the bottom bracket to the centre of the saddle clamp, and the X component was the horizontal distance from the bottom bracket to the centre of the saddle clamp. One position of each factor was commercial, and the others were random deviations. Each random saddle X and Y position was independently randomly sampled within +/- 10% for X and 5% for Y of the commercial position. A shorter range of 5% was used for saddle Y because higher ranges were not anatomically possible. The random crank arm length was either 2.5 mm longer or shorter than commercial (172.5 mm). Randomization of the saddle and crank arm positions are described in Supplemental 1. For every position, effective seat tube angle (θ_{ST}) was calculated as $arctan(\frac{Y}{X})$, and represents a seat tube from the bottom bracket to the X-Y position of the saddle clamp on the fit-bike. Effective seat tube angle takes saddle fore-aft into account, where a steeper effective seat tube is equivalent to a more anterior saddle position. The effective saddle height was calculated by summing the crank arm length, the bottom bracket to saddle clamp distance ($a = \sqrt{X^2 + Y^2}$), and the saddle clamp to top of saddle distance. The saddle clamp to top of saddle distance was calculated along the line of the effective seat tube as:

$saddle\ clamp\ to\ top\ of\ saddle\ distance = \frac{saddle\ vertical\ depth}{sin(\theta_{ST})}$ where the saddle vertical depth was the vertical height from the centre of the saddle rails to the top of the

saddle. The effective saddle height was equivalent to the distance from the top of the saddle to the pedal spindle using a straight line that intersected the bottom bracket (Figure 2-1). In each position, handlebars were moved to maintain their position relative to the saddle.

2.2.2.6 Data Processing

All marker data were manually assessed to ensure accuracy (Cortex, Motion Analysis Corporation, Santa Rosa, CA, USA). Marker data were filtered using a second order dual low-pass 6 Hz Butterworth filter. The cut-off frequency retained 95% of the signal [27]. Data from the commercial trial were used to fit a functional knee centre using the symmetric centre of rotation estimation method constrained to the midpoint of the mediolateral axis of the femur[28]. Hip joint centres were calculated using the Harrington method[29]. Joint centres and anatomical markers were used to scale a 16 degrees of freedom (DOF) lower-body musculoskeletal model (provided by Lai and colleagues: simtk.org/projects/model-high-flex) using OpenSim[30,31]. The lower-body model subtalar joint was locked, reducing the original model from 18 to 16 DOF. Using the lower-body model, joint angles were calculated for the final minute of each bout using inverse kinematics[31]. The final minute reflected steady state [32].

Kinematic data from the right and left legs were divided into individual revolutions, starting and ending when the pedal was in the 12 o'clock position, as tracked by 4

retroreflective markers permanently affixed to each pedal. 12 o'clock was defined as the instant when the centre of the pedal was at a local maximum. Minimum and maximum knee and hip flexion were obtained from every revolution. For each posture, kinematic outcomes were calculated as the average across all revolutions.

2.2.3 Statistical Analysis

To predict saddle height from bicycle geometry (θ_{ST}) and anthropometrics, a 10-fold cross-validation and a Least Absolute Shrinkage and Selection Operator (LASSO) regression was used. This machine learning approach identified the combination of predictors that had the smallest testing error[33]. We then identified the final model coefficients for the identified predictors using least squares regression with all available data.

2.2.3.1 Least Absolute Shrinkage and Selection Operator

LASSO is a form of least squares regression that is done through regularization using a tuneable parameter ($\boldsymbol{\lambda}$), that when set to 0 is equivalent to least squares regression. As $\boldsymbol{\lambda}$ is increased unimportant beta coefficients are reduced to zero, therefore simplifying model selection. The optimal $\boldsymbol{\lambda}$ was identified using a grid-search that identified the smallest out-of-sample mean squared error (MSE) using 10-fold cross-validation. That is, in each of the 10 cross-validation steps, a different $1/10^{th}$ of the data was withheld (not used to fit the model). The fitted model was then used to predict the outcomes of the

withheld data and the MSE was calculated. This was repeated for every tested $\boldsymbol{\lambda}$. The $\boldsymbol{\lambda}$ with the smallest average MSE (across the 10 cross-validation steps) was the optimal $\boldsymbol{\lambda}$. The final predictors are those that had non-zero beta coefficients at the optimal $\boldsymbol{\lambda}$. Potential predictors were: minimum knee flexion, seat tube angle, lateral malleolar height, femoral condyle height, greater trochanter height, inseam, foot length, and body height, squared terms for each predictor, and the interactions between minimum knee flexion and seat tube angle with each of the anthropometrics.

2.2.3.2 Least Squares Regression

Final model coefficients were determined by fitting the final predictors using linear least squares and a cluster-robust variance matrix to account for non-independence between measurements of the same participant. The final model was fit using all data. The model was checked for assumptions of linear regression. To assess generalizability, the final regression model was re-fit using 10-fold cross-validation where each iteration estimated the root mean squared error (RMSE) on withheld data.

2.2.3.3 Secondary Analysis

To address the secondary questions, we replaced minimum knee flexion with each of the respective kinematic variables. Also, sex and its interaction with each predictor were added as covariates in each model and a backwards deletion method determined if sex added to the model[33].

LASSO regression was run using Scikit-learn version 0.22 for Python 3.7. All other statistical analyses were conducted using statsmodels version 0.11 for Python 3.7.

2.3 Results

Forty-one healthy adults (17 women, 24 men) completed the study. Analyses include data from 40 participants (Table 2-1); data from one man (38 years, 27.9 kg/m^2) was excluded due to error in their recorded saddle positions. Participants reported Lower Extremity Functional Scale scores (mean 79.2, standard deviation (SD) 1.9) consistent with no impairment[23]. The distributions, including the mean, standard deviation, skew, and kurtosis, of minimum and maximum knee and hip flexion angles, normalized saddle heights ($normalized\ saddle\ height\ =\ saddle\ height/inseam$) and seat tube angles obtained over all bicycling bouts are included in **Supplemental 2**.

Table 2-1. Participant demographics including age, body mass index (BMI), height, body mass, foot length, lateral malleolar height, lateral femoral epicondyle height, greater trochanter height, and inseam height, as well as cadence and power during the bicycling activity. Demographics are presented for Women, Men, and All participants. Each cell is formatted as: mean (standard deviation) [range].

	Women (n=17)	Men (n=23)	All (n=40)
Age (y)	27.7 (6.0) [23 – 42]	29.2 (8.0) [19 – 44]	28.6 (7.2) [19 - 44]
BMI (kg/m^2)	23.7 (3.1) [20.1 – 31.6]	24.5 (2.2) [20.3 – 30.5]	24.2 (2.6) [20.1 - 31.6]
Height (m)	1.66 (0.07) [1.57 – 1.78]	1.80 (0.08) [1.65 – 1.95]	1.74 (0.10) [1.57 – 1.95]
Body Mass (kg)	65.5 (9.3) [54.2 – 92.1]	79.6 (10.0) [63.2 – 102.3]	73.6 (11.9) [54.2 – 102.3]
Foot Length (cm)	24.3 (1.3) [21.7 – 26.6]	27.2 (1.2) [24.8 – 28.9]	26.0 (1.9) [21.7 – 28.9]
Lateral Malleolar Height (cm)	6.9 (0.7) [5.5 – 8.0]	7.4 (0.6) [6.5 – 9.0]	7.2 (0.7) [5.5 – 9.0]
Lateral Femoral Epicondyle Height (cm)	47.3 (3.0) [41.0 – 52.5]	52.1 (2.7) [47.5 – 57.5]	50.0 (3.7) [41.0 – 57.5]
Greater Trochanter Height (cm)	84.8 (5.4) [75 – 93.5]	92.4 (5.1) [82.0 – 102.0]	89.2 (6.4) [75.0 – 102.0]
Inseam Height (cm)	78.7 (4.9) [67.8 – 85.8]	84.3 (5.2) [73.3 – 96.0]	81.9 (5.7) [67.8 – 96.0]
Cadence (rpm)	81.2 (6.1) [70 – 92]	85.5 (6.6) [72 - 99]	83.7 (6.7) [70 - 99]
Power (W)	86.4 (30.6) [40 – 146]	140.4 (37.7) [75 - 216]	117.4 (43.8) [40 - 216]

The model that best predicted saddle height using minimum knee flexion (R^2=0.97, RMSE=1.15 cm; Table 2-2) is presented in Equation 2-1. The RMSE from cross validation was 1.16 cm; a small reduction in RMSE from cross validation indicates there was little to no overfitting. Higher inseam, steeper (higher) seat tube angles and lower knee flexion angle resulted in a higher predicted saddle height (Figure 2-2). Saddle height was most sensitive to changes in inseam, where a 1 SD change in inseam resulted in a 5.96 cm change in saddle height.

Equation 2-1

$$Saddle\ Height = 7.41 + 0.82(inseam) - 0.10(min\ knee\ \theta) + 0.003(seat\ tube\ \theta)(inseam)$$

Table 2-2. Results of the regression models predicting saddle height based on minimum knee flexion angle (top model), and maximum knee flexion (bottom model).

	Coefficient	**Standard Error**	**P-value**
Minimum knee flexion model **R^2** = 0.97; **F** = 1232 (3, 1390)			
Intercept	7.41	2.27	0.001
Inseam	0.82	0.08	<0.001
Minimum knee flexion	-0.10	0.02	<0.001
Seat tube angle X inseam	0.003	0.001	0.001
Maximum knee flexion model **R^2** = 0.97; **F** = 994.2 (3, 1390)			
Intercept	41.63	5.48	<0.001
Inseam	0.78	0.08	<0.001
Maximum knee flexion	-0.25	0.04	<0.001
Seat tube angle X inseam	0.0020	0.001	0.058

When creating secondary models to predict saddle height, neither minimum nor maximum hip flexion added to the model fit. Therefore, only maximum knee flexion results are presented. Equation 2-2 shows the maximum knee flexion model (R^2=0.97, RMSE=1.15 cm; Table 2-2). The RMSE of predicted saddle height for the cross validation (1.19 cm) demonstrates negligible overfitting. Lower specified maximum knee flexion, higher inseam, and steeper seat tube angles resulted in higher predicted saddle height (Figure 2-2).

Equation 2-2

$$Saddle\ Height\ =\ 41.63\ +\ 0.78(inseam)\ -\ 0.25(max\ knee\ \theta)\ +\ 0.002(seat\ tube\ \theta)(inseam)$$

Sex did not add to the minimum or maximum knee flexion angle models.

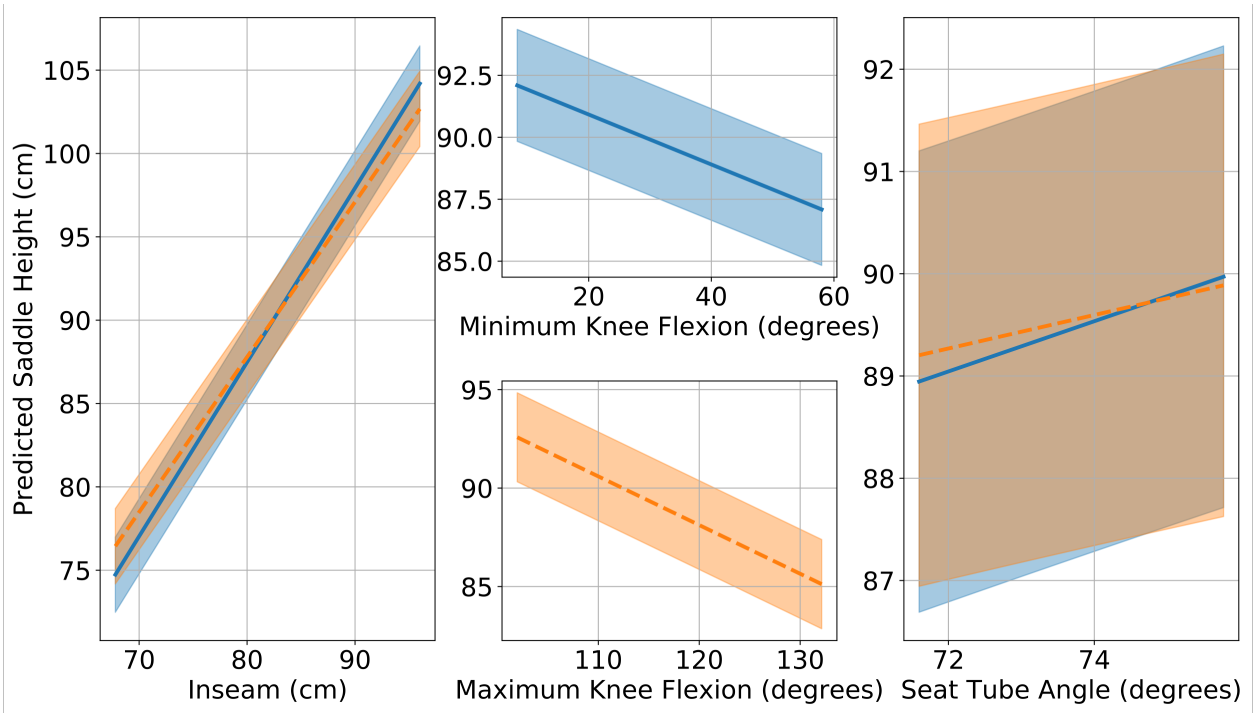


Figure 2-2. Line graphs with prediction bands demonstrating the change in predicted saddle height (cm) resulting from changes in inseam, minimum knee flexion, maximum knee flexion, and seat tube angle. Blue fitted lines and shaded areas are for the minimum knee flexion model. Orange dashed lines and shaded areas are for the maximum knee flexion model. Saddle height was positively related to inseam and seat tube angle, and inversely related to minimum and maximum knee flexion. Saddle height is most sensitive to changes in inseam.

2.4 Discussion

This study generated two novel equations to predict saddle height from inseam, the desired knee flexion angle (minimum or maximum) and the effective seat tube angle of the bicycle. As demonstrated by identical R^2 and RMSE, the minimum and maximum knee flexion models were not meaningfully different. Pragmatically, the model including minimum knee flexion angle, an industry standard for bicycle-fit, is likely the most useful. This is the first study to develop a saddle height prediction equation that enables

specification of user-defined joint kinematics, while accounting for bicycle geometry. Our study has direct practical implications for bicycle-fit among healthy adults.

We produced the first experimentally-derived equation that accounts for rider and bicycle geometry to yield a desired saddle position. Compared to the Hamley and Thomas[17] method, which was derived as the discrete saddle position that minimized time to exhaustion, the current study produced a smaller magnitude beta coefficient (0.82 vs 1.09) for inseam. For a given increase in inseam, our equation predicts a smaller increase in saddle height, likely due to the inclusion of multiple predictors. The significant intercept indicates that the single predictor methods of predicting saddle height (trochanteric[18], ischial tuberosity[19], LeMond[4]) are likely oversimplified and inadequate for people with anthropometrics outside of those collected in the original study. This point is highlighted in work by Millour and colleagues who found that when using inseam based methods to set saddle height, those with short inseams (children) and those with longer than average inseams had knee flexion angles outside of the recommended range[20]. Both bicycle geometry (seat tube angle) and desired kinematics (minimum or maximum knee flexion) improve saddle height prescription. The fitted beta coefficients more closely align with findings from Ferrer-Roca and colleagues[11]. Ferrer-Roca and colleagues improved upon the single-predictor models by including knee flexion, though their equation was derived based on between-subject differences in a single bout of bicycling at the participant's habitual bicycle-fit. The current study used a fully-crossed random assignment of 18 bicycling postures to capture within-participant relationships.

Furthermore, the current study used machine-learning (LASSO regression) to identify the optimal combination of factors for robust estimates of saddle height.

Importantly, the equations derived from this study provide an accurate solution for positioning saddle height without the need for measuring knee kinematics. Measuring knee kinematics requires one to choose between highly accurate and costly approaches that require specialized training (e.g., motion capture camera system, video analysis) versus practical, low cost but less accurate approaches (e.g., static goniometry). There is currently no consensus on the best method. The most common practical approaches extract: (i) knee angle when the pedal is at the bottom dead centre (6 o'clock)[12,34], (ii) knee angle when the crank is aligned with the seat tube[11,20], and (iii) minimum knee angle of the revolution[35]. Methods (i) and (ii) can be executed both statically and dynamically. Method (i) is flawed because rotation of all bike components (seat and handlebars) about the bottom bracket results in the same geometry of the rider, though shifted relative to the bicycle. This approach is likely to conserve the overall joint kinematic pattern, though this pattern would likely occur at a different crank arm angle. Method (ii) accounts for gross bicycle geometry (seat tube angle) and measures joint angles when the pedals are theoretically furthest from the saddle. However, Method (ii) does not account for horizontal saddle adjustment [1], changes in handlebar position, and rider position on the saddle[5]. Method (iii) directly measures the smallest knee flexion angle in the revolution. In practical settings, use of Method (iii) is likely limited by access to technology.

In the minimum knee flexion equation (Equation 2-1), the positive relationship between the interaction of seat tube angle and inseam to saddle height is not intuitive. Individually, inseam and seat tube angle are positively related to saddle height (Figure 2-2). The interaction indicates that their individual effects are multiplicative. Yet, it is unclear how steeper seat tube angles predicted higher saddle heights. Insight can be provided by a study of 14 competitive triathletes, which found that changing seat tube angle did not affect ankle, knee or hip kinematics.[35] However, steeper seat tube angles resulted in greater anterior pelvic tilt, perhaps by re-orienting the body with respect to gravity. It is possible that changes to pelvic tilt in the sagittal plane moved the hip centre of rotation relative to the bottom bracket, requiring a higher saddle height for a given knee flexion angle and inseam.

The predictive ability of the minimum and maximum knee flexion models are similar. Therefore, we recommend the minimum knee flexion model because it is a standard in bicycling[4]. However, the maximum knee flexion equation may be relevant for patellofemoral ailments, where minimizing patellar compression is hypothesised to be achieved with higher saddle height and a smaller effective seat tube angle (more posterior saddle)[2,36]. However, a recent report indicated that seat tube angle had little/no effect on patellofemoral loading[9]. Higher saddle height does reduce knee flexion[34] throughout the pedal revolution, and under constant quadriceps force, patellofemoral joint reaction force theoretically increases proportional to knee flexion to a point[37]. Musculoskeletal modelling[9,30] and electromyography studies[15] show peak quadriceps forces and

activations occur when the pedal is between 9 and 3 o'clock, the same region in which graphs of knee flexion by crank arm angle show the knee is most flexed[38]. Recent work also shows that patellofemoral joint reaction forces are more sensitive to maximum than to minimum knee flexion angle[10]. Therefore, maximum knee flexion may ultimately prove to be more important than minimum knee flexion for patellofemoral ailments.

Sex did not modify bicycle saddle height predictive equations. This finding is in contrast to the only previous study to compare bicycle fit between men and women, which found that using existing inseam based equations based on men (1.77 (0.4) m)[11], men achieved knee flexion angles within the recommended range, while women had lower than recommended knee flexion[39]. However, the men (1.79 (0.065) m) were on average taller than the women (1.63 (0.049) m)[39]. In contrast, equations generated from the current study included men and women from 1.57 to 1.95 m tall. The women in this sample covered the 25th to >95th percentile heights of women over 20 years of age, and the men covered from <10th to > 95th percentile of heights (McDowell et al., 2008). More data is needed on shorter women (<1.57 m), however these equations should generalize to a wider range of the healthy adult population.

2.4.1 Limitations

Data were collected while participants bicycled using flat pedals with straps and running shoes, which may not be applicable to clipless systems. Participant anthropometrics were

obtained while participants were barefoot and shoe thickness may have added to the error in estimated saddle height. It is possible that some participants fatigued. The bicycling conditions accumulated to ~ 1 hour of cycling. To minimize the potential impact of fatigue on the findings, participants were encouraged to rest between bouts of cycling and the order of conditions was randomized. To enable maximum utility of these equations, future research should identify the optimal knee flexion angle (minimum or maximum) for specific ailments, such as patellofemoral pain.

2.5 Conclusions

The current investigation gives clinicians and bike shops a method to accurately prescribe bicycle saddle height with knowledge of the desired minimum (or maximum) knee angle, inseam length and bicycle seat tube angle. Separate equations were determined for minimum and maximum knee flexion angles. Due to comparable predictive ability between the minimum and maximum knee flexion angle equations, we suggest use of the minimum knee flexion angle equation as it is the standard in the bicycling community. The derived equations may be generalized to riders of both sexes.

Acknowledgements

We would like to thank Emily Wiebenga for help collecting data. Anthony A. Gatti was supported by an Ontario Graduate Scholarship and The Arthritis Society. Monica R. Maly is supported by The Arthritis Society Career Development Award, funded by the

Canadian Institutes of Health Research-Institution of Musculoskeletal Health and Arthritis. This research was funded by an NSERC Discovery grant (Monica R. Maly: 353715), the Canadian Foundation for Innovation Leaders Opportunity Fund, and the Ministry of Research and Innovation – Ontario Research Fund (Monica R. Maly & Peter J. Keir).

Disclosure of Interest

The authors report no conflict of interest.

2.6 References

1. De Bernardo N, Barrios C, Vera P, Laíz C, Hadala M. Incidence and risk for traumatic and overuse injuries in top-level road cyclists. *Journal of Sports Sciences*. 2012;30(10):1047-1053. doi:10.1080/02640414.2012.687112

2. Silberman MR. Bicycling Injuries: *Current Sports Medicine Reports*. 2013;12(5):337-345. doi:10.1249/JSR.0b013e3182a4bab7

3. Dettori NJ, Norvell DC. Non-traumatic bicycle injuries : a review of the literature. *Sports Med*. 2006;36(1):7-18.

4. Bini RR, Hume PA, Croft JL. Effects of bicycle saddle height on knee injury risk and cycling performance. *Sports Medicine*. 2011;41(6):463–476.

5. Bini RR, Hume PA, Lanferdini FJ, Vaz MA. Effects of moving forward or backward on the saddle on knee joint forces during cycling. *Physical Therapy in Sport*. 2013;14(1):23-27. doi:10.1016/j.ptsp.2012.02.003

6. Ericson MO, Nisell R. Patellofemoral joint forces during ergometric cycling. *Physical Therapy*. 1987;67(9):1365–1369.

7. Ericson MO, Nisell R. Tibiofemoral joint forces during ergometer cycling. *The American journal of sports medicine*. 1986;14(4):285–290.

8. Johnston TE, Baskins TA, Koppel RV, Oliver SA, Stieber DJ, Hoglund LT. The Influence of Extrinsic Factors on Knee Biomechanics During Cycling: A Systematic Review of the Literature. *International Journal of Sports Physical Therapy*. 2017;12(7):1023-1033. doi:10.26603/ijspt20171023

9. Menard M, Domalain M, Decatoire A, Lacouture P. Influence of saddle setback on knee joint forces in cycling. *Sports Biomechanics*. Published online June 19, 2018:1-13. doi:10.1080/14763141.2018.1466906

10. Gatti AA, Keir PJ, Noseworthy MD, Beauchamp MK, Maly MR. Hip and ankle kinematics are the most important predictors of knee joint loading during

bicycling. *Journal of Science and Medicine in Sport*. Published online July 2020. doi:10.1016/j.jsams.2020.07.001

11. Ferrer-Roca V, Roig A, Galilea P, García-López J. Influence of saddle height on lower limb kinematics in well-trained cyclists: Static vs. dynamic evaluation in bike fitting. *The Journal of Strength & Conditioning Research*. 2012;26(11):3025–3029.

12. Fonda B, Sarabon N, Li F-X. Validity and reliability of different kinematics methods used for bike fitting. *J Sports Sci*. 2014;32(10):940-946. doi:10.1080/02640414.2013.868919

13. Priego Quesada JI, Pérez-Soriano P, Lucas-Cuevas AG, Salvador Palmer R, Cibrián Ortiz de Anda RM. Effect of bike-fit in the perception of comfort, fatigue and pain. *Journal of Sports Sciences*. 2017;35(14):1459-1465. doi:10.1080/02640414.2016.1215496

14. Stratford PW. Getting More from the Literature: Estimating the Standard Error of Measurement from Reliability Studies. *Physiotherapy Canada*. 2004;56(01):027. doi:10.2310/6640.2004.15377

15. Jorge M, Hull ML. Analysis of EMG measurements during bicycle pedalling. *Journal of biomechanics*. 1986;19(9):683–694.

16. Krevolin JL, Pandy MG, Pearce JC. Moment arm of the patellar tendon in the human knee. *Journal of Biomechanics*. 2004;37(5):785-788. doi:10.1016/j.jbiomech.2003.09.010

17. Hamley EJ, Thomas V. Physiological and postural factors in the calibration of the bicycle ergometer. *J Physiol (Lond)*. 1967;191(2):55P-56P.

18. Nordeen-Snyder KS. The effect of bicycle seat height variation upon oxygen consumption and lower limb kinematics. *Med Sci Sports*. 1977;9(2):113-117.

19. Shennum PL, deVries HA. The effect of saddle height on oxygen consumption during bicycle ergometer work. *Med Sci Sports*. 1976;8(2):119-121.

20. Millour G, Duc S, Puel F, Bertucci W. Comparison of two static methods of saddle height adjustment for cyclists of different morphologies. *Sports Biomechanics*. Published online January 28, 2019:1-16. doi:10.1080/14763141.2018.1556324

21. Peveler WW, Pounders JD, Bishop PA. Effects of saddle height on anaerobic power production in cycling. *J Strength Cond Res*. 2007;21(4):1023-1027. doi:10.1519/R-20316.1

22. *Pre-Screening for Physical Activity: Get Active Questionnaire*. Canadian Society for Exercise Physiology https://store.csep.ca/pages/getactivequestionnaire

23. Wang Y-C, Hart DL, Stratford PW, Mioduski JE. Clinical interpretation of a lower-extremity functional scale–derived computerized adaptive test. *Physical therapy*. 2009;89(9):957–968.

24. Beekley MD, Brechue WF, Dehoyos DV, Garzarella L, Werber-Zion G, Pollock* ML. Cross-validation of the YMCA submaximal cycle ergometer test to predict VO2max. *Research quarterly for exercise and sport*. 2004;75(3):337–342.

25. Peveler WW. Effects of saddle height on economy in cycling. *The Journal of Strength & Conditioning Research*. 2008;22(4):1355–1359.

26. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*. 2001;37(1):153–156.

27. Sinclair J, Taylor PJ, Hobbs SJ. Digital Filtering of Three-Dimensional Lower Extremity Kinematics: an Assessment. *Journal of Human Kinetics*. 2013;39(1):25-36. doi:10.2478/hukin-2013-0065

28. Ehrig RM, Taylor WR, Duda GN, Heller MO. A survey of formal methods for determining the centre of rotation of ball joints. *Journal of Biomechanics*. 2006;39(15):2798-2809. doi:10.1016/j.jbiomech.2005.10.002

29. Harrington ME, Zavatsky AB, Lawson SEM, Yuan Z, Theologis TN. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on

magnetic resonance imaging. *Journal of Biomechanics*. 2007;40(3):595-602. doi:10.1016/j.jbiomech.2006.02.003

30. Lai AKM, Arnold AS, Wakeling JM. Why are Antagonist Muscles Co-activated in My Simulation? A Musculoskeletal Model for Analysing Human Locomotor Tasks. *Annals of Biomedical Engineering*. 2017;45(12):2762-2774. doi:10.1007/s10439-017-1920-7

31. Delp SL, Anderson FC, Arnold AS, et al. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering*. 2007;54(11):1940-1950. doi:10.1109/TBME.2007.901024

32. Gatti AA, Maly MR. Accuracy of estimates of cumulative load during a confined activity: bicycling. *International Biomechanics*. 2019;6(1):66-74. doi:10.1080/23335432.2019.1642141

33. Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer; 2009.

34. Bini RR, Hume PA, Kilding AE. Saddle height effects on pedal forces, joint mechanical work and kinematics of cyclists and triathletes. *European Journal of Sport Science*. 2014;14(1):44-52. doi:10.1080/17461391.2012.725105

35. Silder A, Gleason K, Thelen DG. Influence of Bicycle Seat Tube Angle and Hand Position on Lower Extremity Kinematics and Neuromuscular Control: Implications for Triathlon Running Performance. *Journal of Applied Biomechanics*. 2011;27(4):297-305. doi:10.1123/jab.27.4.297

36. Silberman MR, Webner D, Collina S, Shiple BJ. Road bicycle fit. *Clinical Journal of Sport Medicine*. 2005;15(4):271–276.

37. Loudon JK. Biomechanics and Pathomechanics of the Patellofemoral Joint. *IJSPT*. 2016;11(6):820-830.

38. Bini RR, Tamborindeguy AC, Mota CB. Effects of saddle height, pedaling cadence, and workload on joint kinetics and kinematics during cycling. *J Sport Rehabil*. 2010;19(3):301-314.

39. Encarnación-Martínez A, Ferrer-Roca V, García-López J. Influence of Sex on Current Methods of Adjusting Saddle Height in Indoor Cycling: *Journal of Strength and Conditioning Research*. Published online June 2018:1. doi:10.1519/JSC.0000000000002689

2.7 Supplemental

2.7.1 Supplemental 1

Participants completed 18 bicycle conditions. The conditions included all possible combinations of 3 saddle X (horizontal) positions, 3 saddle Y (vertical) positions, and 2 crank arm lengths.

Saddle X	**Saddle Y**	**Crank Arm Length**
Commercial	Commercial	Commercial (172.5 mm)
Random X position 1 within +/- 10% of commercial	Random Y position 1 within +/- 5% of commercial	Random crank arm position (170 or 175 mm)
Random X position 2 within +/- 10% of commercial	Random Y position 2 within +/- 5% of commercial	

The 18 positions ensured that participants completed all possible combinations of each level (3 x 3 x 2 = 18). Random positions for X and Y were random independent samples and were completed independently for each participant. In particular, each mm increment within the defined range (e.g., for the saddle X position: from commercial – 10% to commercial + 10%) had equal probabilities of being sampled.

Once the random Saddle X, Saddle Y, and crank arm lengths were sampled, their orders within each group (Saddle X, Saddle Y, and crank arm) were randomized. The following table is an example of one possible random order of the positions:

Position	**Saddle X**	**Random Order Saddle X**
1	Commercial	Random X position 2 within +/- 10% of commercial
2	Random X position 1 within +/- 10% of commercial	Random X position 1 within +/- 10% of commercial
3	Random X position 2 within +/- 10% of commercial	Commercial

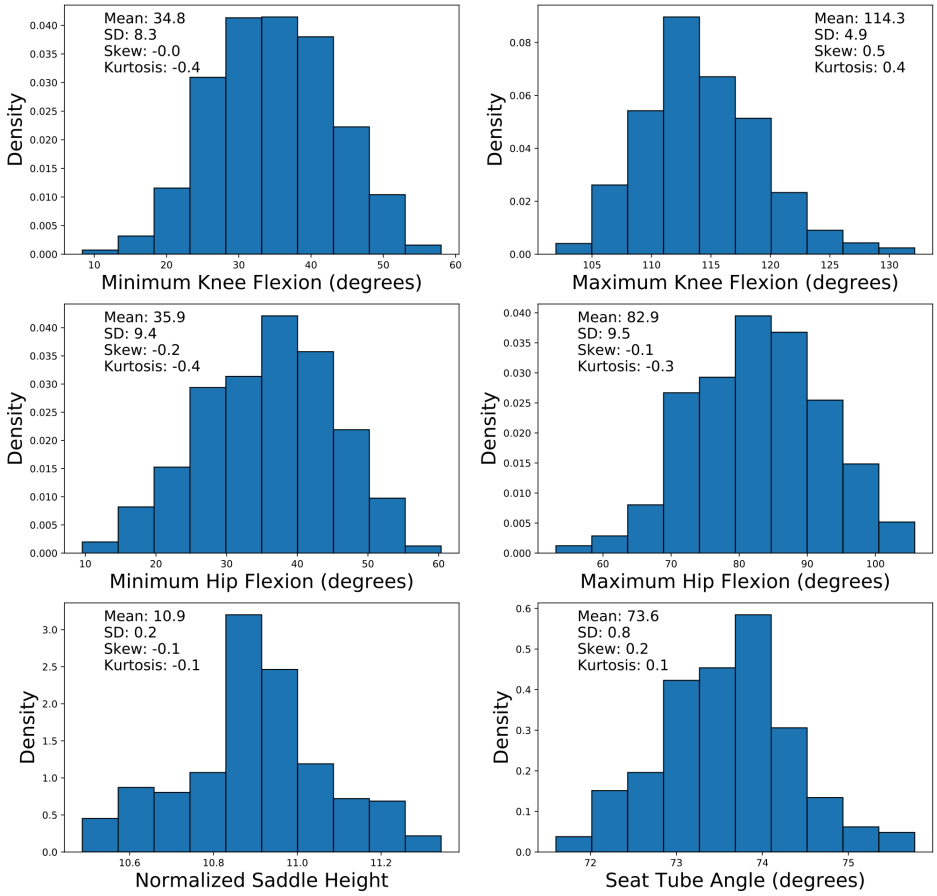
After re-ordering, the positions (commercial or random) were equally likely to be in any of the positions (1, 2, or 3 for Saddle X and Saddle Y, and 1 or 2 for Crank arm length).

The following table provides the randomly ordered positions as completed in this study. All trials that included the re-ordered crank arm position 1 were completed first; then all of the re-ordered crank arm position 2 were completed second. Blocking crank arm lengths (as well as Saddle X and Saddle Y) was completed for feasibility.

Collection number	**Crank arm length**	**Saddle X**	**Saddle Y**
1	Position 1	Position 1	Position 1
2	Position 1	Position 1	Position 2
3	Position 1	Position 1	Position 3
4	Position 1	Position 2	Position 1
5	Position 1	Position 2	Position 2
6	Position 1	Position 2	Position 3
7	Position 1	Position 3	Position 1
8	Position 1	Position 3	Position 2
9	Position 1	Position 3	Position 3
10	Position 2	Position 1	Position 1
11	Position 2	Position 1	Position 2
12	Position 2	Position 1	Position 3
13	Position 2	Position 2	Position 1
14	Position 2	Position 2	Position 2
15	Position 2	Position 2	Position 3
16	Position 2	Position 3	Position 1
17	Position 2	Position 3	Position 2
18	Position 2	Position 3	Position 3

2.7.2 Supplemental 2

Histograms of minimum and maximum knee and hip flexion, as well as normalized saddle height (normalized saddle height = saddle height / inseam) and seat tube angle for all bicycling postures of all participants. The histograms demonstrate that the breadth of recommended bicycling postures were encapsulated by the study randomization protocol. Each figure includes the mean, standard deviation (SD), skew, and kurtosis of the distribution. Density is the normalized height of each bin such that the total area of all bins sums to 1.0.



3 Hip and ankle kinematics are the most important predictors of knee joint loading during bicycling

This paper was accepted by the Journal of Science and Medicine in Sport and is currently in-press (https://doi.org/10.1016/j.jsams.2020.07.001). Included is the final submitted version of the manuscript.

Hip and ankle kinematics are the most important predictors of knee joint loading during bicycling

Journal of Science and Medicine in Sport

[A]Anthony A. Gatti, MSc, [B]Peter J. Keir, PhD, [C,D]Michael D. Noseworthy, PhD, [A,E]Marla K. Beauchamp, PT PhD, [A,B,F]Monica R. Maly, PT PhD

McMaster University
1280 Main St. W.
Hamilton, ON, Canada,
L8S 3R8
[A]School of Rehabilitation Sciences, [B]Department of Kinesiology, [C]School of Biomedical Engineering, [D]School of Electrical and Computer Engineering, and [E]Department of Medicine
University of Waterloo
200 University Ave,
Waterloo, ON, Canada,
N2L 3G1
[F]Department of Kinesiology

Corresponding Author:
Monica R. Maly
University of Waterloo
200 University Ave,
Waterloo, Ontario, Canada,
N2L 3G1
Fax: +1 519 885 4070
Email: mrmaly@uwaterloo.ca

Word count: 3000
Abstract word count: 249
Number of tables: 1
Number of figures: 2

Abstract

Objectives: To assess the effect of ankle, knee, and hip kinematics on patellofemoral and tibiofemoral joint reaction forces (JRF) during bicycling. Secondarily, to assess if sex, horizontal saddle position, or crank arm length were related to JRFs, after accounting for kinematics.

Design: Experimental cross-sectional study.

Methods: Forty healthy adults (mean (SD); 28.6 (7.2) years, 24.2 (2.6) kg/m^2, 17 women) bicycled under 18 bicycling positions. One position used commercial guidelines and 17 randomly deviated from commercial. Resultant patellofemoral as well as compressive and shear tibiofemoral JRFs were calculated. Linear mixed-effects models with a random intercept of leg-nested-in-participant were used.

Results: Patellofemoral resultant forces were most sensitive to all joint kinematics (i.e., sensitivity was defined as the slope of single predictor models); all JRFs were least sensitive to minimum knee flexion. Tibiofemoral compression was predicted by minimum hip flexion and maximum ankle dorsiflexion (R^2=0.90). Tibiofemoral shear (R^2=0.86) and the resultant patellofemoral JRF (R^2=0.90) were predicted by minimum hip flexion, maximum ankle dorsiflexion, minimum knee flexion, and the interaction between minimum hip flexion and minimum knee flexion. Adding sex as a factor improved fit of all models. This sex-effect was driven by differences in cycling intensity, reflected by the tangential crank arm force. Horizontal saddle position and crank arm length were not related to JRFs.

Conclusions: Optimizing joint kinematics should be the primary goal of bicycle-fit. JRFs were least sensitive to the current gold standard for assessing bicycle-fit, minimum knee flexion. Bicycle-fit is of particular importance for those working at high intensities.

Keywords: Biomechanical Phenomena; Injuries; Knee; Musculoskeletal Modeling

Practical Implications

- The amount of flexion at the hip and dorsiflexion at the ankle were most important for predicting forces inside of the knee.
- Men had higher forces inside of their knee than women; these differences were attributed to men (on average) working at a higher intensity.
- The effect of intensity indicates that bicycle-fit is likely more important for individuals cycling at higher intensities.
- Bicycle crank arm length and horizontal alignment of the bicycle saddle did not affect forces inside of the knee.

3.1 Introduction

More than 40% of Canadians report having bicycled in the past 12-months[1]. Bicycling is utilized for weight management[2] and rehabilitation of many chronic conditions, including knee osteoarthritis because the seat off-loads lower extremity joints[3]. However, >51% of cyclists report overuse injuries[4], 1/3 of which occur at the knee[5]. Anterior knee pain is the most common reason cyclists seek medical care[4].

Common bicycle-fit aims to prevent and alleviate knee injuries with two primary recommendations. The first is to set the saddle position to elicit minimum knee flexion of 25-40°. Higher knee flexion is hypothesized to minimize iliotibial-band and hamstrings injuries, while lower knee flexion is hypothesized to minimize patellofemoral/anterior knee pain[4,6]. The second recommendation is to set the horizontal alignment of the saddle using the knee-over-pedal-spindle (KOPS) methodology[4,6]. KOPS recommends the anterior knee be at, or behind, the pedal spindle when the pedal is at 3 o'clock[4].

Only a small body of literature has tested the knee flexion and KOPS approaches. A mathematical model comparing three bicycle saddle heights (commercial recommendation, +3%, – 3%), using data from nine men, showed saddle height did not affect tibiofemoral and patellofemoral joint reaction forces (JRF)[7]. Menard and colleagues used a musculoskeletal model to investigate the effect of saddle horizontal positions (recommended, – 10%, + 10%) on tibiofemoral and patellofemoral joint loading

in ten cyclists[8]. Saddle setback did not affect patellofemoral forces, but a posterior saddle position increased tibiofemoral compression. These studies included only males, compared discrete positions that do not reflect the continuous relationship between bicycle-fit and joint loading, and only assessed a single facet of saddle position, either saddle height or set-back. Complicating things further, the saddle height axis (straight line from bottom bracket to top of saddle) and set-back axis (horizontal) are not perpendicular and are therefore inter-related. Thus, the impact of each of saddle height, set-back, and their interaction on knee JRFs remains unclear.

Joint kinematics likely drive the relationship between saddle position and joint loading. For example, lower saddle height increases knee flexion[9]. Greater knee flexion increases patellofemoral reaction forces (given an equivalent quadriceps force)[10]. Making matters worse, in the knee flexion range observed during cycling[11], higher knee flexion decreases the quadriceps moment arm thereby requiring higher muscle forces to produce a given extension moment[12]. Maximum knee flexion may be particularly important to patellofemoral JRFs because peak quadriceps forces and activations occur while the pedal is near the top of the revolution[8], where the knee is most flexed[13]. Cartilage surface pressures are likely more appropriate measures of joint loading than JRFs as joint congruence changes with knee flexion and may therefore dissipate a given JRF over a larger surface area[10]. Nevertheless, no work has investigated which kinematic outcomes are most important for predicting knee JRFs.

The JRFs may also be influenced by crank arm length, that is, the radius around which the bicycle pedals rotate. Research to-date has produced conflicting results, where one study showed that longer crank arms decreased knee extension moment and power, while increasing hip extension power[14]; and another study reported that increased crank arm length increased knee moments[15,16].

The purpose of this study was to assess the effect of ankle, knee, and hip kinematics on patellofemoral and tibiofemoral JRFs during bicycling in healthy adults. The secondary purpose was to assess if sex, saddle horizontal position, or crank arm length further explained variance in JRFs, after accounting for joint kinematics. We hypothesized that higher ankle, knee and hip flexion would be positively related to patellofemoral and tibiofemoral reaction forces. Since we believed that JRFs would be driven primarily by differences in kinematics, we hypothesized that sex, saddle horizontal position, and crank arm length would not explain additional variance in any JRFs.

3.2 Methods

This cross-sectional study employed a fully-crossed design, using random assignment of 18 bicycle postures. All participants provided written, informed consent. This study was approved by the Hamilton Integrated Research Ethics Board.

Healthy adults (18-45 years) with a range of cycling experience from recreational urban cycling to regional age-group level athletes in mountain biking, cyclocross, and triathlon were recruited. Participants were excluded if it was unsafe to exercise based on the Get Active Questionnaire (GAQ)[17], or if they self-reported injury to the lower extremity during the past 3-months, rheumatoid or osteoarthritis, gout, unstable angina, or neurologic conditions. Descriptive variables included the Lower Extremity Functional Scale[18], age, sex, height, and body mass index (BMI).

Participants attended one visit and donned shorts, running shoes, and were either shirtless or wore a sports bra. The protocol included three steps: (i) commercial bicycle-fit, (ii) incremental cycling protocol to determine the experimental exercise intensity, and (iii) experimental bicycling protocol including 18 bouts of three-minute bicycling.

To scale a musculoskeletal model and track kinematics at the ankle, knee, and hip, participants were outfitted with 40 markers attached to anatomical locations (Supplemental 1). Four markers were located on each bicycle pedal. Marker data were collected at 112.5 Hz using 12 high-speed infrared cameras (Motion Analysis Corporation, Santa Rosa, CA). Synchronous pedal kinetics were collected at 450 Hz using pedals that measure three-dimensional forces and moments (Science To Practice, Ljubljana, Slovenia). The pedals can be outfitted with multiple clip interfaces; in the current study a flat pedal with Velcro straps was utilized. Marker positions and pedal

kinetics were filtered using a second-order low pass dual pass Butterworth filter at a cut-off frequency of 6 Hz.

Functional knee joint centres were calculated from one bicycling trial using the symmetric centre of rotation estimation method[19]. Hip joint centres were calculated using the Harrington method[20]. A 16 degree-of-freedom lower body musculoskeletal model actuated by 86 musculotendon actuators was scaled to the anatomical markers using OpenSim[21]. Joint kinematics were calculated by minimizing the squared error between experimental and model markers. Ankle knee and hip moments were calculated using inverse dynamics. Kinematic outcomes (minimum and maximum angles) were extracted from each revolution in the final minute[22]; outcomes were calculated as the mean from all revolutions. Intra- and inter-rater standard error of measurement for kinematic waveforms was <2.6°, and for lower extremity joint moments was <0.08 Nm/kg[23].

Static optimization that accounts for muscle coordination while minimizing the sum of muscle activations squared was used to determine the muscle forces needed to produce joint moments from inverse dynamics[24]. The muscle coordination schema was based on weighting constants that minimize the error compared to *in vivo* JRFs from an instrumented total knee replacement[24]. Supplemental 2 provides the static optimization protocol and a comparison to *in vivo* JRFs, during bicycling, from an instrumented knee replacement[25].

The primary outcomes, patellofemoral and tibiofemoral JRFs, were calculated using the Joint Reaction tool in OpenSim[24]. This tool calculates the resultant force not accounted for in the model, thus incorporating bone-on-bone and ligamentous forces. Because the model defines transverse and coronal plane knee kinematics as a function of knee flexion, only sagittal plane forces were included. The sagittal plane resultant of the patellofemoral JRF was defined in the patellar coordinate system (Supplemental 1). Tibiofemoral JRFs were defined in the tibial coordinate system[24] (Supplemental 1) and include a compressive and anterior/posterior shear component. The tibiofemoral compressive component is primarily comprised of "bone-on-bone" forces, where compression of the femur into the tibia is negative. Tibiofemoral shear forces primarily account for forces by ligamentous structures and patellofemoral reaction forces[24]; posterior forces on the femur were positive.

The reaction forces were extracted using all revolutions from the last minute of each bout[22]. All data were divided into individual pedal revolutions starting and ending with the pedal at 12 o'clock and were normalized to 101 points. A median ensemble curve was generated for each condition and reaction force. The peak of each ensemble curve was the primary outcome.

The commercial bicycle-fit included setting the saddle position to 109% of inseam[6,26] with an effective seat tube angle of 73.7° (mean of 2018 Trek road / city bicycles), and a crank arm length of 172.5mm on a commercial fit-bike (Fit-bike Pro 1; Purely Custom, USA). Handlebars were positioned based on comfort, after initially being set to level with the saddle, and at a horizontal distance that achieved 70-90° between the torso and upper arms.

Bicycling intensity of the experimental bouts, in cadence (revolutions per minute, RPM) and in power (watts, W), was determined during an incremental bicycling protocol based on the YMCA submaximal VO_2 test[27]. Participants initially cycled at a resistance of ~four newton-metres at a self-selected cadence (30-45W, dependent on cadence). Every three minutes, resistance was increased by one gear (4^{th} order exponential). Participant heart rate (HR) was recorded every minute (Polar T31, Kempele, Finland). If HR differed by more than five beats between consecutive measurements, the stage was extended by one minute until steady-state was achieved. The protocol continued until HR was within 70-75% of age-predicted maximum (208-0.7*Age)[28]. A HR range of 70-75% of maximum was chosen to elicit a moderate intensity that improved function and symptoms in a sample with knee osteoarthritis[3,29]. The average power output and cadence of the final stage were used for the experimental bicycling bouts.

Each participant completed 18 three-minute bicycling bouts, each at the cadence and power derived from the incremental protocol, in random order. Bicycle positions included all combinations of three vertical saddle positions (Y), three horizontal saddle positions (X), and two crank arm lengths ($3 \times 3 \times 2 = 18$). The same bicycle saddle (Selle Royal, Seta, 143 mm width) was used for all positions. The three saddle X and three saddle Y positions each included one position that was based on a commercial bicycle-fit; and the other two were random independent samples within +/-10% (X) and +/-5% (Y) of commercial. Y deviations only ranged 5% as larger deviations were not anatomically possible. The commercial crank arm length was 172.5 mm, the second crank arm length was randomly chosen to be either 2.5 mm longer (175 mm) or shorter (170 mm) than commercial. In each position, handlebars were moved to maintain the same position relative to the saddle. Randomization of positioning was performed independently for each participant (Supplemental 3). Randomized positions were used to achieve sampling of the continuous relationship between factors.

To answer the primary question, mixed-effects models were used. First, to identify the independent effect of kinematic variables on joint loading, six models that included only one predictor (maximum and minimums of knee flexion/extension, hip flexion/extension, ankle dorsi/plantar flexion) were generated for each of the three kinetic outcomes (patellofemoral resultant, tibiofemoral compressive and shear forces). The beta coefficients (slope) between each kinematic and kinetic measurement represent the sensitivity of the kinetic measurement to independent changes in the kinematic

measurement expressed in N/°. Second, the most parsimonious models that predicted the kinetic outcomes were generated using forward selection where a predictor was deemed to improve a model if it improved model fit measured using the likelihood ratio test ($p<0.05$) and had a significant beta coefficient ($p<0.05$). If multiple predictors improved the model, the one with the highest model log likelihood was selected.

To answer the secondary research questions, each of the three secondary-predictors of interest (sex, normalized saddle horizontal position, and crank arm length) were independently added to the most parsimonious models. Interactions between secondary-predictors and base-predictors were tested using forward selection. Model comparisons were conducted using the likelihood ratio test.

All models included data from both legs. To satisfy assumptions, all models used a random intercept with random effects of legs-nested-within-participants. A rounded assessment of goodness-of-fit for mixed models was provided using the marginal (fixed factors) as well as conditional (fixed and random factors) coefficients of determination (R^2)[30]. Analyses were performed using the StatsModels package (Python 3.7).

3.3 Results

Forty-one adults (age: 28.6 (7.2) years, BMI: 24.2 (2.6) kg/m^2, height: 1.74 (0.09) m, 17 women) completed the study. Analyses were conducted on 40 participants; data from one

participant (male, 38 years, 27.9 kg/m^2) was excluded due to error in recording saddle positions. Women were on average shorter than men (1.66 (0.07) versus 1.80 (0.08) m, $p<0.001$), with no difference in BMI (23.7 (3.1) versus 24.5 (2.2) kg/m^2, $p=0.345$). Participants had Lower Extremity Functional Scale scores reflecting excellent function (mean 79.2, SD 1.9)[18]. Participants bicycled with a breadth of power outputs (mean 117.4 W, SD 43.8), and narrower range of cadences (mean 83.7 RPM, SD 6.7). Men cycled at systematically higher cadences (85.5 (6.6) versus 81.2 (6.1) RPM, $p=0.040$) and powers (140 .4 (37.7) versus 86.4 (30.6), $p<0.001$). Supplemental 4 shows the distributions of each kinematic outcome over bicycling bouts.

The sensitivity of the three JRF outcomes to each kinematic measurement, represented by the single predictor model beta coefficients, is presented in Figure 3-1. The resultant patellofemoral JRF was most sensitive to joint kinematics, while the tibiofemoral shear force was least sensitive. All reaction forces were least sensitive to changes in minimum knee flexion.

The best kinematic models for predicting each of the JRFs are presented in Table 3-1 and visualized in Figure 3-2. For the compressive component of the tibiofemoral reaction force, only minimum hip and maximum ankle dorsiflexion were included (marginal R^2: 0.14, conditional R^2: 0.90). For the shear component of the tibiofemoral reaction force (marginal R^2: 0.11, conditional R^2: 0.86), and the resultant patellofemoral JRF (marginal

R^2: 0.15, conditional R^2: 0.90), minimum hip flexion, maximum ankle dorsiflexion, minimum knee flexion, and the interaction of minimum knee flexion and minimum hip flexion were included.

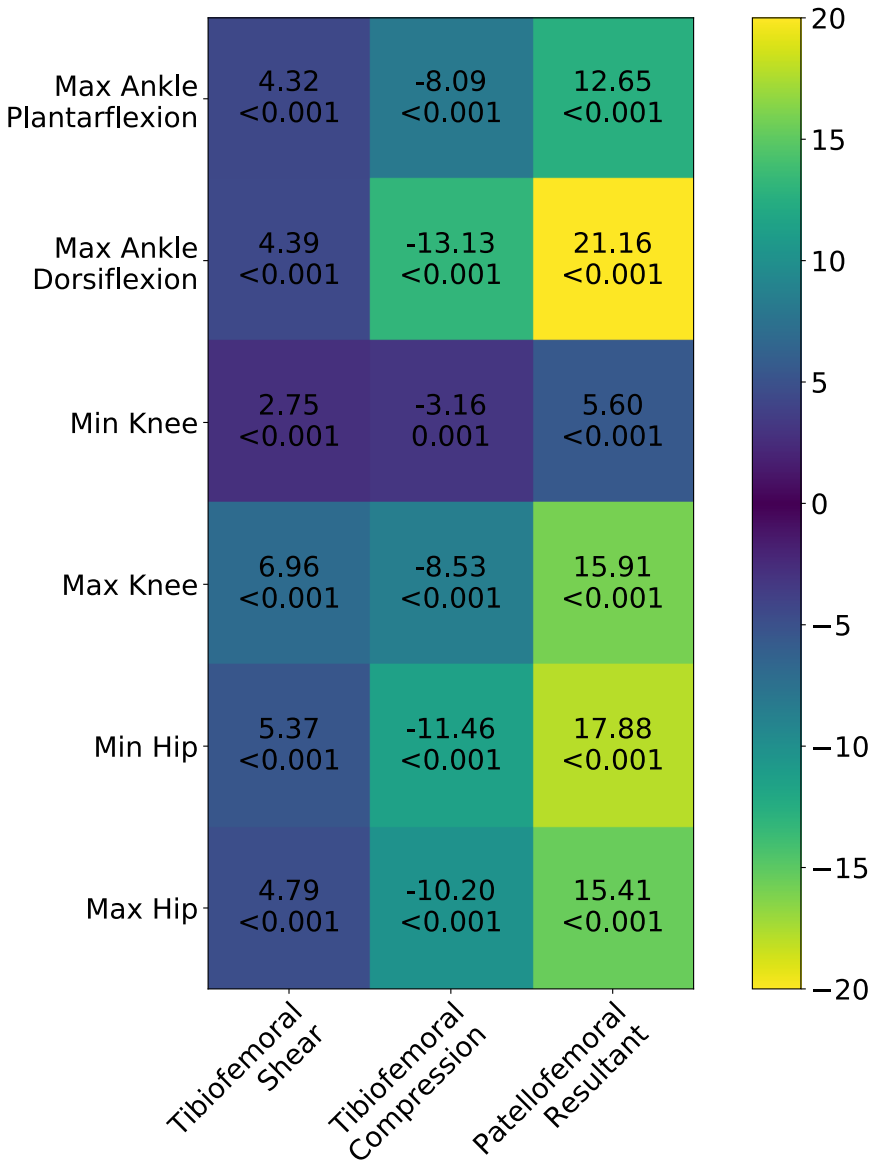


Figure 3-1. Visual representation of the independent relationship between each kinematic outcome and the three joint reaction force outcomes. Each square includes the beta coefficient from a single predictor mixed effects model on top and the significance of that beta coefficient on the bottom. The beta coefficients can be interpreted as the change in newtons of the coinciding joint reaction force for a 1° increase in the kinematic outcome of interest. The cells are coloured according to the absolute magnitude of the beta coefficient. More yellow (brighter) indicates higher sensitivity of joint reaction forces to the kinematic outcome, more purple (darker) indicates lower sensitivity.

For the secondary analyses, crank arm length and saddle horizontal alignment did not add to any model (Supplemental 5). Sex contributed to every model (Table 3-1, Figure 3-2). The sex-effect was likely driven by the differences in power and cadence between the sexes.

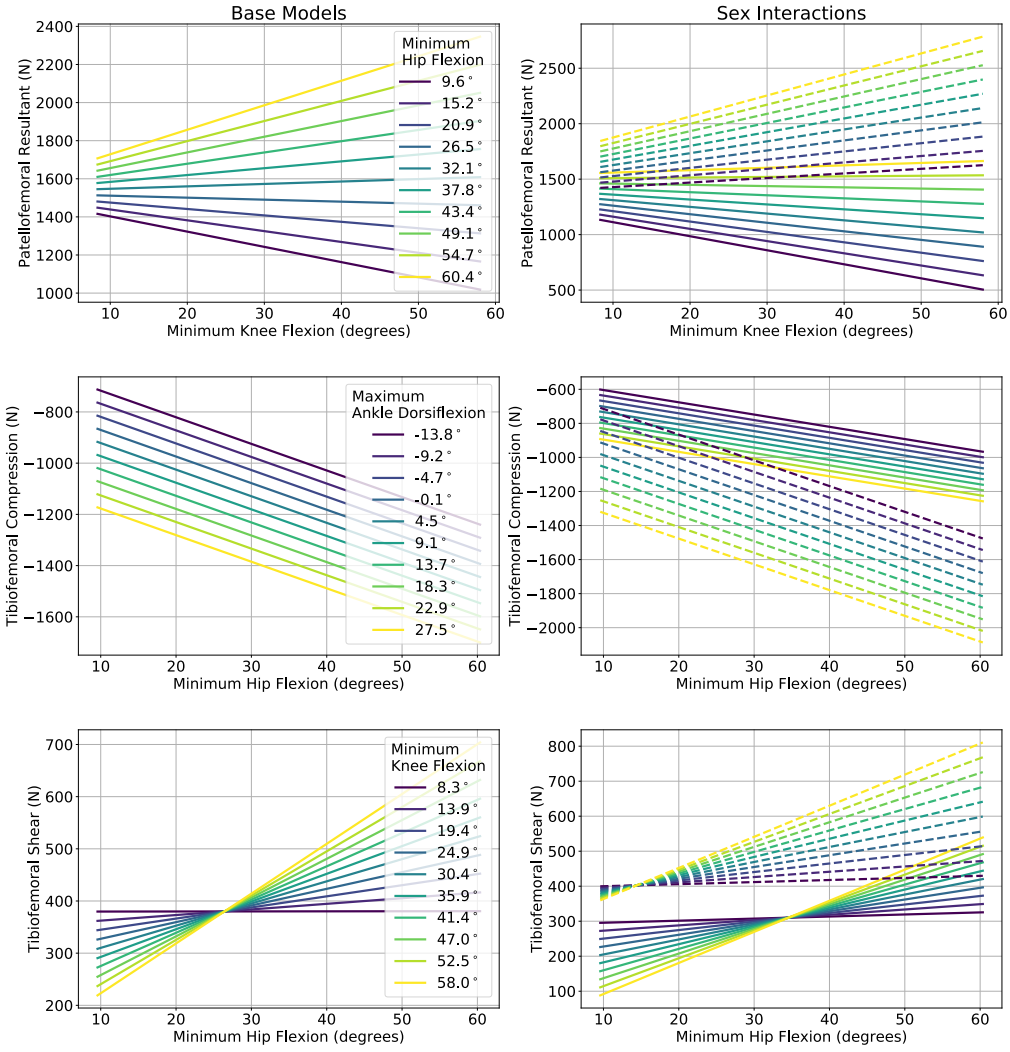


Figure 3-2. Visual representation of the base models (left column) as well as the sex interactions (right column) for predicting joint reaction forces. All sex models (right column) follow their equivalent base model (left column) legend. In the sex interaction plots (right column) solid lines represent women and dashed lines represent men.

Because we hypothesized the sex effects were driven by intensity differences, base models were re-run including the crank arm tangential force ("tangential force" = $\frac{power\ (w)}{cadence(revolutions\ per\ second\ )\times\ 2\ \pi\ crank\ arm\ length\ (m)}$) as a covariate (mean difference 31.3 N, $p<0.001$). For the tibiofemoral compressive force, the tangential force, and the interaction of the tangential force with each of minimum hip and maximum ankle dorsiflexion were included (marginal R^2: 0.45, conditional R^2: 0.90; Supplemental 6). For the tibiofemoral shear force (marginal R^2: 0.33, conditional R^2: 0.87), and the patellofemoral JRF (marginal R^2: 0.45, conditional R^2: 0.90), tangential force, and the interaction of tangential force with minimum knee flexion as well as maximum ankle dorsiflexion were included (Supplemental 6). Interactions resulted in a multiplicative effect.

Table 3-1. Results from the base models and the base plus sex models used to predict resultant patellofemoral, as well as compressive and shear tibiofemoral forces. Bolded p-values indicate significance at $p<0.05$.

Model	Predictor	Beta Coefficient	Standard error	p-value
Base Models				
Patellofemoral Resultant Marginal R^2: 0.15 Conditional R^2: 0.90	Intercept	1307.41	158.22	**<0.001**
	Minimum hip flexion	2.31	3.56	**0.516**
	Maximum ankle dorsiflexion	18.21	2.05	**<0.001**
	Minimum knee flexion	-11.94	3.98	**0.003**
	Minimum hip flexion X minimum knee flexion	0.41	0.10	**<0.001**
Tibiofemoral Compression Marginal R^2: 0.14 Conditional R^2: 0.90	Intercept	-767.20	66.40	**<0.001**
	Minimum hip flexion	-10.39	0.98	**<0.001**
	Maximum ankle dorsiflexion	-11.14	1.39	**<0.001**
Tibiofemoral Shear Marginal R^2: 0.11 Conditional R^2: 0.86	Intercept	391.73	45.28	**<0.001**
	Minimum hip flexion	-1.59	1.07	0.139
	Maximum ankle dorsiflexion	3.60	0.62	**<0.001**
	Minimum knee flexion	-5.07	1.20	**<0.001**
	Minimum knee flexion X minimum hip flexion	0.19	0.03	**<0.001**
Base Models + Sex				
Patellofemoral Resultant Marginal R^2: 0.46 Conditional R^2: 0.87	Intercept	1082.65	163.74	**<0.001**
	Minimum hip flexion	5.89	3.53	0.096
	Maximum ankle dorsiflexion	11.84	2.74	**<0.001**
	Minimum knee flexion	-15.45	3.97	**<0.001**
	Minimum knee flexion X minimum hip dorsiflexion	0.29	0.10	**0.005**
	Sex	38.43	162.80	0.813
	Sex X minimum knee flexion	16.78	2.64	**<0.001**
	Sex X maximum ankle dorsiflexion	13.05	3.29	**0.001**
Tibiofemoral Compression Marginal R^2: 0.47 Conditional R^2: 0.90	Intercept	-629.88	75.56	**<0.001**
	Minimum hip flexion	-7.19	1.24	**<0.001**
	Maximum ankle dorsiflexion	-7.04	1.88	**<0.001**
	Sex	-138.45	107.71	0.199
	Sex X minimum hip flexion	-7.92	1.95	**<0.001**
	Sex X maximum ankle dorsiflexion	-7.74	2.72	**0.004**
Tibiofemoral Shear Marginal R^2: 0.43 Conditional R^2: 0.87	Intercept	322.22	45.42	**<0.001**
	Minimum hip flexion	-0.80	1.08	0.456
	Maximum ankle dorsiflexion	1.86	0.83	**0.025**
	Minimum knee flexion	-5.77	1.20	**<0.001**
	Minimum knee flexion X minimum hip flexion	0.17	0.03	**<0.001**
	Sex	49.06	41.51	0.237
	Sex X minimum knee flexion	3.37	0.80	**<0.001**
	Sex X maximum ankle dorsiflexion	3.24	1.19	**0.007**

3.4 Discussion

This study is the first to provide empirical data showing that appropriate bicycle-fit should aim to optimize joint kinematics at the hip, ankle and knee. Higher flexion at any of the lower extremity joints resulted in higher JRFs for both the tibiofemoral and, to a larger degree, patellofemoral joint. After accounting for kinematics, saddle horizontal alignment and crank arm length were not related to JRFs. Further, sex improved prediction of all JRFs, though this effect was explained by sex-differences in bicycling intensity, reflected by crank arm tangential forces. These findings are particularly important for individuals cycling at higher intensities.

Knee JRFs were most sensitive to hip flexion and to ankle dorsiflexion, suggesting these are important targets during bicycle-fit. Ankle dorsiflexion may be the most sensitive because it reflects how flexed the lower extremities were as a whole. Bi-articular gastrocnemius muscles may also play an important role. Gastrocnemius are activated during the downstroke[31,32]. However, the gastrocnemius flexes the knee, and there is a net knee extension moment during this same phase[31]; therefore, increased gastrocnemius activation requires a concomitant increase in knee extensors to maintain the same net knee extension moment.

Hip, knee and ankle kinematics have more important influences on JRFs than saddle position or crank arm length. After accounting for joint kinematics, saddle horizontal

position and crank arm length did not add to the models. In contrast, Menard and colleagues found that moving the saddle posteriorly increased tibiofemoral compression[8]. The contradictory findings are likely attributed to inclusion of joint kinematics. Knee and hip flexion increase with increasing crank arm length[14], potentially explaining why inclusion of kinematics makes crank arm length unimportant.

Sex was a predictor in all models. This relationship reflected differences in crank arm tangential force between the sexes. Similarly, adults with force-measuring joint replacements that produced higher power outputs or lower cadences (which both increase tangential crank arm forces) demonstrated larger tibiofemoral JRFs[25]. Not only do JRFs increase with increasing tangential forces, but they have a multiplicative effect, whereby a given change in kinematics has a more pronounced effect on JRFs when the crank arm tangential force is higher. The magnitude of change highlights the importance of bicycle-fit for individuals bicycling at high intensities.

Higher flexion at any of the hip, knee or ankle increased both patellofemoral and tibiofemoral JRFs, a finding that supports the recommendation of a higher saddle height to reduce joint loading[4]. However, a saddle height that is too high may elicit iliotibial band impingement[4] around 30° of knee flexion[33,34] or cause the pelvis to rock from side-to-side [4]. Therefore, when aiming to reduce patellofemoral and/or tibiofemoral loading (e.g., patellofemoral and tibiofemoral osteoarthritis, anterior knee pain) the highest

feasible saddle height is recommended. A good upper limit would be a minimum knee flexion at 30°. Another metric commonly monitored is excessive lateral pelvic tilt (i.e., pelvis rocking side-to-side)[4].

This study is not without limitations. JRFs were calculated based on assumptions of physiologic parameters. However, comparisons to *in vivo* measurements during bicycling show strong agreement (Supplemental 2), particularly for peak forces. Participants were healthy adults between 18-44 years old who completed all bicycling bouts using flat pedals with a Velcro strap and running shoes and therefore findings may not generalize to other samples, or to clipless pedals and rigid soled cycling shoes. Finally, one saddle was used. It is possible that different individuals require different saddle shapes or sizes. Future studies should explore clipless pedals and compare different saddles.

3.5 Conclusion

Higher lower extremity joint flexion increased tibiofemoral and, to a greater degree, patellofemoral JRFs. These findings highlight that kinematics of all lower-body joints should be considered during bicycle-fit. After accounting for joint kinematics, saddle horizontal alignment and crank arm length were not related to joint loading, showing that bicycle-fit should continue to focus on joint kinematics. Changes in JRFs were more sensitive to kinematics in men compared to women. This sex-effect reflects differences in

bicycling intensity, indicating that changes in bicycle-fit are more important for those producing higher power output.

Practical Implications

- The amount of flexion at the hip and dorsiflexion at the ankle were most important for predicting forces inside of the knee.
- Men had higher forces inside of their knee than women; these differences were attributed to men on average working at a higher intensity.
- The effect of intensity indicates that bicycle-fit is more important for individuals cycling at higher intensities.
- Bicycle crank arm length and horizontal alignment of the bicycle saddle did not affect forces inside of the knee.

3.6 References

1 Ramage-Morin PL. *Statistics Canada, Health Reports, Cycling in Canada*. 2017.

2 Bianco A, Bellafiore M, Battaglia G, et al. The effects of indoor cycling training in sedentary overweight women. *J Sports Med Phys Fitness* 2010; 50(2):159–165.

3 Salacinski AJ, Krohn K, Lewis SF, et al. The Effects of Group Cycling on Gait and Pain-Related Disability in Individuals With Mild-to-Moderate Knee Osteoarthritis: A Randomized Controlled Trial. *J Orthop Sports Phys Ther* 2012; 42(12):985–995. Doi: 10.2519/jospt.2012.3813.

4 Silberman MR. Bicycling Injuries: *Curr Sports Med Rep* 2013; 12(5):337–345. Doi: 10.1249/JSR.0b013e3182a4bab7.

5 De Bernardo N, Barrios C, Vera P, et al. Incidence and risk for traumatic and overuse injuries in top-level road cyclists. *J Sports Sci* 2012; 30(10):1047–1053. Doi: 10.1080/02640414.2012.687112.

6 Bini RR, Hume PA, Croft JL. Effects of bicycle saddle height on knee injury risk and cycling performance. *Sports Med* 2011; 41(6):463–476.

7 Tamborindeguy AC, Rico Bini R. Does saddle height affect patellofemoral and tibiofemoral forces during bicycling for rehabilitation? *J Bodyw Mov Ther* 2011; 15(2):186–191. Doi: 10.1016/j.jbmt.2009.07.009.

8 Menard M, Domalain M, Decatoire A, et al. Influence of saddle setback on knee joint forces in cycling. *Sports Biomech* 2018:1–13. Doi: 10.1080/14763141.2018.1466906.

9 Bini RR, Hume PA, Kilding AE. Saddle height effects on pedal forces, joint mechanical work and kinematics of cyclists and triathletes. *Eur J Sport Sci* 2014; 14(1):44–52. Doi: 10.1080/17461391.2012.725105.

10 Loudon JK. Biomechanics and Pathomechanics of the Patellofemoral Joint. *IJSPT* 2016; 11(6):820–830.

11 Silder A, Gleason K, Thelen DG. Influence of Bicycle Seat Tube Angle and Hand Position on Lower Extremity Kinematics and Neuromuscular Control: Implications for Triathlon Running Performance. *J Appl Biomech* 2011; 27(4):297–305. Doi: 10.1123/jab.27.4.297.

12 Krevolin JL, Pandy MG, Pearce JC. Moment arm of the patellar tendon in the human knee. *J Biomech* 2004; 37(5):785–788. Doi: 10.1016/j.jbiomech.2003.09.010.

13 Bini RR, Tamborindeguy AC, Mota CB. Effects of saddle height, pedaling cadence, and workload on joint kinetics and kinematics during cycling. *J Sport Rehabil* 2010; 19(3):301–314.

14 Barratt PR, Martin JC, Elmer SJ, et al. Effects of Pedal Speed and Crank Length on Pedaling Mechanics during Submaximal Cycling: *Med Sci Sports Exerc* 2016; 48(4):705–713. Doi: 10.1249/MSS.0000000000000817.

15 Ferrer-Roca V, Bescós R, Roig A, et al. Acute effects of small changes in bicycle saddle height on gross efficiency and lower limb kinematics. *J Strength Cond Res*

Natl Strength Cond Assoc 2014; 28(3):784–791. Doi: 10.1519/JSC.0b013e3182a1f1a9.

16 Johnston TE, Baskins TA, Koppel RV, et al. The Influence of Extrinsic Factors on Knee Biomechanics During Cycling: A Systematic Review of the Literature. *Int J Sports Phys Ther* 2017; 12(7):1023–1033. Doi: 10.26603/ijspt20171023.

17 Petrella AFM, Gill DP, Petrella RJ. Evaluation of the Get Active Questionnaire in community-dwelling older adults. *Appl Physiol Nutr Metab* 2018; 43(6):587–594. Doi: 10.1139/apnm-2017-0489.

18 Wang Y-C, Hart DL, Stratford PW, et al. Clinical interpretation of a lower-extremity functional scale–derived computerized adaptive test. *Phys Ther* 2009; 89(9):957–968.

19 Ehrig RM, Taylor WR, Duda GN, et al. A survey of formal methods for determining the centre of rotation of ball joints. *J Biomech* 2006; 39(15):2798–2809. Doi: 10.1016/j.jbiomech.2005.10.002.

20 Harrington ME, Zavatsky AB, Lawson SEM, et al. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. *J Biomech* 2007; 40(3):595–602. Doi: 10.1016/j.jbiomech.2006.02.003.

21 Delp SL, Anderson FC, Arnold AS, et al. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Trans Biomed Eng* 2007; 54(11):1940–1950. Doi: 10.1109/TBME.2007.901024.

22 Gatti AA, Maly MR. Accuracy of estimates of cumulative load during a confined activity: bicycling. *Int Biomech* 2019; 6(1):66–74. Doi: 10.1080/23335432.2019.1642141.

23 Kainz H, Graham D, Edwards J, et al. Reliability of four models for clinical gait analysis. *Gait Posture* 2017; 54:325–331. Doi: 10.1016/j.gaitpost.2017.04.001.

24 Steele KM, DeMers MS, Schwartz MH, et al. Compressive tibiofemoral force during crouch gait. *Gait Posture* 2012; 35(4):556–560. Doi: 10.1016/j.gaitpost.2011.11.023.

25 Kutzner I, Heinlein B, Graichen F, et al. Loading of the Knee Joint During Ergometer Cycling: Telemetric In Vivo Data. *J Orthop Sports Phys Ther* 2012; 42(12):1032–1038. Doi: 10.2519/jospt.2012.4001.

26 Hamley EJ, Thomas V. Physiological and postural factors in the calibration of the bicycle ergometer. *J Physiol* 1967; 191(2):55P-56P.

27 Peveler WW. Effects of saddle height on economy in cycling. *J Strength Cond Res* 2008; 22(4):1355–1359.

28 Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol* 2001; 37(1):153–156.

29 Garber CE, Blissmer B, Deschenes MR, et al. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise. *Med Sci Sports Exerc* 2011; 43(7):1334–1359. Doi: 10.1249/MSS.0b013e318213fefb.

30 Nakagawa S, Schielzeth H. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol Evol* 2013; 4(2):133–142. Doi: 10.1111/j.2041-210x.2012.00261.x.

31 Gregor RJ, Cavanagh PR, LaFortune M. Knee flexor moments during propulsion in cycling—a creative solution to Lombard's paradox. *J Biomech* 1985; 18(5):307–316.

32 Sanderson DJ, Martin PE, Honeyman G, et al. Gastrocnemius and soleus muscle length, velocity, and EMG responses to changes in pedalling cadence. *J Electromyogr Kinesiol* 2006; 16(6):642–649. Doi: 10.1016/j.jelekin.2005.11.003.

33 Farrell KC, Reisinger KD, Tillman MD. Force and repetition in cycling: possible implications for iliotibial band friction syndrome. *The Knee* 2003; 10(1):103–109.

34 Holmes JC, Pruitt AL, Whalen NJ. Iliotibial band syndrome in cyclists. *Am J Sports Med* 1993; 21(3):419–424. Doi: 10.1177/036354659302100316.

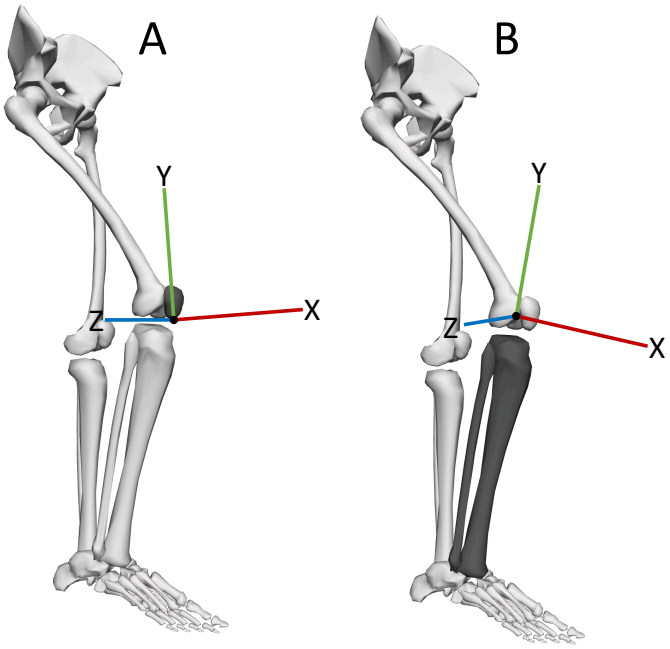
3.7 Supplemental

3.7.1 Supplemental 1.

The following table describes the location of retroreflective markers used to scale a participant-specific musculoskeletal model, and to determine joint kinematics. Markers listed for the thigh, shank, and foot were included on both legs. A total of 40 markers were used.

Segment	Markers
Pelvis	Right anterior superior iliac spine Left anterior superior iliac spine Right posterior superior iliac spine Left posterior superior iliac spine
Thigh	Greater trochanter Lateral femoral condyle Medial femoral condyle Rigid cluster of 4 markers attached to the lateral thigh
Shank	Tibial tuberosity Lateral malleolus Medial malleolus Rigid cluster of 4 markers attached to the lateral shank
Foot	Calcaneus 1st metatarsal head 5th metatarsal head Lateral midfoot

The following shows the patellar (A) and tibial (B) coordinate systems of the musculoskeletal model. Axes lines originate at the origin and point in the positive direction; X (red) is anterior/posterior, Y (green) is inferior/superior, Z (blue) is medial/lateral. Patellofemoral forces were expressed in the patellar coordinate system; the resultant patellofemoral force was in the plane created by X and Y. Tibiofemoral forces were expressed in the tibial coordinate system; Y is compression and X is shear.



The joints of the above musculoskeletal model have the following degrees of freedom.

Joint	Rotational Degrees of Freedom	Translational Degrees of Freedom	Total Degrees of Freedom	Description
Pelvis	3	3	6	The pelvis was free to translate and rotate about all axes.
Hip	3	0	3	The hip had a fixed joint centre but was free to rotate about all 3 axes.
Knee	1	0	1	The knee consisted of a single degree of freedom that was a function of knee flexion. Rotations and translations occurred about the other axes but were prescribed based on knee flexion.
Ankle	1	0	1	The ankle only allowed for ankle plantar/dorsiflexion.

3.7.2 Supplemental 2.

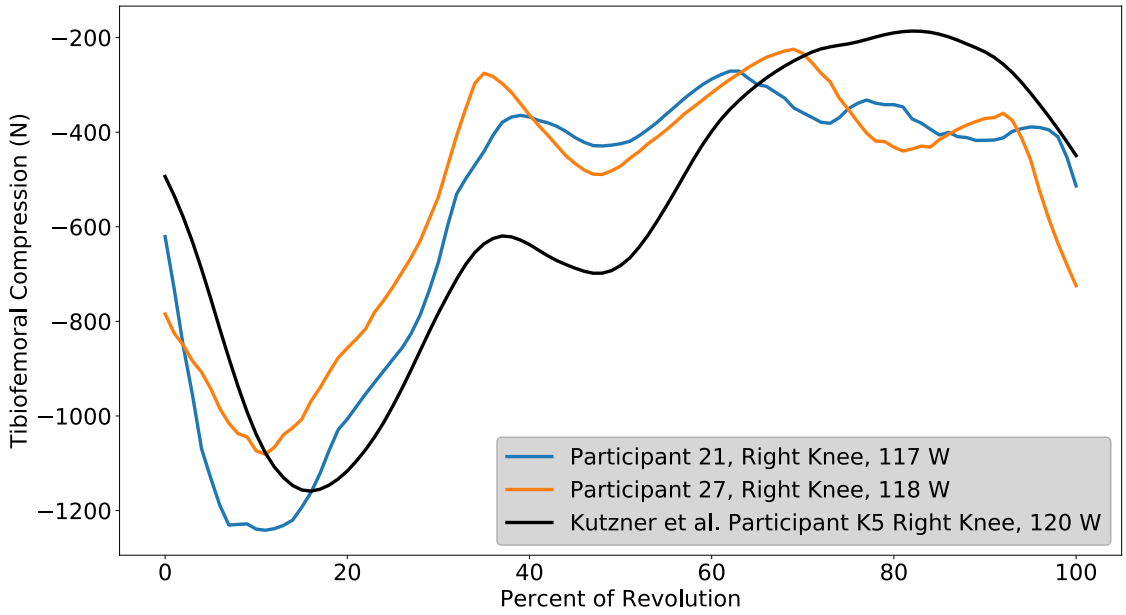
Static optimization minimized the function:

$$min \sum_{i=1}^{N} c_i a_i^2 \quad (1)$$

where N is the number of muscles, a_i is the activation level of muscle i and c_i is a weighting constant for muscle i. Previously reported weighting constants determined to minimize the error between in vivo joint reaction forces, from an instrumented total knee replacement, and model predicted joint reaction forces were used[3]. Weighting constants of 1 are equivalent to regular static optimization. Weighting constants less than one will promote activation of the muscle, and higher than one will penalize activation of that muscle. All muscles had a weighting constant of 1, except for the hamstrings (semimembranosus, semitendinosus, biceps femoris long head, and biceps femoris short head) and gastrocnemius (medial and lateral) muscles which had weightings of three and seven, respectively[3].

The compressive component of the tibiofemoral joint reaction force from this investigation was compared to the compressive component of the tibiofemoral joint reaction force measured using an instrument knee replacement[1]. All force curves are ensembles that were interpolated to 101 points (0-100%). The participant with the instrumented joint replacement was bicycling at a power output of 120W (black line), whereas two representative participants from the present study were bicycling at 117W (blue) and 118W (orange), under their commercial bicycle setup conditions. The vertical

compressive force is negative because forces are expressed in the tibial coordinate system (superior is positive). The x-axis shows the percent of the revolution, with 0% when the crank arm is at the top-dead-centre position (12 o'clock) and 50% is the bottom-dead-centre (6 o'clock). Within the first 20% of the revolution, the magnitude and timing of the first peak aligns. A second peak in the instrumented knee occurs at about 50% of the revolution; this peak is also apparent in the modelled reaction forces, though of a lower magnitude. It is important to note that the instrumented knee appears to reduce tibiofemoral joint reaction forces during the upswing phase of the revolution (65-90%). These differences may be attributed to differences in the cycling posture, differences between the samples, or those caused by the knee replacement itself.



3.7.3 Supplemental 3.

The experimental protocol included 18 bicycle conditions. These conditions comprised of all combinations of 3 saddle horizontal (X) positions, 3 saddle vertical (Y) positions, and 2 crank arm lengths.

Saddle Horizontal Position (X)	**Saddle Vertical Position (Y)**	**Crank Arm Length**
Commercial	Commercial	Commercial (172.5 mm)
Random X position 1 within +/- 10% of commercial	Random Y position 1 within +/- 5% of commercial	Random crank arm position (170 or 175 mm)
Random X position 2 within +/- 10% of commercial	Random Y position 2 within +/- 5% of commercial	

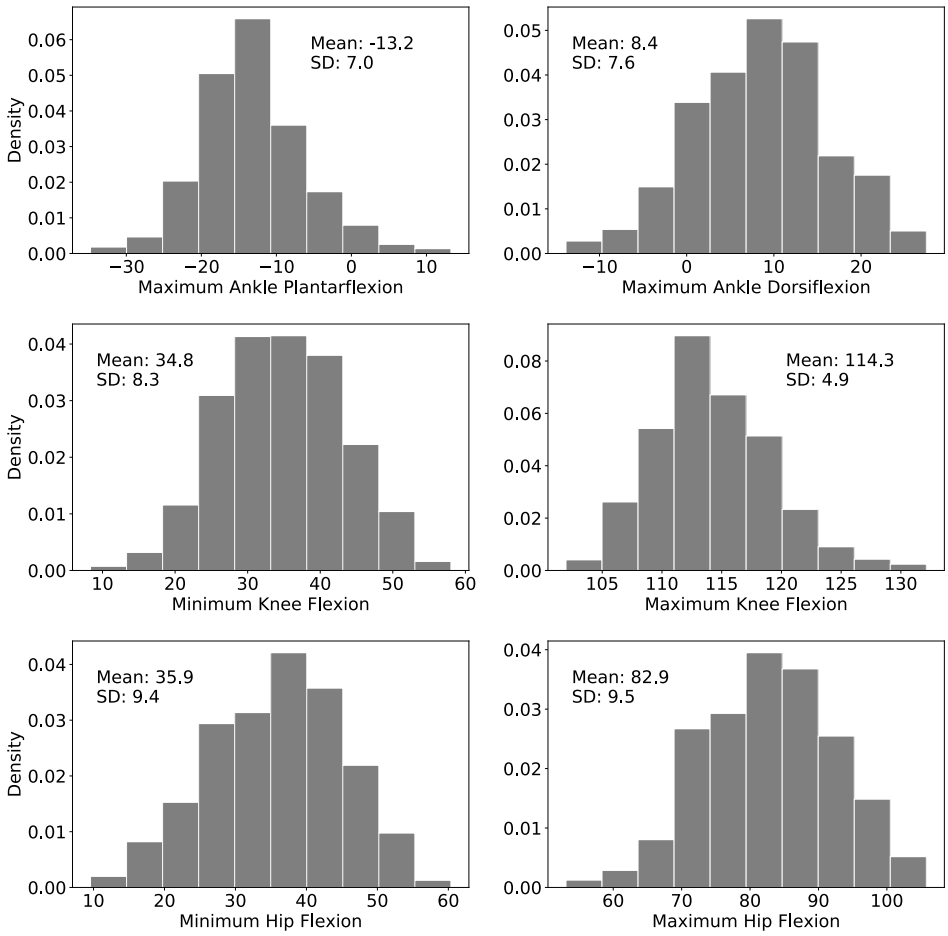
The 18 positions ensured that participants completed all possible combinations of each level (3 x 3 x 2 = 18). Randomly allocated positions for X and Y were independent samples where each mm increment within the defined range (e.g., for the saddle horizontal position: from commercial – 10% to commercial + 10%) had an equal probability of being sampled. This approach resulted in every participant having one position that was commercial (cycling position 1 below). The order in which cycling positions were tested was randomized for each participant.

Cycling Position	**Saddle horizontal**	**Saddle vertical**	**Crank arm length**
1	Commercial	Commercial	Commercial
2	Commercial	Commercial	Random position 1
3	Commercial	Random position 1	Commercial
4	Commercial	Random position 1	Random position 1
5	Commercial	Random position 2	Commercial
6	Commercial	Random position 2	Random position 1
7	Random position 1	Commercial	Commercial
8	Random position 1	Commercial	Random position 1
9	Random position 1	Random position 1	Commercial
10	Random position 1	Random position 1	Random position 1
11	Random position 1	Random position 2	Commercial
12	Random position 1	Random position 2	Random position 1
13	Random position 2	Commercial	Commercial
14	Random position 2	Commercial	Random position 1

15	Random position 2	Random position 1	Commercial
16	Random position 2	Random position 1	Random position 1
17	Random position 2	Random position 2	Commercial
18	Random position 2	Random position 2	Random position 1

3.7.4 Supplemental 4.

Histograms of ankle, knee, and hip kinematics from all participants and all bicycling postures. Each panel includes the mean and standard deviation of the distribution. Density is the normalized height of each bin such that the total area of all bins sums to 1.0.



3.7.5 Supplemental 5.

Results from the base + crank arm length model and base + saddle horizontal position model. Crank arm length and horizontal saddle position were not significant predictors in any models. Bolding indicates significance at $p<0.05$.

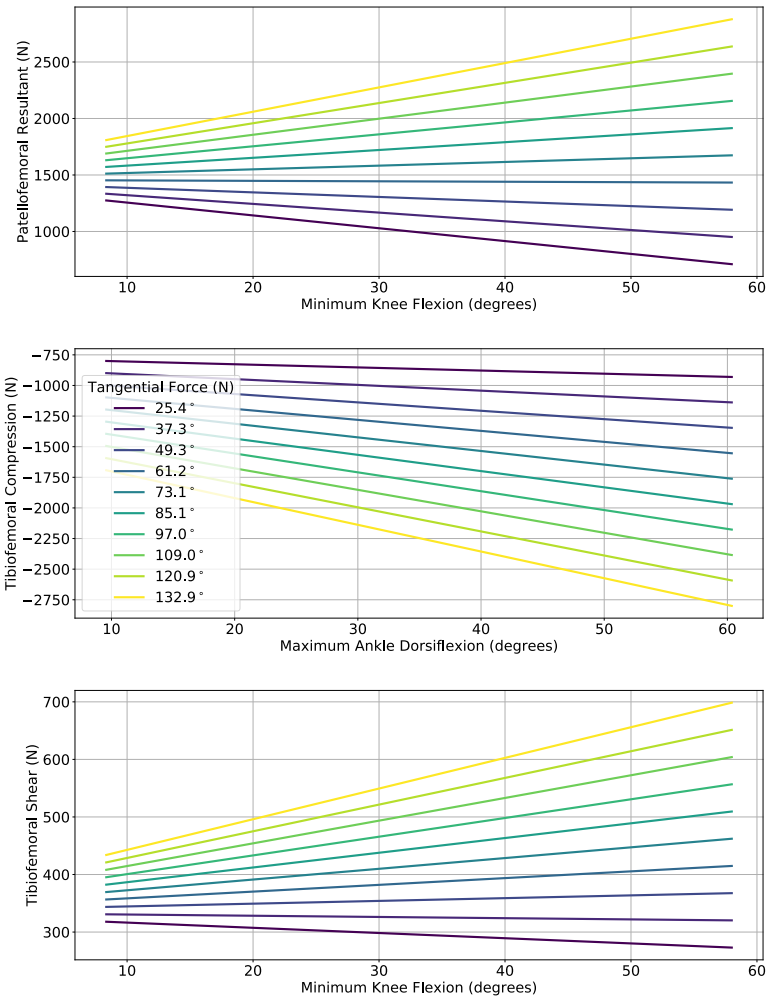
Model	Predictor	Beta Coefficient	Standard error	p-value
Base Models + Crank Arm Length				
Patellofemoral Resultant Marginal R^2: 0.15 Conditional R^2: 0.90	Intercept	1892.37	808.05	**0.019**
	Minimum hip flexion	2.37	3.55	0.505
	Maximum ankle dorsiflexion	18.08	2.06	**<0.001**
	Minimum knee flexion	-12.05	3.98	**0.002**
	Minimum knee flexion X minimum hip flexion	0.41	0.10	**<0.001**
	Crank arm length	-33.78	45.76	0.460
Tibiofemoral Compression Marginal R^2: 0.14 Conditional R^2: 0.90	Intercept	-1154.99	542.76	**0.033**
	Minimum hip flexion	-10.42	0.98	**<0.001**
	Maximum ankle dorsiflexion	-11.06	1.39	**<0.001**
	Crank arm length	22.51	31.26	0.472
Tibiofemoral Shear Marginal R^2: 0.11 Conditional R^2: 0.86	Intercept	89.66	243.42	0.713
	Minimum hip flexion	-1.62	1.07	0.132
	Maximum ankle dorsiflexion	3.66	0.62	**<0.001**
	Minimum knee flexion	-5.02	1.20	**<0.001**
	Minimum knee flexion X minimum hip flexion	0.19	0.03	**<0.001**
	Crank arm length	17.44	13.81	0.207
Base Models + Saddle Horizontal Position				
Patellofemoral Resultant Marginal R^2: 0.15 Conditional R^2: 0.90	Intercept	1226.04	220.03	**<0.001**
	Minimum hip flexion	2.23	3.56	0.531
	Maximum ankle dorsiflexion	18.23	2.05	**<0.001**
	Minimum knee flexion	-11.74	4.00	**0.003**
	Minimum knee flexion X minimum hip flexion	0.41	0.10	**<0.001**
	Horizontal saddle position	34.10	64.11	0.595
Tibiofemoral Compression Marginal R^2: 0.15 Conditional R^2: 0.90	Intercept	-752.01	116.47	**<0.001**
	Minimum hip flexion	-10.38	0.98	**<0.001**
	Maximum ankle dorsiflexion	-11.13	1.39	**<0.001**
	Horizontal saddle position	-6.81	42.92	0.874
Tibiofemoral Shear Marginal R^2: 0.11 Conditional R^2: 0.86	Intercept	338.24	64.66	**<0.001**
	Minimum hip flexion	-1.64	1.07	0.126
	Maximum ankle dorsiflexion	3.61	0.62	**<0.001**
	Minimum knee flexion	-4.94	1.21	**<0.001**
	Minimum knee flexion X minimum hip flexion	0.19	0.03	**<0.001**
	Horizontal saddle position	22.42	19.35	0.247

3.7.6 Supplemental 6.

This supplemental document includes a table and a figure that present the findings of the relationship between crank arm tangential force, joint kinematics, and knee joint reaction forces. The following table summarizes the results of the base + crank arm tangential force models used to predict resultant patellofemoral, as well as compressive and shear tibiofemoral forces. Bolding in the table indicates significance at $p<0.05$.

Model	Predictor	Beta Coefficient	Standard error	p-value
Base Models + Crank Arm Tangential Force				
Patellofemoral Resultant Marginal R^2: 0.45 Conditional R^2: 0.90	Intercept	1173.27	263.08	**<0.001**
	Minimum hip flexion	5.15	3.50	0.140
	Maximum ankle dorsiflexion	-5.97	5.67	0.293
	Minimum knee flexion	-29.24	4.83	**<0.001**
	Minimum hip flexion X minimum knee flexion	0.28	0.10	**0.006**
	Tangential force	-0.34	3.00	0.909
	Tangential force X minimum knee flexion	0.31	0.05	**<0.001**
	Tangential force X maximum ankle dorsiflexion	0.33	0.07	**<0.001**
Tibiofemoral Compression Marginal R^2: 0.45 Conditional R^2: 0.90	Intercept	-690.87	177.46	**<0.001**
	Minimum hip flexion	2.31	3.22	0.474
	Maximum ankle dorsiflexion	1.97	3.90	0.613
	Tangential force	-0.20	2.35	0.931
	Tangential force X minimum hip flexion	-0.18	0.04	**<0.001**
	Tangential force X maximum ankle flexion	-0.18	0.05	**<0.001**
Tibiofemoral Shear Marginal R^2: 0.33 Conditional R^2: 0.87	Intercept	376.56	73.81	**<0.001**
	Minimum hip flexion	-0.97	1.07	0.362
	Maximum ankle dorsiflexion	-3.72	1.72	**0.031**
	Minimum knee flexion	-8.26	1.47	**<0.001**
	Minimum knee flexion X minimum hip flexion	0.16	0.03	**<0.001**
	Tangential force	-0.25	0.83	0.766
	Tangential force X maximum ankle dorsiflexion	0.10	0.02	**<0.001**
	Tangential force X minimum knee flexion	0.06	0.02	**<0.001**

The following figure visually demonstrates the interactions between the crank arm tangential force and minimum knee flexion on the resultant patellofemoral reaction force, the crank arm tangential force and maximum ankle dorsiflexion on tibiofemoral compression, and the crank arm tangential force and minimum knee flexion on tibiofemoral shear. Each line represents a different tangential force. All panels follow the same legend.



4 Automatic Cartilage and Bone Segmentation using Multi-Stage Convolutional Neural Networks: Data from the Osteoarthritis Initiative

This paper was submitted to the journal Medical Image Analysis on October 8th 2020 and transferred to the journal Computerized Medical Imaging and Graphics on October 14th 2020.

Automatic Cartilage and Bone Segmentation using Multi-Stage Convolutional Neural Networks: Data from the Osteoarthritis Initiative

Computerized Medical Imaging and Graphics

[a,b,*]Anthony A. Gatti, [a,c]Monica R. Maly, PT PhD

[a]School of Rehabilitation Sciences, McMaster University, 1280 Main St. W., L8S 4L8, Hamilton, Ontario, Canada

[b]NeuralSeg Ltd., Hamilton, Ontario, Canada

[c]Department of Kinesiology, University of Waterloo, Waterloo, Canada

*Corresponding Author

Anthony A. Gatti

Email addresses: anthony@neurlaseg.com (A.A. Gatti), mrmaly@uwaterloo.ca (M.R. Maly)

Keywords: Cartilage, Deep learning, Magnetic Resonance Imaging, Osteoarthritis, Segmentation

Abstract

Presented is a method of automatically segmenting bone and cartilage from magnetic resonance imaging (MRI) images using a multi-stage convolutional neural network framework. Stage 1 coarsely segments images outputting probability maps of each voxel belonging to eight classes of interest: 4 cartilage, 3 bone, 1 background. Stage 2 segments overlapping sub-volumes that include the probability maps outputted from Stage 1, concatenated to raw image data. Using 6-fold cross-validation, this framework was tested on two datasets: 176 images from 88 individuals in the Osteoarthritis Initiative (OAI) and 60 images from 15 healthy young men. The framework produces the best reported segmentation accuracies (Dice similarity coefficient; DSC) on the OAI dataset for the femoral (0.907) and medial tibial cartilage (0.876). It also produced the second-best accuracies for lateral tibial cartilage (0.913), and patellar cartilage accuracy of 0.840. The healthy cartilage accuracies are the best reported to-date (femoral = 0.938, medial tibial = 0.911, lateral tibial = 0.930, patellar = 0.956). Average surface distances are less than in-plane resolution. Segmentations take 91 +/- 11 s. The framework learns how to automatically segment cartilage and bone from MR images acquired with two different sequences, producing efficient, accurate quantifications at varying disease severities.

4.1 Introduction

Knee osteoarthritis (OA) is a chronic joint disease estimated to affect more than 7% of Americans (Deshpande et al., 2016). While OA damages all joint tissues, cartilage degeneration is the hallmark (Creamer and Hochberg, 1997; Kraus et al., 2015). As such, magnetic resonance imaging (MRI) is recommended for assessing structural OA by quantifying cartilage outcomes (Conaghan et al., 2011; Hunter et al., 2015; Peterfy et al., 2006). Medial joint cartilage is of particular interest due to a higher prevalence of medial knee osteoarthritis (Metcalfe et al., 2012). Cartilage localization must be highly accurate to enable specific and sensitive measurements necessary for identifying disease modifying OA drugs in clinical trials and for clinical usage.

A major hurdle to quantifying cartilage outcomes from MRI scans is the resources necessary for tissue segmentation. Manual segmentation is the gold standard for segmenting cartilage from MRI images, however it takes hours to perform, requires highly specialized knowledge, and suffers from inter- and intra- rater errors (Duryea et al., 2007; Pedoia et al., 2016; Shim et al., 2009). As a result, manual segmentation is not a realistic strategy to manage large volumes of data necessary to understand knee OA progression or to evaluate changes in joint structure in clinical trials. For example, the Osteoarthritis Initiative (OAI), a private-public collaboration by the National Institutes of Health (NIH), acquired serial MRI scans of thousands to study disease progression

(Peterfy et al., 2008). Not surprisingly, quantitative cartilage outcomes are unavailable for most OAI scans.

Automated extraction of cartilage outcomes is necessary to enable efficient analysis. Conventional methods include statistical shape modelling (SSM), active appearance models (AAM), atlas-based algorithms, and traditional machine learning (support vector machines, k-nearest neighbours). These methodologies require 10-minutes to 48 hours (Ahn et al., 2016; Dodin et al., 2010; Fripp et al., 2010; Shan et al., 2014) and currently fall short of human accuracies (Ahn et al., 2016; Dam et al., 2015; Dodin et al., 2010; Fripp et al., 2010; Prasoon et al., 2013; Shan et al., 2014; Tamez-Pena et al., 2012; Wang et al., n.d.; Yin Yin et al., 2010). Convolutional neural networks (CNN) have become the state-of-the-art method for many tasks in computer vision (Shen et al., 2017), and have shown promise in cartilage segmentation (Ambellan et al., 2019; Liu, 2018; Norman et al., 2018; Zhou et al., 2018).

Major barriers in segmenting 3D medical data using deep learning are the limited memory available on graphics processing units (GPUs) and small samples. The standard 3D CNN trained on a 12GB GPU typically handle images with dimensions of 128x128x64 (Milletari et al., 2016) with a batchsize of only one. The batchsize is the number of examples passed through the network before updating the weights (Hastie et al., 2009). Images from the OAI commonly used for cartilage quantification

(384x384x160) would thus require downsampling to less than 5% of their original size. Deep learning methods have drastically reduced segmentation times, with the fastest algorithms taking <15 s but still fall short of traditional automated methods in terms of accuracy (Liu, 2018; Norman et al., 2018). Meanwhile, Ambellan and colleagues showed that a combination of SSMs with deep learning had promise, matching or exceeding traditional methodologies in accuracy, with a segmentation time of 10 min per knee (Ambellan et al., 2019). Nevertheless, heterogeneity introduced by osteophytes required manual definitions in this SSM, making transfer to new tissues more challenging. We require methods that better balance speed and accuracy to enable wide-spread usage of cartilage analyses.

Multiple methods have been proposed to overcome the memory constraints. Original methods used sub-volume or patch-based methods (Yu et al., 2017). More recently, multi-stage methods have been proposed, using stage one to localize and stage two to produce a higher resolution segmentation, or to vote using machine learning methods other than deep learning. For example, the SSM plus deep learning method by Ambellan and colleagues leverages stage one to localize the bone surface using a CNN, refines the CNN segmentation using an SSM and then does sub-volume segmentation using a smaller 3D-CNN over the bone surface (Ambellan et al., 2019). The authors trained a different network for each cartilage tissue. In applications to segment abdominal tissues, Zhu and colleagues (2019) use a patch-based method. First, to identify a bounding box of the tissue of interest, the whole image was segmented in large patches (64^3 pixels), then

smaller patches re-segmented this smaller region. The authors indicated that the larger patches improve specificity while the smaller patches improve sensitivity (Zhu et al., 2019). Roth and colleagues (2018) use a similar approach but without a bounding box; instead this approach identified a new region of interest of arbitrary shape from stage one and then re-segmented this smaller region using a second network in stage two (Roth et al., 2018). Subsequently, Roth and colleauges use holistically nested CNNs to segment all three orthogonal planes of an image in 2D and then aggregate the 2D predictions using a random forest classifier (Roth et al., 2018). Similarly, Pang et al. used a CNN at stage one and then refined the segmentation in stage two using a custom conditional random field (Pang et al., 2019).

The purpose of this study is to define a general multi-stage fully CNN framework for segmenting cartilage and bone from knee MRI data that overcomes memory constraints, while balancing computation time and accuracy when segmenting a variety of sequences and disease severities. Our method builds upon previous segmentation literature. In our multi-stage framework, Stage 1 uses multiple CNNs to segment the image coarsely and Stage 2 conducts a fine segmentation, where the CNN takes inputs of both raw image data as well as the probabilities outputted from Stage 1. Providing probabilities gives global context to the second stage that segments image sub-volumes in 3D; this global context has not been used in previously proposed methods. We describe the segmentation algorithm, then train and test the algorithm on images acquired from two MRI vendors using different pulse sequences, and in samples with and without OA.

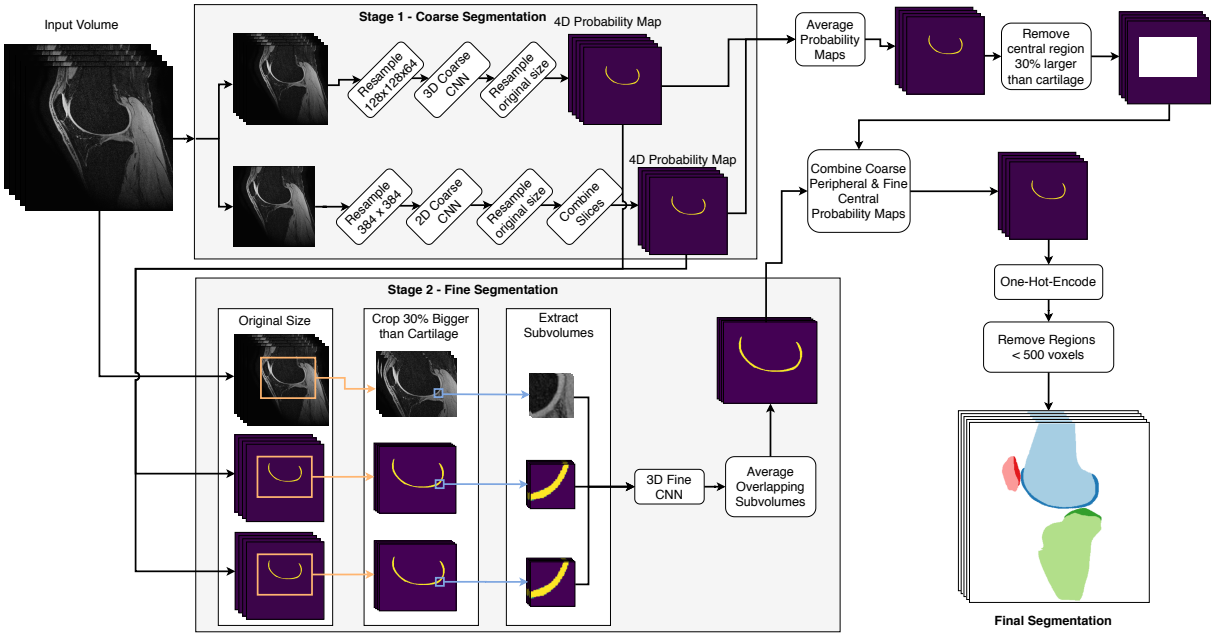


Figure 4-1. Visualization of the algorithm. Stage 1 coarsely segments the input image using two convolutional neural networks (CNNs). Probability outputs of Stage 1 are combined with the raw image to perform full resolution segmentation of sub-volumes using a third CNN in Stage 2. Stage 2 segmentations from the main tissues, or classes, of interest are combined with coarse segmentations at the periphery to produce the final segmentation.

4.2 Methods

4.2.1 Segmentation Framework

The framework (Figure 4-1) uses two stages of CNNs. Stage 1 performs two coarse segmentations (one 2D, one 3D CNN) of the entire MRI volume. Each outputs a 4D probability map, with the indices along the last (4^{th}) dimension contain the probabilities for each voxel belonging to each of the eight classes of interest. Classes are the different categories that each voxel can be categorized to; here 7 anatomical tissues and the background. Stage 2 segments overlapping sub-volumes of the image, taking inputs of

full resolution image data and priors (probability maps) outputted from Stage 1. The segmented sub-volumes are then combined to create a full-resolution segmentation.

4.2.1.1 Stage 1

Two networks that share the same U-Net style architecture segmented the background and 7 tissues of interest: femur, tibia, patella, femoral cartilage, medial tibial cartilage, lateral tibial cartilage, and patellar cartilage from a single MRI volume (Ronneberger et al., 2015). The first had an input shape of 128x128x64 and an output shape of 128x128x64x8 and segmented an entire image in a single pass. The second network segmented an individual slice at an image size of 384x384 and had an output of 384x384x8. The last dimension of the two networks was length 8 to accommodate the 8 classes (7 tissues, 1 background). The 2D network was applied to each slice, then the slices were combined to create the full 4D segmentation. The results of the 2D and 3D analyses were resampled to be the same shape as the original image. For both networks, prior to segmentation, the whole 3D image was normalized to have mean 0 and unit variance by subtracting the mean and dividing by the standard deviation of the pixel intensities.

4.2.1.1.1 Stage 1 Network

The coarse network used for Stage 1 is described in Figure 4-2. Like U-Net, the coarse network included long residual connections from the compression to decompression

branches. It also included short residual connections at each level of the compression and decompression branches (Milletari et al., 2016; Yu et al., 2017). All short residual connections used a summation methodology and all long residual connections used concatenation (Figure 4-2)(Kayalibay et al., 2017). For every step down the compression branch, the image dimension was compressed to half its previous size using a stride of 2 with a regular convolution. For every step up the decompression branch, the image dimension was doubled by using a stride of 2 with a transpose convolution. The number of filters for each convolution is included in Figure 4-2. Filters included on the decompression branch mirrored the compression branch. Convolutions throughout the network comprised of 5x5 (2D) or 3x3x3 (3D) convolutional filters, followed by batch normalization, dropout (probability = 0.2), and a parametric rectified linear unit (PReLU) (He et al., 2015). In addition to the traditional compression and decompression branches, we included a form of deep supervision inspired by Kayalibay et al., 2017. The deep supervision more directly passed data from the deep layers directly to the final output, using the same number of filters as classes. However, we used a PReLU activation instead of a logistic or Softmax function as used previously (Kayalibay et al., 2017). The final convolution of the network used a Softmax activation which gives probabilities of each voxel belonging to each of the 8 classes (4 cartilage, 3 bone, 1 background).

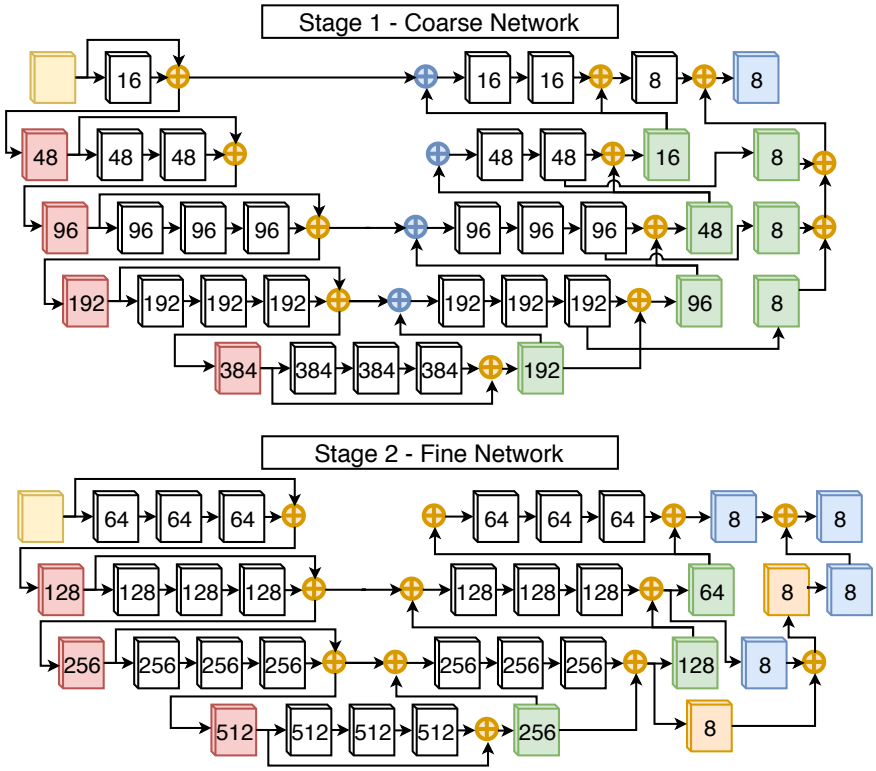


Figure 4-2. Visual depiction of the Stage 1 (coarse) and Stage 2 (fine) segmentation network architectures. Yellow cubes represent inputs, white are regular convolutions with parametric rectified linear unit (PReLU) activations, red are down convolutions (2x2x2 stride) with PReLU activations, green are transpose (up) convolutions (2x2x2) with PReLU activations, orange are transpose convolutions (2x2x2) with Softmax activations, blue are regular convolutions with a 1x1x1 filter and Softmax activations. Orange circles represent addition and blue circles represent concatenation. The number of filters in a convolution is printed on the face of each cube.

4.2.1.1.2 Stage 1 Training

During training, the 3D network used batch sizes of 1 and the 2D network used batch sizes of 16. Image augmentation was implemented, including 50% probability of flipping along the slice-axis (3D network only), random height and width shift within +/- 20%, and random in-plane rotation within +/- 6 degrees. The Adam optimizer (Kingma and Ba,

2017) with a learning rate of 10^{-4} and early stopping was implemented. Early stopping was invoked after 10 consecutive epochs with less than 0.02 improvement in the loss function, when evaluated on validation data. Learning rate and early stopping criteria were determined using a grid-search in a previous study (Gatti, 2018). The loss function was a custom generalized Dice similarity coefficient (gen-DSC). In the following, (1) is the traditional DSC, and (2), which utilizes (1) in its definition, is the gen-DSC:

$$DSC\ (X_n, Y_n)\ =\ \frac{2\ |X_n \cap Y_n|}{|X_n|\ +\ |Y_n|} \tag{1}$$

$$Gen - DSC\ =\ \sum_{n=1}^{N} -1\ *\ DSC\ (X_n, Y_n) \tag{2}$$

where N is the number of classes, X_n is the predicted probability map for class n, and Y_n is the ground truth (manual) segmentation for class n. Under this paradigm, the best possible gen-DSC score was -8 (the negative of the number of classes).

4.2.1.2 Stage 2

This single 3D CNN (Figure 4-2) segmented sub-volumes, sized 32x32x32 of the original image. This network took a 4-dimensional input, where the first 3 dimensions were the physical dimensions of the sub-volume, and the 4th dimension was length 17; 1 for the original image data, 8 for the probabilities outputted from the coarse 3D based segmentation from Stage 1, and 8 for the probabilities outputted from the coarse 2D based segmentation from Stage 1. The output was also 4-dimensional, with the 4th dimension being equal to eight, the number of classes. The whole image was normalized before sub-volumes were extracted. For efficiency, a region 30% bigger than the region

containing all cartilage, determined using results from Stage 1, was extracted and segmented (Figure 4-3). From this extracted space, sub-volumes that overlap by 50% in each of the three dimensions were sampled. Using such a scheme resulted in a total of 8 predictions for all but the most peripheral voxels.

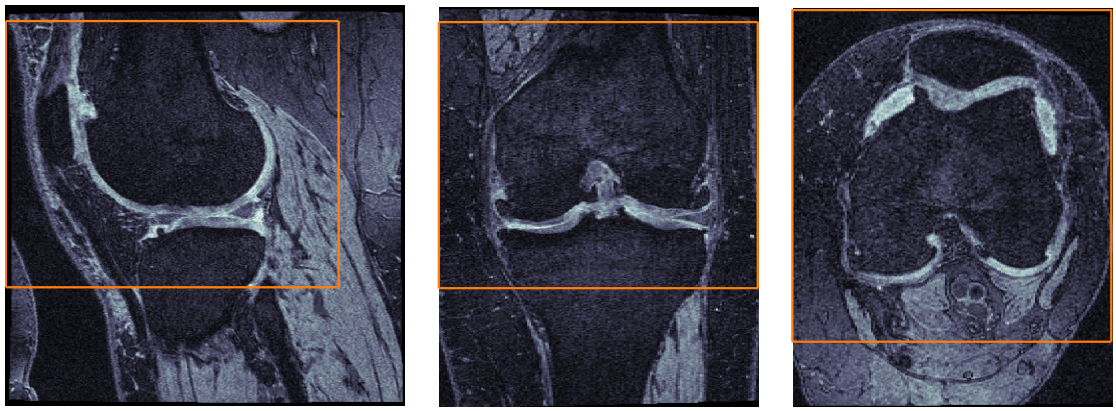


Figure 4-3. Example of the region that was extracted (orange) then segmented in Stage 2. The region is 30% larger than the cartilage segmentations and includes all cartilage as well as the primary bone areas of interest, e.g., locations of osteophyte formation. Osteophytes are bony growths at the joint margin characteristic of osteoarthritis (OA). The example image is a participant from the OA dataset.

4.2.1.2.1 Stage 2 Network

This fine network (Figure 4-2**)** used convolutions of 3x3x3 throughout. The same as the Stage 1 coarse network, the final convolution filter used a Softmax activation to give probabilities of each voxel belonging to the classes of interest. The fine network had differences compared to the coarse network. The fine network included 3 down convolutions (as opposed to 4), a greater number of filters (64 at the first level, doubling

after each down convolution), and used addition for long and short residual connections. The greater number of filters at the first level enabled high level feature extraction that is not possible in the coarse network due to memory constraints. The shallower depth enabled bigger batch sizes. Finally, the deep supervision branch used a Softmax activation as was originally designed (Kayalibay et al., 2017).

4.2.1.2.2 Stage 2 Training

During training, every 33 epochs, 1000 sub-volumes were selected from each image using stratified random sampling. Stratified random sampling ensured that, at a minimum, each tissue was included in the following percentages of sub-volumes: femur 9%, tibia 9%, patella 7%, femoral cartilage 22%, medial tibial cartilage 15%, lateral tibial cartilage 15%, patellar cartilage 15%, and background 8%. A batch-size of 8 and the Adam optimizer with a learning rate of 10^{-3} with early stopping was implemented. For the fine network, early stopping was invoked after 10 consecutive epochs with less than 0.00025 improvement when tested on the validation data. These parameters were determined using a grid-search.

To enable efficient training of the class-imbalanced tissues, a custom weighted generalized DSC (weighted-DSC) was utilized. Class-imbalance refers to the unequal occurrences of the different classes (that is, tissues) within a knee MRI scan. The weighted-DSC multiplied a DSC per class by a weighting factor. The background class

weighting factor was always 1. The weighting factors for the non-background classes summed to 7, equal to the number of non-background classes. For each new sub-volume, the weightings for the non-background classes were re-calculated to distribute the total available weighting over the non-background classes. The weighting for each sub-volume was distributed based on the percentage of non-background voxels (from both the predicted segmentation and the ground truth), which each class occupied. For example, if the sub-volume only contained one class in both the prediction and ground truth segmentations, then that class received a weighting of 7 (1.0x7) and all others 0 (0.0x7); if there were only two classes in the sub-volume, one in 25% of the voxels and the other in 75%, then they would have weightings of 1.75 (0.25x7) and 5.25 (0.75x7), respectively, with all other classes being weighted at 0 (0.0x7). Because predictions were not binary, the number of voxels for any one class was calculated by summing the probabilities of each voxel belonging to that class. In the following, (3) describes the calculation for the number of non-background-voxels (NBV), (4) describes the proportion of NBVs for a given class (pNBV), and (5) describes the full weighted-DSC calculation:

$$NBV = \sum_{n=2}^{N}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} X_{i,j,k,n} + Y_{i,j,k,n} \tag{3}$$

$$pNBV(n) = \frac{1}{NBV}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} X_{i,j,k,n} + Y_{i,j,k,n} \tag{4}$$

$$Weighted - DSC = \frac{-DSC\ (X_1, Y_1) + \sum_{n=2}^{N}\left(-(N-1) \times pNBV(n) \times DSC\ (X_n, Y_n)\right)}{N} \tag{5}$$

where N is the number of classes, and 1 is the background class; I is image dimension 1, J is image dimension 2, and K is image dimension 3; X is the predicted probability map,

and Y is the ground truth probability map; the first 3 dimensions of X and Y are the image dimensions and the 4th dimension is the segmentation class. This approach efficiently learned to segment the tissues of interest and outperformed the generalized-DSC described in Equation 2 used in previous implementations of this algorithm (Gatti, 2018).

4.2.1.3 Compiling Final Segmentations

To create full-resolution segmentation probability maps, first the full-sized 2D and 3D segmentation probabilities from Stage 1 were averaged, yielding the average-coarse-segmentation. From Stage 2, the segmentation probabilities for each class from overlapping sub-volumes were averaged, yielding the average-fine-segmentation. The final segmentation consisted of the average-fine-segmentation in the region it was extracted from, and the average-coarse-segmentation in the area outside this region. This scheme efficiently segmented the full image, with high resolution for cartilage and bone approximating the articular surfaces.

Each voxel was classified according to the class that it had the highest probability of belonging to. To remove potentially spurious segmentations, any connected regions of a class that were <500 connected voxels were labelled as background.

4.2.2 Testing Accuracy and Segmentation Times

To test the accuracy and segmentation times of the framework, we conducted a 6-fold cross-validation. The cross-validation used data from two samples: (1) 176 knee MRI scans from 88 individuals enrolled in the OAI, and (2) 60 MRI scans from 15 healthy young men.

4.2.2.1 Osteoarthritis Sample

Data were from individuals in the OAI with doubtful (grade 1; n=2), minimal (grade 2; n=31), moderate (grade 3, n=52), and severe (grade 4; n=3) OA severity defined on the gold standard Kellgren and Lawrence (KL) classification system (Kellgren and Lawrence, 1957). Images were acquired using a 3T Siemens MRI scanner at one of four sites (Peterfy et al., 2008). All images were acquired with the same 3D sagittal water excited Dual Echo in the Steady State (DESS) sequence with in-plane resolution of 0.365x0.456 mm and resampled to be 0.365x0.365 mm, with a slice thickness of 0.7 mm, and no gaps (Peterfy et al., 2008). Manual segmentations of cartilage were performed for the OAI by an industrial partner (iMorphics)(Williams et al., 2010). As well, manual segmentations of the 3 bones (femur, tibia, patella) were conducted by one researcher (AAG).

4.2.2.2 Healthy Sample

In a previous study, (Gatti et al., 2017) healthy participants completed 2 MRI visits, obtaining 2 sets of knee MRI scans per visit (4 image sets per participant). Images were acquired using a 3T GE Discovery MR750 (GE Healthcare, Milwaukee WI) using a 3D

sagittal fat saturated fast spoiled gradient recalled sequence (FSPGR), in-plane resolution of 0.3125x0.3125 mm, slice thickness of 1.0 mm, and no gaps (Gatti et al., 2017). Manual cartilage and bone segmentations were conducted by one researcher (AAG).

4.2.2.3 Experiments

To conduct the 6-fold cross-validation, 236 MRI volumes from 103 participants were split into 6 partitions by participant. The 6 partitions are described in Table 4-1. Five to ten cross-validation folds balance the trade-off between bias and variance (Hastie et al., 2009). During each cross-validation step, one unique partition of the data was held out for final testing. Of the remaining five partitions, one was used for validation and four were used to train all stages of the algorithm. The validation partition was used to identify when training was complete using the aforementioned early stopping criteria. Assessments of segmentation quality were performed for five cartilage classes (femoral, medial tibial, lateral tibial, all tibial, patellar) and three bones (femur, tibia, patella).

Table 4-1. Partitions used for cross-validation including the number (n) of participants with osteoarthritis (OA), the number of healthy participants, the total number of participants, and the total number of images per partition.

	Partitions					
	1	**2**	**3**	**4**	**5**	**6**
n OA	15	15	15	15	14	15
n healthy	2	2	2	3	3	3
n total	17	17	17	18	17	18
n total images	38	38	38	42	40	42

Quality assessments were run on the entire image volume, as well as on cropped image volumes that were 30% bigger than the region containing all cartilage, with cropping performed based on the reference segmentations. Assessments were performed for all final testing data, and separately for the validation data used for early stopping.

Assessment of segmentation accuracies was performed using the DSC (Equation 1), the volume difference (VD; (6)), and the average surface distance (ASD; (7)) (Ambellan et al., 2019; Dam et al., 2015; Desai et al., n.d.; Liu, 2018). The DSC and average surface distance are symmetric; volume difference is a percent difference relative to the reference volume (Ambellan et al., 2019). Time to complete all segmentation steps was recorded. Volume difference and average surface distance were defined as follows:

$$VD = 100\,\frac{|X_n| - |Y_n|}{|Y_n|} \tag{6}$$

$$ASD = \frac{\sum_{i=1}^{n_{\partial X}} \min_{y \in \partial Y} \|x_i - y\|_2 + \sum_{j=1}^{n_{\partial Y}} \min_{x \in \partial X} \|y_j - x\|_2}{n_{\partial X} + n_{\partial Y}} \quad (7)$$

where X_n is the predicted segmentation for tissue n, and Y_n is the ground truth, ∂X and ∂Y are the boundary voxels of segmentations X and Y, $n_{\partial X}$ and $n_{\partial Y}$ are the number of boundary voxels for ∂X and ∂Y, $|\cdot|$ is the volume, and $\|\cdot\|_2$ the Euclidean distance.

To explore the role of disease severity and cartilage morphometry on cartilage segmentation accuracies, two analyses were performed. In the OA sample, we explored whether cartilage DSC was dependent on radiographic disease severity (KL grade), using one-way analysis of variance (ANOVA). In all participants, we determined whether the DSC was dependent on tissue volume using linear regression; separate regression models were run for the OA and healthy samples. These secondary analyses were run independently for each cartilage tissue of interest.

Finally, an ablation study is performed to determine how the Stage 2 (fine stage) adds to prediction accuracy. Using only the testing data, we report the segmentation accuracies (DSC, volume difference, and average surface distance) separately for the two Stage 1 networks, as well as a simple average of the two Stage 1 networks. All post-processing was implemented the same as the full model (4.2.1.3); e.g., the segmentation was one-hot-encoded and islands less than 500 connected voxels were removed. These data enable identification of if and how much Stage 2 improves segmentations.

All experiments were conducted using a virtual machine with 12 CPU threads, 78 GB of RAM, and an NVIDIA Tesla V100 GPU on the Google Cloud Platform. Deep learning was done using Keras (Chollet, 2015) with a Tensorflow backend in Python.

4.3 Results

Summary statistics of the segmentation accuracies, by sample, are presented in Table 4-2 and graphically displayed in Figure 4-4. An example of the automated cartilage segmentations as well as segmentation surface errors are presented in Figure 4-7. For the testing data of the OA sample, the mean (SD) DSCs were: femoral cartilage 0.907 (0.023); medial tibial cartilage 0.876 (0.042); lateral tibial cartilage 0.913 (0.026); all tibial cartilage 0.897 (0.026); and patellar cartilage 0.840 (0.128). DSC distributions by KL grade (Figure 4-5**)** show small decreases in accuracy with worse disease severity. In the OA sample, mean volume differences were systematically larger using the proposed methodology for the femoral and tibial cartilage, and smaller for the patellar cartilage, compared to the reference. The mean and median average surface distance for all OA segmentations were less than in-plane resolution (0.365 mm). The OA sample segmentation times were on average 91.4 (9.6) s.

Table 4-2. Segmentation results presented separately for the osteoarthritis (OA) and healthy samples. Results are presented for the testing (held out), validation (used for early stopping), and testing images assessed only in the cropped segmentation region. Results include the Dice similarity coefficient (DSC), the volume difference (VD), and the average surface distance (ASD). Results are presented as median; mean ± standard deviation. -- means that the results for the testing cropped region were identical to when assessed for the testing full image.

	Cartilage					Bone		
	Femoral	**Medial Tibial**	**Lateral Tibial**	**All Tibial**	**Patellar**	**Femur**	**Tibia**	**Patella**
	OA							
	Testing							
DSC	0.910; 0.907 ± 0.023	0.885; 0.876 ± 0.042	0.918; 0.913 ± 0.026	0.901; 0.897 ± 0.026	0.875; 0.840 ± 0.128	0.990; 0.989 ± 0.006	0.991; 0.987 ±0.022	0.988; 0.986 ± 0.007
VD	1.60; 1.82 ± 6.29	-0.17; 1.37 ± 12.19	0.74; 0.37 ± 6.49	0.17; 0.50 ± 7.11	-1.32; -0.62 ± 29.31	0.36; 0.35 ± 0.85	0.57; 0.16 ± 3.63	-0.13; -0.15 ± 0.97
ASD	0.151; 0.174 ± 0.088	0.161; 0.229 ± 0.218	0.127; 0.145 ± 0.059	0.151; 0.187 ± 0.117	0.197; 0.354 ± 0.640	0.152; 0.179 ± 0.143	0.171; 0.247 ± 0.392	0.068; 0.093 ± 0.133
	Validation							
DSC	0.912; 0.908 ± 0.022	0.885; 0.877 ± 0.041	0.921; 0.914 ± 0.024	0.902; 0.898 ± 0.025	0.876; 0.842 ± 0.126	0.990; 0.989 ± 0.005	0.991; 0.988 ± 0.020	0.987; 0.985 ± 0.007
VD	1.68; 1.56 ± 6.08	-0.06; 0.97 ± 11.80	0.08; 0.54 ± 6.74	-0.10; 0.42 ± 7.15	-1.28; 0.88 ± 23.90	0.31; 0.32 ± 0.75	0.55; 0.16 ± 3.36	-0.11; -0.09 ± 1.18
ASD	0.151; 0.172 ± 0.087	0.158; 0.238 ± 0.255	0.126; 0.145 ± 0.060	0.148; 0.189 ± 0.132	0.193; 0.358 ± 0.691	0.151; 0.170 ± 0.108	0.172; 0.254 ± 0.354	0.068; 0.106 ± 0.205
	Testing - Cropped Region of Interest							
DSC	--	--	--	--	--	0.993; 0.992 ± 0.005	0.994; 0.991 ± 0.024	--
VD	--	--	--	--	--	-0.06; -0.10 ± 0.74	-0.10; -0.50 ± 3.86	-0.14; -0.14 ± 0.99
ASD	--	--	--	--	--	0.080; 0.102 ± 0.131	0.063; 0.096 ± 0.184	0.068; 0.091 ± 0.130
	Healthy							
	Testing							
DSC	0.942; 0.938 ± 0.015	0.913; 0.911 ± 0.015	0.931; 0.930 ± 0.011	0.922; 0.922 ± 0.011	0.958; 0.956 ± 0.013	0.984; 0.984 ± 0.003	0.981; 0.978 ± 0.009	0.984; 0.981 ± 0.007
VD	-0.41; -0.45 ± 2.02	-2.86; -1.66 ± 5.41	-1.01; -1.30 ±3.56	-1.61; -1.53 ± 3.47	0.83; 0.13 ± 2.82	1.11; 1.05 ± 1.01	1.56; 1.62 ± 1.19	-0.16; 0.81 ± 2.06
ASD	0.081; 0.088 ± 0.019	0.125; 0.134 ± 0.037	0.124; 0.127 ± 0.029	0.131; 0.131 ± 0.024	0.062; 0.068 ± 0.021	0.230; 0.240 ± 0.059	0.274; 0.317 ± 0.166	0.091; 0.113 ± 0.048
	Validation							
DSC	0.943; 0.938 ± 0.016	0.914; 0.910 ± 0.016	0.934; 0.930 ± 0.012	0.924; 0.922 ± 0.012	0.958; 0.956 ± 0.012	0.984; 0.984 ± 0.003	0.982; 0.979 ± 0.007	0.983; 0.980 ± 0.007
VD	-0.23; -0.17 ± 2.13	-3.17; -2.36 ± 5.52	-1.18; -1.45 ± 3.78	-2.34; -1.91 ± 3.54	0.21; 0.23 ± 2.84	0.99; 1.03 ± 0.75	1.69; 1.59 ± 1.13	0.00; 0.87 ± 2.23
ASD	0.082; 0.088 ± 0.019	0.123; 0.135 ± 0.041	0.121; 0.127 ± 0.031	0.127; 0.131 ± 0.028	0.059; 0.067 ± 0.020	0.227; 0.239 ± 0.062	0.262; 0.326 ± 0.198	0.091; 0.115 ± 0.051
	Testing - Cropped Region of Interest							
DSC	--	--	--	--	--	0.991; 0.990 ± 0.002	0.992; 0.991 ± 0.003	--
VD	--	--	--	--	--	-0.056; 0.096 ± 0.672	0.021; 0.045 ± 0.452	--
ASD	--	--	--	--	--	0.101; 0.108 ± 0.032	0.078; 0.086 ± 0.038	--

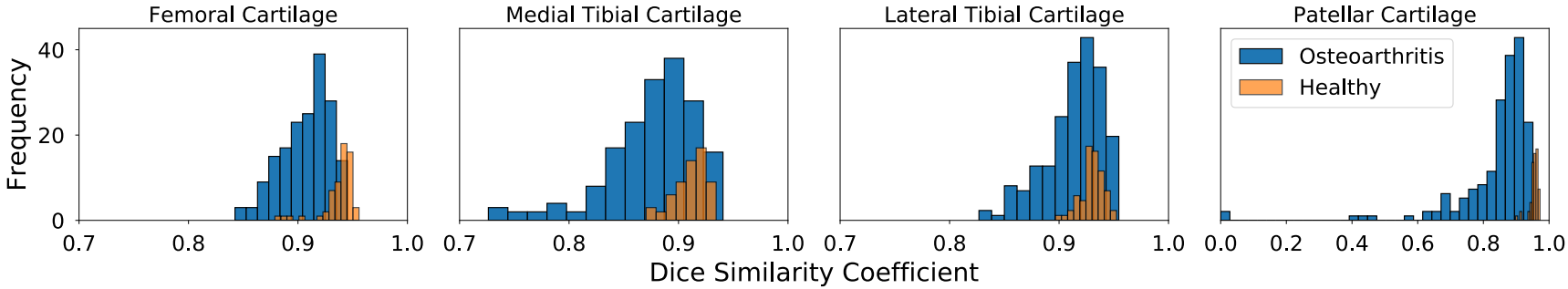


Figure 4-4. Histograms of the Dice similarity coefficients (DSC) from testing data for the four cartilage classes of interest (femoral, patellar, medial tibial, lateral tibial). Blue is the osteoarthritis (OA) sample, orange is the healthy sample.

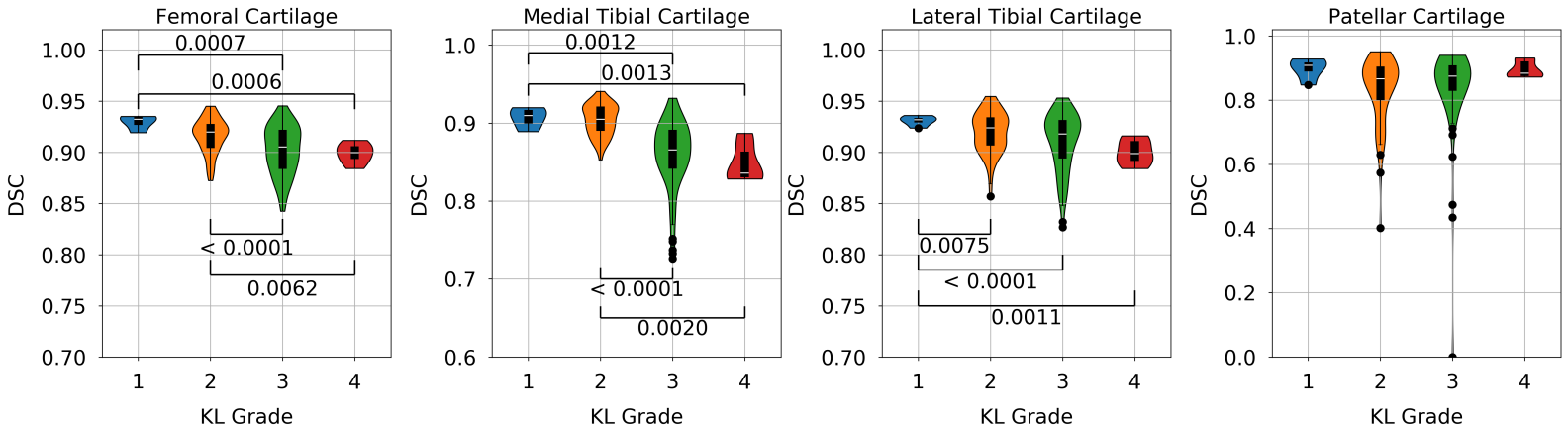


Figure 4-5. Violin plots of Dice similarity coefficient (DSC) by Kellgren and Lawrence (KL) disease severity for individuals with osteoarthritis (OA). One-way analyses of variance were run to compare DSCs across KL grades. Post-hoc significant p-values, at Bonferroni corrected p <0.0125, are included.

In the healthy sample, the mean (SD) DSC values were: femoral cartilage 0.938 (0.015); medial tibial cartilage 0.911 (0.015); lateral tibial cartilage 0.930 (0.011); all tibial cartilage 0.922 (0.011); and patellar cartilage 0.956 (0.013). The average surface distance for all cartilage classes in the healthy sample were <0.135 mm, which is less than one-half of the in-plane resolution (0.3125 mm). The average surface distance for the bone segmentations had higher errors (femur: 0.240 mm, tibia: 0.317 mm, patella: 0.113 mm), however, when only analyzing the cropped region, the errors for the femur (0.108 mm) and tibia (0.086 mm) dropped to <1/3 of the in-plane resolution. Patellar errors were unaffected by the cropping. Segmentation times for the healthy sample were on average 91.4 (10.8) s.

Dependence of DSC on cartilage volume (Figure 4-6**)** was observed in femoral cartilage in the healthy sample (R^2=0.30) and patellar cartilage in the OA sample (R^2=0.33). All other R^2 for correlations between cartilage DSC and volume were <0.14, indicating negligible volume dependence.

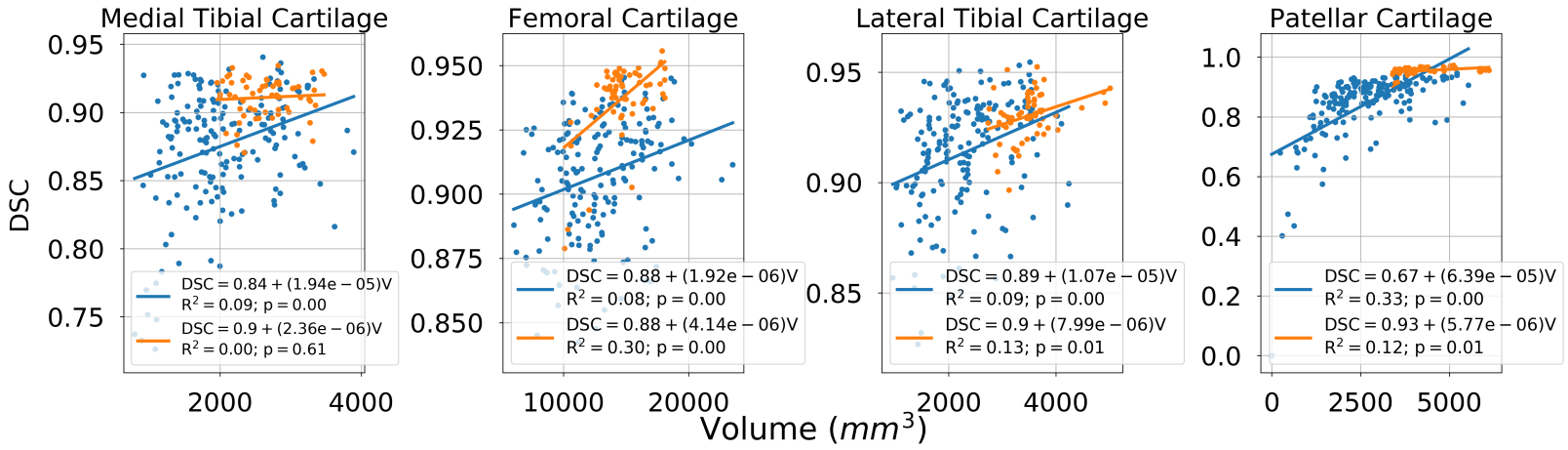


Figure 4-6. Scatter plots and fitted regression lines predicting Dice similarity coefficient (DSC) from cartilage volume. Blue is for individuals with osteoarthritis (OA) and orange for healthy knees. The equation, its R^2 and significance are presented for each fitted line in the respective legends.

The ablation study showed that, other than for volume difference in the medial tibial and lateral tibial cartilage of the healthy group, the full algorithm (Stage 2) outperformed all Stage 1 (2D, 3D, average) segmentations for all metrics (Table 4-3). In the exception, (tibial cartilage regions), the Stage 1 2D model had a small overestimation of volume (medial tibial cartilage: 0.87, lateral tibial cartilage: 0.24) whereas the full model had a small underestimation (medial tibial cartilage: -1.66, lateral tibial cartilage: -1.30). Broadly, the Stage 1 2D model outperformed the Stage 1 3D model, and the segmentation produced by their average.

Comparing the Stage 1 2D model to the full model (Stage 2), in the OA sample, the full model outperformed the Stage 1 2D model in terms of DSC by 0.002 to 0.007 across all tissues. Similarly, in the bones of the healthy knees the full model outperformed in terms of DSC by 0.001 to 0.002. However, in the healthy cartilage there was a bigger difference with the full model outperforming the Stage 1 2D model by 0.016 to 0.022. In terms of average surface distance, the full model performed considerably better than the Stage 1 2D model producing average surface distances between 6.9% (OA medial tibial cartilage) and 42.6% (OA patella) better. The average surface distance had the biggest difference in the bones, where the full model outperformed the Stage 1 2D model by 22.6% (healthy tibia) to 42.6% (OA patella). In cartilage regions the full model outperformed the Stage 1 2D model by 6.9% (medial tibial cartilage) to 10.5% (lateral tibial cartilage) for the OA knees and by 14.2% (lateral tibial cartilage) to 21.4% (femoral cartilage) in the healthy knees.

The Stage 1 (coarse) 3D model was the worst in every region and outcome for both healthy and OA knees. The average prediction (average of Stage 1 2D and 3D predictions) outperformed the Stage 1 2D model for average surface distance of all OA bones (femur 0.251mm versus 0.258mm; tibia 0.308mm versus 0.322mm; patella 0.132mm vs 0.162mm), and effectively tied the Stage 1 2D model for DSC in those same regions. With the exception of these OA bones, the average predictions had outcomes that fell in-between the Stage 1 2D and 3D models.

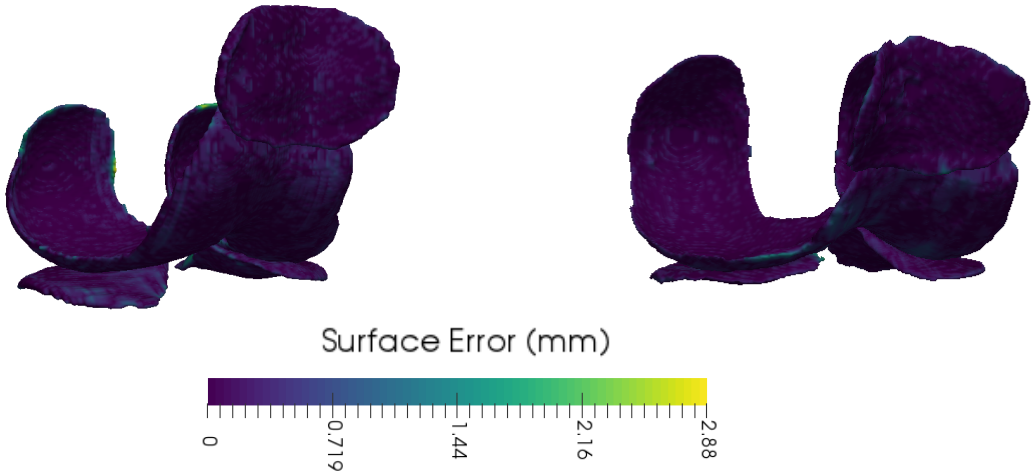


Figure 4-7. Three-dimensional visualizations of automated cartilage segmentations from a healthy participant (left) and one in the osteoarthritis sample (right). The surfaces are coloured based on the surface distance to the reference (manual) segmentations. The majority of the errors occur on the cartilage peripheries where partial-volume effects are common.

Table 4-3. Results of the ablation study analyzing the prediction accuracies of the Stage 1 (coarse) segmentation networks on the testing (held-out) data. The Dice similarity coefficient (DSC), volume difference (VD), and average surface distance (ASD) are presented for each region. Results are presented for the Stage 1 (coarse) 2D model, the Stage 1 (coarse) 3D model, and the average of the Stage 1 2D and 3D models. Each cell includes the mean ± standard deviation. Separately for the osteoarthritis (OA) and healthy datasets, for each tissue and each outcome, the best outcome is bolded; in the event of a tie the prediction with the lower standard deviation is bolded. The OA dataset includes n=88 people and n=176 images, the healthy dataset includes n=15 people and n=60 images. * indicates that the metric outperformed the full model (Stage 2).

Model	Region	DSC	VD	ASD	DSC	VD	ASD
		OA			**Healthy**		
Stage 1 coarse 2D model							
	Femoral cartilage	**0.900 ± 0.025**	**6.59 ± 6.95**	**0.193 ± 0.104**	**0.916 ± 0.013**	**2.32 ± 2.37**	**0.112 ± 0.018**
	Medial tibial cartilage	**0.871 ± 0.045**	**6.13 ± 13.89**	**0.246 ± 0.242**	**0.889 ± 0.015**	**0.87 ± 5.64 ***	**0.164 ± 0.039**
	Lateral tibial cartilage	**0.908 ± 0.028**	**5.05 ± 7.67**	**0.162 ± 0.067**	**0.914 ± 0.010**	**0.24 ± 3.44 ***	**0.148 ± 0.029**
	Patellar cartilage	**0.835 ± 0.129**	**7.30 ± 31.77**	**0.390 ± 0.698**	**0.938 ± 0.012**	**2.44 ± 4.19**	**0.095 ± 0.023**
	Femur bone	**0.986 ± 0.007**	**0.58 ± 1.06**	0.258 ± 0.262	**0.982 ± 0.003**	**1.07 ± 0.93**	**0.260 ± 0.049**
	Tibia bone	0.985 ± 0.023	**0.18 ± 3.88**	0.322 ± 0.459	**0.977 ± 0.008**	**1.66 ± 1.29**	**0.328 ± 0.156**
	Patella bone	**0.980 ± 0.008**	**0.35 ± 1.23**	0.162 ± 0.259	**0.979 ± 0.009**	**1.16 ± 2.32**	**0.144 ± 0.055**
Stage 1 coarse 3D model							
	Femoral cartilage	0.774 ± 0.033	19.43 ± 13.25	0.413 ± 0.091	0.812 ± 0.020	12.76 ± 8.82	0.298 ± 0.030
	Medial tibial cartilage	0.735 ± 0.061	19.89 ± 21.72	0.536 ± 0.617	0.789 ± 0.020	5.08 ± 9.57	0.358 ± 0.171
	Lateral tibial cartilage	0.787 ± 0.043	15.53 ± 15.65	0.387 ± 0.118	0.829 ± 0.019	3.91 ± 5.78	0.364 ± 0.191
	Patellar cartilage	0.705 ± 0.140	36.21 ± 62.44	0.672 ± 0.905	0.855 ± 0.024	8.42 ± 9.73	0.276 ± 0.038
	Femur bone	0.962 ± 0.018	2.32 ± 3.40	0.576 ± 0.277	0.960 ± 0.008	4.12 ± 2.22	0.600 ± 0.139
	Tibia bone	0.965 ± 0.015	1.15 ± 3.36	0.649 ± 0.610	0.960 ± 0.013	3.30 ± 3.04	0.706 ± 0.516
	Patella bone	0.947 ± 0.019	2.21 ± 4.28	0.434 ± 0.790	0.951 ± 0.007	3.24 ± 2.65	0.315 ± 0.067
Average of Stage 1 models							
	Femoral cartilage	0.889 ± 0.026	8.88 ± 7.61	0.208 ± 0.097	0.914 ± 0.013	3.58 ± 2.59	0.115 ± 0.018
	Medial tibial cartilage	0.854 ± 0.049	9.32 ± 15.75	0.270 ± 0.248	0.885 ± 0.017	1.34 ± 5.75	0.170 ± 0.039
	Lateral tibial cartilage	0.894 ± 0.030	7.04 ± 9.10	0.183 ± 0.066	0.909 ± 0.011	0.54 ± 3.05	0.161 ± 0.032
	Patellar cartilage	0.820 ± 0.132	14.19 ± 38.96	0.438 ± 0.824	0.935 ± 0.015	3.84 ± 5.41	0.101 ± 0.026
	Femur bone	0.985 ± 0.007	0.83 ± 1.23	**0.251 ± 0.233**	0.981 ± 0.003	1.58 ± 1.07	0.279 ± 0.061
	Tibia bone	**0.985 ± 0.022**	0.32 ± 3.76	**0.308 ± 0.442**	0.976 ± 0.009	2.03 ± 1.34	0.336 ± 0.165
	Patella bone	0.979 ± 0.009	0.57 ± 1.34	**0.132 ± 0.094**	0.974 ± 0.007	1.58 ± 2.27	0.149 ± 0.060

4.4 Discussion

The presented work describes and tests a novel framework for segmenting cartilage and bone from knee MRI scans using only CNNs. When benchmarked on the OAI dataset, the results demonstrate the best reported DSC and average surface distance for femoral and medial tibial cartilage, and the second best for lateral tibial cartilage (surpassed by one algorithm that was tested on data from 9 individuals (Table 4-4) (Ambellan et al., 2019; Desai et al.,; Gaj et al., 2020; Panfilov et al., 2019). Amongst algorithms with predictions on the whole OAI dataset, we present the best accuracies for all cartilage regions (Table 4-4)(Ambellan et al., 2019). The results on the healthy dataset are the best cartilage segmentations reported to-date. The full model outperformed the Stage 1 (coarse) predictions in overlap (DSC) and particularly surface distance (average surface distance) metrics. The segmentations produced were amongst the fastest times reported (Ambellan et al., 2019; Gaj et al., 2020; Liu et al., 2017; Norman et al., 2018). The proposed algorithm demonstrates an ability to learn, without human intervention, how to automatically segment cartilage and bone from both sagittal FSPGR and DESS MRI sequences on individuals with and without OA. This work enables an efficient quantification of cartilage outcomes for research (basic science, clinical trials) and clinical usage.

Table 4-4. Segmentation accuracies and sample sizes of algorithms trained and tested on the OAI iMorphics dataset. The table includes the Dice similarity coefficient (DSC) as well as the testing datasets sample size in-terms of people and images included. The rows are presented in descending order of sample size. The highest accuracy for each tissue is bolded; in the event of a tie the algorithm tested on the largest sample of data is bolded. -- denotes no data available.

	Testing Sample Size (participants, images)	Femur	Tibia	Medial Tibia	Lateral Tibia	Patella
Ours	88 (176)	**0.907**	**0.897**	**0.876**	0.913	0.840
Ambellan et al. 2019	88 (176)	0.893	--	0.860	0.902	--
Panfilov et al. 2019	36 (72)	0.907	0.897	--	--	**0.871**
Desai et al. 2020	14 (28)	0.9	0.89	--	--	0.86
Gaj et al. 2020	9 (18)	0.897	--	0.861	**0.918**	0.842

The proposed framework failed to match the best cartilage segmentation accuracies for patellar cartilage, when benchmarked on the OAI dataset, achieving mean DSC of 0.840 and mean average surface distance of 0.354 mm. Three algorithms beat the current implementation (Table 4-4)(Desai et al., n.d.; Gaj et al., 2020; Panfilov et al., 2019). All of these previous models tested their results on relatively smaller samples all using less than half of the sample included in this study (Table 4-4). As can be seen in Figure 4-4, there is a long tail of patellar cartilage DSCs on the OAI dataset, indicating a few poor performances reduced accuracy. Figure 4-6 shows that these poor results occurred in knees with low patellar cartilage volume. It is possible that previous algorithms were tested on sub-samples (Desai et al., n.d.; Gaj et al., 2020; Panfilov et al., 2019) that did not include these images of knees with low patellar cartilage volume (Figure 4-6). There are a few potential reasons for this volume dependence. First, as thin structures like cartilage decrease in volume, a greater proportion of the voxels are found on the

boundaries, where the majority of errors occur. Second, in low volume segmentations the cartilage itself is sparse, and at times disconnected, likely introducing greater error in the manual segmentations (Figure 4-8). Still, the framework proposed by Panfilov et al. produced excellent DSCs for patellar cartilage. The authors used a 2D approach with a deeper network and more filters than the current implementation. A bigger 2D coarse network may also improve the current frameworks results. Finally, it should be noted that segmentation of patellar cartilage in the healthy sample produced the highest accuracy cartilage segmentations we are aware of. It is possible that this discrepancy could be attributed to differences in the FSPGR versus DESS sequences or differences between the young healthy versus OA samples.

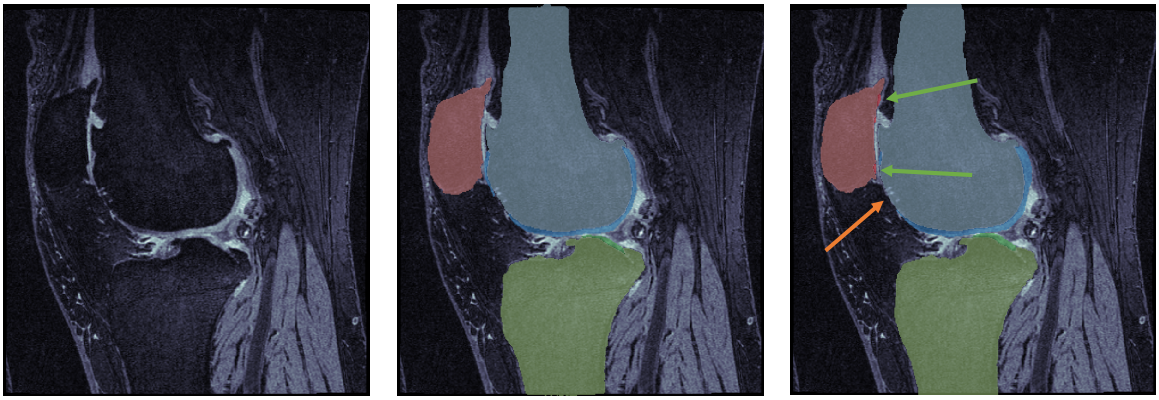


Figure 4-8. Raw image (left) and the manual (middle) and predicted (right) segmentations of a participant with OA from the Osteoarthritis Initiative. Femur (light blue), femoral cartilage (blue), tibia (light green), tibial cartilage (green), patella (light red), and patellar cartilage (red) are overlaid on the raw image. Of note is the region on the trochlea where the framework correctly omitted cartilage (orange arrow). It is also interesting to note that the predicted regions of patellar cartilage (green arrows) are plausibly correct, and therefore potentially indicate an instance of outperforming the reference segmentation.

When benchmarked on the whole OAI segmentation dataset, the current implementation produced bone DSCs for the femur (0.989) and tibia (0.987) that were better than those reported by Ambellan et al. (2019). In terms of average surface distance, when considering the whole image, the current implementation was slightly worse than Ambellan and colleagues (femur: 0.179 mm vs 0.17 mm; tibia: 0.25 mm vs. 0.18 mm), however, the current implementation only produced coarse segmentations for the periphery of the images. When considering only the high-resolution area, the average surface distance for the current implementation drops considerably to be 0.102 mm and 0.096 mm for the femur and tibia respectively. The DSC also improved considerably for the cropped regions, yielding DSCs of 0.993 and 0.994 for the femur and tibia. These cropped regions include osteophytes and all primary classes of interest (Figure 4-3). The comparison to Ambellan et al. is noteworthy because they used multiple stages of segmentations as well as SSMs to control bone shape and minimize errors. The comparable errors between the two approaches indicate that a fully-learned CNN approach does not require the regularization imposed by a SSM.

Segmentation times (~1.5 min) fall between extremely fast algorithms (<15 s) that produce lower-accuracy results (Liu, 2018; Norman et al., 2018), and slower methods (10+ min) (Ambellan et al., 2019) that produce comparable results. One algorithm of note by Gaj et al. reports segmentation times of less than 1 min with comparable segmentation accuracies to the current framework (Gaj et al., 2020). The work presented by Gaj et al.

was tested on MRI scans from 9 individuals and thus inherently has greater variability associated with the DSC point estimate.

In the current implementation, the relatively small OA and healthy datasets create potential short-comings. In particular, probabilities outputted from Stage 1 act as priors in Stage 2 and we trained Stage 1 and Stage 2 using the same images. With more data, Stage 1 and 2 would ideally be trained using different sets of data, letting Stage 2 better learn how to weight priors. In addition to more data, it is possible that different network architectures or loss functions may improve results. For example, a deeper network, such as that by Panfilov et al. (Panfilov et al., 2019) applied to the Stage 1 (coarse) 2D network, may have yielded better performance. Or, since the Stage 1 (coarse) 3D network performed worst, it is possible that using two or three coarse 2D networks trained along orthogonal planes at Stage 1 may be better. Another potential alternative would be to train Stage 1 using multiple similar networks that have different loss functions. For example, one network could be trained to maximize volume overlap using DSC, and other networks could be trained based on surface errors (Caliva et al., 2019). This scheme could provide better information for Stage 2 to predict the final segmentation.

To date, the gold standard of cartilage segmentation is manual. The only work to test manual segmentations tested the inter-radiologist DSC of all tibial, femoral, and patellar cartilage as one segmented region with one label (Shim et al., 2009). That work showed a

mean DSC of 0.878 for 10 individuals from the OAI with primarily mild OA (KL0: 1, KL1: 3, KL2: 4, KL3: 2). The cartilage DSCs for 3 of 4 individual compartments (femur, medial tibial, lateral tibia) from the current framework, and from many published results using deep learning (Ambellan et al., 2019; Desai et al., n.d.; Gaj et al., 2020; Panfilov et al., 2019), surpass this level (Shim et al., 2009). The only class that falls short is the patellar cartilage with a DSC of 0.84-0.87 from the 4 best algorithms. When assessing the DSC of all femoral, tibial, and patellar cartilage as one label on predictions from the current framework, the mean (SD) DSC of the OAI dataset was 0.901 (0.023). By these standards, many currently published deep learning algorithms are more accurate than the current gold-standard, manual segmentation. This may also be why many algorithms are converging to similar accuracies (Table 4-4) as we may be approaching the highest possible metrics when the benchmark is manual segmentation; this is highlighted by the plausibly correct segmentations predicted in Figure 4-8 where the manual segmentations said there was no cartilage. These performances are also achieved in timescales measured in seconds (Liu, 2018; Norman et al., 2018) or minutes (Ambellan et al., 2019; Gaj et al., 2020) as opposed to hours (Shim et al., 2009). Nonetheless, algorithms make mistakes (Ambellan et al., 2019; Gaj et al., 2020). Spot-checking segmentations more likely to include errors (low cartilage volume, high KL) could enable efficient analysis of data while maintaining high fidelity. Checking segmentations for accuracy is recommended for individual analysis such as that required by clinical workflows. However, analyses of large datasets allow small errors to be overcome by group statistics.

4.5 Conclusions

The proposed multi-stage CNN segmentation framework provides excellent accuracies when segmenting knee cartilage from OAI DESS images in an average of 1.5 min. Cartilage segmentation accuracies on the healthy dataset are the best reported to-date. The segmentation framework is flexible and fully learns from provided examples, therefore showing promise for segmenting other musculoskeletal tissues in future work. Furthermore, a single framework was trained to handle knees across the OA severity spectrum, and from different MRI vendors and sequences. Together, these results demonstrate an ability to efficiently analyze cartilage outcomes for basic science, clinical trials, and clinical usage.

Acknowledgment

We would like to acknowledge Google for providing cloud compute credits used to conduct the experiments. The Osteoarthritis Initiative (OAI) is a public-private partnership funded by the National Institutes of Health (NIH) and private partners including Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. This manuscript was prepared using an OAI public use data set and does not reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

Funding

A. A. Gatti was supported by an Ontario Graduate Scholarship, The Arthritis Society, and a Mitacs Accelerate Entrepreneur award. M.R. Maly holds The Arthritis Society Stars Mid-Career Development Award funded by the Canadian Institutes of Health Research-Institute of Musculoskeletal Health and Arthritis and an NSERC Discovery grant that supported this work (MRM: 353715).

Conflicts of interest

A. A Gatti is the founder of NeuralSeg, Ltd. There are no other conflicts of interest to disclose.

4.6 References

Ahn, C., Bui, T.D., Lee, Y., Shin, J., Park, H., 2016. Fully automated, level set-based segmentation for knee MRIs using an adaptive force function and template: data from the osteoarthritis initiative. BioMedical Engineering OnLine 15. https://doi.org/10.1186/s12938-016-0225-7

Ambellan, F., Tack, A., Ehlke, M., Zachow, S., 2019. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the Osteoarthritis Initiative. Medical Image Analysis 52, 109–118. https://doi.org/10.1016/j.media.2018.11.009

Caliva, F., Iriondo, C., Martinez, A.M., Majumdar, S., Pedoia, V., 2019. Distance Map Loss Penalty Term for Semantic Segmentation. arXiv:1908.03679 [cs, eess].

Chollet, F., 2015. Keras.

Conaghan, P.G., Hunter, D.J., Maillefert, J.F., Reichmann, W.M., Losina, E., 2011. Summary and recommendations of the OARSI FDA osteoarthritis Assessment of Structural Change Working Group. Osteoarthritis and Cartilage 19, 606–610. https://doi.org/10.1016/j.joca.2011.02.018

Creamer, P., Hochberg, M.C., 1997. Osteoarthritis. Lancet 350, 503–508. https://doi.org/10.1016/S0140-6736(97)07226-7

Dam, E.B., Lillholm, M., Marques, J., Nielsen, M., 2015. Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from

the osteoarthritis initiative. J Med Imaging (Bellingham) 2, 024001. https://doi.org/10.1117/1.JMI.2.2.024001

Desai, A.D., Caliva, F., Iriondo, C., Khosravan, N., Mortazi, A., Jambawalikar, S., Torigian, D., Ellerman, J., Akcakaya, M., Bagci, U., Tibrewala, R., n.d. The International Workshop on Osteoarthritis Imaging Knee MRI Segmentation Challenge: A Multi-Institute Evaluation and Analysis Framework on a Standardized Dataset. *ArXiv:2004.14003 [Cs, Eess]*. http://arxiv.org/abs/2004.14003

Deshpande, B.R., Katz, J.N., Solomon, D.H., Yelin, E.H., Hunter, D.J., Messier, S.P., Suter, L.G., Losina, E., 2016. Number of Persons With Symptomatic Knee Osteoarthritis in the US: Impact of Race and Ethnicity, Age, Sex, and Obesity: Symptomatic Knee OA in the US. Arthritis Care & Research. https://doi.org/10.1002/acr.22897

Dodin, P., Pelletier, J., Martel-Pelletier, J., Abram, F., 2010. Automatic Human Knee Cartilage Segmentation From 3-D Magnetic Resonance Images. IEEE Transactions on Biomedical Engineering 57, 2699–2711. https://doi.org/10.1109/TBME.2010.2058112

Duryea, J., Neumann, G., Brem, M.H., Koh, W., Noorbakhsh, F., Jackson, R.D., Yu, J., Eaton, C.B., Lang, P., 2007. Novel fast semi-automated software to segment cartilage for knee MR acquisitions. Osteoarthritis and Cartilage 15, 487–492. https://doi.org/10.1016/j.joca.2006.11.002

Fripp, J., Crozier, S., Warfield, S.K., Ourselin, S., 2010. Automatic Segmentation and Quantitative Analysis of the Articular Cartilages From Magnetic Resonance Images of the Knee. IEEE Transactions on Medical Imaging 29, 55–64. https://doi.org/10.1109/TMI.2009.2024743

Gaj, S., Yang, M., Nakamura, K., Li, X., 2020. Automated cartilage and meniscus segmentation of knee MRI with conditional generative adversarial networks. Magnetic Resonance in Medicine 84, 437–449. https://doi.org/10.1002/mrm.28111

Gatti, A.A., 2018. NEURALSEG: state-of-the-art cartilage segmentation using deep learning–analyses of data from the osteoarthritis initiative. Presented at the Osteoarthritis Research Society International. https://doi.org/10.1016/j.joca.2018.02.110

Gatti, A.A., Noseworthy, M.D., Stratford, P.W., Brenneman, E.C., Totterman, S., Tamez-Peña, J., Maly, M.R., 2017. Acute changes in knee cartilage transverse relaxation time after running and bicycling. Journal of Biomechanics 53, 171–177. https://doi.org/10.1016/j.jbiomech.2017.01.017

Hastie, T., Tibshirani, R., Friedman, J.H., 2009. The elements of statistical learning: data mining, inference, and prediction, 2nd ed. ed, Springer series in statistics. Springer, New York, NY.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, in: 2015 IEEE

International Conference on Computer Vision (ICCV). Presented at the 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, Santiago, Chile, pp. 1026–1034. https://doi.org/10.1109/ICCV.2015.123

Hunter, D.J., Altman, R.D., Cicuttini, F., Crema, M.D., Duryea, J., Eckstein, F., Guermazi, A., Kijowski, R., Link, T.M., Martel-Pelletier, J., Miller, C.G., Mosher, T.J., Ochoa-Albíztegui, R.E., Pelletier, J.-P., Peterfy, C., Raynauld, J.-P., Roemer, F.W., Totterman, S.M., Gold, G.E., 2015. OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis. Osteoarthritis and Cartilage 23, 698–715. https://doi.org/10.1016/j.joca.2015.03.012

Kayalibay, B., Jensen, G., van der Smagt, P., 2017. CNN-based Segmentation of Medical Imaging Data. arXiv:1701.03056 [cs].

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

Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs]. http://arxiv.org/abs/1412.6980

Kraus, V.B., Blanco, F.J., Englund, M., Karsdal, M.A., Lohmander, L.S., 2015. Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. Osteoarthritis and Cartilage 23, 1233–1241. https://doi.org/10.1016/j.joca.2015.03.036

Liu, F., 2018. SUSAN: segment unannotated image structure using adversarial network. Magnetic Resonance in Medicine. https://doi.org/10.1002/mrm.27627

Liu, F., Zhou, Z., Jang, H., Samsonov, A., Zhao, G., Kijowski, R., 2017. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging: Deep Learning Approach for Segmenting MR Image. Magnetic Resonance in Medicine. https://doi.org/10.1002/mrm.26841

Metcalfe, A.J., Andersson, M.L., Goodfellow, R., Thorstensson, C.A., 2012. Is knee osteoarthritis a symmetrical disease? Analysis of a 12 year prospective cohort study. BMC Musculoskeletal Disorders 13. https://doi.org/10.1186/1471-2474-13-153

Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. arXiv:1606.04797v1.

Norman, B., Pedoia, V., Majumdar, S., 2018. Use of 2D U-Net Convolutional Neural Networks for Automated Cartilage and Meniscus Segmentation of Knee MR Imaging Data to Determine Relaxometry and Morphometry. Radiology 172322. https://doi.org/10.1148/radiol.2018172322

Panfilov, E., Tiulpin, A., Klein, S., Nieminen, M.T., Saarakkala, S., 2019. Improving Robustness of Deep Learning Based Knee MRI Segmentation: Mixup and Adversarial Domain Adaptation, in: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Presented at the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), IEEE, Seoul, Korea (South), pp. 450–459. https://doi.org/10.1109/ICCVW.2019.00057

Pang, S., Du, A., He, X., Díez, J., Orgun, M.A., 2019. Fast and Accurate Lung Tumor Spotting and Segmentation for Boundary Delineation on CT Slices in a Coarse-to-Fine Framework, in: Gedeon, T., Wong, K.W., Lee, M. (Eds.), Neural Information Processing, Communications in Computer and Information Science. Springer International Publishing, Cham, pp. 589–597. https://doi.org/10.1007/978-3-030-36808-1_64

Pedoia, V., Majumdar, S., Link, T.M., 2016. Segmentation of joint and musculoskeletal tissue in the study of arthritis. Magnetic Resonance Materials in Physics, Biology and Medicine 29, 207–221. https://doi.org/10.1007/s10334-016-0532-9

Peterfy, C., Woodworth, T., Altman, R., 2006. Workshop for Consensus on Osteoarthritis Imaging: MRI of the knee. Osteoarthritis and Cartilage 14, 44–45. https://doi.org/10.1016/j.joca.2006.02.025

Peterfy, C.G., Schneider, E., Nevitt, M., 2008. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis and Cartilage 16, 1433–1441. https://doi.org/10.1016/j.joca.2008.06.016

Prasoon, A., Igel, C., Loog, M., Lauze, F., Dam, E.B., Nielsen, M., 2013. Femoral cartilage segmentation in knee MRI scans using two stage voxel classification, in: Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. IEEE, pp. 5469–5472.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Roth, H.R., Oda, H., Zhou, X., Shimizu, N., Yang, Y., Hayashi, Y., Oda, M., Fujiwara, M., Misawa, K., Mori, K., 2018. An application of cascaded 3D fully convolutional networks for medical image segmentation. Computerized Medical Imaging and Graphics 66, 90–99. https://doi.org/10.1016/j.compmedimag.2018.03.001

Shan, L., Zach, C., Charles, C., Niethammer, M., 2014. Automatic atlas-based three-label cartilage segmentation from MR knee images. Medical Image Analysis 18, 1233–1246. https://doi.org/10.1016/j.media.2014.05.008

Shen, D., Wu, G., Suk, H.-I., 2017. Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering 19, 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

Shim, H., Chang, S., Tao, C., Wang, J.-H., Kwoh, C.K., Bae, K.T., 2009. Knee cartilage: efficient and reproducible segmentation on high-spatial-resolution MR images with the semiautomated graph-cut algorithm method. Radiology 251, 548–556.

Tamez-Pena, J.G., Farber, J., Gonzalez, P.C., Schreyer, E., Schneider, E., Totterman, S., 2012. Unsupervised Segmentation and Quantification of Anatomical Knee Features: Data From the Osteoarthritis Initiative. IEEE Transactions on

Biomedical Engineering 59, 1177–1186. https://doi.org/10.1109/TBME.2012.2186612

Wang, Q., Wu, D., Lu, L., Liu, M., Boyer, K.L., Zhou, S.K., n.d. Semantic Context Forests for Learning-Based Knee Cartilage Segmentation in 3D MR Images 10.

Williams, T.G., Holmes, A.P., Bowes, M., Vincent, G., Hutchinson, C.E., Waterton, J.C., Maciewicz, R.A., Taylor, C.J., 2010. Measurement and visualisation of focal cartilage thickness change by MRI in a study of knee osteoarthritis using a novel image analysis tool. The British Journal of Radiology 83, 940–948. https://doi.org/10.1259/bjr/68875123

Yin Yin, Xiangmin Zhang, Williams, R., Xiaodong Wu, Anderson, D.D., Sonka, M., 2010. LOGISMOS—Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces: Cartilage Segmentation in the Knee Joint. IEEE Transactions on Medical Imaging 29, 2023–2037. https://doi.org/10.1109/TMI.2010.2058861

Yu, L., Yang, X., Chen, H., Qin, J., Heng, P.-A., 2017. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images.

Zhou, Z., Zhao, G., Kijowski, R., Liu, F., 2018. Deep convolutional neural network for segmentation of knee joint anatomy. Magnetic Resonance in Medicine 80, 2759–2770. https://doi.org/10.1002/mrm.27229

Zhu, Z., Xia, Y., Xie, L., Fishman, E.K., Yuille, A.L., 2019. Multi-Scale Coarse-to-Fine Segmentation for Screening Pancreatic Ductal Adenocarcinoma. arXiv:1807.02941 [cs].

5 Investigating Acute Changes in Osteoarthritic Cartilage by Integrating Joint Reaction Forces and Statistical Shape Models: Data from the Osteoarthritis Initiative

The following manuscript has been prepared for submission to the Journal of Biomechanics.

Investigating Acute Changes in Osteoarthritic Cartilage by Integrating Joint Reaction Forces and Statistical Shape Models of Bone: Data from the Osteoarthritis Initiative

Journal of Biomechanics

[1,5]Anthony A. Gatti, MSc, [2]Peter J. Keir, PhD, [3,4]Michael D. Noseworthy, PhD, [1,2,6]Monica R. Maly, PT PhD

[1]School of Rehabilitation Sciences, [2]Department of Kinesiology, [3]School of Biomedical Engineering, and [4]School of Electrical and Computer Engineering, McMaster University, Hamilton, Canada
[5]NeuralSeg Ltd., Hamilton, Canada
[6]Department of Kinesiology, University of Waterloo, Waterloo, Canada

Corresponding Author:
Monica R. Maly
University of Waterloo
200 University Ave,
Waterloo, Ontario, Canada,
N2L 3G1
Fax: +1 519 885 4070
Email: mrmaly@uwaterloo.ca

Abstract word count: 249
Manuscript word count: 3498
Figure: 9
Table: 1
Keywords: Statistical parametric mapping, statistical shape model, joint reaction forces, cartilage, magnetic resonance imaging, osteoarthritis, T_2, aerobic exercise

Abstract

Purpose: This proof-of-principle study integrates measures of joint reaction forces (JRF) and bone shape to assess the acute changes to cartilage caused by walking and cycling in women with symptomatic knee osteoarthritis. **Methods:** Sixteen women with symptomatic knee OA were recruited. All completed biomechanical assessments to estimate JRFs during walking and cycling. Subsamples had magnetic resonance imaging (MRI) performed immediately before and after a 25-minute walking (n=7; mean (SD): 62.2(4.9) kg, 28.1(5.9) kg/m^2) and/or cycling (n=9; mean (SD): 62.9(5.7) kg, 28.6(5.5) kg/m^2) activity. MRI scans were obtained to assess cartilage shape and composition (T_2 relaxation time). Bone shape was quantified using a statistical shape model (SSM), which was built using MRI scans from 100 age, height, body mass, and OA-severity matched women from the Osteoarthritis Initiative. Cartilage change caused by activity and correlations between cartilage change with JRFs and SSM features were quantified using statistical parametric mapping (SPM). **Results:** Cartilage thickness and T_2 decreased in the tibial plateau after activity. On the femur, T_2 decreased or increased in different regions, dependent on the activity. Femur and tibia SSMs had features characteristic of osteoarthritis severity. Greater tibiofemoral JRF was associated with more cartilage deformation on the lateral femoral condyle after walking. Knees more consistent with osteoarthritis resulted in smaller decreases in tibial cartilage thickness. **Conclusions:** Walking and cycling caused distinct patterns of cartilage deformation dependent on knee JRFs and bone morphology. For the first time, these results show that cartilage deformation is dependent on bone shapes and JRFs in vivo.

5.1 Introduction

The integrated joint system (IJS) theory (Edd et al., 2018) suggests that joint disease disrupts homeostasis between joint mechanics and tissue characteristics. IJS states that relationships between mechanics and tissue characteristics vary across the disease spectrum. For example, in healthy joints, greater cartilage thickness and quality (denser proteoglycans) are related to larger joint loads (Koo et al., 2011; Scanlan et al., 2013; Souza et al., 2012) and cartilage surface pressures (Van Rossom et al., 2017), indicating adaptation to the mechanics. Yet, in early OA, there is little to no correlation between joint loads and cartilage thickness (Blazek et al., 2014; Erhart-Hledik et al., 2015; Vanwanseele et al., 2010). In late OA, there is a negative correlation between loading and cartilage thickness, suggesting loading exacerbates pathology (Erhart-Hledik et al., 2015; Maly et al., 2015).

There is little work evaluating how mechanics influence the acute response of cartilage to loading. Higher peak (Boocock et al., 2009) but not cumulative (Gatti et al., 2017) loads contribute to greater cartilage deformation, measured as thickness and T_2 change, in healthy adults. T_2 is a magnetic resonance imaging (MRI) measure of mobile water in cartilage (Choi and Gold, 2011). Based on the triphasic theory of cartilage, water is redistributed away from areas of load application (Lu and Mow, 2008), decreasing T_2 in these regions. Under static load, those with OA trended towards greater deformation, suggesting poorer ability to attenuate load (Cotofana et al., 2011). Quantification of the acute response of cartilage to physiologic loads across the disease spectrum is needed.

Bone has an important relationship with cartilage and mechanics. Higher bone mineral density is associated with greater joint loading regardless of knee health, indicating that bone mineral density adapts to mechanics (Edd et al., 2018; Teichtahl et al., 2015). In healthy knees, higher bone mineral density is positively related to cartilage thickness (Babel et al., 2020). In OA, the opposite is true: higher bone mineral density is related to thinner cartilage (Omoumi et al., 2019). IJS theory suggests cartilage thickness and bone mineral density reflect concurrent adaptive responses to increased loads, until cartilage can no longer adapt, the homeostatic relationship is broken, and a degenerative process is initiated.

Bone shape is another critical consideration in joint health. Three-dimensional bone shape quantified using statistical shape models (SSMs) is associated with structural disease severity (Barr et al., 2015). Neogi and colleagues predicted radiographic knee OA progression using SSMs of bone; OA was characterized by widening and flattening of the medial joint surfaces, with a ridge of osteophytic growth (Neogi et al., 2013). Biomechanical modeling, exploring the effect of bone shape on joint mechanics (Clouthier et al., 2019), suggests that OA bone shapes (Neogi et al., 2013) increased medial knee contact pressures, providing a mechanical pathway to cartilage degeneration (Clouthier et al., 2019). It remains unclear whether relationships between cartilage, bone shape and joint mechanics exist *in vivo*.

This proof-of-principle study integrates biomechanics, SSMs, and cartilage outcomes in an experimental study of joint tissues and mechanical loading *in vivo*. In women with symptomatic knee OA, the primary objectives were to determine (1) whether cartilage thickness and T_2 changed after activity; and (2) whether changes in cartilage thickness and T_2 correlated with joint reaction forces (JRFs). We also conducted exploratory analyses of bone shape as it relates with cartilage and JRFs to identify (3) whether bone shape features correlate with changes in cartilage thickness and T_2; and (4) relationships between tibiofemoral and patellofemoral JRFs from gait with bone shape features. We hypothesized that (1) cartilage thickness and T_2 would decrease in the medial and lateral tibia, the weightbearing regions of the femur, and the trochlea; and (2) absolute tibiofemoral and patellofemoral JRFs would be inversely related to cartilage change in these same regions. Due to the lack of *a priori* knowledge of the SSM shape features, objectives 3 and 4 were exploratory.

5.2 Methods

5.2.1 Participants

We recruited sixteen women >50 years of age with symptomatic knee OA, according to the American College of Rheumatology criteria (Altman et al., 1986), and Lower Extremity Functional Scale (LEFS) scores between 30 and 71 (Kennedy et al., 2011). Exclusion included any of: rheumatoid arthritis, gout, unstable angina, acute lower-limb injury in the preceding 3-months, and contraindications to MRI. The Get Active Questionnaire (GAQ) was used to ensure exercise was safe (*Pre-Screening for Physical*

Activity: Get Active Questionnaire, n.d.). Descriptive statistics included age, height, body mass, body mass index, inseam, 25-minute walking speed, cycling cadence and power output, and LEFS scores.

5.2.2 Visits

This experimental study included three visits: one for biomechanical analyses and two to obtain MRI scans immediately before and after cycling and walking. MRI visits were scheduled in block-randomized order.

5.2.2.1 Biomechanics visit

The biomechanics visit included four steps: First, we standardized walking speed to a Froude number (Fr) of 0.25, which produces ~3% cartilage strain after 25 min of walking (Paranjape et al., 2019). The Fr, a unitless scalar, normalizes walking speed to leg length where $walking\ speed\ = \sqrt{(Fr)\ (leg\ length;\ m)\ (gravity;\ 9.81\ m/s^2)}$. Participants completed a 100 m walk at their fastest self-selected speed. If their fastest speed was below the Fr-defined speed, the lower speed was used. Second, bicycle-fit using an inseam-based equation and handlebars set based on rider preference (Gatti et al., 2019, 2020) was determined. Third, an incremental cycling protocol (Beekley et al., 2004) to determine moderate cycling intensity (i.e., steady-state heart rate between 70-75% of age-predicted maximum (208-0.7*Age) (Gatti et al., 2020; Tanaka et al., 2001)). Last,

collection of motion and force data during gait and cycling using the bicycle-fit and intensity from these protocols.

5.2.2.2 Magnetic resonance imaging (MRI) visits

Visits two and three required participants to arrive at 8 am and rest, laying supine, for 30 min (Gatti et al., 2017). Then, participants underwent pre-activity scanning, followed by a 25 min activity (walking or cycling) and finally post-activity MRI scans. Activities were performed at the intensities determined at the biomechanics visit.

5.2.3 Patellofemoral and tibiofemoral joint reaction forces

JRFs were modeled from kinematic and kinetic data collected during walking and cycling. Participants wore 40 markers attached to anatomical locations (Supplemental 1). Marker data were collected at 112.5 Hz (Motion Analysis Corporation, Santa Rosa, CA). During walking, synchronous force plate data was collected at 1125 Hz (OR6-7, AMTI, MA, USA). During cycling, synchronous three-dimensional pedal forces and moments were collected at 450 Hz (Science To Practice, Ljubljana, Slovenia). Marker positions and kinetic data were filtered using a second-order low-pass dual-pass Butterworth filter with a cut-off frequency of 6 Hz. Motion and force data were processed as previously published (Gatti et al., 2020).

Functional knee joint centres were fit from cycling data (Ehrig et al., 2006). Hip joint centres were determined using the Harrington method (Harrington et al., 2007). A 16 degrees-of-freedom lower-body model with 86 musculotendon actuators was scaled with Opensim (Delp et al., 2007). Joint kinematics and dynamics were calculated using inverse methods (Delp et al., 2007). Static optimization with muscle weightings validated using *in vivo* JRFs estimated muscle forces (Gatti et al., 2020; Steele et al., 2012). Using the OpenSim Joint Reaction Tool, the patellofemoral JRF was calculated as the sagittal plane resultant force, and the tibiofemoral JRF as the compressive component of the tibiofemoral JRF defined along the long axis of the tibia (Steele et al., 2012). Reaction forces were extracted using all revolutions from the last minute of each cycling bout and 5 gait trials (Gatti and Maly, 2019a). Data were divided into individual pedal revolutions or stance phases to generate median ensemble curves for each reaction force and activity; the absolute peak JRF was extracted from each ensemble.

5.2.4 Magnetic resonance imaging data acquisition and analysis

MR images were acquired using a dedicated quadrature transmit and 16-channel receive knee coil array (Invivo Corp) in a 3-Tesla GE Discovery MR750 (GE Healthcare). Two sequences were collected before and after activity. Following a 3-plane localizer scan a commercial sagittal multi-echo spin echo (MESE) sequence (CartiGram; GE Healthcare) for T_2 mapping (Gatti et al., 2017) (relaxation time (TR) 2450ms, 8 echo times (TE) at multiples of 6.312 ms, in-plane resolution 0.625x0.625 mm, slice thickness 3 mm, slice spacing 1mm) was acquired. For segmentation, a 3D fat-saturated T1-weighted sagittal

fast spoiled gradient recalled (fSPGR) sequence was acquired (TR 17.388 ms, TE 5.832 ms, in-plane resolution 0.3125x0.3125 mm, slice thickness 1 mm, slice spacing 0 mm). Pre-activity, the fSPGR scan was collected first and the MESE second. Post-activity, the MESE was acquired first.

5.2.4.1 Calculation of cartilage thickness and T_2

fSPGR images were segmented using a convolutional neural network (CNN), (Gatti, 2018; Gatti and Maly, 2019b) then manually checked (Fedorov et al., 2012). Bone and cartilage surfaces were created by applying a Gaussian filter to the binary mask from each tissue (bone $\boldsymbol{\sigma}^2$ = 1.0 mm, cartilage $\boldsymbol{\sigma}^2$ = 0.625 mm) followed by surface extraction using the marching cubes algorithm (Lorensen and Cline, n.d.; Mun and Kim, 2017). Bone surfaces were resampled to have 10,000 vertices using voronoi clustering (Valette et al., 2008; Valette and Chassery, 2004). Cartilage thickness was calculated for each bone vertex by projecting vectors normal to the surface and calculating the 3D Euclidean distance from the bone to articular surface. T_2 maps were created by fitting a mono-exponential decay curve (Equation 3) to the signal intensities (SI) collected at the 8 TEs using a Levenberg-Marquardt algorithm, where the intercept is equivalent to the mobile proton density (PD) (Gatti et al., 2017). T_2 map post-processing excluded voxels with T_2 > 100ms, and R^2 <0.7 (Gatti et al., 2017). Furthermore, T_2 values were assigned to the bone meshes by projecting vectors normal to the surface. When vectors intersected cartilage, all T_2 cartilage voxels along the vector were averaged and assigned to the vertex (MacKay et al., 2020).

Equation 3 $$SI(TE) = PD \times e^{-TE/T_2}$$

5.2.4.2 Statistical shape modelling

The SSM was built using scans acquired from 113 women. Of these, 13 scans were acquired from the women we recruited, and 100 from the Osteoarthritis Initiative (OAI) https://nda.nih.gov/oai (Peterfy et al., 2008). From the OAI, stratified sampling was used across 5 Kellgren-Lawrence (KL) grades of severity (Kellgren and Lawrence, 1957). For each KL grade, 20 sets of demographics were randomly sampled from a multivariate normal distribution based on demographics (age, height, and body mass) of the 13 participants recruited for this study. The OAI participant that best matched each set of demographics using a z-score normalized root mean squared (RMS) error was included. For each of these OAI participants, baseline knee images (sagittal dual echo in the steady state (DESS) with in-plane resolution 0.365x0.365 mm, slice thickness 0.7 mm and no gaps (Peterfy et al., 2008)) were extracted and segmented using the CNN.

The SSM was built in four steps. (i) For each segmentation, a 3D surface with 10,000 vertices was generated. (ii) A reference knee, with the smallest RMS surface error to all other knees after rigid plus scaling registration using the iterative closest point algorithm, was identified; all knees were registered to it. (iii) Point correspondences between the reference and every other bone were found using the Feature Oriented Correspondence

using Spectral Regularization algorithm with minimum and maximum curvature as the features (Lombaert et al., 2013; Pedoia et al., 2015). (iv) Singular value decomposition was used to find the principal modes of variation of these correspondences. SSM mode scores were generated by projecting each mesh onto the mode vectors. SSM mode 1 explains the most variance, each subsequent mode explains less variance than the previous one. We tested correlations between cartilage outcomes and bone shape using SSM mode 1; and relationships between knee JRFs and the first three SSM modes.

5.2.5 Statistical analysis

5.2.5.1 Statistical parametric mapping

Statistical parametric mapping (SPM) was used to perform statistical analyses continuously over the cartilage surfaces. Cluster-wise inference was performed using a permutation method (10,000 permutations) to minimize the risk of committing a Type I error and assumptions involved with identifying field smoothness and stationarity (Cox et al., 2017; Eklund et al., 2016; Legendre et al., 1998; Nichols and Holmes, 2002). Test statistics (z-statistics for one-sample difference tests, and t-statistics for correlations) were computed for every permutation. Clusters above thresholds ($p=0.05$, 0.025, 0.01, 0.005, and 0.001) were identified. The maximum contiguous cluster area for every permutation and threshold was recorded. Significance of a cluster at a given threshold was identified based on its position in the permutation distribution. Clusters with $p<0.05$ at 2 or more threshold levels are presented.

For objective 1, we identified whether significant clusters of T_2 or thickness change occurred using a one-sample difference test. For objective 2, we identified whether significant clusters of correlation between T_2 or thickness change and either patellofemoral or tibiofemoral JRFs existed. In our exploratory analysis, we determined whether significant clusters of correlation between T_2 or thickness change and the tibia or femur SSM modes 1 existed. SPM analyses were conducted for: only walking data, only cycling data, and data from walking and cycling combined.

5.2.5.2 Regression analyses

For objective 4, we computed simple linear regressions between each of tibiofemoral compression and the resultant patellofemoral JRF during gait with each of the first 3 SSM modes.

5.3 Results

5.3.1 Participants

Data collection was interrupted by COVID-19 restrictions. Sixteen women were recruited and scheduled. Of these, three completed all visits (biomechanics and 2 MRI). Ten participants completed the biomechanics visit and at least one MRI visit (n=4 walking visit, n=6 cycling visit). Demographics are included in Table 5-1. Three of 13 participants (only 1 that completed the walking MRI visit) were unable to maintain the

Fr-defined walking speed (Table 5-1). These participants were negligibly slower (by 0.001 to 0.045 m/s) than the Fr-defined speed.

Table 5-1. Demographics (mean ± standard deviation) of women that completed just the cycling MRI visit, just the walking MRI visit, or both MRI visits. LEFS = Lower Extremity Functional Scale.

	cycling MRI visit (n=6)	Walking MRI visit (n=4)	Both MRI visits (n=3)
Age (years)	61.7 ± 6.5	59.8 ± 5.7	65.3 ± 3.5
Height (m)	1.60 ± 0.04	1.66 ± 0.04	1.56 ± 0.03
Mass (kg)	76.1 ± 10.4	81.2 ± 10.4	64.2 ± 19.9
Body mass index (kg/m^2)	29.8 ± 4.0	29.5 ± 4.1	26.2 ± 7.6
Inseam length (m)	0.75 ± 0.03	0.77 ± 0.03	0.73 ± 0.01
Walking speed (m/s)	1.42 ± 0.04	1.46 ± 0.04	1.39 ± 0.04
Cadence (rpm)	67.5 ± 6.1		70.3 ± 5.5
Power (watts)	52.2 ± 21.6		52.7 ± 5.7
LEFS score	58.7 ± 8.8	70.0 ± 0.8	57.0 ±2.7
n maintained Froude speed	4	4	2

5.3.2 Cartilage response to activity

Cycling, as well as walking and cycling analyzed together, reduced cartilage thickness on the lateral tibial plateau (Figure 5-1). Cycling reduced T_2 on the medial tibial plateau (Figure 5-1) and increased T_2 along the lateral condyle. On the lateral femoral condyle, cycling also caused a notable region of large, but not significant, decrease in T_2 adjacent to the region of significant increase (Figure 5-3). Walking decreased T_2 at the lateral trochlea (Figure 5-2).

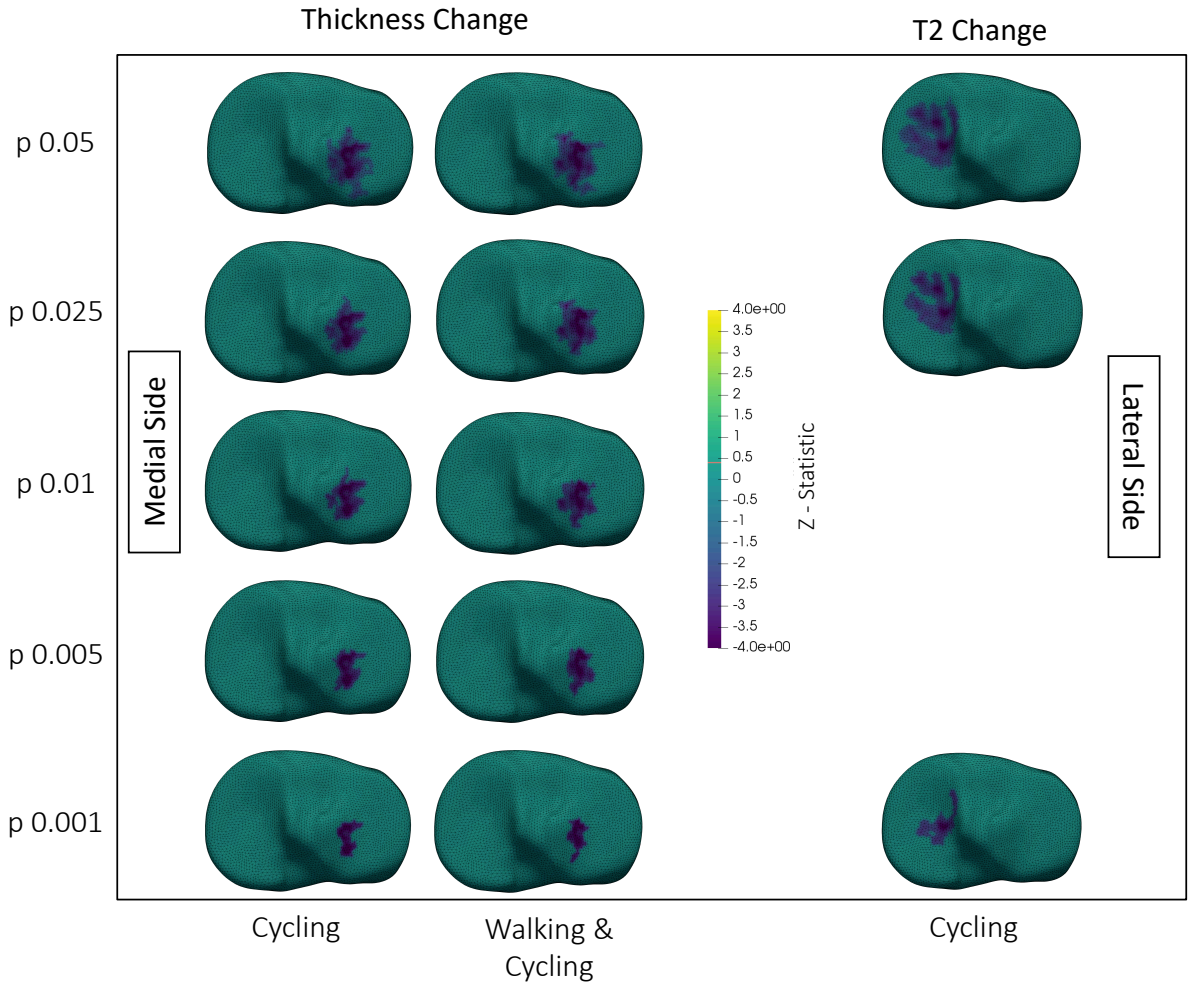


Figure 5-1. Superior view showing changes in tibial cartilage thickness (left two columns), and T_2 (right column). The top is anterior, medial and lateral sides are labelled in the figure. Each tibial plateau shows the region of significant change after thresholding at the p-value on the y-axis. Walking & Cycling indicates that data from walking and cycling were combined in the analysis. Surfaces are colored by the z-statistic associated with the change, where purple indicates a decrease in T_2 or thickness. Spaces with no surface mesh indicate no significant clusters. No column is provided for walking data, because there were no significant changes when analyzing walking alone.

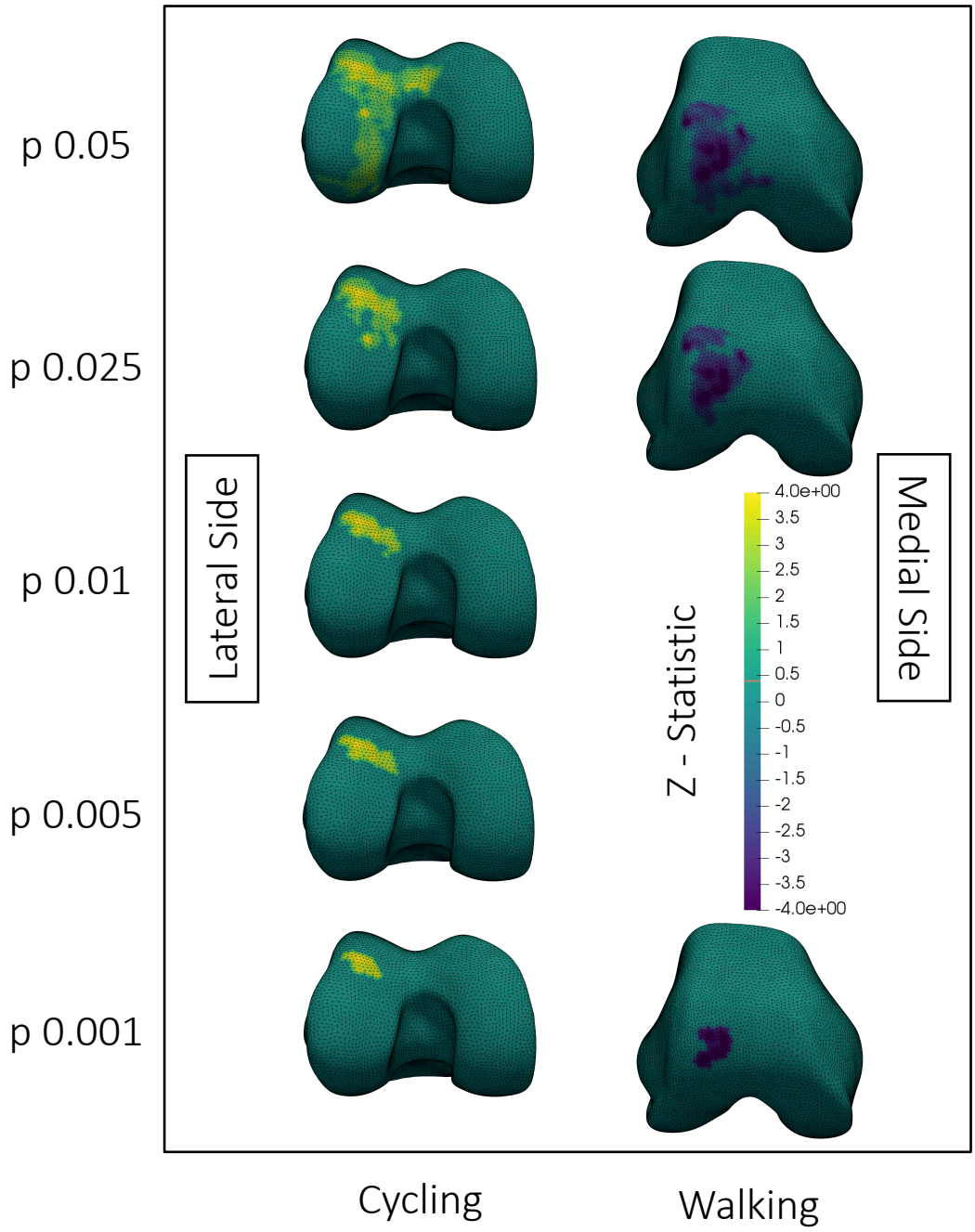


Figure 5-2. Inferior view (left column) and antero-inferior view (right column) of the changes in femoral cartilage T_2 after walking or cycling. Medial and lateral sides are labelled in the figure. Each femoral surface mesh shows the region of significant change after thresholding at the p-value on the y-axis. Surfaces are colored by the z-statistic associated with the change, where purple indicates a decrease in T_2 or thickness and yellow indicates an increase. Spaces with no surface mesh indicate no significant clusters. All surface meshes are of the right knee.

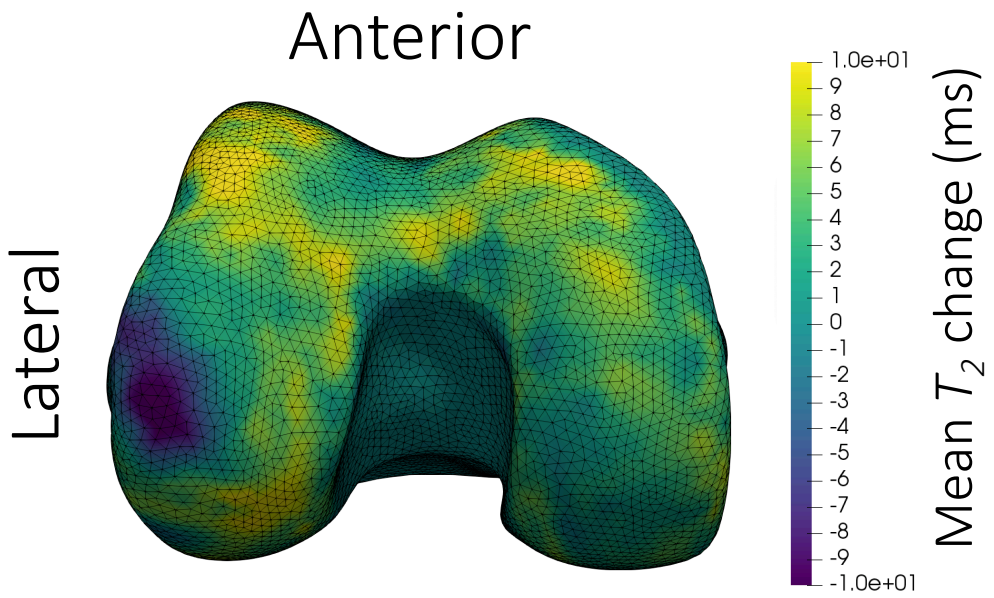


Figure 5-3. Inferior view of the mean change in femoral cartilage T_2 after cycling for 25 min at a moderate 70-75% of maximum heart rate. The colormap shows decreases in T_2 as purple and increases as yellow. The surface mesh is of a right knee; the anterior and lateral sides are labelled in the figure. A region of large magnitude decrease in T_2 can be observed on the lateral weight-bearing femur, with primarily increased or negligible change in T_2 throughout the rest of the surface.

5.3.3 Relationships between joint reaction forces and cartilage change

A greater tibiofemoral compressive JRF during walking resulted in greater reduction in lateral femoral cartilage thickness after walking (Figure 5-4). On the lateral trochlea and the medial femoral condyle, greater tibiofemoral compression produced a greater increase in T_2 after cycling.

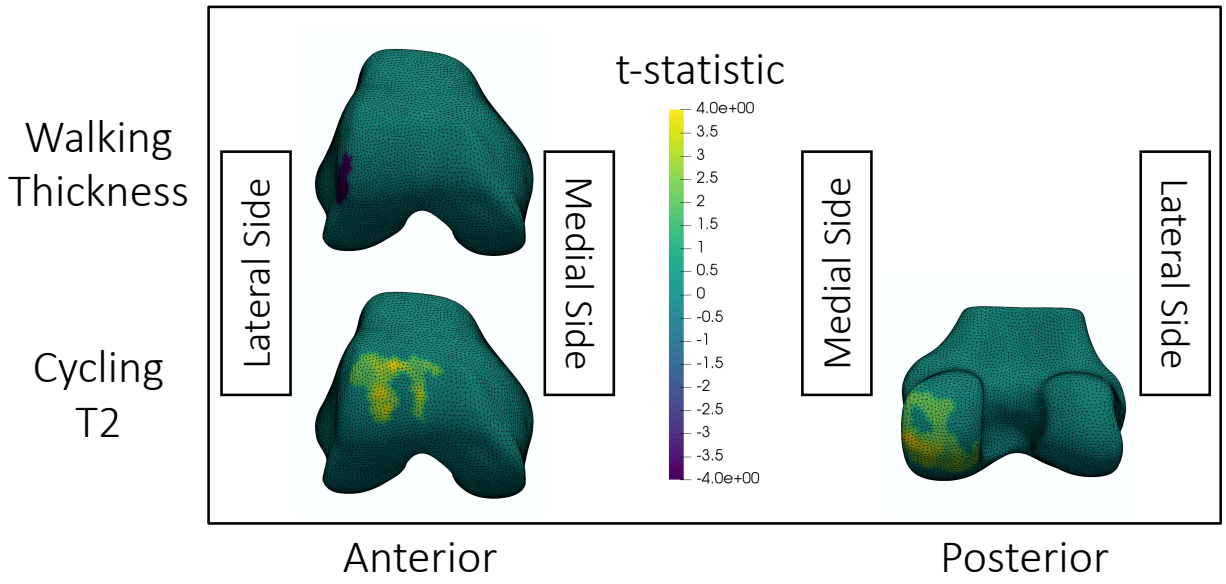


Figure 5-4. Visualization of statistically significant clusters of correlations between the tibiofemoral compressive joint reaction force (JRF) and cartilage change after walking or cycling. The top row shows correlations between changes in cartilage thickness after walking with peak tibiofemoral compression. The bottom row shows correlations between changes in T_2 after cycling with peak tibiofemoral compression. The colormap shows t-statistics for the correlations, where yellow indicates positive correlations and purple indicates negative correlations. All meshes are of a right knee; the medial and lateral sides are labelled in the figure.

5.3.4 Statistical shape model

Increased SSM mode 1 for both the tibia and femur were indicative of increased OA severity. Tibia mode 1 (Figure 5-5) had three main features. First, viewed anteriorly, increasing values corresponded with an increased height of the medial plateau. Second, when viewed superiorly, increasing values corresponded to widening the anterior lateral border of the lateral tibial plateau. Last, when viewed posteriorly, increasing values corresponded with greater concavity between the shaft and the tibial plateau indicating relative broadening.

Femur mode 1 (Figure 5-5) revealed one major feature: increasing values corresponded with an increased bone-cartilage-interface (BCI) surface area. This feature was best observed in the inferior view where the femoral condyles broaden. This broadening corresponded with a narrower intercondylar notch observed inferiorly and posteriorly. From an anterior view, increasing femur mode 1 corresponded with a ridge of osteophytes at the cartilage borders of the trochlea and weightbearing femur.

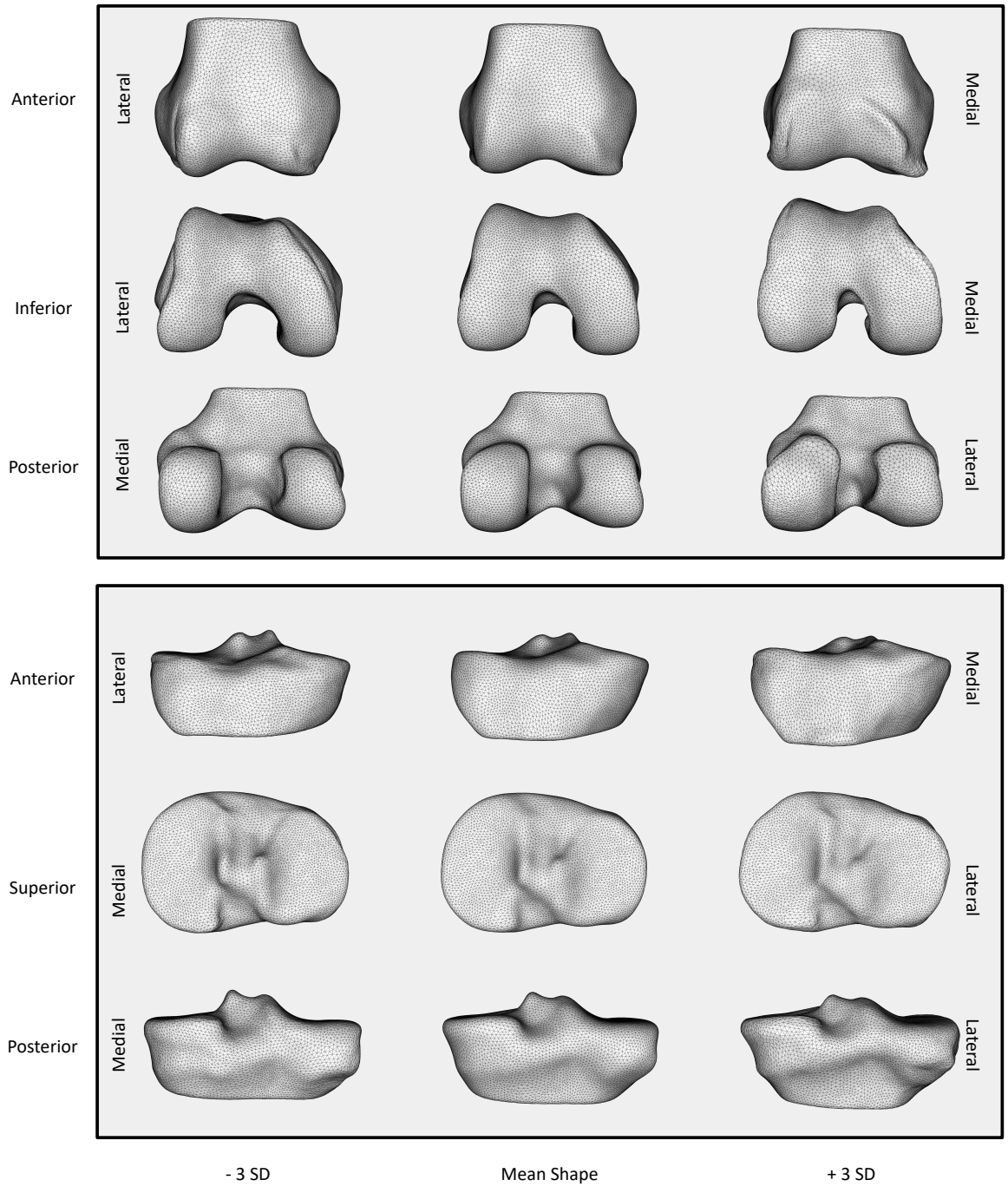


Figure 5-5. Visualization of mode 1 of the femur (top) and tibia (bottom) statistical shape models (SSM). The y-axis identifies the perspective displayed for that row. The x-axis identifies where on the continuum of the SSM mode the visualization falls from – 3 standard deviations (SD) to + 3 SDs. The medial and lateral sides for each row are presented on the right and left borders.

5.3.5 Relationships between statistical shape model and cartilage change

5.3.5.1 Cartilage thickness change and bone SSM correlation

Analysing walking data only, increased tibia and femur modes 1 (greater OA severity) corresponded with smaller reductions in lateral tibial cartilage thickness after walking (Figure 5-6). Meanwhile, analysing walking and cycling together, we noted increased femur mode 1 (greater BCI surface area) corresponded with smaller reductions in medial tibial cartilage thickness (Figure 5-6).

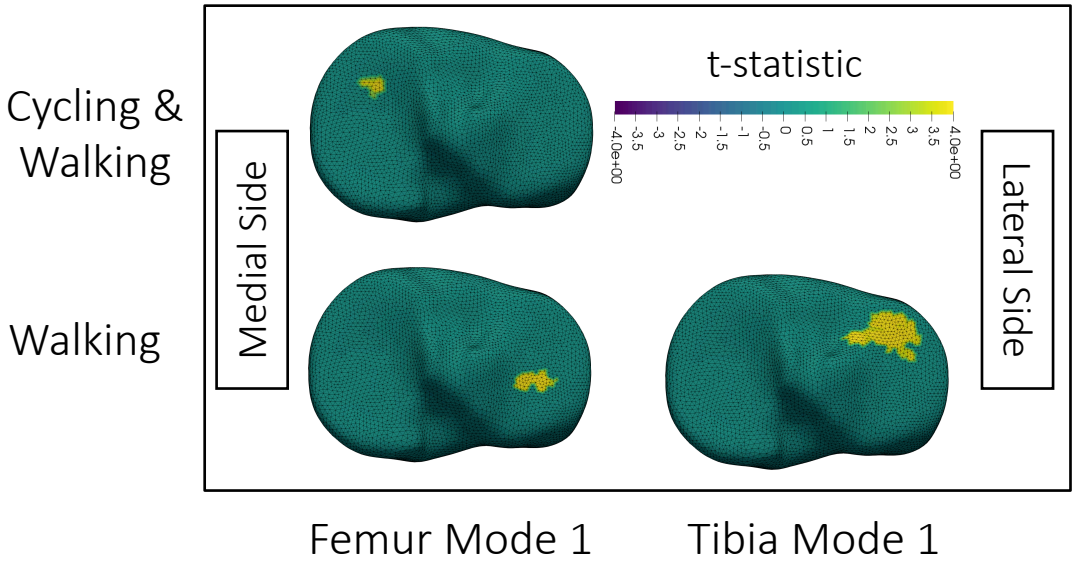


Figure 5-6. Visualization of clusters of significant correlation obtained after thresholding the t-statistic at p=0.01. Correlations are between change in tibial cartilage thickness and the statistical shape model mode identified on the x-axis for data analyzed from activities listed on the y-axis. Cycling & Walking indicates that cycling and walking were analyzed together. No cycling column is included because no significant clusters existed for cycling data alone. The colormap shows the t-statistic for correlations, with positive t-statistics, and thus correlations, in yellow. The top is anterior. Medial and lateral sides are labelled in the figure. All visualizations are of the right knee. For femur mode 1 Cycling & Walking, the same significant cluster existed at thresholds of p=0.01, 0.005, and 0.001. For femur mode 1 Walking data, the same significant cluster existed at thresholds of p=0.01 and 0.005. For tibia mode 1 Walking data, the same significant cluster existed at p=0.01 and 0.005.

5.3.5.2 Cartilage T_2 change and bone SSM correlation

Correlations between T_2 change and SSM features occurred for the femur (Figure 5-7) and tibia (Figure 5-8). With the analysis of walking and cycling together, greater increases in T_2 on the medial femoral condyle occurred with higher femur SSM mode 1 (Figure 5-7). Following analysis of both activities, femoral and tibial bones with higher modes 1 showed smaller decreases in T_2 on the tibial plateaus (Figure 5-8). When analyzing cycling, higher femoral SSM mode 1 resulted in smaller decreases in T_2 on the lateral tibia (Figure 5-8).

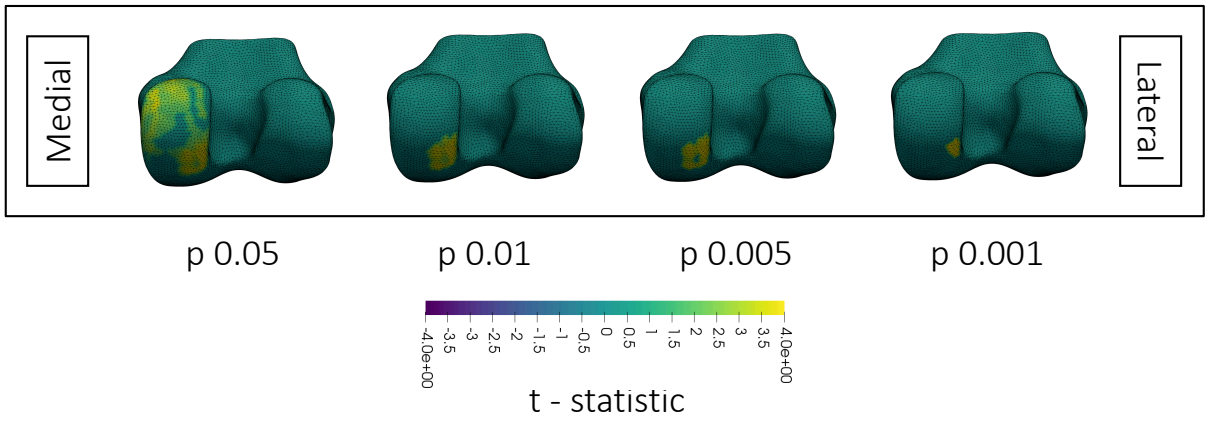


Figure 5-7. Visualization of clusters of significant correlation between change in femoral cartilage T_2 and femoral statistical shape model mode 1 when analyzing data from walking and cycling together. Each femoral surface mesh shows the region of significant correlation after thresholding at the p-value on the x-axis. Surfaces are colored by the t-statistic associated with the change, where yellow indicates a positive t-statistic and correlation. All surface meshes are of the right knee; the medial and lateral sides are labelled in the figure.

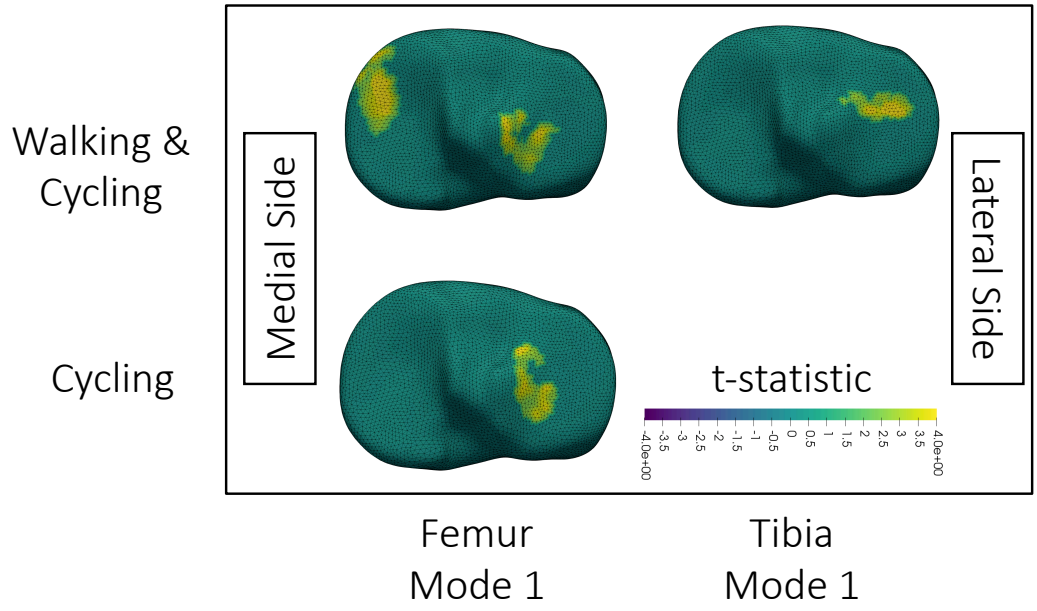


Figure 5-8. Visualization of clusters of significant correlation between change in tibial cartilage T_2 and statistical shape model features when thresholded at p=0.025. The x-axis identifies what statistical shape model mode was used in the correlation. The y-axis identifies what data was included in the analysis. The medial and lateral sides are labelled in the figure. Surfaces are colored by the t-statistic associated with the correlation, where yellow indicates a positive t-statistic and correlation. All surface meshes are of the right knee. The walking and cycling clusters were significant at every threshold level (p=0.05, 0.025, 0.01, 0.005, 0.001). The cycling only cluster was significant at thresholds p=0.05 and 0.025.

5.3.6 Relationship between joint reaction forces and bone SSM features

Femur SSM mode 1 negatively related to the patellofemoral resultant JRF (intercept = 555.7, beta = -151.2, R^2 = 0.4, p=0.021) (Figure 5-9). Femur mode 1 explained 27% of the variance in tibiofemoral compression (intercept=2017.1, beta=-203.0, R^2=0.27, p=0.067) (Figure 5-9). All other correlations had R^2<0.18 and p>0.14.

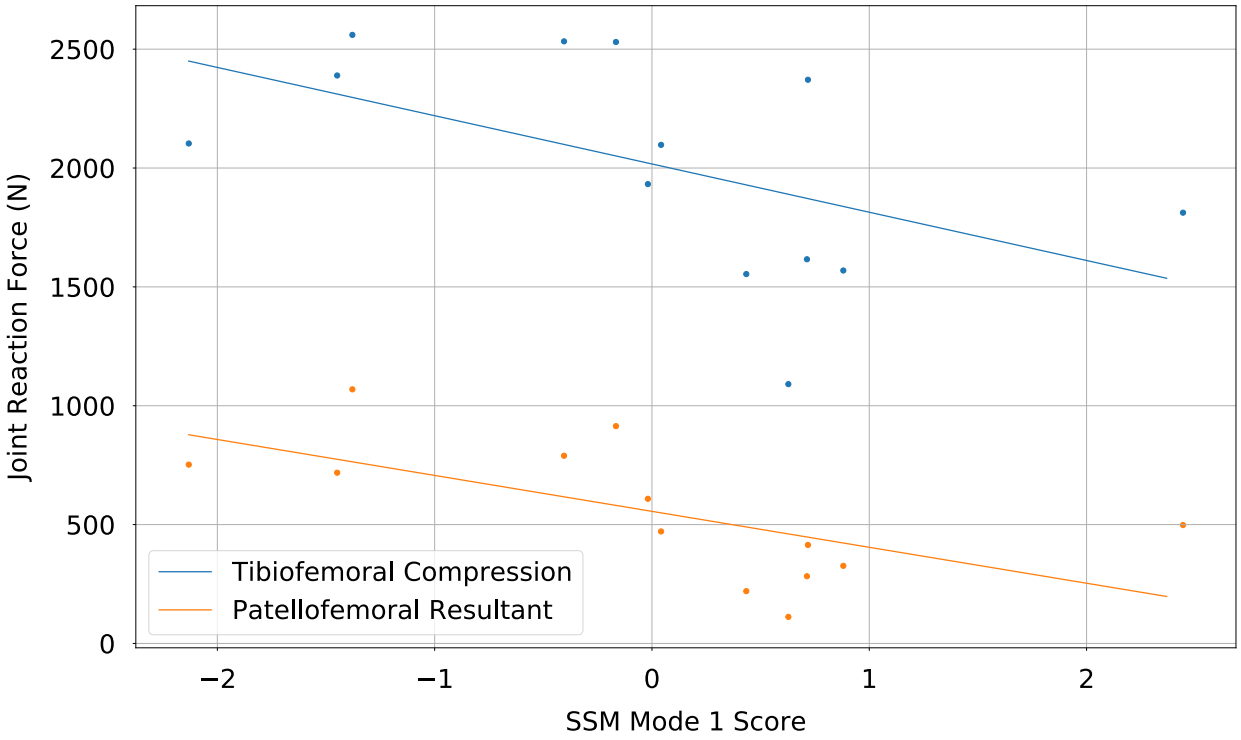


Figure 5-9. Visualization of the relationships between femur statistical shape model (SSM) mode 1 with each of the resultant patellofemoral joint reaction force (orange) and the compressive tibiofemoral joint reaction force (blue).

5.4 Discussion

This work uniquely integrated biomechanical outcomes, measures of cartilage thickness and quality, and bone shape to explore the acute responses of osteoarthritic cartilage to physical activity. Walking and cycling caused distinct patterns of cartilage deformation. JRFs correlated with localized regions of cartilage change. The bone SSMs identified characteristics of OA (Barr et al., 2016; Neogi et al., 2013; Pedoia et al., 2015) that associated with the cartilage changes resulting from activity. Further, femurs with more severe OA had lower knee JRFs during gait. Together these findings show that, among

women with symptomatic knee OA, weight-bearing cartilage T_2 and thickness decrease in response to loading; however, greater disease severity and differences in knee JRFs modulate individual responses.

Femoral cartilage deformation was associated with JRFs. This finding corroborates a previous study that found lateral femoral cartilage deformation was related to tibiofemoral compression (Boocock et al., 2009). After walking, we found a cluster on the lateral trochlea that was thinned more with greater tibiofemoral compression. We also found greater increases in T_2 on the posterior medial condyle with greater tibiofemoral compression (Figure 5-4). Increased T_2 likely reflects mobile water redistribution away from regions of compression; (Choi and Gold, 2011; Eckstein et al., 1999; Lu and Mow, 2008) with more redistribution occurring with greater JRFs.

Tibial plateau T_2 and thickness decreases following loading observed in this sample (Figure 5-1) coincide with the regions of peak cartilage surface pressures during gait in previous work (Clouthier et al., 2019; Van Rossom et al., 2017). After walking, trochlear T_2 decreased, as hypothesized. However, after cycling, T_2 on the lateral condyle increased (Figure 5-2). The increased T_2 may reflect water redistribution away from adjacent regions of insignificant T_2 decrease (Figure 5-3). The larger area of T_2 increase, compared to the region of decrease, may be why the region of decrease was not statistically significant. Along the same line of thought, larger regions in different

compartments, could explain why tibial changes in T_2 and thickness occurred in only one compartment. These findings highlight that SPM may avoid challenges inherent in mean values of anatomically-defined cartilage subregions.

Femur SSM mode 1 shows key features of OA. This single mode closely matches the OA versus non-OA axis of the 70-dimensional SSM by Neogi and colleagues, which was also built from OAI data (Neogi et al., 2013). Our model, and previously published ones (Barr et al., 2016; Neogi et al., 2013) included broadening of the distal femoral cartilage surfaces, increased osteophytes, and narrowing of the intercondylar notch. Consistency between studies signifies robustness of these features.

The acute response of cartilage to a standardized load depends on the SSM score. Since the SSM reflected features of OA, this finding supports the IJS theory that relationships between cartilage, mechanics, and bone differ over the disease spectrum. For example, at the lateral tibial plateau, T_2 decreased less with increasing tibia mode 1, likely reflecting OA severity. In this case, we propose broadening of the plateau may distribute forces over a larger area, reducing pressure leading to smaller cartilage changes. Alternatively, more OA knees may have had thinner cartilage in those locations and therefore less to deform. At the posterior medial femoral condyle, increased OA severity lead to greater increases in T_2 after cycling. In this case, more severe OA knees with pathologic

cartilage likely had greater fluid redistribution, which is congruent with the triphasic theory of cartilage (Choi and Gold, 2011; Eckstein et al., 1999; Lu and Mow, 2008).

Knees more characteristic of OA had smaller JRFs. Previous work supports our finding, showing that when walking at a constant speed (1.11 m/s) individuals with OA have lower compressive forces during gait than their healthy counterparts (Henriksen et al., 2006). These works highlight that JRFs during gait change with OA progression.

5.4.1 Limitations

Due to collection restrictions imposed by COVID-19 this study was under-sampled. Permutation methods ensured accurate p-values for these participants. Point correspondences between 3D surfaces may affect the SSM. Alignment of SSM mode 1 with previous models indicates convergent validity.

5.5 Conclusion

In women with symptomatic knee OA, walking and cycling caused distinct patterns of cartilage deformation that were dependent on knee JRFs and bone morphology. Individuals with femurs more consistent with OA had smaller tibiofemoral JRFs. This work supports the IJS theory that relationships between bone, cartilage, and mechanics change over the disease process and shows that bone shape should be considered an important factor in these relationships.

Conflict of Interest

AAG is the founder of NeuralSeg Ltd. a provider of medical image segmentation and analysis. No other authors have any declarations or conflicts.

Acknowledgements

We would like to acknowledge Emily Wiebenga and Elora Brenneman Wilson for assistance in data collection. We are very grateful to the Laboratory Manager of the Imaging Research Center at St. Joseph's Healthcare, Hamilton Ontario, Canada, Norm Konyer; as well as the Magnetic Resonance Imaging Technologists Carol Awde and Shannon Faseruk for their fastidious attention and guidance throughout this study. MRM holds the Arthritis Society Stars Mid-Career Development Award funded by the Canadian Institutes of Health Research-Institute of Musculoskeletal Health and Arthritis. This study was supported by a Network for Aging Research Catalyst Grant from the University of Waterloo (MRM, MDN), a Mitacs Accelerate (AAG, MRM: IT14939) a Natural Sciences and Engineering Research Council of Canada Discovery Grant (MRM #353715) and the Canadian Foundation for Innovation Leaders Opportunity Fund and the Ministry of Research and Innovation – Ontario Research Fund (MRM, PJK).

5.6 References

Altman, R., Asch, E., Bloch, D., Bole, G., Borenstein, D., Brandt, K., Christy, W., Cooke, T.D., Greenwald, R., Hochberg, M., others, 1986. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis & Rheumatism 29, 1039–1049.

Babel, H., Omoumi, P., Andriacchi, T.P., Jolles, B.M., Favre, J., 2020. New insight on the subchondral bone and cartilage functional unit: Bone mineral density and cartilage thickness are spatially correlated in non-osteoarthritic femoral condyles. Osteoarthritis and Cartilage Open 100079. https://doi.org/10.1016/j.ocarto.2020.100079

Barr, A.J., Campbell, T.M., Hopkinson, D., Kingsbury, S.R., Bowes, M.A., Conaghan, P.G., 2015. A systematic review of the relationship between subchondral bone features, pain and structural pathology in peripheral joint osteoarthritis. Arthritis Research & Therapy 17. https://doi.org/10.1186/s13075-015-0735-x

Barr, A.J., Dube, B., Hensor, E.M.A., Kingsbury, S.R., Peat, G., Bowes, M.A., Sharples, L.D., Conaghan, P.G., 2016. The relationship between three-dimensional knee MRI bone shape and total knee replacement?a case control study: data from the Osteoarthritis Initiative. Rheumatology 55, 1585–1593. https://doi.org/10.1093/rheumatology/kew191

Beekley, M.D., Brechue, W.F., Dehoyos, D.V., Garzarella, L., Werber-Zion, G., Pollock*, M.L., 2004. Cross-validation of the YMCA submaximal cycle

ergometer test to predict VO2max. Research quarterly for exercise and sport 75, 337–342.

Blazek, K., Favre, J., Asay, J., Erhart-Hledik, J., Andriacchi, T., 2014. Age and obesity alter the relationship between femoral articular cartilage thickness and ambulatory loads in individuals without osteoarthritis: FEMORAL ARTICULAR CARTILAGE THICKNESS AND AMBULATORY LOADS. Journal of Orthopaedic Research 32, 394–402. https://doi.org/10.1002/jor.22530

Boocock, M., McNair, P., Cicuttini, F., Stuart, A., Sinclair, T., 2009. The short-term effects of running on the deformation of knee articular cartilage and its relationship to biomechanical loads at the knee. Osteoarthritis and Cartilage 17, 883–890. https://doi.org/10.1016/j.joca.2008.12.010

Choi, J.-A., Gold, G.E., 2011. MR Imaging of Articular Cartilage Physiology. Magnetic Resonance Imaging Clinics of North America 19, 249–282. https://doi.org/10.1016/j.mric.2011.02.010

Clouthier, A.L., Smith, C.R., Vignos, M.F., Thelen, D.G., Deluzio, K.J., Rainbow, M.J., 2019. The effect of articular geometry features identified using statistical shape modelling on knee biomechanics. Medical Engineering & Physics 66, 47–55. https://doi.org/10.1016/j.medengphy.2019.02.009

Cotofana, S., Eckstein, F., Wirth, W., Souza, R.B., Li, X., Wyman, B., Hellio-Le Graverand, M.-P., Link, T., Majumdar, S., 2011. In vivo measures of cartilage deformation: patterns in healthy and osteoarthritic female knees using 3T MR

imaging. European Radiology 21, 1127–1135. https://doi.org/10.1007/s00330-011-2057-y

Cox, R.W., Chen, G., Glen, D.R., Reynolds, R.C., Taylor, P.A., 2017. FMRI Clustering in AFNI: False-Positive Rates Redux. Brain Connectivity 7, 152–171. https://doi.org/10.1089/brain.2016.0475

Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G., 2007. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. IEEE Transactions on Biomedical Engineering 54, 1940–1950. https://doi.org/10.1109/TBME.2007.901024

Eckstein, F., Tieschky, M., Faber, S., Englmeier, K.-H., Reiser, M., 1999. Functional analysis of articular cartilage deformation, recovery, and fluid flow following dynamic exercise in vivo. Anatomy and embryology 200, 419–424.

Edd, S.N., Omoumi, P., Andriacchi, T.P., Jolles, B.M., Favre, J., 2018. Modeling knee osteoarthritis pathophysiology using an integrated joint system (IJS): a systematic review of relationships among cartilage thickness, gait mechanics, and subchondral bone mineral density. Osteoarthritis and Cartilage 26, 1425–1437. https://doi.org/10.1016/j.joca.2018.06.017

Ehrig, R.M., Taylor, W.R., Duda, G.N., Heller, M.O., 2006. A survey of formal methods for determining the centre of rotation of ball joints. Journal of Biomechanics 39, 2798–2809. https://doi.org/10.1016/j.jbiomech.2005.10.002

Eklund, A., Nichols, T.E., Knutsson, H., 2016. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the National Academy of Sciences 113, 7900–7905. https://doi.org/10.1073/pnas.1602413113

Erhart-Hledik, J.C., Favre, J., Andriacchi, T.P., 2015. New insight in the relationship between regional patterns of knee cartilage thickness, osteoarthritis disease severity, and gait mechanics. Journal of Biomechanics 48, 3868–3875. https://doi.org/10.1016/j.jbiomech.2015.09.033

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R., 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 30, 1323–1341. https://doi.org/10.1016/j.mri.2012.05.001

Gatti, A., Keir, P., Maly, M., 2019. Predictive Equations to Define an Ideal Bicycle Saddle Height From Simple Static Measurements. Presented at the International Society for Biomechanics.

Gatti, A.A., 2018. NEURALSEG: state-of-the-art cartilage segmentation using deep learning–analyses of data from the osteoarthritis initiative. Presented at the Osteoarthritis Research Society International. https://doi.org/10.1016/j.joca.2018.02.110

Gatti, A.A., Keir, P.J., Noseworthy, M.D., Beauchamp, M.K., Maly, M.R., 2020. Hip and ankle kinematics are the most important predictors of knee joint loading during

bicycling. Journal of Science and Medicine in Sport. https://doi.org/10.1016/j.jsams.2020.07.001

Gatti, A.A., Maly, M.R., 2019a. Accuracy of estimates of cumulative load during a confined activity: bicycling. International Biomechanics 6, 66–74. https://doi.org/10.1080/23335432.2019.1642141

Gatti, A.A., Maly, M.R., 2019b. Fully-automated cartilage segmentation using deep learning – data from the osteoarthritis initiative. Presented at the International Society for Biomechanics, Calgary, Alberta, Canada.

Gatti, A.A., Noseworthy, M.D., Stratford, P.W., Brenneman, E.C., Totterman, S., Tamez-Peña, J., Maly, M.R., 2017. Acute changes in knee cartilage transverse relaxation time after running and bicycling. Journal of Biomechanics 53, 171–177. https://doi.org/10.1016/j.jbiomech.2017.01.017

Harrington, M.E., Zavatsky, A.B., Lawson, S.E.M., Yuan, Z., Theologis, T.N., 2007. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. Journal of Biomechanics 40, 595–602. https://doi.org/10.1016/j.jbiomech.2006.02.003

Henriksen, M., Simonsen, E.B., Alkjær, T., Lund, H., Graven-Nielsen, T., Danneskiold-Samsøe, B., Bliddal, H., 2006. Increased joint loads during walking – A consequence of pain relief in knee osteoarthritis. The Knee 13, 445–450. https://doi.org/10.1016/j.knee.2006.08.005

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

Kennedy, D.M., Stratford, P.W., Robarts, S., Gollish, J.D., 2011. Using outcome measure results to facilitate clinical decisions the first year after total hip arthroplasty. J Orthop Sports Phys Ther 41, 232–239. https://doi.org/10.2519/jospt.2011.3516

Koo, S., Rylander, J.H., Andriacchi, T.P., 2011. Knee joint kinematics during walking influences the spatial cartilage thickness distribution in the knee. Journal of Biomechanics 44, 1405–1409. https://doi.org/10.1016/j.jbiomech.2010.11.020

Legendre, P., Legendre, L., Legendre, L., 1998. Numerical ecology. Elsevier, Amsterdam; New York.

Lombaert, H., Grady, L., Polimeni, J.R., Cheriet, F., 2013. FOCUSR: Feature Oriented Correspondence Using Spectral Regularization--A Method for Precise Surface Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 35, 2143–2160. https://doi.org/10.1109/TPAMI.2012.276

Lorensen, W.E., Cline, H.E., n.d. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. ACM SIGGRAPH Computer Graphics 21, 7.

Lu, X.L., Mow, V.C., 2008. Biomechanics of Articular Cartilage and Determination of Material Properties: Medicine & Science in Sports & Exercise 40, 193–199. https://doi.org/10.1249/mss.0b013e31815cb1fc

MacKay, J.W., Kaggie, J.D., Treece, G.M., McDonnell, S.M., Khan, W., Roberts, A.R., Janiczek, R.L., Graves, M.J., Turmezei, T.D., McCaskie, A.W., Gilbert, F.J.,

2020. Three-Dimensional Surface-Based Analysis of Cartilage MRI Data in Knee Osteoarthritis: Validation and Initial Clinical Application. Journal of Magnetic Resonance Imaging. https://doi.org/10.1002/jmri.27193

Maly, M.R., Acker, S.M., Totterman, S., Tamez-Peña, J., Stratford, P.W., Callaghan, J.P., Adachi, J.D., Beattie, K.A., 2015. Knee adduction moment relates to medial femoral and tibial cartilage morphology in clinical knee osteoarthritis. Journal of Biomechanics 48, 3495–3501. https://doi.org/10.1016/j.jbiomech.2015.04.039

Mun, D., Kim, B.C., 2017. Three-dimensional solid reconstruction of a human bone from CT images using interpolation with triangular Bézier patches. Journal of Mechanical Science and Technology 31, 3875–3886. https://doi.org/10.1007/s12206-017-0732-x

Neogi, T., Bowes, M.A., Niu, J., De Souza, K.M., Vincent, G.R., Goggins, J., Zhang, Y., Felson, D.T., 2013. Magnetic Resonance Imaging-Based Three-Dimensional Bone Shape of the Knee Predicts Onset of Knee Osteoarthritis: Data From the Osteoarthritis Initiative: 3-D Bone Shape Predicts Incident Knee OA. Arthritis & Rheumatism 65, 2048–2058. https://doi.org/10.1002/art.37987

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Human Brain Mapping 15, 1–25. https://doi.org/10.1002/hbm.1058

Omoumi, P., Babel, H., Jolles, B.M., Favre, J., 2019. Relationships between cartilage thickness and subchondral bone mineral density in non-osteoarthritic and severely

osteoarthritic knees: In vivo concomitant 3D analysis using CT arthrography. Osteoarthritis and Cartilage 27, 621–629. https://doi.org/10.1016/j.joca.2018.12.014

Paranjape, C.S., Cutcliffe, H.C., Grambow, S.C., Utturkar, G.M., Collins, A.T., Garrett, W.E., Spritzer, C.E., DeFrate, L.E., 2019. A New Stress Test for Knee Joint Cartilage. Scientific Reports 9. https://doi.org/10.1038/s41598-018-38104-2

Pedoia, V., Lansdown, D.A., Zaid, M., McCulloch, C.E., Souza, R., Ma, C.B., Li, X., 2015. Three-dimensional MRI-based statistical shape model and application to a cohort of knees with acute ACL injury. Osteoarthritis and Cartilage 23, 1695–1703. https://doi.org/10.1016/j.joca.2015.05.027

Peterfy, C.G., Schneider, E., Nevitt, M., 2008. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis and Cartilage 16, 1433–1441. https://doi.org/10.1016/j.joca.2008.06.016

Pre-Screening for Physical Activity: Get Active Questionnaire, n.d. . Canadian Society for Exercise Physiology. https://store.csep.ca/pages/getactivequestionnaire

Scanlan, S.F., Favre, J., Andriacchi, T.P., 2013. The relationship between peak knee extension at heel-strike of walking and the location of thickest femoral cartilage in ACL reconstructed and healthy contralateral knees. Journal of Biomechanics 46, 849–854. https://doi.org/10.1016/j.jbiomech.2012.12.026

Souza, R.B., Baum, T., Wu, S., Feeley, B.T., Kadel, N., Li, X., Link, T.M., Majumdar, S., 2012. Effects of Unloading on Knee Articular Cartilage T1rho and T2 Magnetic Resonance Imaging Relaxation Times: A Case Series. Journal of Orthopaedic & Sports Physical Therapy 42, 511–520. https://doi.org/10.2519/jospt.2012.3975

Steele, K.M., DeMers, M.S., Schwartz, M.H., Delp, S.L., 2012. Compressive tibiofemoral force during crouch gait. Gait & Posture 35, 556–560. https://doi.org/10.1016/j.gaitpost.2011.11.023

Tanaka, H., Monahan, K.D., Seals, D.R., 2001. Age-predicted maximal heart rate revisited. Journal of the American College of Cardiology 37, 153–156.

Teichtahl, A.J., Wluka, A.E., Wijethilake, P., Wang, Y., Ghasem-Zadeh, A., Cicuttini, F.M., 2015. Wolff's law in action: a mechanism for early knee osteoarthritis. Arthritis Research & Therapy 17. https://doi.org/10.1186/s13075-015-0738-7

Valette, S., Chassery, J.-M., 2004. Approximated Centroidal Voronoi Diagrams for Uniform Polygonal Mesh Coarsening. Computer Graphics Forum 23, 381–389. https://doi.org/10.1111/j.1467-8659.2004.00769.x

Valette, S., Chassery, J.-M., Prost, R., 2008. Generic Remeshing of 3D Triangular Meshes with Metric-Dependent Discrete Voronoi Diagrams. IEEE Transactions on Visualization and Computer Graphics 14, 369–381. https://doi.org/10.1109/TVCG.2007.70430

Van Rossom, S., Smith, C.R., Zevenbergen, L., Thelen, D.G., Vanwanseele, B., Van Assche, D., Jonkers, I., 2017. Knee Cartilage Thickness, T1ρ and T2 Relaxation Time Are Related to Articular Cartilage Loading in Healthy Adults. PLOS ONE 12, e0170002. https://doi.org/10.1371/journal.pone.0170002

Vanwanseele, B., Eckstein, F., Smith, R.M., Lange, A.K., Foroughi, N., Baker, M.K., Shnier, R., Fiatarone Singh, M.A., 2010. The relationship between knee adduction moment and cartilage and meniscus morphology in women with osteoarthritis. Osteoarthritis and Cartilage 18, 894–901. https://doi.org/10.1016/j.joca.2010.04.006

5.7 Supplemental

5.7.1 Supplemental 1.

Location of markers used to scale a musculoskeletal model, and to determine joint kinematics. Markers listed for the thigh, shank, and foot were included on the right and left legs.

Segment	**Markers**
Foot	Calcaneus
	1st metatarsal head
	5th metatarsal head
	Lateral side of the midfoot
Shank	Tibial tuberosity
	Lateral malleolus
	Medial malleolus
	Rigid cluster of 4 markers attached to the lateral shank
Thigh	Greater trochanter
	Lateral femoral condyle
	Medial femoral condyle
	Rigid cluster of 4 markers attached to the lateral thigh
Pelvis	Right anterior superior iliac spine
	Left anterior superior iliac spine
	Right posterior superior iliac spine
	Left posterior superior iliac spine

6 Discussion

This thesis consists of three main parts. First, a fully crossed experimental study was designed to thoroughly capture motion and force data while a variety of healthy adults bicycled under random deviations of bicycle saddle position and crank arm length. The second major component described a novel method of segmenting bone and cartilage from knee MRIs and tested its accuracies. The third and final component integrated results from the first and second parts in order to investigate the interplay between bone shape and knee JRFs on the acute changes in knee cartilage T_2 and thickness to the common aerobic activities of walking and bicycling.

6.1 Thesis Overview

This thesis filled four gaps in the literature relevant to bicycling and how it affects knee joint forces and cartilage. The following sections highlight unique aspects of each study relevant to each gap:

1. There is no easy way to prescribe or determine the optimal bicycle saddle position to elicit a desired amount of knee flexion.
2. There are no thorough experimental studies that determine how alterations in bicycle saddle position, and thus lower extremity joint kinematics, affect forces acting inside of the knee.
3. Current methods of segmenting cartilage and bone from MRI for quantitative analyses are slow and prone to error. Faster and more accurate methods are required to enable efficient analysis.

4. We lack an understanding of how knee cartilage affected by OA responds to physical activity. Ultimately, we do not fundamentally understand how two common activities, walking and bicycling, affect the OA knee.

6.1.1 Biomechanical Analyses of Bicycle-Fit and Knee Joint Reaction Forces

Chapter 2 created a novel equation for prescribing bicycle saddle position that elicits a desired knee kinematic; this equation enables easy prescription of joint kinematics, the gold-standard of bicycle-fit, inside the clinic, bicycle-shop, or at home. Chapter 3 confirmed the popular notion that the patellofemoral JRF is most sensitive to joint kinematics. Also, we were the first to show that knee JRFs are least sensitive to the most common bicycle-fit parameter, minimum knee flexion angle (Bini et al., 2011; Holmes et al., 1994; Silberman, 2013), and that the overall most important predictor of knee JRFs was minimum hip flexion angle.

These findings were enabled by a fully-crossed random assignment study design specifically created to sample the breadth of possible saddle positions, continuously. Historically, studies of bicycle-fit have collected data on a small set of finite saddle positions (Bini et al., 2010; Menard et al., 2016, 2018; Tamborindeguy & Bini, 2011). These perturbations to saddle position were typically large, in the range of +/- 10% (Menard et al., 2018). Practically, changes of these magnitudes would almost never be conducted and therefore findings may not translate to practical bicycle-fit adjustments.

Our design randomized each individual participant's horizontal (recommended +/- 10%) and vertical (recommended +/- 5%) saddle positions to be any mm increment within similar ranges. We also tested all possible combinations of changing 3 bicycle-fit parameters (3 saddle X, 3 saddle Y, 2 crank arm length). Because of their study designs, previous studies determine how the saddle position affects JRFs (Menard et al., 2018; Tamborindeguy & Bini, 2011). That is, these studies used designs that tested a few discrete saddle positions and only tested differences using analysis of variance of the discrete positions. By collecting continuous data, we were able to determine the relationship between kinematics and the outcomes of interest (saddle position and knee JRFs). Comparing joint kinematics seems crucial as they have become the standard in bicycle-fit, (Bini et al., 2011; Fonda et al., 2014; Holmes et al., 1994) and are likely the mediators between changes in saddle position and knee JRFs. By collecting continuous data on the largest sample used in any study of this nature, we were able to perform continuous analyses to provide novel equations for bicycle-fit and unique findings of how joint kinematics affect knee JRFs.

It is interesting to note that hip flexion, which was not important when predicting saddle heights, was the most sensitive measurement and appeared in every JRF model. I hypothesize that the statistical approach provides insight into why hip kinematics were identified as important to joint reaction forces. Particularly, I propose that use of mixed-effects models explains hip flexions inclusion in the joint reaction forces analysis but not the saddle height prediction equations. Discussion section 6.2 outlines this rationale. I

also propose that maximum knee flexion angle should continue to be explored, at least in the realm of patellofemoral conditions which are the most common reason for medical care amongst cyclists (Silberman, 2013).

6.1.2 Advances in Cartilage Segmentation from Magnetic Resonance Images

Chapter 4 presents a state-of-the-art multi-stage convolutional neural network framework that uses stage 1 predictions as inputs into stage 2 producing excellent results for segmenting all regions of interest, and in timescales (~1.5 mins) that do not hinder clinical or research timelines.

To enable this testing, I designed a cross-validation study that assessed how the novel algorithm both learns from and predicts the presence and location of cartilage on knees ranging from healthy to severe OA, and from two different MRI vendors and sequences. The previous best method tested on the entirety of the same OAI dataset required a SSM to limit bone errors and localize cartilage. We show that such an SSM is not needed. In fact, the multi-stage convolutional neural network produced better DSC and average surface distance metrics than the SSM plus deep learning method that took 10-minutes to segment a knee. Our approach could just as easily add other anatomical tissues at the knee, or be employed for the hip, brain or any other imaging data, without the need to manually define SSM features. Chapter 4 highlights that the accuracies of this algorithm are better than inter-radiologist accuracies, and thus surpass the gold-standard (Shim et

al., 2009). Automated, fast methods of analyzing data that previously required tedious manual work performed by experts is a trend being facilitated by machine learning that has the potential to have a major impact on biomedical research.

6.1.3 Integrating Bone Shape and Biomechanics to Understand Cartilage Deformation

In Chapter 5, we showed for the first time that in women with symptomatic knee OA, aerobic exercises walking and bicycling cause distinct patterns of cartilage deformation dependent on knee JRFs and bone shape. These findings provide support for the IJS theory, and the methods provide new approaches to integrate the study of cartilage, bone, and mechanics.

To achieve these outcomes, we used information garnered from chapters 2-4 and integrated musculoskeletal modelling, medical imaging, deep learning, and advanced statistical methods to study the acute response of osteoarthritic cartilage to physical activity. To create the SSM used to categorize bone shape, we created and open-sourced two new packages for the Python programming language: cycpd and pyfocusr. To analyze the changes in cartilage continuously over the bone surface using SPM we created and open-sourced a third package (pyKneeSPM). Due to restrictions imposed by COVID-19, this study was performed on a small sample. Nevertheless, the findings support the IJS theory (Edd et al., 2018). Furthermore, the proposed framework is a

strength that may be used in future studies that hope to integrate bone, cartilage, and mechanics.

6.1.4 Key Ideas, Findings, and Concepts

These studies bring about a few important ideas, findings, and concepts. First, important for translating the findings to clinical usage, hip flexion angles were broadly the most important factors in the JRF analyses but did not aid in predicting saddle height. This mis-match leaves the question as to how bicycle-saddle height should be set. I propose that this is largely a result of the study purposes and statistical methods, as well as effects of upper body positioning on hip kinematics. Second, as an alternative target for bicycle-fit in the clinic or bicycle-shop, I propose that research should explore maximum knee flexion angle. Third, the application of machine learning in chapters 2 and 4 highlights the potential importance in enabling biomedical research. Fourth, open-source software produced and maintained by others has been crucial at enabling this research; this thesis also contributes three new packages to the open-source community. Finally, biomechanical and biomedical imaging data have been foundational in the study of OA. Chapter 5 of this thesis provides new methods of thoroughly integrating these data; stronger integration appears integral for progression in applying these fields to the study of OA. In the following sections I elaborate on each of these topics.

6.2 Why Hip Kinematics are Related to Joint Forces but Not Saddle Positioning

Hip kinematics predicted bicycle saddle position, indicating that if saddle position is changed hip kinematics do not change consistently. Conversely, knee kinematics predicted saddle position, indicating that changing saddle position will systematically affect knee kinematics. Yet, hip kinematics were the most important predictors of knee JRFs, and minimum knee flexion was the least. Together, these indicate that saddle position does not seem to affect JRFs, at least not through changes in knee or hip kinematics. Yet, the current consensus is that bicycle-fit, primarily saddle positioning, is the gold standard method of avoiding musculoskeletal injuries, presumably via reductions in mechanical aetiologies such as JRFs (Bini, 2016; Bini et al., 2011; Silberman, 2013). This is particularly the case for patellofemoral pain which is hypothesised to be driven by high patellofemoral compression caused by poor bicycle-fit (Holmes et al., 1994). My hypothesis for why this is the case comes down to the statistical methods used combined with the fact that hip kinematics likely vary more between participants than knee kinematics.

The fully-crossed random assignment study design that contained repeated measurements of the same participants was a strength of this work. Due to the repeated measurements, these data were correlated with one another, which is known to inflate statistical significance (Cameron & Miller, 2015). Two common methods used to handle the correlated measurements for regression analyses are: 1) mixed-effects models that account for the between-subjects variance, (Detry & Ma, 2016) or 2) regular linear regression followed by correction of standard errors after fitting the model (StataCorp LP,

2013). A mixed-effects model accounts for between-subjects variability, enabling modelling of the relationship between the independent and dependent variables within an individual. Whereas, regular linear regression plus standard error correction, typically ignores the between-subjects effects and therefore does not enable within participant inference. Therefore, if there are greater differences between participants than there are within participants, then the regular linear regression model will be fitted to these between-subjects effects missing the within-subject effects entirely.

I believe that the purposeful statistical modelling decisions combined with between-subject differences in hip kinematics are why hip flexion was important to JRFs but not bicycle-fit. Specifically, hip kinematics systematically varied between participants. That is, commercial bicycle saddle position was standardized and used as a reference frame. However, each participant was allowed to select their own preferred handlebar position. A more aggressive handlebar position would result in a more bent over rider and an anteriorly rotated pelvis (Silder et al., 2011). Anterior rotation of the pelvis with respect to gravity will likely increase the hip flexion angles throughout the pedal revolutions, if we maintain all other bicycle-fit parameters (Silder et al., 2011). Therefore, when trying to create one equation to predict saddle height, the between-subjects variability in hip kinematics precluded it being an important overarching predictor of saddle height.

When modelling the relationship between joint kinematics and knee JRFs, the purpose was not to attribute a specific relationship between joint kinematics and knee JRFs. That is, we did not expect that a certain joint angle would produce the same JRF in all participants. This is because other participant specific factors (e.g., power output, cadence) are likely more important for predicting these JRFs. We were instead interested in how changing joint kinematics within an individual affected knee JRFs. To achieve this outcome, we leveraged the fully-crossed study design and mixed-effects models. The fully-crossed design allowed us to garner a breadth of data on how the independent and dependent variables varied as we changed saddle positioning, within a person. Then, the mixed-effects models accounted for the between-subjects effects due to factors like power output and cadence, enabling identification of the within-subject effect of hip kinematics on knee JRFs. Ultimately, between-subjects differences in hip kinematics and appropriate use of modelling techniques were the drivers of the discrepancy between studies.

6.3 Maximum Knee Flexion as a Future Measure of Bicycle-fit

The fact that hip flexion is most important for JRF prediction but is unimportant for predicting saddle height leaves a void as to how saddle positioning should be set. Taking an integrated view, I propose that maximum knee flexion warrants further study. First, separate equations using either minimum or maximum knee flexion contained the same fit parameters (R^2 0.97, RMSE 1.15 cm) and were very close in terms of cross-validation based RMSE (minimum knee flexion: 1.16cm, maximum knee flexion: 1.19 cm).

Second, we showed that the patellofemoral JRF was least sensitive to minimum knee flexion angle (5.6 N/º), however, it was ~3 times more sensitive to maximum knee flexion angle (15.9 N/º). The sensitivity of the patellofemoral JRF to minimum hip flexion (17.9 N/º) was only slightly more than to maximum knee flexion. Therefore, maximum knee flexion angle will enable prescription of bicycle saddle height while concurrently providing insight into the patellofemoral JRF.

In addition to these new findings, there is strong theoretical reasoning why maximum knee flexion angle should be an important predictor of patellofemoral compression. While minimum and maximum knee flexion should, in theory, be strongly correlated, there are bicycle-fit related factors that affect one and not necessarily the other. For example, shorter crank arms which are recommended commercially to reduce patellofemoral JRFs will have asymmetric effects on minimum and maximum knee flexion angles. A focus on maximum knee flexion is important for at least three reasons. First, when bicycling, studies that use electromyography (Jorge & Hull, 1986; Lai et al., 2017) or musculoskeletal modelling (Lai et al., 2017) show that the quadriceps are most activated in the top half of the pedal revolution (9 o'clock to 3 o'clock) and that there are peak patellofemoral compressive forces in this region (Bini, 2020). Furthermore, these studies show minimum values at the bottom of the revolution, where the gold standard bicycle-fit parameter minimum knee flexion is measured (Bini, 2020; Jorge & Hull, 1986; Lai et al., 2017). Second, patellofemoral biomechanics indicate that with greater knee flexion, for a given quadriceps force, there will be higher patellofemoral JRFs

(Figure 6-1). This is because, with increased flexion, the quadriceps/patellar tendon line of action is directed increasingly posteriorly. Therefore, with greater knee flexion, a greater proportion of the quadriceps force compresses the patella into the femur (Hungerfordm & Maureenbarryb, 1979; Nordin & Frankel, 2012). This scenario presumes that the quadriceps force is constant. Third, with increased knee flexion above 45º the quadriceps moment arm decreases, thus theoretically requiring an even greater quadriceps force for a given knee extension moment[§§] (Krevolin et al., 2004). I therefore propose that future research should focus on maximum knee flexion as a key bicycle-fit related parameter relevant to patellofemoral ailments. In the meantime, this thesis highlights equations to predict bicycle saddle height based on a desired maximum knee flexion angle and indicates that less flexion will reduce patellofemoral JRFs.

[§§] Assuming the same extensor moment is needed for a given bicycling intensity (cadence and power).

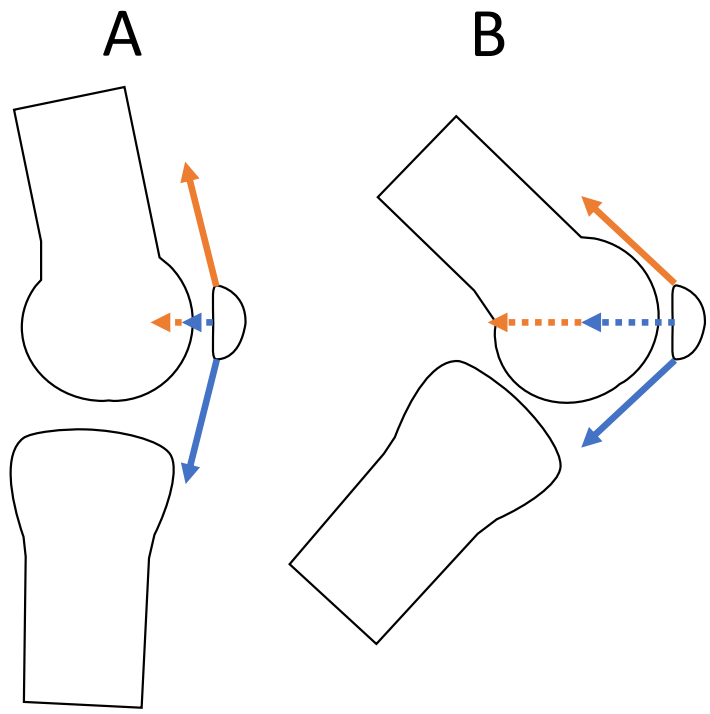


Figure 6-1. Knee flexion angle versus patellar compression. Visual depiction of how changing knee flexion angle changes patellofemoral compression for constant quadriceps tendon (QT; orange solid vector) and patellar tendon (PT; blue solid vector) forces. The solid vectors (blue and orange) are all the same length, representing a constant force applied through each tendon. The dashed lines show the compressive component of the QT and PT forces under two hypothetical knee flexion angles. The dashed lines are placed tip to tail so that the total length of the dashed arrows visually represents the total compressive force. Subfigure A) shows a knee at near full extension and subfigure B) shows a knee flexed to approximately 90 degrees. The figure demonstrates that greater knee flexion, for a constant muscular (tendon) force, greatly increases patellofemoral compression.

6.4 Machine Learning in Biomedical Research

Machine learning, and deep learning, have received excesses of hype over the past decade, particularly to predict health-related outcomes from medical images. Machine learning is a form of applied statistics that emphasizes using computers and is focused

around prediction instead of inference (Goodfellow et al., 2016). Deep learning is a flexible and non-linear type of machine learning (Goodfellow et al., 2016). In the knee OA literature, machine learning methods have been used to directly predict structural disease progression from up to 366 predictors*** (Pedoia et al., 2018). Deep learning has been extensively used to grade KL scores from radiographs (Norman et al., 2019; Thomas et al., 2020; Tiulpin et al., 2018). Tolpadi et al. used deep learning to predict total knee replacement (TKR) from sagittal MRIs combined with clinical and demographics data (Tolpadi et al., 2020). The growth of deep learning, and particularly convolutional neural networks has greatly improved the ability for direct prediction of outcomes from medical images.

While machine learning has improved predictions obtained directly from images, using convolutional neural networks directly on the image does not automatically solve the problem at hand. For example, in models predicting TKR using only imaging data (X-ray or MRI) there is poor sensitivity when predicting TKR in those with KL grades 0 (7.9 to 66.9%), and poor specificity in predicting TKR in those with KL grades 4 (7.2 to

***Predictors included: demographics, patient reported outcomes, qualitative cartilage assessments, biomechanical outcomes, and compositional MRI outcomes.

9.2%)††† (Tolpadi et al., 2020). Furthermore, using these pipelines to directly predict the clinical outcome does not improve our understanding of the underlying pathophysiology. Some work has been done to improve interpretability, particularly of convolutional neural networks. Tolpadi et al. used an occlusion map technique to identify image "hot spots" relevant for predicting TKR (Tolpadi et al., 2020). Saliency maps have also been used to identify image regions pertinent for predicting KL grades (Norman et al., 2019; Thomas et al., 2020; Tiulpin et al., 2018). While occlusion maps highlight areas of interest they do not explain their importance.

While applications of directly using machine learning are attractive, one of the greatest uses of machine learning today may be speeding up previously tedious and costly data processing steps. In the biomechanics field, and in MRI studies of cartilage, small sample sizes are common. Small samples are not by design, but feasibility. Chapters 2 and 3 used biomechanical data collected from 40 participants under 18 bicycling

---

††† Integrating clinical data improves these predictions but is not perfect with moderate (KL 2, 3) and severe (KL 4) knees having sensitivities ranging from 64 to 84% and specificities ranging from 49.6% to 74.7% (Tolpadi et al., 2020).

positions each; this sample size is large relative to the previous work[‡‡‡]. However, sample sizes even larger would provide more precise estimates of the relationships between joint kinematics and JRFs. Larger samples were precluded by the fact that data collections took up to 4.5 hours per participant and data post-processing took longer. Preliminary post-processing took 40 hours for a single participant. This efficiency was drastically improved, but it still took approximately one 8-hour day to manually process motion data from one participant.

Markerless motion analysis systems may provide one solution to improving the efficiency of acquiring and processing biomechanical data. Recently, markerless motion analysis systems have been proposed and made available as open-source (Mathis et al., 2018) and commercial (Theia Markerless, Kingston, Ontario, Canada) software (Kanko, Laende, et al., 2020; Kanko, Strutzenberger, et al., 2020). These software have the potential to reduce data collection and processing times. Aside from the ease of use, if everyone uses the same automated software then research between labs will become more comparable, a difficulty today when different laboratories have different equipment, (Fonda et al., 2014; Kaufman et al., 2016) different protocols for data acquisition, varying

---

[‡‡‡] Previous investigations included 9 or 10 male participants (Bini, 2020; Menard et al., 2018; Tamborindeguy & Bini, 2011) and just 3 conditions each totaling 27-30 trials. This thesis included 40 participants and 18 trials each, totaling 720 trials.

expertise in anatomical palpation and marker placement, (Gorton et al., 2009) different marker sets, (Kaufman et al., 2016) and different post processing methods (Kainz et al., 2017). However, these automated systems do have limitations. While the marker placements may be more consistent, this does not mean that they are correct[§§§] and there will be conditions and scenarios that these systems knowingly, and unknowingly, fail. For example, the system may fail when analyzing gymnastics or other aerial sports where people appear upside down in the frame due to a lack of sufficient training data in these edge cases. Furthermore, there will likely be more than one vendor, posing new issues of comparability between systems. Ultimately, like any tool, researchers will need to be informed and vigilant of the limitations of the systems they are using, which may pose a particular challenge as the technology underpinning the tools becomes more and more specialized.

Specific to this thesis, the segmentation algorithm presented in Chapter 4 enabled automated segmentation of knee MRIs in ~1.5 minutes. This efficiency saved hundreds of hours of manual analysis for the SSM creation and cartilage analysis used to explore

---

[§§§]To date, the commercial Theia Markerless system has been evaluated for test re-test reliability, (Kanko, Laende, et al., 2020) and for validity of simple gait metrics only: gait speed, cadence, step time, stance time, swing time, double-limb support time, step length, stride length, and stride width (Kanko, Strutzenberger, et al., 2020).

the acute response of OA cartilage to walking and bicycling in Chapter 5[****]. The increased capacity of automated methods is highlighted by a recent publication that quantified cartilage T_2 relaxation time for the entire OAI dataset (Razmjoo et al., 2020). The segmentation DSCs in that study were markedly poorer (medial femur: 0.69, lateral femur: 0.69, medial tibia: 0.68, lateral tibia: 0.75, patella: 0.57) than those presented here (femur: 0.91, medial tibial: 0.88, lateral tibia: 0.91, patella: 0.84). Nonetheless, analysis of a dataset with >25,000 knee MRIs enabled confirmation that T_2 does increase with increasing disease severity and that it is predictive of disease progression[††††] (Razmjoo et al., 2020). However, similar to markerless analyses, different algorithms will likely be systematically different and therefore criteria necessary for clinical usage may be system specific. While prediction of new images is relatively fast, training these algorithms is computationally expensive. A single cross-validation fold trained to segment cartilage and bone from MRI images in the experiments of Chapter 4 using the latest Tesla V100

[****] The 132 MRI images used for building the SSM and conducting the analyses on acute changes to cartilage in Chapter 5 were segmented in ~ 198 minutes (3.3 hours) as opposed to the hundreds of hours manual analysis would have taken (~2-4+ hours per knee) (Shim et al., 2009).

[††††] The highest 25% quartile of T_2 had a 5 times higher risk of developing radiographic knee OA (KL >= 2) 2 years later (Razmjoo et al., 2020).

(Nvidia, California, USA) graphics cards[‡‡‡‡] took at least 10 days of continuous training, 60+ days for all folds. These times do not include time to develop and test the methods. Ultimately these tools have the ability to improve and expedite measurements, however, they will have their own sets of limitations that must be considered.

6.5 Open-source Software

This thesis would not be possible without open-source software. I almost exclusively performed data analyses in the Python programming language, which has a strong open-source community. I also leveraged open-source libraries for biomechanical analyses (biomechanical tool kit, OpenSim), image analyses (Insight Toolkit [ITK], Visualization Toolkit [VTK], ScikitImage, pycpd), statistics (StatsModels), machine learning (ScikitLearn, Keras, Tensorflow), and general computation (Numpy, Scipy, Pandas, Jupyter, Matplotlib). In the final chapter of my thesis, where I used SSMs and SPM to assess changes in cartilage from walking and bicycling, I contributed three new open-source packages: cycpd, pyfocusr, and pyKneeSPM.

[‡‡‡‡] Purchase price of a Tesla V100 GPU is $10,000 and cloud rental is $1,800/month. These prices do not include other computer hardware, memory, storage, or central processing unit (CPU).

The three packages that I created leverage one another, and together did all of the computation for the main analysis in Chapter 5. The pyfocusr package[§§§§] was used for registering bones of different participants together, a step that was crucial for creating the SSMs in Chapter 5. The pyKneeSPM package[*****] was designed to conduct SPM analyses on surface meshes, and thus was used for all of the SPM statistical analyses presented in Chapter 5. Registering bones together is necessary for conducing SPM analyses, and therefore the pyKneeSPM package used pyfocusr. Open-source contribution of these packages means that not only will I be able to better reproduce and run these analyses in the future, but others may also use them for their own research, expediting the research process. The pyfocusr algorithm is already being tested by a researcher in bioinformatics[†††††].

[§§§§] pyfocusr (https://github.com/gattia/pyfocusr) is a python implementation of the Feature Oriented Correspondence using Spectral Regularization (FOCUSR) algorithm (Lombaert et al., 2013). FOCUSR registers surface-meshes, e.g., the bones in this thesis. FOCUSR combines mesh spectral coordinates with features such as curvature of each point. These features help improve anatomical correspondence during registration.

[*****] pyKneeSPM (https://github.com/gattia/pyKneeSPM) conducts SPM analyses on 3D mesh data (e.g., bones). pyKneeSPM uses pyfocusr for registration and permutation statistics which are more robust than random field theory analyses (Eklund et al., 2016).

[†††††] https://github.com/gattia/pyfocusr/issues/1

Open-source software is developed and maintained in the open by contributors from anywhere. Independent interaction and review of the codebase is a major asset of open-source software because it identifies key features that users require as well as errors in the code. Once errors are identified, these become public knowledge so that every user can update their code. Key features that are requested by users can easily be added to the development path or can be carried out by any user if they require the feature faster. For example, I contributed "bug fixes" to the Keras deep learning library‡‡‡‡‡ that fixed and enabled 3D transpose convolutions which were crucial for implementation of the convolutional neural networks used for the segmentation algorithm in Chapter 4. In a closed-source ecosystem these features would only be available when it became a priority of the developers, not the users.

‡‡‡‡‡ https://github.com/keras-team/keras-contrib/pull/67

Existing open-source projects help inspire new software. I used the open-source pycpd[§§§§§] package as the basis for my first open-source library cycpd[******]. The reason for creating cycpd was that the pyfocusr algorithm I created and used for registering meshes together relied on the coherent point drift (CPD) algorithm, but the open-source pycpd version that was available was too slow. cycpd therefore uses the basic structure of pycpd developed by Siavash Khallaghi except some of the key components were re-written in Cython, a bridge between C code and Python. Cython was used to speed up (3x) iterative parts of the CPD algorithm. I also added the low-rank option for deformable registrations into pycpd[††††††]. Low-rank further sped up non-rigid registration (9x) and enabled better regularization of non-rigid registration methods. I have since contributed these non-rigid methods back to the pycpd library thus continuing its development and progress.

[§§§§§] pycpd (https://github.com/siavashk/pycpd) by Siavash Khallaghi is based on the Coherent Point Drift (CPD) algorithm originally written in Matlab (Myronenko & Xubo Song, 2010).

[******]https://github.com/gattia/cycpd

[††††††]Low-rank reduces the size of a linear system of equations that are solved iteratively during registration, thus reducing memory use and computation time.

I have not only used open-source software, but I have also learned immensely from this body of work. Open-source software provides users with a tool and shows how that tool was built. Reading code, contributing to other codebases, and generally being involved in the open-source data analysis community has helped hone my computer science and data science skills. For example, I have learned coding style best-practices, how to use version control tools such as git that are crucial for maintaining all software (open or closed), and how to deploy cloud resources for enabling massively distributed computation.

6.6 Integrating Biomechanics & Medical Imaging

Biomechanics and medical imaging are often combined to determine how biomechanical measures are related to OA changes in cartilage or bone cross-sectionally and longitudinally (Andriacchi, 2009; Bennell et al., 2011; Brisson et al., 2017; Chehab et al., 2014; Koo & Andriacchi, 2007; Maly et al., 2013, 2015). These studies typically assess some mechanical outcome, commonly the KAM, and determine how it relates to cross-sectional cartilage health, (Maly et al., 2015) or longitudinal changes in cartilage (Bennell et al., 2011). Medical images have also been used to extract information that then informs generic biomechanical models. For example, Harrington and colleagues used MRI data to formulate equations for determining the location of hip joint centers that could be applied during biomechanical gait analyses (Harrington et al., 2007). While these studies have been important in improving biomechanical models and building the

links between mechanics and actual changes to the joint, they often lack specificity of joint mechanics or cartilage outcomes.

Recent work has begun to change this lack of specificity. The Concurrent Optimization of Muscle Activations and Kinematics (COMAK) routine created by Colin Smith (Smith, 2017; Smith et al., 2018) is one example that enables quantification of cartilage surface pressures continuously over the cartilage surface using a biomechanical model. This predicts contact pressures of the articular cartilage of subject-specific models. However, it is more frequently used to run simulation analyses where parameters of a model of a single participant are perturbed and the resulting changes in cartilage pressures are monitored. For example, comparisons between healthy knees, anterior cruciate ligament (ACL) deficient knees, menisci deficient knees, and ACL and menisci deficient knees‡‡‡‡‡‡ (Smith et al., 2019). These simulations were performed on gait data from a single participant. In another study, Clouthier and colleagues used this same COMAK routine to investigate bone shape as an independent variable and the resulting cartilage surface pressures as well as arthromechanics as a response (Clouthier et al., 2019). Instead of directly running simulations for each participant for whom MRI data was available, Clouthier and colleagues created a SSM of the knee and used it to vary knee

‡‡‡‡‡‡ The knees deficient of ACL or meniscus were created by omitting these structures from a healthy knee.

joint morphology. This team also used gait data from a single participant for these simulations. Another study used the COMAK routine to estimate cartilage surface pressures during gait analysis for 15 health participants and compared these pressures to cartilage thicknesses acquired from MRI (Van Rossom et al., 2017). However, this study still used a generic musculoskeletal model for the COMAK routine as opposed to a subject-specific knee model to match the MRI data.

In this thesis, I proposed novel methods of analyzing the MRI and biomechanics data statistically using SPM conducted on the bone surfaces and open-sourced these methods as pyfocusr, and pyKneeSPM. In recent months (early 2020), Smith has open-sourced the COMAK routine (https://github.com/clnsmith/opensim-jam/) within the Joint and Articular Mechanics (JAM) OpenSim plugin. By combining the methods used in this thesis for statistical analysis (SPM, pyKneeSPM) and those presented previously by Clouthier and Smith for calculating surface pressures (COMAK, JAM) we have the ability, for the first time, to quantify cartilage surface pressures and directly determine how they are related to cartilage deformation at those same surface locations, *in vivo*. In addition to the *in vivo* studies, we can use SPM to expand the simulation studies conducted by Smith and Clouthier to now apply the perturbations (e.g., SSM variations or remove ACL) to gait or other biomechanical data collected from multiple participants. By including more participants for the biomechanical (e.g., gait) analyses the results of these simulations are more likely to generalize, as opposed to being highly specific to one person. New research designs enabled by combinations of these methods open a new

door to investigating the intersection of bone, cartilage and mechanics in the study of knee OA.

6.7 Limitations

The studies presented in this thesis were not without limitations. Chapters 2 and 3, which created equations to predict saddle position and explored relationships between joint kinematics and knee JRFs, only collected and analyzed data while participants wore running shoes. Running shoes were chosen because exercise amongst the general population, particularly those with musculoskeletal ailments such as knee OA were the primary motivators of this work. However, due to use of running shoes, these results may not generalize to cyclists who use cycling shoes and clipless systems. The initial study design intended to get participants who had their own cycling shoes to repeat half of the cycling bouts wearing those shoes. However, the duration of the existing protocol (upwards of 4.5 hours per participant), and the likelihood that participants would be systematically fatigued during the collections with cycling shoes precluded this being feasible.

Chapters 3 and 5 used JRFs in their analyses, these studies were limited as we were only able to estimate JRFs, and not measure them. With currently available technology, measurement of JRFs *in vivo* is only possible in those with instrumented prostheses and therefore we cannot be certain of the validity of these estimates. However, Chapter 3

Supplemental 2 (3.7.2) compared the JRFs estimated for two participants in this thesis to measured *in vivo* JRFs of a male participant cycling under similar conditions (power, cadence). This comparison showed good agreement, particularly at the regions of interest (peak forces), and the force profiles followed similar patterns.

Chapter 4, which describes and validates the segmentation framework was developed and tested on two MRI vendors and sequences. However, only healthy young adults were tested on one vendor/sequence (GE, FSPGR), and middle-aged or older adults across the spectrum of knee health were tested on the other (Siemens, DESS). Continued validation is necessary on other sequences, and over more knee conditions and participant ages.

Finally, Chapter 5 which explored the relationship between SSMs, JRFs, and changes in cartilage after walking and cycling was completed in a relatively small sample (n=16). Due to interruptions in data collections imposed by COVID-19 only ½ of the planned MRI collections were obtained. To improve robustness of the reported statistical findings permutation style statistical analyses were used (Nichols & Holmes, 2002).

6.8 Future Directions

This thesis highlights a few key areas of potential future research. As identified in the limitations, we are uncertain how cycling with clipless pedals will affect the proposed equations to predict saddle height and relationships between joint kinematics and JRFs. A future study could compare joint kinematics and knee JRFs in riders under multiple saddle configurations and using both running shoes and clipless pedals. Dependent on the findings of such an investigation it might be: (i) necessary to repeat the bicycle-fit related investigations of this thesis with clipless pedals, or (ii) possible to correct the findings for systematic differences in clipless systems versus running shoes, or (iii) that the results can reasonably be applied to clipless conditions directly.

As indicated in this discussion, the combination of the COMAK routine created by Colin Smith (Smith et al., 2018) and software open-sourced as part of this thesis (pyKneeSPM, pyfocusr) enable new ways of exploring the intersection of cartilage, bone, and mechanics to understand knee OA pathophysiology. First, I propose a study to acquire gait data of multiple participants, instead of just one as has been done previously for COMAK studies (Clouthier et al., 2019; Smith et al., 2019). Then, using the COMAK routine, changes in cartilage surface pressures of each participant could be simulated for multiple joint conditions (e.g., presence/absence of ACL, posterior cruciate ligament, and meniscus) or over a range of bone shapes using SSMs. Of particular interest would be to vary knee shape along principle modes of variation that are indicative of knee OA. The acquired data on cartilage surface pressures from multiple subjects under multiple

simulated joint conditions could then be analyzed using pyKneeSPM to statistically test if any differences in cartilage surface pressures exist between these knee conditions. Findings from such a study design would likely be more generalizable than the n=1 studies performed previously.

A second potential investigation that uses COMAK and pyKneeSPM could collect gait data as well as MRI data pre and post walking from multiple subjects at multiple walking speeds. With these data, the COMAK OpenSim knee model could be morphed to be subject specific, thus enabling prediction of cartilage surface pressures for each participant's specific joint morphology and their multiple gait speed trials. Changes in cartilage thickness, at the different speeds, could then be calculated continuously over the bone surface as presented in this thesis. Finally, pyKneeSPM could be used to determine whether the cartilage surface pressures predicted by the COMAK routine correlate spatially over the cartilage surface with the observed cartilage deformations. Such an investigation would provide predictive validity for the COMAK routine and greatly bolster evidence of its efficacy.

The segmentation framework described and validated in Chapter 4 produces excellent segmentations as measured using the DSC and average surface distance. However, studies of cartilage are primarily concerned with measures of cartilage thickness and volume. Findings from this thesis show that the algorithm is able to detect small changes

in cartilage from a single bout of loading, indicating that it is sensitive to change. Nonetheless, the next steps in establishing this pipeline for research and clinical usage are to determine reproducibility for measures of cartilage thickness. These reproducibility measures could be established continuously over the cartilage surface similar to how analyses were conducted in the final study of this thesis. In addition, reproducibility should be established for average thickness measurements in the key cartilage regions of interest, e.g., medial tibia, lateral tibia, medial weight-bearing femur, and lateral weight-bearing femur.

Last, maximum knee flexion angle was proposed in this discussion as a key target for bicycle-fit. This thesis provides evidence that maximum knee flexion angle is predictive of knee JRFs and that it can be used to predict saddle height. Future work is needed to establish the optimal range of maximum knee flexion angles. Separate lines of work could be conducted to determine optimal ranges for performance, and for injury avoidance. For example, a study exploring the mean and range of maximum knee flexion angles amongst high level athletes could provide initial guidance for performance-based recommendations.

6.9 Conclusion

This thesis advances knowledge and research methods at the intersection of bicycling, knee OA, cartilage, bone, and computational methods. Thesis studies on biomechanical

analyses of bicycling help improve our understanding of how bicycle-fit, primarily joint kinematics, affects joint forces and also provides news equations to predict saddle height for these same kinematics. In this thesis discussion I highlight why I believe that maximum knee flexion angle, which is rarely if ever talked about in the bicycling literature or cycling community, is likely an important variable for patellofemoral related bicycling ailments and should be a target for use by clinicians and everyday cyclists. A novel convolutional neural network framework for segmenting cartilage and bone from MRI data was presented. These segmentation methods promise to improve accuracy of extracted data while concurrently speeding up processing times therefore enabling better powered studies. Finally, integrating statistical shape models and new knowledge and methods from the previous chapters was performed in order to study how bone shape and knee JRFs affect the acute response of OA cartilage to physical activity. Chapter 5 shows for the first time that the acute response of cartilage to loading is dependent on knee OA severity (reflected in a SSM) and that knee JRFs are inversely related to knee OA severity. To accomplish these outcomes, Chapter 5 produced, and open-sourced, statistical and computational methods of combining and analyzing biomechanical data with joint anatomy and anatomical tissue compositions acquired from medical images. These methods hold great potential for evoking novel and specific insights that have not previously been possible and therefore open the door to better understanding the intersection of bone, cartilage, and mechanics in knee OA pathophysiology.

7 References

The following references are for Chapters 1 (Introduction) and 6 (Discussion). All other references are contained within the respective chapter.

Ahn, C., Bui, T. D., Lee, Y., Shin, J., & Park, H. (2016). Fully automated, level set-based segmentation for knee MRIs using an adaptive force function and template: Data from the osteoarthritis initiative. *BioMedical Engineering OnLine*, *15*(1). https://doi.org/10.1186/s12938-016-0225-7

Akella, S. V., Reddy Regatte, R., Gougoutas, A. J., Borthakur, A., Shapiro, E. M., Kneeland, J. B., Leigh, J. S., & Reddy, R. (2001). Proteoglycan-induced changes in T1ρ-relaxation of articular cartilage at 4T. *Magnetic Resonance in Medicine*, *46*(3), 419–423.

Alkatan, M., Baker, J. R., Machin, D. R., Park, W., Akkari, A. S., Pasha, E. P., & Tanaka, H. (2016). Improved Function and Reduced Pain after Swimming and Cycling Training in Patients with Osteoarthritis. *The Journal of Rheumatology*, *43*(3), 666–672. https://doi.org/10.3899/jrheum.151110

Altman, R., Asch, E., Bloch, D., Bole, G., Borenstein, D., Brandt, K., Christy, W., Cooke, T. D., Greenwald, R., Hochberg, M., & others. (1986). Development of criteria for the classification and reporting of osteoarthritis: Classification of osteoarthritis of the knee. *Arthritis & Rheumatism*, *29*(8), 1039–1049.

Ambellan, F., Tack, A., Ehlke, M., & Zachow, S. (2019). Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional

neural networks: Data from the Osteoarthritis Initiative. *Medical Image Analysis*, *52*, 109–118. https://doi.org/10.1016/j.media.2018.11.009

Anderson, F. C., & Pandy, M. G. (2001). Static and dynamic optimization solutions for gait are practically equivalent. *Journal of Biomechanics*, *34*(2), 153–161. https://doi.org/10.1016/S0021-9290(00)00155-X

Andriacchi, T. P. (2009). Gait Mechanics Influence Healthy Cartilage Morphology and Osteoarthritis of the Knee. *The Journal of Bone and Joint Surgery (American)*, *91*(Supplement_1), 95. https://doi.org/10.2106/JBJS.H.01408

Antony, B., Ding, C., Stannus, O., Cicuttini, F., & Jones, G. (2011). Association of Baseline Knee Bone Size, Cartilage Volume, and Body Mass Index with Knee Cartilage Loss Over Time: A Longitudinal Study in Younger or Middle-aged Adults. *The Journal of Rheumatology*, *38*(9), 1973–1980. https://doi.org/10.3899/jrheum.101309

Arden, N., & Nevitt, M. (2006). Osteoarthritis: Epidemiology. *Best Practice & Research Clinical Rheumatology*, *20*(1), 3–25. https://doi.org/10.1016/j.berh.2005.09.007

Babel, H., Omoumi, P., Andriacchi, T. P., Jolles, B. M., & Favre, J. (2020). New insight on the subchondral bone and cartilage functional unit: Bone mineral density and cartilage thickness are spatially correlated in non-osteoarthritic femoral condyles. *Osteoarthritis and Cartilage Open*, 100079. https://doi.org/10.1016/j.ocarto.2020.100079

Balamoody, S., Williams, T. G., Wolstenholme, C., Waterton, J. C., Bowes, M., Hodgson, R., Zhao, S., Scott, M., Taylor, C. J., & Hutchinson, C. E. (2013). Magnetic resonance transverse relaxation time T2 of knee cartilage in osteoarthritis at 3-T: A cross-sectional multicentre, multivendor reproducibility study. *Skeletal Radiology*, *42*(4), 511–520. https://doi.org/10.1007/s00256-012-1511-5

Barr, A. J., Campbell, T. M., Hopkinson, D., Kingsbury, S. R., Bowes, M. A., & Conaghan, P. G. (2015). A systematic review of the relationship between subchondral bone features, pain and structural pathology in peripheral joint osteoarthritis. *Arthritis Research & Therapy*, *17*(1). https://doi.org/10.1186/s13075-015-0735-x

Barr, A. J., Dube, B., Hensor, E. M. A., Kingsbury, S. R., Peat, G., Bowes, M. A., Sharples, L. D., & Conaghan, P. G. (2016). The relationship between three-dimensional knee MRI bone shape and total knee replacement?a case control study: Data from the Osteoarthritis Initiative. *Rheumatology*, *55*(9), 1585–1593. https://doi.org/10.1093/rheumatology/kew191

Bennell, K. L., Bowles, K.-A., Wang, Y., Cicuttini, F., Davies-Tuck, M., & Hinman, R. S. (2011). Higher dynamic medial knee load predicts greater cartilage loss over 12 months in medial knee osteoarthritis. *Annals of the Rheumatic Diseases*, *70*(10), 1770–1774. https://doi.org/10.1136/ard.2010.147082

Bianco, A., Bellafiore, M., Battaglia, G., Paoli, A., Caramazza, G., Farina, F., & Palma, A. (2010). The effects of indoor cycling training in sedentary overweight women. *The Journal of Sports Medicine and Physical Fitness*, *50*(2), 159–165.

Bini, R. R. (2016). The need for a link between bike fitting and injury risk. *Journal of Science and Cycling*, 2.

Bini, R. R. (2020). Influence of saddle height in 3D knee loads commuter cyclists: A statistical parametric mapping analysis. *Journal of Sports Sciences*, 1–14. https://doi.org/10.1080/02640414.2020.1816289

Bini, R. R., Hume, P. A., & Croft, J. L. (2011). Effects of bicycle saddle height on knee injury risk and cycling performance. *Sports Medicine*, *41*(6), 463–476.

Bini, R. R., Hume, P. A., & Kilding, A. E. (2014). Saddle height effects on pedal forces, joint mechanical work and kinematics of cyclists and triathletes. *European Journal of Sport Science*, *14*(1), 44–52. https://doi.org/10.1080/17461391.2012.725105

Bini, R. R., Tamborindeguy, A. C., & Mota, C. B. (2010). Effects of saddle height, pedaling cadence, and workload on joint kinetics and kinematics during cycling. *Journal of Sport Rehabilitation*, *19*(3), 301–314.

Birtwhistle, R., Morkem, R., Peat, G., Williamson, T., Green, M. E., Khan, S., & Jordan, K. P. (2015). Prevalence and management of osteoarthritis in primary care: An epidemiologic cohort study from the Canadian Primary Care Sentinel Surveillance Network. *CMAJ Open*, *3*(3), E270–E275. https://doi.org/10.9778/cmajo.20150018

Blagojevic, M., Jinks, C., Jeffery, A., & Jordan, K. P. (2010). Risk factors for onset of osteoarthritis of the knee in older adults: A systematic review and meta-analysis. *Osteoarthritis and Cartilage*, *18*(1), 24–33. https://doi.org/10.1016/j.joca.2009.08.010

Bombardier, C., Hawker, G., & Mosher, D. (2011). *The Impact of Arthritis in Canada: Today and Over the Next 30 Years*. Arthritis Alliance of Canada.

Braun, H. J., & Gold, G. E. (2012). Diagnosis of osteoarthritis: Imaging. *Bone*, *51*(2), 278–288. https://doi.org/10.1016/j.bone.2011.11.019

Brenneman, E. C., Gatti, A. A., & Maly, M. R. (2019a, May). *Influence of sex on tibiofemoral cartilage response to running in healthy young men and women*. Osteoarthritis Research Society International, Toronto, Ontario, Canada.

Brenneman, E. C., Gatti, A. A., & Maly, M. R. (2019b, August). *Sex does not influence tibiofemoral cartilage response to running in healthy adults*. International Society for Biomechanics, Calgary, Alberta, Canada.

Brenneman, E. C., Kuntz, A. B., Wiebenga, E. G., & Maly, M. R. (2015). A Yoga Strengthening Program Designed to Minimize the Knee Adduction Moment for Women with Knee Osteoarthritis: A Proof-Of-Principle Cohort Study. *PLOS ONE*, *10*(9), e0136854. https://doi.org/10.1371/journal.pone.0136854

Brisson, N. M., Wiebenga, E. G., Stratford, P. W., Beattie, K. A., Totterman, S., Tamez-Pe?a, J. G., Callaghan, J. P., Adachi, J. D., & Maly, M. R. (2017). Baseline knee adduction moment interacts with body mass index to predict loss of medial tibial

cartilage volume over 2.5 years in knee Osteoarthritis: CARTILAGE VOLUME LOSS IN osteoarthritis. *Journal of Orthopaedic Research*. https://doi.org/10.1002/jor.23564

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *ArXiv:2005.14165 [Cs]*. http://arxiv.org/abs/2005.14165

Buckwalter, J. A., Mankin, H. J., & Grodzinsky, A. J. (2005). Articular cartilage and osteoarthritis. *Instructional Course Lectures*, *54*, 465–480.

Burr, D. B., & Gallant, M. A. (2012). Bone remodelling in osteoarthritis. *Nature Reviews Rheumatology*, *8*(11), 665–673. https://doi.org/10.1038/nrrheum.2012.130

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Cha, J. G., Lee, J. C., Kim, H. J., Han, J. K., Lee, E. H., Kim, Y. D., & Jeon, C. H. (2012). Comparison of MRI T2 Relaxation Changes of Knee Articular Cartilage before and after Running between Young and Old Amateur Athletes. *Korean Journal of Radiology*, *13*(5), 594. https://doi.org/10.3348/kjr.2012.13.5.594

Chakravarty, E. F., Hubert, H. B., Lingala, V. B., Zatarain, E., & Fries, J. F. (2008). Long Distance Running and Knee Osteoarthritis. *American Journal of Preventive Medicine*, *35*(2), 133–138. https://doi.org/10.1016/j.amepre.2008.03.032

Chang, A. H., Moisio, K. C., Chmiel, J. S., Eckstein, F., Guermazi, A., Prasad, P. V., Zhang, Y., Almagor, O., Belisle, L., Hayes, K., & Sharma, L. (2015). External knee adduction and flexion moments during gait and medial tibiofemoral disease progression in knee osteoarthritis. *Osteoarthritis and Cartilage*. https://doi.org/10.1016/j.joca.2015.02.005

Chehab, E. F., Favre, J., Erhart-Hledik, J. C., & Andriacchi, T. P. (2014). Baseline knee adduction and flexion moments during walking are both associated with 5?year cartilage changes in patients with medial knee osteoarthritis. *Osteoarthritis and Cartilage*, *22*(11), 1833–1839. https://doi.org/10.1016/j.joca.2014.08.009

Cheung, R. T. H., Ho, K. K. W., Au, I. P. H., An, W. W., Zhang, J. H. W., Chan, Z. Y. S., Deluzio, K., & Rainbow, M. J. (2018). Immediate and short-term effects of gait retraining on the knee joint moments and symptoms in patients with early tibiofemoral joint osteoarthritis: A randomized controlled trial. *Osteoarthritis and Cartilage*, *26*(11), 1479–1486. https://doi.org/10.1016/j.joca.2018.07.011

Choi, J.-A., & Gold, G. E. (2011). MR Imaging of Articular Cartilage Physiology. *Magnetic Resonance Imaging Clinics of North America*, *19*(2), 249–282. https://doi.org/10.1016/j.mric.2011.02.010

Clouthier, A. L., Smith, C. R., Vignos, M. F., Thelen, D. G., Deluzio, K. J., & Rainbow, M. J. (2019). The effect of articular geometry features identified using statistical shape modelling on knee biomechanics. *Medical Engineering & Physics*, *66*, 47–55. https://doi.org/10.1016/j.medengphy.2019.02.009

Collins, A. T., Kulvaranon, M. L., Cutcliffe, H. C., Utturkar, G. M., Smith, W. A. R., Spritzer, C. E., Guilak, F., & DeFrate, L. E. (2018). Obesity alters the in vivo mechanical response and biochemical properties of cartilage as measured by MRI. *Arthritis Research & Therapy*, *20*(1). https://doi.org/10.1186/s13075-018-1727-4

Creaby, M. W. (2015). It's not all about the knee adduction moment: The role of the knee flexion moment in medial knee joint loading. *Osteoarthritis and Cartilage*, *23*(7), 1038–1040. https://doi.org/10.1016/j.joca.2015.03.032

Creamer, P., & Hochberg, M. C. (1997). Osteoarthritis. *Lancet*, *350*(9076), 503–508. https://doi.org/10.1016/S0140-6736(97)07226-7

Dam, E. B., Lillholm, M., Marques, J., & Nielsen, M. (2015). Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from the osteoarthritis initiative. *Journal of Medical Imaging (Bellingham, Wash.)*, *2*(2), 024001. https://doi.org/10.1117/1.JMI.2.2.024001

Damm, P., Kutzner, I., Bergmann, G., Rohlmann, A., & Schmidt, H. (2017). Comparison of in vivo measured loads in knee, hip and spinal implants during level walking. *Journal of Biomechanics*, *51*, 128–132. https://doi.org/10.1016/j.jbiomech.2016.11.060

Davis, M. A., Ettinger, W. H., & Neuhaus, J. M. (1990). Obesity and osteoarthritis of the knee: Evidence from the National Health and Nutrition Examination Survey (NHANES I). *Seminars in Arthritis and Rheumatism*, *20*(3 Suppl 1), 34–41.

De Bernardo, N., Barrios, C., Vera, P., Laíz, C., & Hadala, M. (2012). Incidence and risk for traumatic and overuse injuries in top-level road cyclists. *Journal of Sports Sciences*, *30*(10), 1047–1053. https://doi.org/10.1080/02640414.2012.687112

Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., Guendelman, E., & Thelen, D. G. (2007). OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering*, *54*(11), 1940–1950. https://doi.org/10.1109/TBME.2007.901024

Delp, S. L., Loan, J. P., Hoy, M. G., Zajac, F. E., Topp, E. L., & Rosen, J. M. (1990). An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering*, *37*(8), 757–767.

DeMers, M. S., Pal, S., & Delp, S. L. (2014). Changes in tibiofemoral forces due to variations in muscle activity during walking: TIBIOFEMORAL FORCES AND MUSCLE ACTIVITY. *Journal of Orthopaedic Research*, *32*(6), 769–776. https://doi.org/10.1002/jor.22601

Desai, A., Caliva, F., Iriondo, C., Khosravan, N., Mortazi, A., Jambawalikar, S., Torigian, D., Ellermann, J., Akcakaya, M., Bagci, U., Tibrewala, R., Flament, I.,

O'Brian, M., Majumdar, S., Perslev, M., Pai, A., Igel, C., Dam, E., Gaj, S., … Chaudhari, A. (2020). A multi-institute automated segmentation evaluation on a standard dataset: Findings from the international workshop on osteoarthritis imaging segmentation challenge. *Osteoarthritis and Cartilage*, *28*, S304–S305. https://doi.org/10.1016/j.joca.2020.02.477

Deshpande, B. R., Katz, J. N., Solomon, D. H., Yelin, E. H., Hunter, D. J., Messier, S. P., Suter, L. G., & Losina, E. (2016). Number of Persons With Symptomatic Knee Osteoarthritis in the US: Impact of Race and Ethnicity, Age, Sex, and Obesity: Symptomatic Knee OA in the US. *Arthritis Care & Research*. https://doi.org/10.1002/acr.22897

Detry, M. A., & Ma, Y. (2016). Analyzing Repeated Measurements Using Mixed Models. *JAMA*, *315*(4), 407. https://doi.org/10.1001/jama.2015.19394

Dettori, N. J., & Norvell, D. C. (2006). Non-traumatic bicycle injuries: A review of the literature. *Sports Medicine (Auckland, N.Z.)*, *36*(1), 7–18.

Dodin, P., Pelletier, J., Martel-Pelletier, J., & Abram, F. (2010). Automatic Human Knee Cartilage Segmentation From 3-D Magnetic Resonance Images. *IEEE Transactions on Biomedical Engineering*, *57*(11), 2699–2711. https://doi.org/10.1109/TBME.2010.2058112

Dore, D. A., Winzenberg, T. M., Ding, C., Otahal, P., Pelletier, J.-P., Martel-Pelletier, J., Cicuttini, F. M., & Jones, G. (2013). The association between objectively measured physical activity and knee structural change using MRI. *Annals of the*

Rheumatic Diseases, *72*(7), 1170–1175. https://doi.org/10.1136/annrheumdis-2012-201691

Eckstein, F. (2005). In vivo cartilage deformation after different types of activity and its dependence on physical training status. *Annals of the Rheumatic Diseases*, *64*(2), 291–295. https://doi.org/10.1136/ard.2004.022400

Eckstein, Felix, Tieschky, M., Faber, S., Englmeier, K.-H., & Reiser, M. (1999). Functional analysis of articular cartilage deformation, recovery, and fluid flow following dynamic exercise in vivo. *Anatomy and Embryology*, *200*(4), 419–424.

Edd, S. N., Omoumi, P., Andriacchi, T. P., Jolles, B. M., & Favre, J. (2018). Modeling knee osteoarthritis pathophysiology using an integrated joint system (IJS): A systematic review of relationships among cartilage thickness, gait mechanics, and subchondral bone mineral density. *Osteoarthritis and Cartilage*, *26*(11), 1425–1437. https://doi.org/10.1016/j.joca.2018.06.017

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905. https://doi.org/10.1073/pnas.1602413113

Esculier, J.-F., Jarrett, M., Krowchuk, N. M., Rauscher, A., Wiggermann, V., Taunton, J. E., Wilson, D. R., Gatti, A. A., & Hunt, M. A. (2019). Cartilage recovery in runners with and without knee osteoarthritis: A pilot study. *The Knee*. https://doi.org/10.1016/j.knee.2019.07.011

Fang, Y., Fitzhugh, E. C., Crouter, S. E., Gardner, J. K., & Zhang, S. (2016). Effects of Workloads and Cadences on Frontal Plane Knee Biomechanics in Cycling: *Medicine & Science in Sports & Exercise*, *48*(2), 260–266. https://doi.org/10.1249/MSS.0000000000000759

Farrokhi, S., Jayabalan, P., Gustafson, J. A., Klatt, B. A., Sowa, G. A., & Piva, S. R. (2017). The influence of continuous versus interval walking exercise on knee joint loading and pain in patients with knee osteoarthritis. *Gait & Posture*, *56*, 129–133. https://doi.org/10.1016/j.gaitpost.2017.05.015

Felson, D. T. (2006). Osteoarthritis of the knee. *New England Journal of Medicine*, *354*(8), 841–848.

Felson, D. T., Anderson, J. J., Naimark, A., Walker, A. M., & Meenan, R. F. (1988). Obesity and knee osteoarthritis: The Framingham Study. *Annals of Internal Medicine*, *109*(1), 18–24.

Felson, D. T., Lawrence, R. C., Dieppe, P. A., Hirsch, R., Helmick, C. G., Jordan, J. M., Kington, R. S., Lane, N. E., Nevitt, M. C., & Zhang, Y. (2000). Osteoarthritis: New insights. Part 1: the disease and its risk factors. *Annals of Internal Medicine*, *133*(8), 635–646.

Felson, D. T., Niu, J., Yang, T., Torner, J., Lewis, C. E., Aliabadi, P., Sack, B., Sharma, L., Guermazi, A., Goggins, J., & Nevitt, M. C. (2013). Physical activity, alignment and knee osteoarthritis: Data from MOST and the OAI. *Osteoarthritis and Cartilage*, *21*(6), 789–795. https://doi.org/10.1016/j.joca.2013.03.001

Felson, D. T., Zhang, Y., Anthony, J. M., Naimark, A., & Anderson, J. J. (1992). Weight loss reduces the risk for symptomatic knee osteoarthritis in women. The Framingham Study. *Annals of Internal Medicine*, *116*(7), 535–539.

Ferrer-Roca, V., Roig, A., Galilea, P., & García-López, J. (2012). Influence of saddle height on lower limb kinematics in well-trained cyclists: Static vs. dynamic evaluation in bike fitting. *The Journal of Strength & Conditioning Research*, *26*(11), 3025–3029.

Finucane, F. M., Sharp, S. J., Purslow, L. R., Horton, K., Horton, J., Savage, D. B., Brage, S., Besson, H., De Lucia Rolfe, E., Sleigh, A., Martin, H. J., Aihie Sayer, A., Cooper, C., Ekelund, U., Griffin, S. J., & Wareham, N. J. (2010). The effects of aerobic exercise on metabolic risk, insulin sensitivity and intrahepatic lipid in healthy older people from the Hertfordshire Cohort Study: A randomised controlled trial. *Diabetologia*, *53*(4), 624–631. https://doi.org/10.1007/s00125-009-1641-z

Fonda, B., Sarabon, N., & Li, F.-X. (2014). Validity and reliability of different kinematics methods used for bike fitting. *Journal of Sports Sciences*, *32*(10), 940–946. https://doi.org/10.1080/02640414.2013.868919

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold. *Magnetic*

Resonance in Medicine, *33*(5), 636–647. https://doi.org/10.1002/mrm.1910330508

Fransen, M., McConnell, S., Harmer, A. R., Van der Esch, M., Simic, M., & Bennell, K. L. (2015). Exercise for osteoarthritis of the knee. In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/14651858.CD004376.pub3

Fripp, J., Crozier, S., Warfield, S. K., & Ourselin, S. (2010). Automatic Segmentation and Quantitative Analysis of the Articular Cartilages From Magnetic Resonance Images of the Knee. *IEEE Transactions on Medical Imaging*, *29*(1), 55–64. https://doi.org/10.1109/TMI.2009.2024743

Friston, K. J., Worsley, K. J., Frackowiak, R. S., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, *1*(3), 210–220.

Gaj, S., Yang, M., Nakamura, K., & Li, X. (2020). Automated cartilage and meniscus segmentation of knee MRI with conditional generative adversarial networks. *Magnetic Resonance in Medicine*, *84*(1), 437–449. https://doi.org/10.1002/mrm.28111

Gardner, J. K., Zhang, S., Liu, H., Klipple, G., Stewart, C., Milner, C. E., & Asif, I. M. (2015). Effects of toe-in angles on knee biomechanics in cycling of patients with medial knee osteoarthritis. *Clinical Biomechanics*, *30*(3), 276–282. https://doi.org/10.1016/j.clinbiomech.2015.01.003

Gatti, A. A. (2015). *THE IN VIVO RESPONSE OF KNEE ARTICULAR CARTILAGE TO RUNNING AND BICYCLING*. McMaster University.

Gatti, A. A., & Maly, M. R. (2018, August). *PIXELWISE CHANGES IN PATELLAR CARTILAGE AFTER RUNNING AND BICYCLING*. Canadian Society for Biomechanics, Halifax, Nova Scotia, Canada.

Gatti, A. A., Noseworthy, M. D., Stratford, P. W., Brenneman, E. C., Totterman, S., Tamez-Pena, J. G., & Maly, M. R. (2017). Acute Changes in Knee Cartilage T2 After Running and Bicycling. *Journal of Biomechanics*, *In-press*.

Gatti, A. A., Stratford, P. W., Noseworthy, M. D., & Maly, M. R. (2016, March). *DEPTH DEPENDENT RESPONSE OF TIBIAL CARTILAGE TO PHYSICAL ACTIVITY*. Canadian Society of Biomechanics.

Gerus, P., Sartori, M., Besier, T. F., Fregly, B. J., Delp, S. L., Banks, S. A., Pandy, M. G., D'Lima, D. D., & Lloyd, D. G. (2013). Subject-specific knee joint geometry improves predictions of medial tibiofemoral contact forces. *Journal of Biomechanics*, *46*(16), 2778–2786. https://doi.org/10.1016/j.jbiomech.2013.09.005

Glyn-Jones, S., Palmer, A. J. R., Agricola, R., Price, A. J., Vincent, T. L., Weinans, H., & Carr, A. J. (2015). Osteoarthritis. *The Lancet*, *386*(9991), 376–387. https://doi.org/10.1016/S0140-6736(14)60802-3

Goldring, M. B., & Goldring, S. R. (2010). Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis: Articular cartilage and subchondral bone.

Annals of the New York Academy of Sciences, *1192*(1), 230–237. https://doi.org/10.1111/j.1749-6632.2009.05240.x

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org

Gorton, G. E., Hebert, D. A., & Gannotti, M. E. (2009). Assessment of the kinematic variability among 12 motion analysis laboratories. *Gait & Posture*, *29*(3), 398–402. https://doi.org/10.1016/j.gaitpost.2008.10.060

Guccione, A. A., Felson, D. T., Anderson, J. J., Anthony, J. M., Zhang, Y., Wilson, P. W., Kelly-Hayes, M., Wolf, P. A., Kreger, B. E., & Kannel, W. B. (1994). The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *American Journal of Public Health*, *84*(3), 351–358.

Guermazi, A., Roemer, F. W., Burstein, D., & Hayashi, D. (2011). Why radiography should no longer be considered a surrogate outcome measure for longitudinal assessment of cartilage in knee osteoarthritis. *Arthritis Research & Therapy*, *13*(6), 247. https://doi.org/10.1186/ar3488

Hamner, S. R., Seth, A., & Delp, S. L. (2010). Muscle contributions to propulsion and support during running. *Journal of Biomechanics*, *43*(14), 2709–2716. https://doi.org/10.1016/j.jbiomech.2010.06.025

Harrington, M. E., Zavatsky, A. B., Lawson, S. E. M., Yuan, Z., & Theologis, T. N. (2007). Prediction of the hip joint centre in adults, children, and patients with

cerebral palsy based on magnetic resonance imaging. *Journal of Biomechanics*, *40*(3), 595–602. https://doi.org/10.1016/j.jbiomech.2006.02.003

Hayashi, D., Roemer, F. W., & Guermazi, A. (2018). Imaging of osteoarthritis—Recent research developments and future perspective. *The British Journal of Radiology*, 20170349. https://doi.org/10.1259/bjr.20170349

Hinterwimmer, S., Krammer, M., Kr□tz, M., Glaser, C., Baumgart, R., Reiser, M., & Eckstein, F. (2004). Cartilage atrophy in the knees of patients after seven weeks of partial load bearing. *Arthritis & Rheumatism*, *50*(8), 2516–2520. https://doi.org/10.1002/art.20378

Holmes, J. C., Pruitt, A. L., & Whalen, N. J. (1994). Lower extremity overuse in bicycling. *Clinics in Sports Medicine*, *13*(1), 187–205.

Hungerfordm, D., & Maureenbarryb, A. N. D. (1979). *Biomechanics of the Patellofemoral Joint*. *144*, 7.

Hunter, D. J., Altman, R. D., Cicuttini, F., Crema, M. D., Duryea, J., Eckstein, F., Guermazi, A., Kijowski, R., Link, T. M., Martel-Pelletier, J., Miller, C. G., Mosher, T. J., Ochoa-Albíztegui, R. E., Pelletier, J.-P., Peterfy, C., Raynauld, J.-P., Roemer, F. W., Totterman, S. M., & Gold, G. E. (2015). OARSI Clinical Trials Recommendations: Knee imaging in clinical trials in osteoarthritis. *Osteoarthritis and Cartilage*, *23*(5), 698–715. https://doi.org/10.1016/j.joca.2015.03.012

Hunter, D. J., Nevitt, M., Lynch, J., Kraus, V. B., Katz, J. N., Collins, J. E., Bowes, M., Guermazi, A., Roemer, F. W., & Losina, E. (2016). Longitudinal validation of periarticular bone area and 3D shape as biomarkers for knee OA progression? Data from the FNIH OA Biomarkers Consortium. *Annals of the Rheumatic Diseases*, *75*(9), 1607–1614. https://doi.org/10.1136/annrheumdis-2015-207602

Janocha, K., & Czarnecki, W. M. (2017). On Loss Functions for Deep Neural Networks in Classification. *ArXiv:1702.05659 [Cs]*. http://arxiv.org/abs/1702.05659

Jordan, J. L., Holden, M. A., Mason, E. E., & Foster, N. E. (2010). Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/14651858.CD005956.pub2

Jorge, M., & Hull, M. L. (1986). Analysis of EMG measurements during bicycle pedalling. *Journal of Biomechanics*, *19*(9), 683–694.

Kainz, H., Graham, D., Edwards, J., Walsh, H. P. J., Maine, S., Boyd, R. N., Lloyd, D. G., Modenese, L., & Carty, C. P. (2017). Reliability of four models for clinical gait analysis. *Gait & Posture*, *54*, 325–331. https://doi.org/10.1016/j.gaitpost.2017.04.001

Kanko, R., Laende, E., Selbie, S., & Deluzio, K. (2020). *Inter-session repeatability of* Theia3D *markerless motion capture gait kinematics* [Preprint]. Biophysics. https://doi.org/10.1101/2020.06.23.155358

Kanko, R., Strutzenberger, G., Brown, M., Selbie, S., & Deluzio, K. (2020). *Assessment of spatiotemporal gait parameters using a deep learning algorithm-based markerless motion capture system* [Preprint]. engrXiv. https://doi.org/10.31224/osf.io/j4rbg

Kaufman, K., Miller, E., Kingsbury, T., Russell Esposito, E., Wolf, E., Wilken, J., & Wyatt, M. (2016). Reliability of 3D gait data across multiple laboratories. *Gait & Posture*, *49*, 375–381. https://doi.org/10.1016/j.gaitpost.2016.07.075

Kellgren, J. H., & Lawrence, J. S. (1957). Radiological assessment of osteo-arthrosis. *Annals of the Rheumatic Diseases*, *16*(4), 494–502.

Kluzek, S., Newton, J. L., & Arden, N. K. (2015). Is osteoarthritis a metabolic disorder? *British Medical Bulletin*, *115*(1), 111–121. https://doi.org/10.1093/bmb/ldv028

Knarr, B. A., & Higginson, J. S. (2015). Practical approach to subject-specific estimation of knee joint contact force. *Journal of Biomechanics*, *48*(11), 2897–2902. https://doi.org/10.1016/j.jbiomech.2015.04.020

Kohn, M. D., Sassoon, A. A., & Fernando, N. D. (2016). Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. *Clinical Orthopaedics and Related Research®*, *474*(8), 1886–1893. https://doi.org/10.1007/s11999-016-4732-4

Koo, S., & Andriacchi, T. P. (2007). A comparison of the influence of global functional loads vs. Local contact anatomy on articular cartilage thickness at the knee.

Journal of Biomechanics, *40*(13), 2961–2966. https://doi.org/10.1016/j.jbiomech.2007.02.005

Krevolin, J. L., Pandy, M. G., & Pearce, J. C. (2004). Moment arm of the patellar tendon in the human knee. *Journal of Biomechanics*, *37*(5), 785–788. https://doi.org/10.1016/j.jbiomech.2003.09.010

Kutzner, I., Trepczynski, A., Heller, M. O., & Bergmann, G. (2013). Knee Adduction Moment and Medial Contact Force – Facts about Their Correlation during Gait. *PLoS ONE*, *8*(12), e81036. https://doi.org/10.1371/journal.pone.0081036

Lai, A. K. M., Arnold, A. S., & Wakeling, J. M. (2017). Why are Antagonist Muscles Co-activated in My Simulation? A Musculoskeletal Model for Analysing Human Locomotor Tasks. *Annals of Biomedical Engineering*, *45*(12), 2762–2774. https://doi.org/10.1007/s10439-017-1920-7

Lee, I.-M., & Skerrett, P. J. (2001). Physical activity and all-cause mortality: What is the dose-response relation? *Medicine and Science in Sports and Exercise*, *33*(6; SUPP), S459–S471.

Lee, Y., Choi, J., & Hwang, N. S. (2018). Regulation of lubricin for functional cartilage tissue regeneration: A review. *Biomaterials Research*, *22*(1). https://doi.org/10.1186/s40824-018-0118-x

Lerner, Z. F., Haight, D. J., DeMers, M. S., Board, W. J., & Browning, R. C. (2014). The Effects of Walking Speed on Tibiofemoral Loading Estimated Via

Musculoskeletal Modeling. *Journal of Applied Biomechanics*, *30*(2), 197–205. https://doi.org/10.1123/jab.2012-0206

Li, G., Lopez, O., & Rubash, H. (2001). Variability of a Three-Dimensional Finite Element Model Constructed Using Magnetic Resonance Images of a Knee for Joint Contact Stress Analysis. *Journal of Biomechanical Engineering*, *123*(4), 341. https://doi.org/10.1115/1.1385841

Lin, Y.-C., Dorn, T. W., Schache, A. G., & Pandy, M. G. (2012). Comparison of different methods for estimating muscle forces in human movement. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, *226*(2), 103–112. https://doi.org/10.1177/0954411911429401

Lo, G. H., Musa, S. M., Driban, J. B., Kriska, A. M., McAlindon, T. E., Souza, R. B., Petersen, N. J., Storti, K. L., Eaton, C. B., Hochberg, M. C., Jackson, R. D., Kwoh, C. K., Nevitt, M. C., & Suarez-Almazor, M. E. (2018). Running does not increase symptoms or structural progression in people with knee osteoarthritis: Data from the osteoarthritis initiative. *Clinical Rheumatology*, *37*(9), 2497–2504. https://doi.org/10.1007/s10067-018-4121-3

Lombaert, H., Grady, L., Polimeni, J. R., & Cheriet, F. (2013). FOCUSR: Feature Oriented Correspondence Using Spectral Regularization--A Method for Precise Surface Matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(9), 2143–2160. https://doi.org/10.1109/TPAMI.2012.276

Longpré, H. S., Brenneman, E. C., Johnson, A. L. M., & Maly, M. R. (2015). Identifying yoga-based knee strengthening exercises using the knee adduction moment. *Clinical Biomechanics*, *30*(8), 820–826. https://doi.org/10.1016/j.clinbiomech.2015.06.007

Lories, R. J., & Luyten, F. P. (2011). The bone–cartilage unit in osteoarthritis. *Nature Reviews Rheumatology*, *7*(1), 43–49. https://doi.org/10.1038/nrrheum.2010.197

Lu, X. L., & Mow, V. C. (2008). Biomechanics of Articular Cartilage and Determination of Material Properties: *Medicine & Science in Sports & Exercise*, *40*(2), 193–199. https://doi.org/10.1249/mss.0b013e31815cb1fc

MacKay, J. W., Kaggie, J. D., Treece, G. M., McDonnell, S. M., Khan, W., Roberts, A. R., Janiczek, R. L., Graves, M. J., Turmezei, T. D., McCaskie, A. W., & Gilbert, F. J. (2020). Three-Dimensional Surface-Based Analysis of Cartilage MRI Data in Knee Osteoarthritis: Validation and Initial Clinical Application. *Journal of Magnetic Resonance Imaging*. https://doi.org/10.1002/jmri.27193

Maly, M. R., Acker, S. M., Totterman, S., Tamez-Peña, J., Stratford, P. W., Callaghan, J. P., Adachi, J. D., & Beattie, K. A. (2015). Knee adduction moment relates to medial femoral and tibial cartilage morphology in clinical knee osteoarthritis. *Journal of Biomechanics*, *48*(12), 3495–3501. https://doi.org/10.1016/j.jbiomech.2015.04.039

Maly, M. R., Robbins, S. M., Stratford, P. W., Birmingham, T. B., & Callaghan, J. P. (2013). Cumulative knee adductor load distinguishes between healthy and

osteoarthritic knees–A proof of principle study. *Gait & Posture*, *37*(3), 397–401. https://doi.org/10.1016/j.gaitpost.2012.08.013

Manal, K., Gardinier, E., Buchanan, T. S., & Snyder-Mackler, L. (2015). A more informed evaluation of medial compartment loading: The combined use of the knee adduction and flexor moments. *Osteoarthritis and Cartilage*, *23*(7), 1107–1111. https://doi.org/10.1016/j.joca.2015.02.779

Mangione, K. K., McCully, K., Gloviak, A., Lefebvre, I., Hofmann, M., & Craik, R. (1999). The effects of high-intensity and low-intensity cycle ergometry in older adults with knee osteoarthritis. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *54*(4), M184-190.

Mansell, J. P., Collins, C., & Bailey, A. J. (2007). Bone, not cartilage, should be the major focus in osteoarthritis. *Nature Clinical Practice Rheumatology*, *3*(6), 306–307. https://doi.org/10.1038/ncprheum0505

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

Mazor, M., Best, T. M., Cesaro, A., Lespessailles, E., & Toumi, H. (2019). Osteoarthritis biomarker responses and cartilage adaptation to exercise: A review of animal and human models. *Scandinavian Journal of Medicine & Science in Sports*. https://doi.org/10.1111/sms.13435

McCarthy, C., Mills, P., Pullen, R., Richardson, G., Hawkins, N., Roberts, C., Silman, A., & Oldham, J. (2004). Supplementation of a home-based exercise programme with a class-based programme for people with osteoarthritis of the knees: A randomised controlled trial and health economic analysis. *Health Technology Assessment*, *8*(46). https://doi.org/10.3310/hta8460

McDowell, M. A., Fryar, C. D., Ogden, C. L., & Flegal, K. M. (2008). *Anthropometric Reference Data for Children and Adults: United States, 2003-2006: (623932009-001)* [Data set]. American Psychological Association. https://doi.org/10.1037/e623932009-001

Menard, M., Domalain, M., Decatoire, A., & Lacouture, P. (2016). Influence of saddle setback on pedalling technique effectiveness in cycling. *Sports Biomechanics*, *15*(4), 462–472. https://doi.org/10.1080/14763141.2016.1176244

Menard, M., Domalain, M., Decatoire, A., & Lacouture, P. (2018). Influence of saddle setback on knee joint forces in cycling. *Sports Biomechanics*, 1–13. https://doi.org/10.1080/14763141.2018.1466906

Messier, S. P., Mihalko, S. L., Legault, C., Miller, G. D., Nicklas, B. J., DeVita, P., Beavers, D. P., Hunter, D. J., Lyles, M. F., Eckstein, F., Williamson, J. D., Carr, J. J., Guermazi, A., & Loeser, R. F. (2013). Effects of Intensive Diet and Exercise on Knee Joint Loads, Inflammation, and Clinical Outcomes Among Overweight and Obese Adults With Knee Osteoarthritis: The IDEA Randomized Clinical Trial. *JAMA*, *310*(12), 1263. https://doi.org/10.1001/jama.2013.277669

Metcalfe, A. J., Andersson, M. L., Goodfellow, R., & Thorstensson, C. A. (2012). Is knee osteoarthritis a symmetrical disease? Analysis of a 12 year prospective cohort study. *BMC Musculoskeletal Disorders*, *13*(1). https://doi.org/10.1186/1471-2474-13-153

Meyer, A. J., D'Lima, D. D., Besier, T. F., Lloyd, D. G., Colwell, C. W., & Fregly, B. J. (2013). Are external knee load and EMG measures accurate indicators of internal knee contact forces during gait? *Journal of Orthopaedic Research*, *31*(6), 921–929. https://doi.org/10.1002/jor.22304

Mosher, Timothy J., Smith, H. E., Collins, C., Liu, Y., Hancy, J., Dardzinski, B. J., & Smith, M. B. (2005). Change in Knee Cartilage T2 at MR Imaging after Running: A Feasibility Study1. *Radiology*, *234*(1), 245–249. https://doi.org/10.1148/radiol.2341040041

Mosher, T.J., Liu, Y., & Torok, C. M. (2010). Functional cartilage MRI T2 mapping: Evaluating the effect of age and training on knee cartilage response to running. *Osteoarthritis and Cartilage*, *18*(3), 358–364. https://doi.org/10.1016/j.joca.2009.11.011

Moyer, R. F., Ratneswaran, A., Beier, F., & Birmingham, T. B. (2014). Osteoarthritis Year in Review 2014: Mechanics – basic and clinical studies in osteoarthritis. *Osteoarthritis and Cartilage*, *22*(12), 1989–2002. https://doi.org/10.1016/j.joca.2014.06.034

Multanen, J., Nieminen, M. T., Häkkinen, A., Kujala, U. M., Jämsä, T., Kautiainen, H., Lammentausta, E., Ahola, R., Selänne, H., Ojala, R., Kiviranta, I., & Heinonen, A. (2014). Effects of High-Impact Training on Bone and Articular Cartilage: 12-Month Randomized Controlled Quantitative MRI Study: EFFECTS OF HIGH-IMPACT TRAINING ON BONE AND ARTICULAR CARTILAGE. *Journal of Bone and Mineral Research*, *29*(1), 192–201. https://doi.org/10.1002/jbmr.2015

Musumeci, G. (2013). The role of lubricin in normal and pathological joint tissue: A contemporary review. *OA Anatomy*, *1*(1). https://doi.org/10.13172/2052-7829-1-1-329

Musumeci, Giuseppe, Aiello, F., Szychlinska, M., Di Rosa, M., Castrogiovanni, P., & Mobasheri, A. (2015). Osteoarthritis in the XXIst Century: Risk Factors and Behaviours that Influence Disease Onset and Progression. *International Journal of Molecular Sciences*, *16*(3), 6093–6112. https://doi.org/10.3390/ijms16036093

Myronenko, A., & Xubo Song. (2010). Point Set Registration: Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(12), 2262–2275. https://doi.org/10.1109/TPAMI.2010.46

Neogi, T., Bowes, M. A., Niu, J., De Souza, K. M., Vincent, G. R., Goggins, J., Zhang, Y., & Felson, D. T. (2013). Magnetic Resonance Imaging-Based Three-Dimensional Bone Shape of the Knee Predicts Onset of Knee Osteoarthritis: Data From the Osteoarthritis Initiative: 3-D Bone Shape Predicts Incident Knee OA. *Arthritis & Rheumatism*, *65*(8), 2048–2058. https://doi.org/10.1002/art.37987

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

Nordin, M., & Frankel, V. H. (Eds.). (2012). *Basic biomechanics of the musculoskeletal system* (4th ed). Wolters Kluwer Health/Lippincott Williams & Wilkins.

Norman, B., Pedoia, V., Noworolski, A., Link, T. M., & Majumdar, S. (2019). Applying Densely Connected Convolutional Neural Networks for Staging Osteoarthritis Severity from Plain Radiographs. *Journal of Digital Imaging*, *32*(3), 471–477. https://doi.org/10.1007/s10278-018-0098-3

Palmer, A. J. R., Brown, C. P., McNally, E. G., Price, A. J., Tracey, I., Jezzard, P., Carr, A. J., & Glyn-Jones, S. (2013). Non-invasive imaging of cartilage in early osteoarthritis. *The Bone & Joint Journal*, *95-B*(6), 738–746. https://doi.org/10.1302/0301-620X.95B6.31414

Panfilov, E., Tiulpin, A., Klein, S., Nieminen, M. T., & Saarakkala, S. (2019). Improving Robustness of Deep Learning Based Knee MRI Segmentation: Mixup and Adversarial Domain Adaptation. *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 450–459. https://doi.org/10.1109/ICCVW.2019.00057

Pataky, T. C. (2010). Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *Journal of Biomechanics*, *43*(10), 1976–1982. https://doi.org/10.1016/j.jbiomech.2010.03.008

Pataky, T. C., Caravaggi, P., Savage, R., Parker, D., Goulermas, J. Y., Sellers, W. I., & Crompton, R. H. (2008). New insights into the plantar pressure correlates of walking speed using pedobarographic statistical parametric mapping (pSPM). *Journal of Biomechanics*, *41*(9), 1987–1994. https://doi.org/10.1016/j.jbiomech.2008.03.034

Pataky, T. C., Koseki, M., & Cox, P. G. (2016). Probabilistic biomechanical finite element simulations: Whole-model classical hypothesis testing based on upcrossing geometry. *PeerJ Computer Science*, *2*, e96. https://doi.org/10.7717/peerj-cs.96

Pedoia, V., Haefeli, J., Morioka, K., Teng, H.-L., Nardo, L., Souza, R. B., Ferguson, A. R., & Majumdar, S. (2018). MRI and biomechanics multidimensional data analysis reveals R_2 -$R_{1\rho}$ as an early predictor of cartilage lesion progression in knee osteoarthritis: Multidimensional Data Analysis to Study OA. *Journal of Magnetic Resonance Imaging*, *47*(1), 78–90. https://doi.org/10.1002/jmri.25750

Pedoia, V., Lansdown, D. A., Zaid, M., McCulloch, C. E., Souza, R., Ma, C. B., & Li, X. (2015). Three-dimensional MRI-based statistical shape model and application to a cohort of knees with acute ACL injury. *Osteoarthritis and Cartilage*, *23*(10), 1695–1703. https://doi.org/10.1016/j.joca.2015.05.027

Pedoia, V., Majumdar, S., & Link, T. M. (2016). Segmentation of joint and musculoskeletal tissue in the study of arthritis. *Magnetic Resonance Materials in*

Physics, Biology and Medicine, *29*(2), 207–221. https://doi.org/10.1007/s10334-016-0532-9

Prasoon, A., Igel, C., Loog, M., Lauze, F., Dam, E. B., & Nielsen, M. (2013). Femoral cartilage segmentation in knee MRI scans using two stage voxel classification. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, 5469–5472. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6610787

Priego Quesada, J. I., Pérez-Soriano, P., Lucas-Cuevas, A. G., Salvador Palmer, R., & Cibrián Ortiz de Anda, R. M. (2017). Effect of bike-fit in the perception of comfort, fatigue and pain. *Journal of Sports Sciences*, *35*(14), 1459–1465. https://doi.org/10.1080/02640414.2016.1215496

Pugh, J. W., Radin, E. L., & Rose, R. M. (1974). Quantitative Studies of Human Subchondral Cancellous Bone: ITS RELATIONSHIP TO THE STATE OF ITS OVERLYING CARTILAGE. *The Journal of Bone & Joint Surgery*, *56*(2), 313–321. https://doi.org/10.2106/00004623-197456020-00010

Radin, E. L., & Rose, R. M. (1986). Role of subchondral bone in the initiation and progression of cartilage damage. *Clinical Orthopaedics and Related Research*, *213*, 34–40.

Ramage, L., Nuki, G., & Salter, D. M. (2009). Signalling cascades in mechanotransduction: Cell-matrix interactions and mechanical loading.

Scandinavian Journal of Medicine & Science in Sports, *19*(4), 457–469. https://doi.org/10.1111/j.1600-0838.2009.00912.x

Ramage-Morin, P. L. (2017). *Statistics Canada, Health Reports, Cycling in Canada*. https://www150.statcan.gc.ca/n1/pub/82-003-x/2017004/article/14788/tbl/tbl01-eng.htm

Raya, J. G., Dietrich, O., Horng, A., Weber, J., Reiser, M. F., & Glaser, C. (2009). T_2 measurement in articular cartilage: Impact of the fitting method on accuracy and precision at low SNR. *Magnetic Resonance in Medicine*, NA-NA. https://doi.org/10.1002/mrm.22178

Razmjoo, A., Caliva, F., Lee, J., Liu, F., Joseph, G. B., Link, T. M., Majumdar, S., & Pedoia, V. (2020). T_2 analysis of the entire osteoarthritis initiative dataset. *Journal of Orthopaedic Research*. https://doi.org/10.1002/jor.24811

Roelker, S. A., Caruthers, E. J., Hall, R. K., Pelz, N. C., Chaudhari, A. M. W., & Siston, R. A. (2020). Effects of Optimization Technique on Simulated Muscle Activations and Forces. *Journal of Applied Biomechanics*, *36*(4), 259–278. https://doi.org/10.1123/jab.2018-0332

Safiri, S., Kolahi, A.-A., Smith, E., Hill, C., Bettampadi, D., Mansournia, M. A., Hoy, D., Ashrafi-Asgarabad, A., Sepidarkish, M., Almasi-Hashiani, A., Collins, G., Kaufman, J., Qorbani, M., Moradi-Lakeh, M., Woolf, A. D., Guillemin, F., March, L., & Cross, M. (2020). Global, regional and national burden of osteoarthritis 1990-2017: A systematic analysis of the Global Burden of Disease

Study 2017. *Annals of the Rheumatic Diseases*, annrheumdis-2019-216515. https://doi.org/10.1136/annrheumdis-2019-216515

Salacinski, A. J., Krohn, K., Lewis, S. F., Holland, M. L., Ireland, K., & Marchetti, G. (2012). The Effects of Group Cycling on Gait and Pain-Related Disability in Individuals With Mild-to-Moderate Knee Osteoarthritis: A Randomized Controlled Trial. *Journal of Orthopaedic & Sports Physical Therapy*, *42*(12), 985–995. https://doi.org/10.2519/jospt.2012.3813

Saxby, D. J., Bryant, A. L., Modenese, L., Gerus, P., Killen, B. A., Konrath, J., Fortin, K., Wrigley, T. V., Bennell, K. L., Cicuttini, F. M., Vertullo, C., Feller, J. A., Whitehead, T., Gallie, P., & Lloyd, D. G. (2016). Tibiofemoral Contact Forces in the Anterior Cruciate Ligament–Reconstructed Knee: *Medicine & Science in Sports & Exercise*, *48*(11), 2195–2206. https://doi.org/10.1249/MSS.0000000000001021

Seth, A., Hicks, J. L., Uchida, T. K., Habib, A., Dembia, C. L., Dunne, J. J., Ong, C. F., DeMers, M. S., Rajagopal, A., Millard, M., Hamner, S. R., Arnold, E. M., Yong, J. R., Lakshmikanth, S. K., Sherman, M. A., Ku, J. P., & Delp, S. L. (2018). OpenSim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. *PLOS Computational Biology*, *14*(7), e1006223. https://doi.org/10.1371/journal.pcbi.1006223

Shan, L., Zach, C., Charles, C., & Niethammer, M. (2014). Automatic atlas-based three-label cartilage segmentation from MR knee images. *Medical Image Analysis*, *18*(7), 1233–1246. https://doi.org/10.1016/j.media.2014.05.008

Sharif, B., Kopec, J., Bansback, N., Rahman, M. M., Flanagan, W. M., Wong, H., Fines, P., & Anis, A. (2015). Projecting the direct cost burden of osteoarthritis in Canada using a microsimulation model. *Osteoarthritis and Cartilage*, *23*(10), 1654–1663. https://doi.org/10.1016/j.joca.2015.05.029

Shim, H., Chang, S., Tao, C., Wang, J.-H., Kwoh, C. K., & Bae, K. T. (2009). Knee Cartilage: Efficient and Reproducible Segmentation on High-Spatial-Resolution MR Images with the Semiautomated Graph-Cut Algorithm Method 1. *Radiology*, *251*(2), 548–556.

Silberman, M. R. (2013). Bicycling Injuries: *Current Sports Medicine Reports*, *12*(5), 337–345. https://doi.org/10.1249/JSR.0b013e3182a4bab7

Silder, A., Gleason, K., & Thelen, D. G. (2011). Influence of Bicycle Seat Tube Angle and Hand Position on Lower Extremity Kinematics and Neuromuscular Control: Implications for Triathlon Running Performance. *Journal of Applied Biomechanics*, *27*(4), 297–305. https://doi.org/10.1123/jab.27.4.297

Smith, C. R. (2017). *Simulating the effects of Anterior Cruciate Ligament Injury and Treatment on Cartilage Loading during Walking*. University of Wisconsin-Madison.

Smith, C. R., Brandon, S. C. E., & Thelen, D. G. (2019). Can altered neuromuscular coordination restore soft tissue loading patterns in anterior cruciate ligament and menisci deficient knees during walking? *Journal of Biomechanics*, *82*, 124–133. https://doi.org/10.1016/j.jbiomech.2018.10.008

Smith, C. R., Won Choi, K., Negrut, D., & Thelen, D. G. (2018). Efficient computation of cartilage contact pressures within dynamic simulations of movement. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, *6*(5), 491–498. https://doi.org/10.1080/21681163.2016.1172346

Souza, R. B., Baum, T., Wu, S., Feeley, B. T., Kadel, N., Li, X., Link, T. M., & Majumdar, S. (2012). Effects of Unloading on Knee Articular Cartilage T1rho and T2 Magnetic Resonance Imaging Relaxation Times: A Case Series. *Journal of Orthopaedic & Sports Physical Therapy*, *42*(6), 511–520. https://doi.org/10.2519/jospt.2012.3975

Souza, R. B., Kumar, D., Calixto, N., Singh, J., Schooler, J., Subburaj, K., Li, X., Link, T. M., & Majumdar, S. (2014). Response of knee cartilage T1rho and T2 relaxation times to in vivo mechanical loading in individuals with and without knee osteoarthritis. *Osteoarthritis and Cartilage*. https://doi.org/10.1016/j.joca.2014.04.017

StataCorp LP. (2013). *Stata user's guide*. StataCorp LP.

Steele, K. M., DeMers, M. S., Schwartz, M. H., & Delp, S. L. (2012). Compressive tibiofemoral force during crouch gait. *Gait & Posture*, *35*(4), 556–560. https://doi.org/10.1016/j.gaitpost.2011.11.023

Stürmer, T., Günther, K. P., & Brenner, H. (2000). Obesity, overweight and patterns of osteoarthritis: The Ulm Osteoarthritis Study. *Journal of Clinical Epidemiology*, *53*(3), 307–313.

Subburaj, K., Kumar, D., Souza, R. B., Alizai, H., Li, X., Link, T. M., & Majumdar, S. (2012). The Acute Effect of Running on Knee Articular Cartilage and Meniscus Magnetic Resonance Relaxation Times in Young Healthy Adults. *The American Journal of Sports Medicine*, *40*(9), 2134–2141. https://doi.org/10.1177/0363546512449816

Tamborindeguy, A. C., & Bini, R. R. (2011). Does saddle height affect patellofemoral and tibiofemoral forces during bicycling for rehabilitation? *Journal of Bodywork and Movement Therapies*, *15*(2), 186–191. https://doi.org/10.1016/j.jbmt.2009.07.009

Tamez-Pena, J. G., Farber, J., Gonzalez, P. C., Schreyer, E., Schneider, E., & Totterman, S. (2012). Unsupervised Segmentation and Quantification of Anatomical Knee Features: Data From the Osteoarthritis Initiative. *IEEE Transactions on Biomedical Engineering*, *59*(4), 1177–1186. https://doi.org/10.1109/TBME.2012.2186612

Tang, X., Wang, S., Zhan, S., Niu, J., Tao, K., Zhang, Y., & Lin, J. (2016). The Prevalence of Symptomatic Knee Osteoarthritis in China: Results From the China Health and Retirement Longitudinal Study: SYMPTOMATIC KNEE OSTEOARTHRITIS IN CHINA. *Arthritis & Rheumatology*, *68*(3), 648–653. https://doi.org/10.1002/art.39465

Thelen, D. G., & Anderson, F. C. (2006). Using computed muscle control to generate forward dynamic simulations of human walking from experimental data. *Journal of Biomechanics*, *39*(6), 1107–1115. https://doi.org/10.1016/j.jbiomech.2005.02.010

Thelen, D. G., Anderson, F. C., & Delp, S. L. (2003). Generating dynamic simulations of movement using computed muscle control. *Journal of Biomechanics*, *36*(3), 321–328. https://doi.org/10.1016/S0021-9290(02)00432-3

Thomas, K. A., Kidziński, Ł., Halilaj, E., Fleming, S. L., Venkataraman, G. R., Oei, E. H. G., Gold, G. E., & Delp, S. L. (2020). Automated Classification of Radiographic Knee Osteoarthritis Severity Using Deep Neural Networks. *Radiology: Artificial Intelligence*, *2*(2), e190065. https://doi.org/10.1148/ryai.2020190065

Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P., & Saarakkala, S. (2018). Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-20132-7

Tolpadi, A. A., Lee, J. J., Pedoia, V., & Majumdar, S. (2020). Deep Learning Predicts Total Knee Replacement from Magnetic Resonance Images. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-63395-9

Trepczynski, A., Kutzner, I., Bergmann, G., Taylor, W. R., & Heller, M. O. (2014). Modulation of the Relationship Between External Knee Adduction Moments and Medial Joint Contact Forces Across Subjects and Activities: EAM as a Measure of In Vivo Contact Forces. *Arthritis & Rheumatology*, *66*(5), 1218–1227. https://doi.org/10.1002/art.38374

Urquhart, D. M., Tobing, J. F. L., Hanna, F. S., Berry, P., Wluka, A. E., Ding, C., & Cicuttini, F. M. (2011). What Is the Effect of Physical Activity on the Knee Joint? A Systematic Review: *Medicine & Science in Sports & Exercise*, *43*(3), 432–442. https://doi.org/10.1249/MSS.0b013e3181ef5bf8

Van Ginckel, A., Baelde, N., Almqvist, K. F., Roosen, P., McNair, P., & Witvrouw, E. (2010). Functional adaptation of knee cartilage in asymptomatic female novice runners compared to sedentary controls. A longitudinal analysis using delayed Gadolinium Enhanced Magnetic Resonance Imaging of Cartilage (dGEMRIC). *Osteoarthritis and Cartilage*, *18*(12), 1564–1569. https://doi.org/10.1016/j.joca.2010.10.007

Van Rossom, S., Smith, C. R., Zevenbergen, L., Thelen, D. G., Vanwanseele, B., Van Assche, D., & Jonkers, I. (2017). Knee Cartilage Thickness, T1ρ and T2

Relaxation Time Are Related to Articular Cartilage Loading in Healthy Adults. *PLOS ONE*, *12*(1), e0170002. https://doi.org/10.1371/journal.pone.0170002

Walter, J. P., D'Lima, D. D., Colwell, C. W., & Fregly, B. J. (2010). Decreased knee adduction moment does not guarantee decreased medial contact force during gait. *Journal of Orthopaedic Research*, *28*(10), 1348–1354. https://doi.org/10.1002/jor.21142

Wang, Q., Wu, D., Lu, L., Liu, M., Boyer, K. L., & Zhou, S. K. (2014). *Semantic Context Forests for Learning-Based Knee Cartilage Segmentation in 3D MR Images*. 10.

Wang, Y., Nguyen, U. D. T., Lane, N. E., Lu, N., Wei, J., Lei, G., Zeng, C., & Zhang, Y. (2020). Knee osteoarthritis, potential mediators, and risk of all-cause mortality: Data from the Osteoarthritis Initiative. *Arthritis Care & Research*. https://doi.org/10.1002/acr.24151

Wáng, Y.-X. J., Zhang, Q., Li, X., Chen, W., Ahuja, A., & Yuan, J. (2015). T1ρ magnetic resonance: Basic physics principles and applications in knee and intervertebral disc imaging. *Quantitative Imaging in Medicine and Surgery*, *5*(6), 858–885. https://doi.org/10.3978/j.issn.2223-4292.2015.12.06

Wang, Z., Ai, S., Tian, F., Liow, M. H. L., Wang, S., Zhao, J., & Tsai, T.-Y. (2020). Higher Body Mass Index Is Associated With Biochemical Changes in Knee Articular Cartilage After Marathon Running: A Quantitative T2-Relaxation MRI Study. *Orthopaedic Journal of Sports Medicine*, *8*(8), 232596712094387. https://doi.org/10.1177/2325967120943874

Warburton, D. E. R. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Association Journal*, *174*(6), 801–809. https://doi.org/10.1503/cmaj.051351

Williams, P. T. (2013). Effects of Running and Walking on Osteoarthritis and Hip Replacement Risk: *Medicine & Science in Sports & Exercise*, *45*(7), 1292–1297. https://doi.org/10.1249/MSS.0b013e3182885f26

Xia, Y. (2000). Magic-Angle Effect in Magnetic Resonance Imaging of Articular Cartilage: A Review. *Investigative Radiology*, *35*(10), 602–621. https://doi.org/10.1097/00004424-200010000-00007

Zajac, F. E. (1989). Muscle and tendon: Properties, models, scaling, and application to biomechanics and motor control. *Critical Reviews in Biomedical Engineering*, *17*(4), 359–411.