NORTHWESTERN UNIVERSITY

Active Model-Based Inference for Muscle Strength Diagnostics

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Mechanical Engineering

By

Rebecca Abbott

EVANSTON, ILLINOIS

March 2021

ABSTRACT

Active Model-Based Inference for Muscle Strength Diagnostics

Rebecca Abbott

Muscle strength assessment is a standard part of any clinical evaluation. Due to the kinematic and muscular redundancy in the human musculoskeletal system, muscle strength can not be measured directly in vivo. Clinicians utilize specific postures and forces to bias the muscles of interest, then infer the muscle strength from an indirect measurement. This is particularly challenging in the cervical spine, where more than 25 muscle pairs with multiple attachments and multi-planar actions act across 7 intervertebral joints. As a result, there is currently no method for measuring individual neck muscle strengths.

Model-based parameter estimation techniques infer the values of the unobservable quantities by fitting measured data to a model. There is an opportunity to apply and expand these methods to aid clinicians and researchers in evaluation of individual muscle strengths in complex musculoskeletal systems. In a clinical setting, limiting the number of measurements is a priority due to time constraints and patient fatigue and pain. This thesis presents an active learning approach to parameter estimation that utilizes

information theoretic measures to query the clinician for the next measurement that will maximize information gain relative to the unknown parameters, thus avoiding unnecessary measurements.

This thesis is motivated by clinical questions that arose from imaging studies in individuals with persistent whiplash-associated disorders (WAD). Is there a link between compositional changes in the deep cervical extensor muscles and motor dysfunction in WAD? We begin with two studies confirming and expanding upon the compositional muscle changes and exploring the potential biomechanical consequences of those changes using a musculoskeletal model of the neck. The results illustrate the role of the deep extensors in multi-directional neck strength and provide simulated evidence of altered motor control patterns in WAD.

The lack of an existing method for non-invasive measurement of individual muscle strengths motivates the main contribution of this thesis, a novel framework for active model-based Bayesian inference for muscle strength diagnostics. We demonstrate the utility of the approach for estimating individual neck muscle strengths from multi-directional isometric neck strength measurements from 5 healthy participants. A framework for a clinician-in-the-loop muscle strength estimator is presented, where in addition to accurately estimating the strength of the deep cervical extensor muscles, the algorithm predicts the benefit of taking additional measurements and selects the most informative next measurement to take. The framework is a step towards producing a clinically translatable test for individual neck muscle strengths.

Acknowledgements

Many people have supported my academically and personally through throughout this DPT/PhD journey. Starting with my two primary advisors, Jim and Todd. I would like to thank Todd for guidance and his flexibility and patience as I moved through the phases of the dual degree program, welcoming me back into the lab between and after clinical phases. I have a deep appreciation for the consistent support Jim has provided, always with an open door and consideration of me as a full person, not just a graduate student. And for teaching the importance of collaboration across disciplines and across continents. I would also like to thank Ana Maria, who was so generous with her time and attention, providing consistent positive support particularly towards the end of this journey when I needed it most.

Of course I must acknowledge Jules Dewald, without whom this dual degree DPT/PhD program would not exist. His unwavering support of the DPT/PhD students and genuine excitement about our research and professional growth is a crucial component of our success. And the rest of the NUPT family, instructors, colleagues, and staff who supported me through the DPT program and beyond, continuing to provide mentorship and guidance with clinical and life questions.

To my past and current lab mates, thank you for your friendship and collaboration. To Mark and Andrew of our small but mighty NIRL lab, I couldn't dream of a better group to work alongside and will always think back fondly on our productive NUBREW

meetings. To my LIMS, NxR, CRB lab mates, thank you for welcoming me back in to the group so readily each time I reappeared from my clinical phases and for your patience with talking through my research questions and providing valuable insight.

My fellow DPT/PhD students provided support throughout the years. A special shout out to my cohort Emma and Steve, who accompanied Gromit and me on several holiday road trips out east. Emma has been such a consistent and stable friend, between willingness to help with experiments, walking Gromit, or lending an ear as a friend. To Nayo, Lindsay, and Rachel, my office mates and friends who were always there to support me as a friend and colleague.

I would also like to thank my Chicago family - Ashley and Tucker, Margaret Ann and Dan, Elyse, Meg, and Maya who have been there for me through this journey. My Boston/Tufts family - Andrea and family, Betsy and family, Tofias family, Daniela, Jeff and Melinda, Laura, Katy, Pat, and the others who were my rocks for many years and continue to support and encourage me.

Thank you to my quarantine buddy, Gromit, for making sure I got up and out of the house every morning. Without you this last year would have been quite lonely, and I would have gotten much more sleep.

I also thank my funding sources, including the NIH NIBIB T32-EB009406 training grant that made the DPT/PhD Eng. program possible. As well as the NIH R01 5R01HD079076-03, and two grants through National Science Foundation: NSF 1637764 and NSF 1837515.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Kinematic and muscle redundancy in the human musculoskeletal system makes it difficult to isolate individual muscles for strength testing. As part of an evaluation, clinicians utilize their knowledge of anatomy and kinesiology to problem solve and rule in or out involvement of muscles through the application of various tests. Manual muscle tests use specific joint angles and applied forces to bias the muscle of interest and have been developed for some muscles. Due to the complex musculature and partial observability of the cervical spine, specific tests for individual neck muscle strength have not been developed and it is particularly difficult to develop an intuition for ways to bias an individual neck muscle.

There is an opportunity for musculoskeletal model based parameter estimation algorithms to assist clinicians and researchers in muscle strength diagnostics in these complex musculoskeletal systems. Musculoskeletal models can capture complicated biomechanical relationships and allow for computation of internal loads that can not be measured in vivo, like muscle forces. When combined in a Bayesian framework for parameter estimation, the domain knowledge of the clinician and the mathematical representation in the model sum to provide a more informative estimate of individual muscle strengths.

A parameter estimation algorithm attempts to infer the value of unknown parameters, such as muscle strength, from a set of measured data, such as end effector forces. In a clinical setting, limiting the number of measurements is a priority because of time

limitations and patient fatigue and pain. This thesis uses an active learning approach to parameter estimation to addresses this challenge, where the algorithm queries the clinician for the next measurement that provides the greatest expected information gain. *The focus of this thesis is towards developing a clinically translatable framework for clinician-in-the-loop individual muscle strength estimation utilizing non-invasive measurements in a complex musculoskeletal system.*

1.1. Main Contributions

This thesis is motivated by a clinical question requiring the ability to measure the strength of individual neck muscles. The lack of an existing framework for in vivo measurement of individual muscles strengths presented both a barrier to our progress and an opportunity for a wider contribution to the field of musculoskeletal rehabilitation. We employ an interdisciplinary approach, integrating clinical expertise, quantitative magnetic resonance imaging, musculoskeletal modeling, Bayesian inference, optimal experimental design, and information theory. The result is an algorithmic framework utilizing model-based Bayesian inference for individual muscle strength estimation with automated test sequence selection to maximize expected information gain and minimize the number of measurements.

1.1.1. Deep cervical extensor muscle fat infiltration in WAD and its potential biomechanical effects

Despite the personal and societal cost of the condition [21, 14], clinicians and researchers have been unable to agree upon an underlying etiology or develop consistently successful

treatments for moderate-severe chronic whiplash-associated disorders (WAD). Quantitative Magnetic Resonance Imaging (MRI) studies have identified increased muscle fat infiltration (MFI) throughout the neck muscles of individuals with persistent WAD, with a particularly high concentration in the deep cervical extensors [22, 26]. If the logical assumption is made that increased MFI indicates muscle atrophy and weakness, an obvious next question is how weakness of this specific muscle group would affect neck and head biomechanics. The complex, redundant kinematic and muscular architecture of the neck makes answering this question difficult. Although the general function of the cervical multifidus and semispinalis cervicis have been theorized based on anatomic and EMG studies [4, 89], the role of the deep extensors and biomechanical effects of muscle weakness in these muscles is unknown. This knowledge could inform diagnostic and treatment strategies in WAD.

This thesis begins to address these questions in Chapter 3 by confirming and expanding upon the findings of increased MFI in the deep extensors in WAD when compared to recovered WAD and healthy controls. We analyze the magnitude and distribution of MFI throughout the deep cervical extensor muscles, both across cervical levels (caudal-cephalad) and within a cervical level (medial-lateral). The finding of a widespread increase in MFI throughout the deep extensors further motivates and informs the design of the study in the next chapter. Chapter 4 utilizes a musculoskeletal model of the neck to determine the role of the deep cervical extensors in producing multi-directional neck strength and the biomechanical effects of weakness of this muscle group. Musculoskeletal models provide a valuable tool for exploring the effects of a changes in a parameter like muscle strength that would be difficult or impossible to study in vivo.

The contributions of chapters 3 and 4 are:

- We map the magnitude and spatial distribution of MFI within the deep cervical extensor muscles in chronic WAD, recovered, and healthy controls.
- We show that the spatial distribution of MFI is similar between groups, but there is an increased magnitude of MFI throughout the deep extensor muscles in the WAD group.
- We illustrate, through simulation with a musculoskeletal model of the neck, how deep extensor weakness affects neck strength and muscle activation patterns in 25 test directions mapped to locations on the skull where resistance can be applied by a clinician
- We provide simulated evidence of altered muscle activation patterns that would occur with weakness of the deep extensors, characterized by increased activation of the longus capitis and superficial extensors, and decreased activation of the suboccipitals.
- We reveal the effect of kinematic and muscle model complexity on simulated neck strength and deep extensor muscle activation.

The work from Chapter 3 is published in [1].

1.1.2. Parameter estimation in a highly redundant musculoskeletal system

The problem of measuring individual muscle strength is difficult for a number of reasons. The core issue is that muscle forces can not be measured directly. The human neck is a complex musculoskeletal structure with 7 intervertebral joints and over 25 pairs of named muscles, most having multiple attachments and crossing multiple joints. Wrench (force

and torque) measurements are taken from the end effector of the neck, the head. The kinematic and muscle redundancy sets up an under-constrained and ill-conditioned inverse problem. An additional complication is that we are interested in the muscle strength, the maximum force that a muscle can produce, not just the contribution of the muscle for a particular end effector measurement.

Muscle parameters used in musculoskeletal models, including maximum isometric force (strength), are typically derived from physiological cross sectional area (PCSA) measurements from anatomic cadaver studies [108, 17]. This method produces a generic set of muscle strengths that can be scaled uniformly based on subject specific measurements, but does not address variations in individual muscle strength [78, 110]. Efforts to determine subject-specific model parameters have been made in the hand and wrist [41, 59, 49, 15]. These studies utilize least squares optimization methods to minimize the difference between model predicted and measured data sets. A bounded, non-linear least squares approach similar to these studies could be a viable approach to the strength estimation problem in the neck, but the data available is often noisy and scarce, setting up an ill-posed optimization problem that is known to give poor results for these standard methods [5, 112]. Statistical Bayesian approaches are well-suited for dealing with non-linear models and naturally include uncertainties such as measurement and modeling errors.

In Chapter 5, we develop a framework for parameter estimation in a highly redundant musculoskeletal system. The novel method developed to fulfill this goal involves (1) collecting neck strength in multiple directions and surface EMG profiles of a subset of neck muscles in vivo, and (2) utilizing a simulation-based Bayesian estimator based on Approximate Bayesian Computation (ABC) with rejection and importance sampling to

infer subject specific individual muscle strength from the collected data. This probabilistic framework provides uncertainties and correlations between parameters, rather than just point estimates, which is beneficial in a system with the potential to have many solutions [19].

The contributions of chapter 5 are as follows:

- We present a novel Bayesian framework for model-based parameter estimation in a highly redundant, partially observable musculoskeletal system.
- We demonstrate the capabilities of the algorithm by applying it to models of varying complexity and both simulated and experimental data sets, including a 24 muscle model and 25 multi-directional measurements.
- We show the feasibility of the approach through experimental data collection of multi-directional neck strength and a subset of neck muscle EMG for $n = 5$ healthy participants.
- We show that the novel framework can identify weakness of the deep cervical extensor muscles when provided a rich enough set of measurement data.

1.1.3. Automated Test Sequence Selection for Efficient Musculoskeletal Parameter Estimation

The full set of 25 test directions included in the experiment in Chapter 5 would be too fatiguing for an individual with chronic WAD and too time consuming for clinicians. The next step in making this framework practical for a clinical application is to reduce the number of measurements required to obtain an informative estimate of individual muscle strengths. Information theoretic measures have been used extensively in the fields of

optimal experimental design [95], search [74], and source localization [111] to make measurement selection decisions based on expected information gain. This type of approach has not been attempted for musculoskeletal evaluation.

In Chapter 6, we present a novel framework for musculoskeletal evaluation that sequentially selects the next measurement that maximizes expected information gain. This extension of the model-based Bayesian estimator from Chapter 5 keeps the clinician in the loop, proposes the next measurement direction, and updates the parameter estimate based on the result of that measurement at each iteration. We demonstrate how this algorithm can provide information to the clinician enabling them to make an informed decision on when to stop taking measurements.

The contributions of this chapter are:

- We develop a novel framework for clinician-in-the-loop automated test sequence selection in musculoskeletal parameter estimation that selects the next measurement to take based on expected information gain
- We demonstrate the utility of the approach by showing with a simulated data set and 5 experimental data sets that the algorithm would provide information allowing the clinician weigh the benefit of taking additional measurements.
- We present an example with an experimental data set where the algorithm enables the number of measurements to be reduced from 7 to 4, without a significant change in the strength estimates.

1.2. Dissertation Outline

Chapter 2 presents background information from the multi-disciplinary array of fields drawn upon in this work, including an introduction to chronic whiplash-associated disorders (WAD), neck anatomy and biomechanics, musculoskeletal modeling, parameter estimation, Bayesian inference, and information theory. The organization of the remaining chapters follows the progression from the clinical motivating question, exploration through musculoskeletal modeling, and finally the development of algorithmic solutions to aid in musculoskeletal based parameter estimation.

We begin in **Chapter 3** by investigating the deep cervical extensor muscle structural changes in order to confirm and expand on the evidence of muscle atrophy that appears to be unique to individuals that progress to chronic WAD. Results confirm an overall increase in muscle fat infiltration (MFI) in these muscles in participants with chronic WAD compared to the recovered and healthy control groups. The results further motivate investigations of multi-directional neck strength and individual muscle strength changes that could result from select atrophy of this group of muscles.

Chapter 4 explores the role of the deep cervical extensor muscles in multi-directional neck strength through musculoskeletal simulation. We illustrate the effects of deep extensor muscle weakness on neck strength and muscle activation patterns. We also evaluate the effect of kinematic and muscle model complexity on simulated neck strength and the role of the deep extensor muscles.

The final chapters develop a framework for estimating individual muscle strengths from isometric neck strength measurements. Due to the high degree of kinematic and muscle

redundancy in the neck, indirect inference is required to estimate the muscle strengths that produce the maximum effort wrench measured at the head.

Chapter 5 develops a parameter estimation algorithm with an Approximate Bayesian Computation (ABC) with rejection and importance sampling approach to infer the strength of individual neck muscles based on a pre-collected data set of multi-directional neck strength and electromyography (EMG) measurements. Feasibility of the approach was demonstrated with experimental measurements from 5 healthy participants with data sets measuring 6-D isometric neck strength in 25 directions with surface EMG recordings of 10 superficial neck muscles. The algorithm was able to identify deep extensor weakness when utilizing the full spatial neck model and 25 test directions with simulated data sets.

Chapter 6 expands on the Bayesian estimator from the previous chapter with a framework for clinician-in-the-loop automated test sequence selection for musculoskeletal parameter estimation. The algorithm selects the next measurement direction that maximizes expected information gain. We demonstrate its utility with a 7 muscle planar neck model and a simulated data set and experimental data sets collected in the previous chapter. We show that the automated test sequence and information provided by the algorithm would allow the clinician to reduce the number of measurements without significant change in individual neck strength estimate results. Reducing the number of measurements is valuable because individuals with persistent WAD would not tolerate a high number of strength measurements and the time available to clinicians for evaluation is limited.

CHAPTER 2

Background

The framework for musculoskeletal parameter estimation developed in this thesis draws from a range of disciplines and research topics. This chapter provides the background and relevant work from these different fields. This includes an introduction to persistent whiplash-associated disorders (WAD) and the associated muscle changes, neck anatomy and biomechanics, musculoskeletal modeling, parameter estimation, Bayesian inference, and information theory.

2.1. Whiplash-Associated Disorders

The development of persistent whiplash-associated disorders (WAD) occurs in as many as 50% of individuals who sustain a whiplash injury in a motor vehicle collision (MVC), contributing to a high economic and social burden [14, 21, 54]. Persistent WAD is associated with a myriad of symptoms, including but not limited to, neck pain and stiffness, headache, weakness, dizziness, and psychological distress [99, 14, 27]. Despite the societal and personal impacts of the condition, the cause of the high rate of transition to chronic pain and disability is unknown. There are generally no relevant pathological findings on conventional imaging or other commonly used clinical tests to explain the severe and persistent symptoms [13, 60, 88].

Quantitative Magnetic Resonance Imaging (MRI) studies have identified increased muscle fat infiltration (MFI) throughout the neck muscles of individuals with persistent

WAD, with a particularly high concentration in the deep cervical extensors [22, 26]. MFI is associated with conditions such as denervation, disuse, and other musculoskeletal injury pathologies like rotator cuff tears [36, 40, 55]. The cause of the increased MFI in the deep cervical extensor muscles and the role that it plays in the maintenance of chronic pain and disability in those with WAD remains unknown. It is reasonable to suspect that a muscle with an increased percentage of fat would have less contractile tissue, indicating muscle atrophy resulting in weakness. Previous research has identified changes in cervical motor system function in WAD, including altered electromyographic (EMG) activity [79, 51, 31], altered motor control [117, 116, 10, 43], and global neck weakness [82, 86]. However, a link between MFI and motor dysfunction in WAD has yet to be established.

2.2. Neck Anatomy and Biomechanics

The human neck is a complex musculoskeletal system that is responsible for supporting the weight of the head, flexible positioning of the sensory organs (eyes, ears, vestibular), and protecting neural and vascular structures that connect the head and body.

2.2.1. Skeletal Anatomy and Biomechanics

Between the skull and the first thoracic vertebra, there are 7 cervical vertebrae. The third through the sixth vertebrae (C3 - C6) are considered "typical" cervical vertabrae. The upper two vertebrae, the atlas (C1) and axis (C2), are "atypical", with geometry that allows for a stable interface for the skull and a large amount of axial rotation. The intervertebral junctions of the typical vertebrae include articulations at the intervertebral

disc, a pair of facet (apophyseal) joints, and uncinate processes. This complex articular interface functions to support compressive loads and guide kinematics along with ligamentous and capsular structures. Although the true kinematics are more complex, intervertebral kinematics are typically described relative to 3 rotational degrees of freedom (DOF) (flexion-extension, axial rotation, and lateral bending) at each intervertebral junction. With the cervical spine's 7 intervertebral joints plus the atlanto-occipital junction (C0-C1), that sums to a total of 24 kinematic degrees of freedom for the head-neck system.

Unless utilizing radiographic imaging, neck kinematics are appreciated by observing movement or posture of the head relative to the thorax. Clinically, neck range of motion (ROM) is typically measured as the change in angle of the head about one of the 3 axes relative to its neutral starting posture [44]. In a lab setting, visual markers or inertial measurement units (IMUs) on the head and thorax (sternum, acromion, and/or spinous process of thoracic vertebra) are often used to capture neck kinematics. With all of these methods, the position or movement of the end effector (the head) is used to represent the composite joint angles of the cervical spine. Considering the head to have 6 DOFs (3 rotational and 3 translational), the mapping from head position and orientation (6 DOF) to intervertebral joint angles (24 DOF) is under-constrained, meaning there are infinitely many neck postures that could produce any given head posture. This concept of kinematic redundancy will reappear throughout this thesis and presents a challenge for musculoskeletal modeling and parameter estimation.

2.2.2. Muscle Anatomy and Biomechanics

The musculature of the human neck is complex and comprised of layers of muscles, most with multiple attachments and multi-planar actions over many joints. Detailed descriptions of neck muscle anatomy can be found in the anatomic study by Kamibayashi & Richmond [56], which is the original source for many of the muscle parameters used in the neck models in this thesis. The muscles of the neck can be divided into a few major groups based on their attachments.

The most superficial muscles, connecting the skull to the shoulder girdle, are the sternocleidomastoid (SCM) and upper trapezius. The SCM is the largest lower cervical flexor (C2-C7), but produces a small extension torque on the upper cervical spine (C0-C2). The next layer of muscles connect the skull to the cervical spine. This intermediate group includes the splenius capitis, semispinalis capitis, and longissimus capitis as extensors, and the longus capitis as a flexor. Although grouped together here, the splenius, semispinalis, and longissimus capitis muscles are oriented in very different directions, leading to different actions for rotation. Another intermediate group of muscles connects the cervical spine to the thorax. This includes the scalenes and levator scapulae. Like the trapezius, the levator scapulae attach to the scapulae, so they are not independent of shoulder movement. The hyoid muscles are a group of anterior neck muscles that connect the mastoid process and mandible to the sternum through the hyoid bone. These muscles are generally considered more important for swallowing and mastication, but they have the potential to produce torques about the cervical spine.

The remaining muscles act locally in the spine, attaching vertebra to vertebra. Deep to their capitis counterparts are the splenius cervicis, semispinalis cervicis, and longissimus

cervicis. The longus colli spans the anterior aspect of the spinal column, acting to flex the cervical spine. These muscles are made up of many distinct fascicles that each span 3-5 vertebrae. Finally, the deepest layer of muscles span only 1-3 vertebrae. The cervical multifidus muscles attach a spinous process of a superior vertebra to a facet of a vertebra 2-5 segments inferior [4]. The interspinales and intertransversarii muscles connect adjacent spinous and transverse processes. The suboccipital muscles include 4 pairs of muscles on the dorsal side of the neck that act locally in the upper cervical spine to produce independent motion of the head on the neck.

The deep cervical extensors mentioned in Section 2.1 on WAD refers to the cervical multifidus and semispinalis cervicis muscles, which originate from a spinous process of superior vertebra and descend laterally to insert onto a transverse process of an inferior vertebra. Throughout this thesis, these two muscles will often be referred to collectively as the deep cervical extensor muscles.

A musculoskeletal system is considered to have muscle redundancy when it has more muscles than kinematic DOFs, leading to an infinite set of muscle activation patterns that can produce any given submaximal task [9]. It is difficult to say precisely how many muscles are part of the head-neck system. While only 16 muscles were named in the paragraphs above, many of the muscles are made up of multiple parts. Of course, each neck muscle comes as a left-right pair. The scalenes are made up of 3 distinct muscles, the anterior, middle, and posterior scalenes. The sternocleidomastoid has two heads but a single cephalad insertion. Others, like the cervical multifidus and semispinalis cervicis, are made up of many fascicular subgroups with independent attachments between vertebrae. In addition to the number of muscles, other factors that influence the level of muscle

redundancy include what level of independent neural control humans have within and between neck muscles, the complexity of the task, and the specific distribution of muscles across DOFs [62, 71]. Like the concept of kinematic redundancy, muscle redundancy will be a recurrent topic in throughout this thesis.

2.2.3. Assessment of Neck Strength

Clinical assessment of neck strength is performed using manual muscle testing (grading of 0 to 5) or dynamometry by applying a force to the head and instructing the patient to hold an isometric posture. These tests do not attempt to isolate individual muscle groups, instead referring to the general actions such as flexion, extension, side-bending, or rotation. Daniels and Worthingham's Muscle Testing book [44] offers distinct tests for capital and cervical extension and flexion strength by adjusting the direction and location of application of resistance by the clinician. They also present a manual muscle test for SCM strength, which involves the patient in supine lifting their head against resistance in a rotated neck posture. Kendall's Muscles Testing and Function book [57] includes manual muscle tests for more than 40 individual upper extremity muscles, but only 3 tests for neck strength, none of which isolate an individual neck muscle group. Other neck muscle assessments include clinical tests for neck extensor endurance [63], neck flexor endurance [42], and deep cervical flexor neuromotor control [52]. Controlled lab studies have utilized six degree of freedom (DOF) load cells and custom helmet fixtures or adjustable pads to interface with the head to measure neck strength [33, 109, 82, 86]. These studies report forces measured at the point of contact with the head or resolve forces into torques about an axis in the lower cervical spine [109, 34]. This choice of axis is somewhat arbitrary and

disregards the ability of the neck to produce a combination of forces and moments in 6 DOF. Although some of these studies have included submaximal bi-planar test directions [109, 34], none have attempted to isolate the strength of individual neck muscles.

2.3. Musculoskeletal Modeling of the Neck

Musculoskeletal models enable us to explore quantities that would otherwise be difficult or impossible to measure in vivo. The following sections derive the computational neck model used throughout this thesis.

2.3.1. Equations of Motion

The generalized equations of motion for a musculoskeletal kinematic chain with m muscles, n joint space DOFs, and d task space DOFs, can be expressed in the torque space as

$$M(q)\ \ddot{q} = R(q)\ [\ F_a(q,\dot{q})\ \alpha + F_p(q)\ ] + V(q,\dot{q}) + G(q) + J^T(q)\ \mathbf{w} \tag{2.1}$$

where:

- $q, \dot{q},$ and $\ddot{q} \in \mathbb{R}^{n\times 1}$: generalized coordinates of the model
- $M(q) \in \mathbb{R}^{n\times n}$: mass-inertia matrix
- $R(q) \in \mathbb{R}^{n\times m}$: moment arm matrix
- $F_a(q,\dot{q}) \in \mathbb{R}^{m\times m}$: diagonal matrix of maximum active muscle force. Maximum isometric muscle force (strength) is defined as the maximum force the muscle can produce at optimal fiber length.
- $\alpha \in \mathbb{R}^{m\times 1}$: muscle activations, with elements ranging from 0 to 1.

- $F_p(q) \in \mathbb{R}^{m\times 1}$: passive muscle forces
- $V(q, \dot{q}) \in \mathbb{R}^{n\times 1}$: joint torques due to Coriolis forces
- $G(q) \in \mathbb{R}^{n\times 1}$: joint torques due to gravity
- $J(q) \in \mathbb{R}^{d\times n}$: Jacobian matrix mapping joint velocities to end effector velocities. The transpose of the Jacobian, J^T, maps end effector wrenches to joint torques.
- $\mathbf{w} \in \mathbb{R}^{d\times 1}$: end effector wrench. In the full spatial context, where $d = 6$, the wrench has components $\mathbf{w} = [M_x, M_y, M_z, F_x, F_y, F_z]^T$. In a planar context, where $d = 3$, the wrench has components $W = [M_z, F_x, F_y]^T$

The generalized equations of motion can be simplified for an isometric task, where $\dot{q} = 0$ and $\ddot{q} = 0$. In this case, $V(q, \dot{q}) = 0$ and $M(q)\ \ddot{q} = 0$. If the neck is in a neutral posture, the passive muscle forces are minimal, so $F_p(q) = 0$. If the head is secured in a measurement device, like the one described in the experiments in later chapters, the effects of gravity may also be ignored, so $G(q) = 0$. The simplified isometric equations of motion are then:

$$R\ F\ \alpha - J^T\ \mathbf{w} = 0 \tag{2.2}$$

The moment arm matrix, R, and Jacobian matrix, J, are still dependent on joint angle, q, and must be updated if posture changes. In the context of this thesis, neck strength is always measured in an upright neutral posture, so R and J remain constant throughout. The maximum active muscle force, $F_a(q, \dot{q})$, has been replaced by F, the maximum isometric muscle (muscle strength) at optimal fiber length.

2.3.2. Feasible Wrench Set

One way to think of multi-directional neck strength is as the feasible wrench set (FWS) of the neck. A feasible wrench set is the set of all possible wrenches the nervous system (or model) can produce at the end effector in a particular posture [105]. The boundaries of this set are the maximum wrench magnitudes in every possible direction. Computational geometry and linear algebra provide a structured analytical method for computing the feasible wrench set, by mapping the vertices of the feasible activation space (a unit positive m-dimensional cube) to the wrench space [107, 62]. This method is limited, however, by the dimensionality of the muscle space and the availability of a direct mapping between activation and wrench space. This direct mapping, $\mathbf{w} = [\ J^{-T}\ R\ F\ ]\ \alpha = A\ \alpha$, relies on our ability to invert the Jacobian matrix.

To be invertible, the Jacobian would need to be square and non-singular. This would mean that $n = d$, or the number of joint DOFs would need to equal the number of task space DOFs. Another option is use the pseudoinverse. However, a pseudoinverse is not unique and introduces potentially false constraints. To avoid these limitations, a modified feasible wrench set can be computed by finding the maximum wrench magnitude that can be produced in a set of pre-defined wrench unit directions. This set of equations forms the basis for the computational neck model introduced in Chapter 4 and used throughout this thesis.

The mathematical model for computing the maximum wrench magnitude, w, and associated muscle activation pattern, α, given a set of individual muscle strengths f and unit wrench direction u is:

$$
\begin{aligned}
M(f,u) = \quad & \\
& \max_{\alpha,\, w} \quad w \\
& \text{s.t.} \quad \begin{bmatrix} R\ \mathrm{diag}(f) & -J^T u \end{bmatrix} \begin{bmatrix} \alpha \\ w \end{bmatrix} = 0, \\
& \qquad 0 \le \alpha \le 1, \\
& \qquad 0 \le w
\end{aligned} \tag{2.3}
$$

2.4. Parameter Estimation

The objective of a parameter estimation problem is to determine the values of some unobservable parameters, θ, based on some observable quantities, x. The mathematical model relates these quantities through a forward problem, $M(\theta) = x$. Parameter estimation methods seek to solve the inverse problem. A wide array of computational techniques have been developed and the choice of technique depends on the structure and well-posedness of the inverse problem.

The mathematical model (Equation 2.3) relating individual muscle strengths, f, to maximum wrench magnitudes, w, is a linear programming optimization. However, the existence of the bounded and strictly positive muscle activation hyperparameters, α, and low rank of both the Jacobian, J, and moment arm matrix, R, make the problem nonlinear, non-unique, and ill-conditioned. Regularization techniques that are commonly used to produce usable solutions to otherwise intractable ill-conditioned problems rely on imposing additional constraints that bias the solution [5]. The Bayesian approach bypasses the need for these regularization techniques by treating all variables as random variables and

allowing for a full probability distribution of solutions, rather than requiring somewhat arbitrary constraints.

2.5. Bayesian Inference

We can think of parameter estimation in the Bayesian framework as a statistical quest for information about the unknown parameter. In the context of the mathematical model (Equation 2.3), there are some directly observable quantities, such as the wrench measurements, but the primary quantities of interest, the muscle strengths, are not observable. The Bayesian approach allows us to extract information about the muscle strengths that are dependent on the observable quantities through the model. Instead of finding a single estimate of the individual muscle strengths, a set of probability distributions is produced that can be used to obtain estimates and a level of uncertainty about the estimates. In this way, the inverse problem is no longer ill-posed, but instead it produces the range of possible answers in the larger space of probability distributions.

Bayes' theorem states that the posterior probability distribution of the unobservable random variable, θ, given the observed measurement $X = x$, is:

$$p(\theta|x) = \frac{p(x|\theta)\ p(\theta)}{p(x)} \tag{2.4}$$

The posterior distribution, $p(\theta|x)$, expresses what we know about θ after the measurement data $X = x$ is observed. Information that is known about the unobservable quantity, θ, can be encoded in the prior density $p(\theta)$. The likelihood function $p(x|\theta)$ relates the variables to the model and indicates how likely it is to observe the data x

under the model with parameters θ. In other words, the likelihood function describes the relationship between the observed measurements and the unknown quantity that we are trying to estimate. For some models, this is simply a forward solution to the problem and can be expressed analytically. In the case of the mathematical model presented in Equation 2.3, the likelihood function can not be specified analytically, so a likelihood-free approach is necessary.

2.5.0.1. Approximate Bayesian Computation. In cases where the likelihood function is intractable, Approximate Bayesian Computation (ABC) is a useful alternative to the traditional Bayesian approach. In ABC, the likelihood function is replaced with a simulation of the model to produce an artificial data set that can be compared to the measured data. Specifically, samples of the unknown parameter, θ, are taken and used to simulate data x_{sim}. The simulated data set is then compared to observed data x by computing some distance metric.

In its simplest form, an ABC rejection sampler [85] for a discrete system (Algorithm 1) involves repeatedly sampling candidate parameter values from the prior distribution, $\theta' \sim f(\theta)$, generating data with the model, $x_{sim} = M(\theta')$, and accepting or rejecting the candidate parameter value based on some distance metric, $\rho(x_{sim}, x_{meas}) \leq \eta$.

Algorithm 1 ABC Rejection Sampler

```
for i = 1 to N do
    repeat
        Sample θ' ~ f(θ)
        Generate data x_sim from M(θ')
    until ρ(x_sim, x_meas) ≤ η
    set θ_i = θ'
end for
```

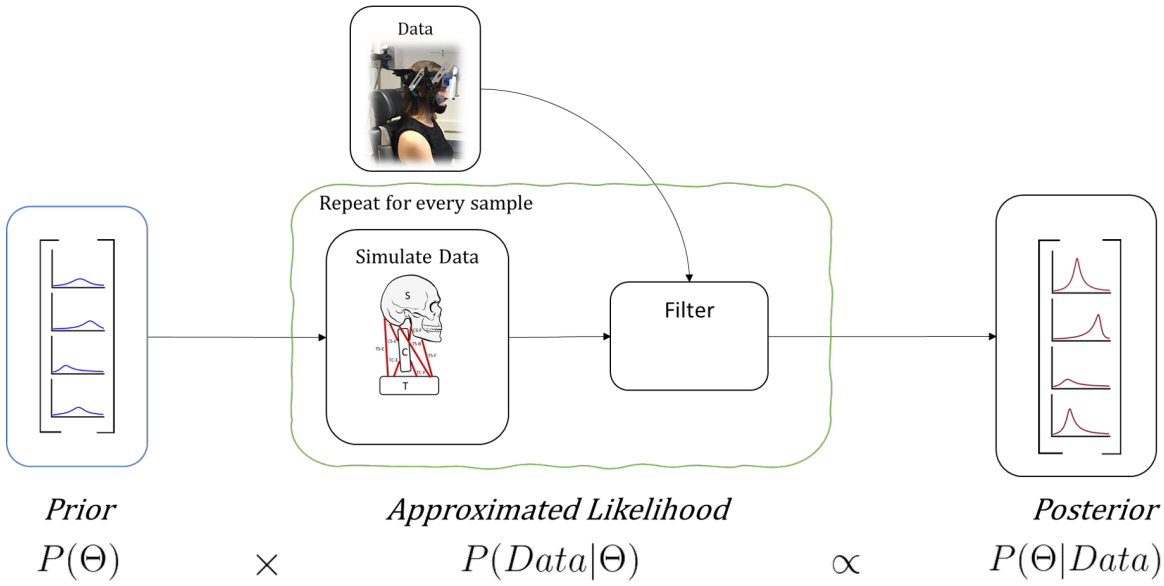


Figure 2.1. **Approximate Bayesian Computation Algorithm**. A high level visualization of an ABC algorithm. Samples are generated from the prior distribution. Data is simulated with the model and compared to the observed data. For a rejection sampler, like Algorithm 1, the filter rejects candidate parameter values that produce data with error larger than some cutoff error.

Many specific extensions of the basic ABC framework have been developed to address challenges with sampling the parameter space and convergence, including MCMC, sequential Monte-Carlo methods importance sampling, and regression adjustments [103, 101, 70].

2.6. Information Theoretic Measures

In robotics, search, and source localization applications, information theoretic measures have been used quite extensively to maximize information gain in a Bayesian framework. Information theory originated from a single paper written in 1948 by Claude Shannon [90] that provides a quantitative metric for measuring the loss of information

during transmission of a communication signal. The work introduces entropy as a metric for encoding the expected uncertainty in a signal. Information theoretic measures have been used extensively in the field of optimal experimental design [95], where the primary objective in the context of parameter estimation is to select an experiment that will produce the best statistical quality of the parameter estimates [77]. One strategy is to maximize the information gain, from prior to posterior. The field of information theory defines self-information of an event x as $I(x) = -\log(p(x))$. $I(x)$ can be thought of as the information content of the random variable X. Shannon entropy [90], the average amount of information in a probability distribution, is computed using Equation 2.6 by taking the expected value of the self-information over all possible values of x. Shannon entropy can also be thought of as the measure of uncertainty of a probability distribution [48].

The expected value of a distribution of a random variable is the weighted sum of the probabilities over the domain of the random variable.

$$\mathbb{E}[X] = \sum_{x} x\ p(x) \tag{2.5}$$

So the average entropy over x is:

$$S(x) = \mathbb{E}(I(x)) = \sum_{i=1}^{N} -p(x_i)\log(p(x_i)) \tag{2.6}$$

where X takes N possible values x_i, $i = 1, ..., N$. Note that $S \geq 0$ and $S = 0$ only if there is no uncertainty in X. The maximum value that S can have is $\ln N$, where all

possible outcomes of X have equal probability (a uniform distribution over the full range of X).

CHAPTER 3

The Geography of Fatty Infiltrates within the Cervical Multifidus and Semispinalis Cervicis in Individuals with Chronic Whiplash-Associated Disorders

3.1. Abstract

We aim to quantify the magnitude and distribution of muscle fat infiltration (MFI) within the cervical multifidus and semispinalis cervicis muscles of individuals with chronic whiplash-associated disorders (WAD) as compared to healthy controls and those who fully recover from a whiplash injury secondary to a motor vehicle collision (MVC).

Previous research has established the presence of increased MFI throughout the entire cervical extensor muscles of WAD patients when compared to healthy controls. These changes appear to be greater in deepest muscles (e.g. multifidus and semispinalis cervicis) than the more superficial muscles. A detailed analysis of the specific distribution of MFI within these deep extensor muscles in WAD, recovered, and control groups may provide foundation investigating specific mechanisms, etiologies, and targets for treatments.

Fifteen participants in 3 Groups: WAD (n=5), Recovered (n=5), and Controls (n=5) were studied using a 3 dimensional (3D) Fat/Water separation magnetic resonance imaging (MRI) sequence. Bilateral measures of cervical multifidus and semispinalis cervicis MFI in four quartiles (Q1 – Q4; medial to lateral) at cervical levels C3 through C7 were included in the analysis. Intra- and Inter-rater reliability was established. A mixed model

analysis was performed to control for covariates, identify interaction effects, and compare MFI distribution between groups.

The Limits of Agreement confirmed strong inter- and intra- rater agreement at all levels (C3-C7). Gender, age, and BMI were identified as significant covariates on MFI. Significant interactions were found between group and muscle quartile ($p < 0.001$) and between muscle quartile and cervical level2 ($p < 0.001$). Pairwise comparisons for intra-quartile MFI between groups revealed a significantly greater MFI in the WAD group when compared to the Recovered and Control groups in Q1 and Q2 ($p < 0.002$) and greater MFI over the recovered group in Q3 and Q4.

This study provides preliminary data mapping the spatial distribution of MFI in the cervical multifidus and semispinalis cervicis muscles in chronic WAD, recovered, and healthy controls. MFI is more concentrated in the medial portion of the muscles in WAD, recovered, and healthy control participants. However, the magnitude of MFI in each quartile is greatest in those with chronic WAD.

3.2. Introduction

Previous studies using magnetic resonance imaging (MRI) have demonstrated widespread fatty infiltrates in the neck extensor [24] and flexor [28] muscles of individuals with chronic whiplash associated disorders (WAD). These high levels of muscle fat infiltration (MFI) were not present in those with chronic non-traumatic neck pain6 or those without a history of neck disorders [24]. While widespread, the greatest magnitude of MFI was consistently observed in the deepest muscular layer of the extensors (e.g. the multifidus and semispinalis cervicis) when compared to the more superficial musculature (e.g.

semispinalis capitis, splenius capitis and upper trapezius) [24, 29]. However, the specific role of MFI in the development and maintenance of chronic WAD is not fully understood [26, 98]. Improvements in our mechanistic understanding of the development of structural changes (e.g. composition and morphology) in the cervical muscles of patients with chronic WAD may shed light on their potential contribution to poor functional recovery.

The extensor and flexor muscles of the human neck are responsible for the majority of the postural stability of the cervical spinal column [81]. The extensors are layered and can be divided into functional groups based on location and attachments. The most superficial extensor muscles bypass the vertebrae to attach the shoulder girdle to the cranium (e.g. upper trapezius). Intermediate extensors (e.g. portions of the splenius and semispinalis capitis), attach the shoulder girdle to individual vertebrae, or individual vertebrae to the cranium. The deep extensors, including the cervical multifidus and the semispinalis cervicis, attach vertebrae directly to other vertebrae. These deep extensors likely play a specific role in segmental support of the cervical spine and fine head/neck postural control [69]. Degeneration of these complex muscles, as the MFI might indicate, may have implications for altered biomechanics of the neck and postural stability in those with chronic WAD. However, the associations between MFI and commonly observed impairments (e.g. proprioceptive functioning) and other functional characteristics, such as stiffness and altered motor control, are yet to be quantified.

Previous animal work suggests specific patterns of MFI are a consequential marker of induced experimental injuries involving the lumbar spine [68]. Clinically, the magnitude and anatomical location of their presence alters plan of care with respect to surgical management of rotator cuff tears [39, 45, 75]. Furthermore, measurement of and histological

confirmation for MFI with a quantitative MRI approach has been produced in both an animal [93] and human model [37].

Accordingly, this preliminary study aims to establish the reliability of an MRI measure for mapping the magnitude and spatial distribution of MFI in the cervical multifidus and semispinalis cervicis and to provide preliminary data for comparing WAD, recovered whiplash, and healthy control groups. Such data is necessary to better understand potential altered biomechanics (e.g. changes in joint torques) and control patterns (e.g. changes in task-dependent muscle activation) that may lead to or perpetuate persistent pain and disability in those with chronic WAD.

3.3. Methods

3.3.1. Study Population

A total of fifteen participants were included in this analysis. Five gender and age matched participants were selected for each group (WAD, Recovered, and Controls) from a concurrent prospective study. The MRI measures for MFI were performed by two-raters that were blind to the clinical severity of each participant. Each participant in the WAD group (n=5) had chronic neck pain (3 months – 5 years post MVC) and a neck disability index (NDI) greater than 30%. The Recovered group (n=5) nominated full resolution of symptoms and were scanned at 3 months post MVC with an NDI score of less than 10%. The Control group (n=5) had no history of neck pain that required treatment in the past 10 years. Additional inclusion criteria included no cervical spine fracture following MVC, no previous neck injury, and no previously diagnosed nervous system disorders. All participants provided informed consent before participating in the study.

3.3.2. MRI Measures and Analysis

A three-dimensional multi-echo gradient echo acquisition was performed to collect the data required for the analysis of phase related to the precessional differences in muscle fat and water. A standard 12-channel head coil and 4-channel neck coil were used as receiver coils to improve signal-to-noise. The axial FLASH dual echo, gradient echo sequence was 4:23 minutes with an in-plane resolution of 0.7mm using a rectangular field of view of 75% and thickness of 3mm and slab oversampling of 22% with 36 partitions to prevent aliasing, TR/TE1/TE2 6.59/2.45/3.68 ms with a FOV of 190 x 320 mm. This scan covered the cephalad portion of C3 through the caudal portion of the C7 vertebral end plate. By collecting data at an echo time when water and fat are in-phase and at an echo time when water and fat are out of phase, two images are produced. The ratio of the pixel intensities within a voxel gives an index of muscle fat infiltration.

The MRI analysis consisted of manually tracing defined regions of interest (ROI) over the bilateral cervical multifidus and semispinalis cervicis muscles on axial MR slices. A customized program was developed using MATLAB (MathWorks^{TM} USA) to quantify the magnitude of MFI in each quartile of the ROI. Quartiles were defined by two dimensional coordinates containing an equal number of pixels running from medial (Quartile 1- Q1) to lateral (Quartile 4 – Q4) based on the orientation of the muscle as viewed in the transverse plane, as shown in Figure 3.1. The measure of fat within the muscle was calculated as the ratio of pixel intensities from the fat and water images:

$$MFI = \frac{IF}{IF + IW}$$

where IF = fat intensity and IW = water intensity

The quartile and whole MFI value at each cervical level (C3-C7) is the average of measurements at 5 axial slices that span the length of the vertebral level. For example, the MFI value of Q2 at cervical level C4 of Subject A, is the average of 5 MFI measurements from sequential axial MR slices that span from the top to the bottom of the C4 vertebral body. An MFI measure for all four quartiles was created in this manner bilaterally at each level for each subject.

To test repeatability, two authors (RA and MH) experienced in this technique, independently traced the ROI for each cervical level of 6 randomly selected participants (4 WADs and 2 Controls) while being blind to the other's assessment.

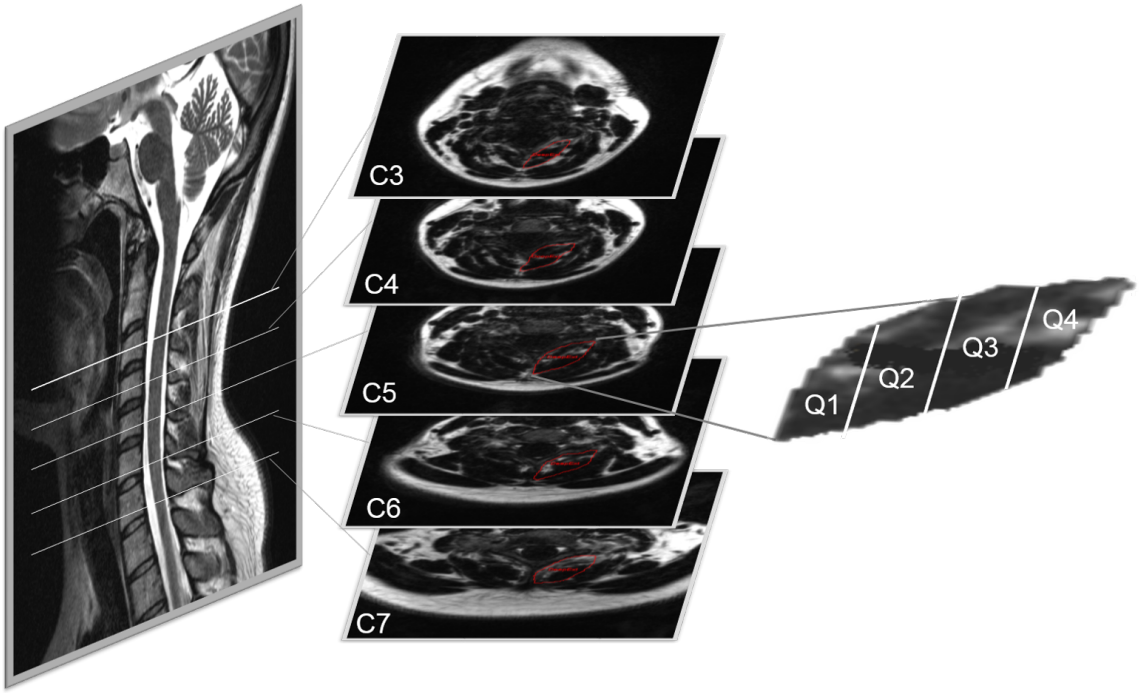


Figure 3.1. Multifidus and Semispinalis Cervicis MFI was analyzed bilaterally (only one side shown above) at 5 cervical levels (C3-C7) in quartiles (Q1 (medial) -Q4 (lateral)).

3.3.3. Statistical Analysis

Intra- and inter- rater reliability of intra-quartile MFI were examined through construction of Bland-Altman plots and calculation of the limits of agreement (LOA) using the method proposed by Bland & Altman [73]. Intra-class correlation coefficients (ICCs) were also calculated for both intra- and inter- rater reliability data using a two-way random effects structure and absolute agreement (ICC[2,1]).

A repeated measures linear mixed model approach was used to investigate differences between groups in the magnitude and spatial distribution of MFI. A within subject factor of level * quartile was used to account for the spatial repetition of measures of MFI across cervical level and muscle quartile. Preliminary investigation of MFI across all cervical levels revealed a non-linear quadratic-type relationship between MFI and cervical level. Thus, the quadratic effect of cervical level (i.e. cervical level) on MFI was entered into the model as a fixed effect. The main and interactive effects of muscle quartile and group (WAD, recovered, control) and the 3-way interactive effect of muscle quartile with group and cervical level were also included in the model. Inter-subject differences in the change in MFI across cervical level (cervical level) and across quartiles were included in the model as random effects. Age, gender, and BMI were entered as covariates given the previous evidence of their effect on MFI. Pairwise comparisons with Bonferroni corrections for multiple comparisons were used to investigate between group differences in intra-quartile MFI. All statistical procedures were carried out using IBM SPSS 22.0 (IBM). For all analyses, the significance level was set to $p \leq 0.05$.

3.4. Results

Both inter- and intra-rater agreement were confirmed with Bland & Altman tests (Figure 3.2 a, b) and ICCs (Figure 3.2 c, d). Bland & Altman plots showed no systematic relationship between MFI values and absolute differences for either inter- or intra-rater data. LOA showed slightly better agreement within (-0.04, 0.06) than between raters (-0.07, 0.07).

ICC values for both inter- (ICC[2,1] = 0.93, 95% CI = 0.90, and 0.94) and intra- (ICC[2,1] = 0.98, 95% CI = 0.97, 0.98) rater reliability showed excellent levels of agreement.

	Gender (Females)	Age (years)	Body Mass Index (BMI)	Neck Disability Index (x/100)
Chronic Whiplash (n=5)	3	30.6 ± 9.0	29.6 ± 4.6	41.2 ± 13.6
Recovered Whiplash (n=5)	3	32.8 ± 7.5	26.4 ± 1.1	1.6 ± 2.6
Healthy Controls (n=5)	3	35.0 ± 8.9	25.8 ± 5.0	

Table 3.1. Demographic data (Mean ± SD)

Mixed model analysis revealed a significant effect of the covariates gender ($F[1, 30.9] = 11.8$, $p = 0.002$) and BMI ($F[1, 30.9] = 28.0$, $p < 0.001$) on MFI. There was no significant effect for age on MFI ($F[1, 30.9] = 1.0$, $p=0.316$). There was a significant interaction between group and muscle quartile ($F[6, 211.7] = 6.4$, $p < 0.001$) and between muscle quartile and cervical level ($F[3, 216.9] = 29.6$, $p < 0.001$) but not between group and cervical level ($F[2, 29.5] = 0.5$, $p = 0.61$). The three-way interaction between group cervical level and muscle quartile was not significant ($F[6, 216.9] = 1.2$, $p = 0.32$ indicating that between-group differences in intra-quartile MFI were consistent across cervical levels.

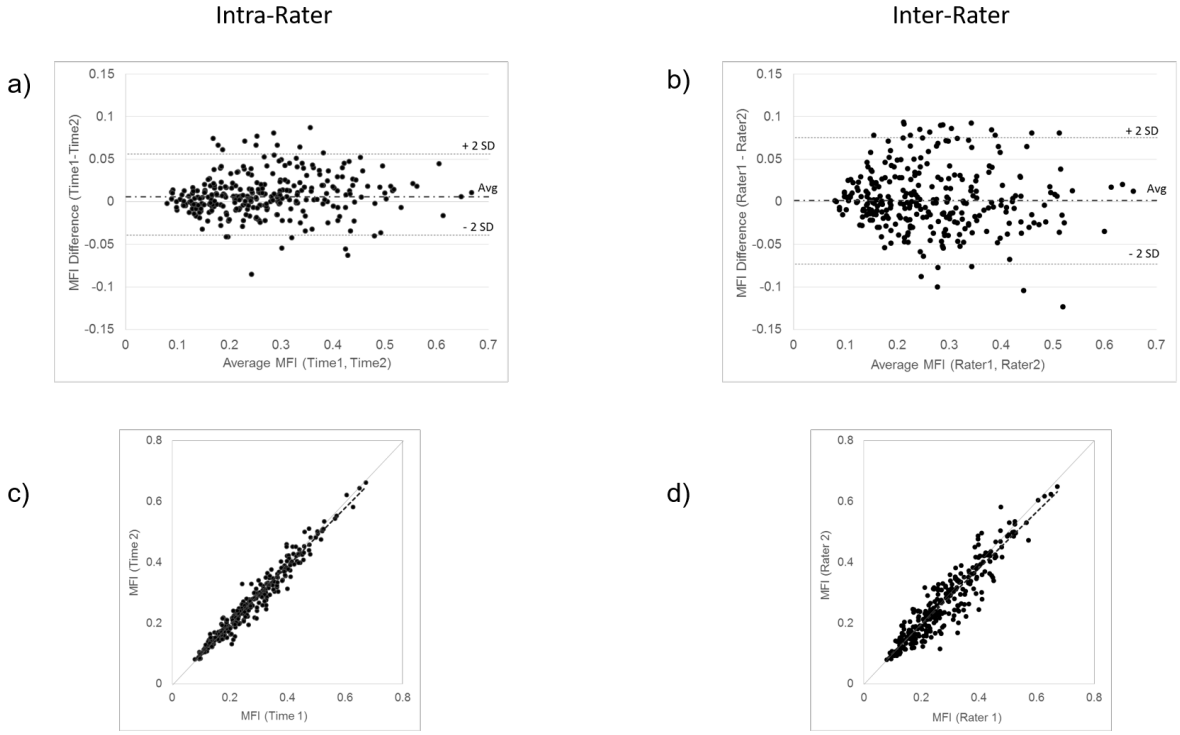


Figure 3.2. Inter- and Intra- rater reliability for the MFI quantification method. a) Bland-Altman Plot for intra-rater reliability (LOA=(-0.04, 0.06)), b) Bland-Altman Plot for inter-rater reliability (LOA=(-0.07, 0.07)), c) Intra-rater correlation (ICC[2,1] = 0.98, 95% CI = 0.97, 0.98), d) Inter-rater correlation (ICC[2,1] = 0.93, 95% CI = 0.90, 0.94).

Pairwise comparisons for intra-quartile fat between study groups are shown in Figure 3.3. The WAD group had significantly higher MFI values than the recovered group in Q1 (df = 45.2, $p < 0.001$), Q2 (df = 46.17, $p < 0.001$) and Q3 (df = 47.6, $p = 0.03$). Significantly higher MFI was also observed in the WAD group in comparison to the control group in Q1 (df = 43.0, $p = 0.002$) and Q2 (df = 43.8, $p = 0.045$). No significant differences between the control and WAD groups were observed in Q3 or Q4. The recovered group showed significantly lower MFI in comparison to the control group in Q1 (df = 48.6, $p = 0.048$).

		Control		Recovered		WAD	
		MFI	SD	MFI	SD	MFI	SD
C7	Q1	0.289	0.079	0.252	0.085	0.381	0.08
	Q2	0.188	0.057	0.146	0.044	0.265	0.098
	Q3	0.148	0.051	0.119	0.027	0.195	0.071
	Q4	0.158	0.06	0.13	0.031	0.197	0.11
	Whole	0.199	0.06	0.164	0.045	0.266	0.08
C6	Q1	0.267	0.064	0.214	0.056	0.363	0.074
	Q2	0.165	0.044	0.131	0.041	0.227	0.073
	Q3	0.112	0.04	0.097	0.022	0.161	0.064
	Q4	0.122	0.048	0.116	0.021	0.158	0.091
	Whole	0.17	0.045	0.142	0.034	0.234	0.071
C5	Q1	0.217	0.073	0.177	0.018	0.313	0.075
	Q2	0.166	0.06	0.141	0.041	0.252	0.057
	Q3	0.113	0.034	0.108	0.016	0.178	0.052
	Q4	0.126	0.031	0.134	0.021	0.176	0.06
	Whole	0.157	0.048	0.141	0.021	0.234	0.048
C4	Q1	0.206	0.049	0.19	0.061	0.305	0.077
	Q2	0.179	0.074	0.155	0.044	0.259	0.074
	Q3	0.129	0.052	0.11	0.021	0.184	0.062
	Q4	0.147	0.041	0.14	0.027	0.214	0.049
	Whole	0.167	0.051	0.15	0.031	0.244	0.056
C3	Q1	0.232	0.064	0.231	0.033	0.332	0.09
	Q2	0.232	0.083	0.22	0.032	0.319	0.096
	Q3	0.218	0.056	0.206	0.046	0.304	0.067
	Q4	0.231	0.056	0.239	0.044	0.325	0.052
	Whole	0.23	0.061	0.226	0.032	0.323	0.065

Table 3.2. Muscle Fat Infiltration by group, level, and quartile

3.5. Discussion

This study provides preliminary data mapping the regional distribution and magnitude of MFI in the deep extensor muscles in a small sample of individuals with chronic WAD, recovered whiplash, and healthy controls. Consistent with previous studies, the WAD group presents with the greatest average MFI throughout the deep extensors when

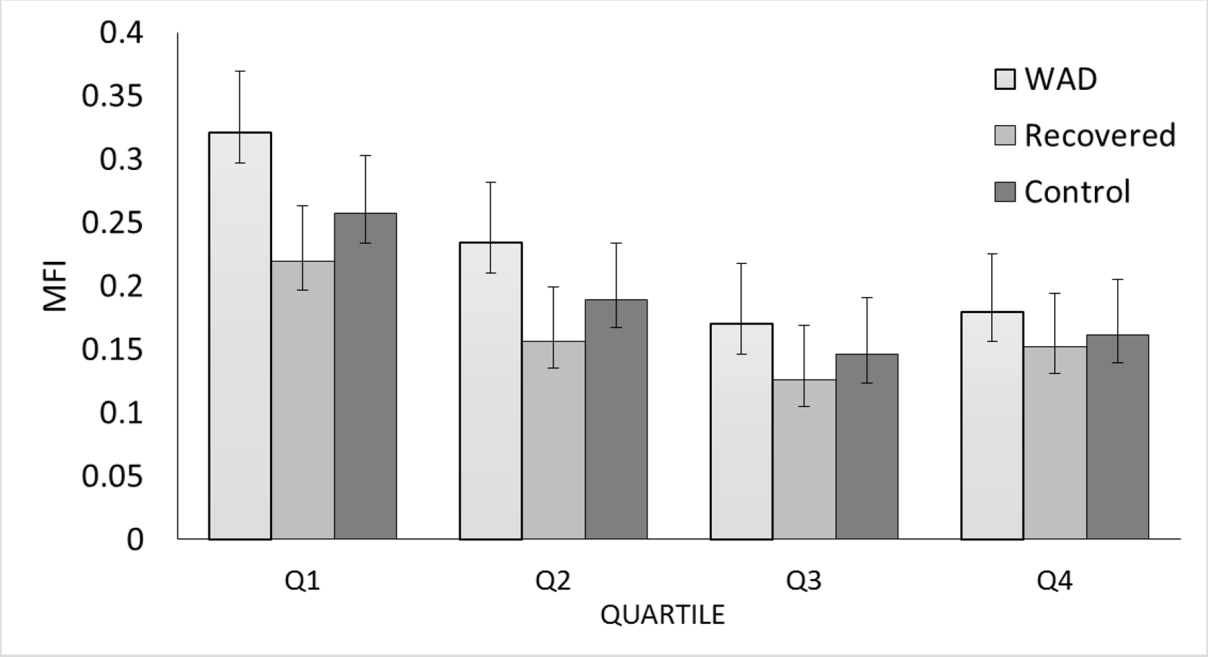


Figure 3.3. Mean MFI for between group comparison at each quartile (Q1-medial, Q4- lateral) averaged over all cervical levels. The values are estimated marginal means and were based on the covariate values of age = 32.8 years, BMI = 27.27, and cervical level = C4. Significance is denoted as * for p¡0.05, ** for p<0.01, and *** for p<0.001.

compared to controls [24] and those that recover from an MVC [26]. Between group comparisons of intra-quartile MFI revealed a significantly greater MFI in Q1 and Q2 in the WAD group compared to both the recovered and control groups. Oddly, intra-quartile MFI was significantly higher in the WAD group compared to the recovered group, but not control group in Q3. Another unexpected result is that the recovered group demonstrated significantly lower MFI in Q1 and Q2 compared to the control group. These two observations are difficult to explain, but may be the result of low subject numbers or traumatic factors [20, 39], such as inflammation [64], that may drive (or inhibit) the development of MFI [98]. Further mechanistic investigation of local and systemic serum quantities of

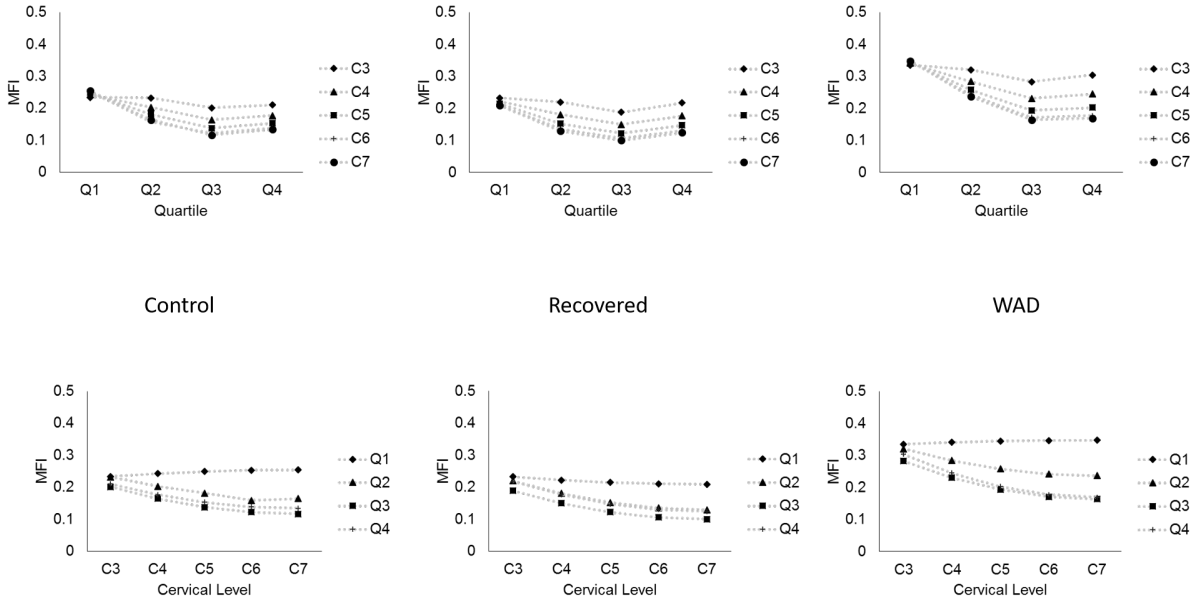


Figure 3.4. Mapping of the MFI across quartiles for controls (a), recovered (b), and WAD (c) subjects and across cervical levels for controls (d), recovered (e), and WAD (f).

inflammatory biomarkers (e.g. TNF-α, C-Reactive Protein, and IL-1β) [98] is warranted as this could help explain the differential response (and presence) of MFI between the recovered and control participants. Also, while we did not collect data regarding the various treatments being received by any of the subjects with whiplash, it remains plausible that the presumed healthier muscle tissue (e.g. less MFI) in recovered participants compared to the controls was, in part, influenced by an intervention effect. This is currently being documented and investigated in a larger prospective study.

Despite low subject numbers, data from the plot of limits of agreement suggest that the MFI quartile analysis is repeatable between and within raters experienced in this method. Because each ROI is considered an independent data point (5 cervical levels x 2 sides = 10 data points per subject) and the images were presented in a randomized

fashion, the sample size is adequate to determine reliability for this measure. According to Walter and colleagues, assuming a minimum agreement (ICC) of 0.80, with a null hypothesis of 0.4, a minimum of 9 independent observations provides an acceptable 5% alpha error rate ($p < 0.05$) and 20% beta error rate [113]. There can be a strong level of confidence in accepting the results of the measure as is demonstrated by the widths of the 95% CI for the WAD group mean inter-rater differences in MFI across all cervical levels.

Also of interest are the interactions between quartile, cervical level, and group. Inter-quartile differences in MFI differed across cervical levels. MFI appears to be relatively equally distributed between quartiles at the highest analyzed cervical level (C3), but at lower cervical levels Q1 is the highest of the quartiles while Q2, Q3, and Q4 decrease or stay the same (See Figure 3.4 a,b,c). This general pattern that is apparent in all groups may be related to the architecture of the deep muscles rather than a true indicator of degeneration at these levels. The second significant interaction reveals that MFI within muscle quartiles differs between the study groups (Figure 3.3). These between-group differences are consistent across cervical levels. While preliminary, the findings support further prospective investigation with a larger cohort that may reveal the development of a characteristic MFI distribution pattern amongst participants with varying levels of pain-related disability. Such work is currently underway, aiming to identify specific injury mechanisms underlying MFI, biomechanical consequences of MFI, and potential treatment targets.

Although higher measurements of total MFI in the multifidus muscles have been attributed to WAD, these current observations suggest the intra-muscular changes may be

more regionally situated than previously thought. The multifidus and semispinalis cervicis muscles have a complex architecture, with fascicles attaching to the spinous process (falling into Q1) of a superior vertebrae and then traveling laterally to attach to the transverse process of an inferior vertebrae (falling into Q4). While the precise biomechanical consequences of the specific spatial patterns of MFI within the deep extensors observed in this study are unknown, several hypotheses can be generated. If local MFI increases translate to deficits in motor function, the location of high MFI in specific neck muscles is relevant to the clinical presentation and the development of biomechanical models of the human head/neck. To our knowledge, the distribution and magnitude of MFI has not factored into the development of and use for current modeling efforts.

The magnitude and location of structural muscle changes may have implications for ongoing neck pain following a MVC, as they could alter the internal forces in joints and muscles. The multifidus and semispinalis cervicis both have attachments to the facet capsules, which have been consistently implicated in the generation of and maintenance for neck pain following whiplash [68, 80, 91, 115]. Treatments targeting facetogenic pain are available. Radiofrequency neurotomy (RFN) interventions have been shown to attenuate the psychophysical signs/symptoms (e.g. thermal/pressure pain thresholds) [93]. However, the long-term biomechanical consequences of RFN are unknown. Equivocal findings with respect to the influence of the lumbar spine muscle system were found by Smuck et al [94]. Further exploration with current quantitative methods are warranted to better understand the biomechanical consequences of RFN [2].

3.6. Conclusions

The results of this study provide preliminary evidence of unique patterns of MFI distribution within the deep extensor muscles of WAD, recovered whiplash, and healthy controls. This provides foundation to explore whether the geography and magnitude of MFI contributes to persistent functional deficits common to some patients with WAD. Further study to verify this pattern and to determine the potential influence of rehabilitative exercise amongst a larger cohort with varying amounts of MFI is warranted.

3.7. Key Points

Findings: The distribution of MFI in the deep cervical extensor muscles differs between groups in a small sample of WAD, recovered whiplash, and control groups.

Implications: Quantifying the magnitude and distribution of MFI in the cervical muscles may elucidate the mechanisms underlying disturbed neuromuscular control of posture and dynamic stability of the head and neck following whiplash injuries. Accordingly, such information may provide foundation for exploring and developing more informed treatment regimens targeting specific muscle groups, improved health, and functional recovery.

Caution: It is necessary to establish a larger dataset of within muscle variation of MFI across all muscles in the cervical spine in participants with WAD before more definitive biomechanical conclusions can be made. Furthermore, a scan-rescan reliability measure is also warranted to account for potential acquisition variability.

CHAPTER 4

The Role of the Deep Cervical Extensor Muscles in Multi-Directional Isometric Neck Strength

In the previous chapter, we confirmed the widespread increase in muscle fat infiltration throughout the deep cervical extensor muscles of individuals with WAD. An obvious next question is, what are the biomechanical consequences of these compositional muscle changes? This chapter utilizes a musculoskeletal model of the neck to explore the role of the deep cervical extensor muscles in multi-directional neck strength and the biomechanical consequences of weakness of this muscle group. Through simulation, we illustrate how deep extensor muscle weakness affects neck strength and muscle activation patterns in 25 test directions mapped to locations on the skull where resistance could be applied by a clinician. The results predict that an individual with decreased deep extensor muscle strength would present with decreased neck strength in the posterior and postero-lateral directions, but would still be able to produce forces in all 25 test directions. We also show that deep extensor weakness could lead to an altered muscle activation pattern characterized by increased activation of the superficial extensors (splenius and semispinalis capitis) and upper cervical flexors (longus capitis) along with a decrease in activation of the suboccipital muscles. These simulation results could help explain some reported motor dysfunction in WAD and warrants further investigation.

4.1. Introduction

Imaging studies have confirmed the development of increased muscle fat infiltration (MFI) in the neck muscles of individuals with chronic whiplash-associated disorders (WAD), with the greatest concentration of MFI found in the deep cervical extensor muscles (multifidus and semispinalis cervicis) [1, 23]. While the role of MFI in the development and maintenance of movement dysfunction and chronic WAD is unknown [117, 100], it is indicative of muscle atrophy, and associated with a decrease in force generating capacity of the deep cervical extensor muscles. Previous studies have shown that individuals with chronic WAD exhibit a marked decrease in isometric neck strength in all cardinal planes (flexion/extension, lateral flexion, and rotation) when compared to healthy controls, with the greatest relative weakness reported in extension [82, 86, 61]. Altered muscle activation patterns in WAD during dynamic tasks are characterized by decreased and delayed anticipatory recruitment of the deep cervical flexors (longus colli and capitis) [51] and increased tonic recruitment of superficial muscles (sternocleidomastoid, upper trapezius, scalenes, and levator scapulae) [53, 31]. In order to establish the link between muscle atrophy of the deep extensors and motor dysfunction in WAD, it is crucial to understand the biomechanical role of the deep extensor muscles within the complex system of the neck. The human neck is structurally complex, with 7 joints and more than 20 muscle pairs with attachments to multiple points that cross one or more joints, acting on one or more degrees of freedom [56]. Kinematic and muscle redundancy prevent the direct computation of individual muscle contribution to neck strength measurements, as any forces and moments measured at the head could be produced by many non-unique combinations of active muscle sets and muscle forces. In this study, we utilize a computational model

of the neck to investigate the effect of deep cervical extensor weakness on the maximum isometric strength of the neck and resultant muscle activation patterns.

To our knowledge, no studies have attempted to identify the influence of deep extensor weakness on neck strength and function in vivo. One barrier is the invasiveness of intramuscular EMG for the deep extensor muscles, which requires inserting the electrode through several layers of muscle and fascia before reaching the muscles. Spatial tuning of neck muscles measured by surface and intramuscular electromyography (EMG) in healthy individuals has given insights into the directional recruitment of specific muscle groups when producing isometric multi-directional forces [110, 58, 34, 11]. An important observation is that the preferred direction of most neck muscles does not align with the muscle's biomechanical line of action, highlighting the complex interdependence of muscles may not be intuitive when balancing torques across multiple joints to produce forces at the head [58].

Clinically, neck strength is tested in the cardinal directions (flexion/extension, lateral flexion, rotation) by the clinician providing manual resistance at points on the head [44]. Unique postures are described to target the capital joint (upper cervical spine) vs. the cervical spine, but the contribution of specific muscle groups is not clear [44]. Experimental studies of isometric neck strength and muscle spatial tuning typically define test directions as torques about 3 axes crossing at the C7-T1 intervertebral joint [109, 34]. This mapping from 6-D measurements at the head to 3-D torques ignores important degrees of freedom available to the neck and head, specifically the differential contribution of torques about the upper vs. lower cervical spine, which would significantly affect muscle contributions. The contribution of particular muscles forces to joint torques and end effector forces can

also be explored using computational models. One way to conceptualize strength of a model is the feasible force set (FFS), defined as the bounds of the biomechanical capability to produce static forces in all directions from its end effector [106]. The FFS has been used to probe the role of particular muscles in producing end effector forces in the finger [104], cat hindlimb [96], and leg [62, 97]. This study presents results of the FFS in 25 unique points and force directions located on landmarks on the skull to account for all 6 degrees of freedom of the head on the neck and better replicate clinical strength tests.

There is no available method for measuring muscle strength directly in humans. Computational models can serve as valuable tools because they allow access to all parameters and state variables, many of which are not manipulable or measurable in the real system, such as muscle force and activation. Neck models in the literature vary considerably in their kinematic and muscle complexity, from planar single joint inverted pendulums with 2 muscles to 48 Degree of Freedom (DOF), 129 muscle element models [17]. It is unclear how much independent control humans have over activation of individual neck muscles. An oversimplified muscle model may over-constrain muscle groups to activate together in a way that removes valid solutions from the solution space, while a model with too many muscle elements may be overly robust to constraints and no longer exhibit the behaviors of the human neck. In the absence of a scientific consensus on how neck muscles are controlled by the nervous system, this study compares models with several levels of kinematic and muscle redundancy. When making decisions on model complexity, there is a trade-off between minimizing complexity to maintain interpretability and having enough complexity for the model to produce realistic behaviors for the specified task. The role of

the deep extensor muscles is likely to be influenced by the level of kinematic and muscle redundancy in the system [97].

The primary objective of this study is to determine the effect of deep extensor weakness on the neck strength and muscle activation patterns when producing forces at 25 locations on the head. A secondary objective is to evaluate the effect of kinematic and muscle model complexity on 1) neck strength and 2) the role of the deep extensor muscles. These aims are achieved with a detailed analysis of muscle contributions to neck strength in simulation using musculoskeletal modeling of the neck. The results of this study can be readily translated to the clinic to support the delivery of tailored physical therapy interventions to target specific neck muscles in WAD.

4.2. Methods

Neck strength is represented in this study as the set of maximum wrench magnitudes generated across 25 pre-defined test directions. The wrench refers to the set of moments and forces measured at the end effector (the head). We examine the role of the deep cervical extensor muscles by varying the deep extensor muscle strength in a musculoskeletal model and simulating the resulting multi-directional neck strength and corresponding changes in muscle activation patterns.

The design choice of computational model can have a significant effect on the outcome and interpretation of a study. The number of kinematic degrees of freedom, individual muscle elements, and the constraints of the task are likely to affect the outcomes of interest: neck strength and muscle activation patterns [97]. To address this issue, we compare neck models with varying kinematic and muscle complexity.

4.2.1. Musculoskeletal Models

The geometry and muscle properties are derived from the generic OpenSim Neck Model developed by Vasavada et al [108] with the addition of the hyoid muscles [78]. The cervical spine is configured in a neutral, upright posture. The OpenSim Matlab API is used to access the moment arm matrix R, maximum isometric forces F, and the skull frame Jacobian J. All computations and simulations are performed in Matlab. The full neck model contains 7 joints (C0-C1 through C6-C7), each with 3 rotational DOF, for a total of 21 DOF. It has 96 individual muscle elements.

4.2.1.1. Kinematic Complexity Variants. Six models with varying kinematic complexity are derived from the full neck model. The three planar variants are the 7 joint (7 DOF), 3 joint (3 DOF), and 2 joint (2 DOF) models. The planar models have their joints oriented in the sagittal plane (flexion/extension). The task space (in coordinates of the skull frame) is 3-D, so that the end effector (skull) wrench is represented in coordinates $w = (M_z, F_x, F_y)$.

The three spatial variants are the 7 joint (21 DOF), 3 joint (9 DOF), and 2 joint (6 DOF) models. Each joint allows 3 rotational DOF (pitch, roll, and yaw). The task space for the spatial models is 6-D, so that the end effector (skull) wrench is represented in coordinates $w = (M_x, M_y, M_z, F_x, F_y, F_z)$.

4.2.1.2. Muscle Complexity Variants. Four variants of muscle complexity are compared for the spatial models. Each muscle group is constrained to activate together. The first variant allows all 96 muscle elements ($m = 96$) to be activated independently. The second variant groups muscles by their sagittal and axial actions across joints in the 3 joint model, for a total of 21 muscle pairs ($m = 42$). For example, muscles that produce a

flexion and ipsilateral rotation moment about C6-C7 are grouped together. Muscles producing torques about multiple joints are grouped separately from those producing torques about individual joints. The third variant groups muscles by only their sagittal action, resulting in 12 muscle pairs ($m = 24$). The fourth variant groups muscles by anatomical name, resulting in 12 muscle pairs ($m = 24$). Planar muscle groupings are the same as the spatial, except with the left and right muscle elements combined, resulting in half of the number of muscle groups. Note that the second variant (sagittal and axial action) is not relevant in the planar models, because the axial plane is removed.

4.2.1.3. Simplified Model Parameters. All model variants are derived from the full 7 joint spatial (d=6) model with $n = 21$ DOF and $m = 96$ independent muscles. The transpose of the frame Jacobian $J^T \in \mathbb{R}^{n \times d}$ projects wrenches from the $d = 6$ DOF task (or wrench) space to torques in $n = 21$ DOF joint space. To obtain the Jacobian for the simplified 3 joint spatial variant model, the rows corresponding to the $n = 9$ joint coordinates are selected, resulting in a new $J^T \in \mathbb{R}^{9 \times 6}$. For the planar models, only the columns that correspond to the task space coordinates (M_z, F_x, F_y) are included, resulting in an $(m \times 3)$ Jacobian transpose.

The moment arm matrix $R \in \mathbb{R}^{n \times m}$ is simplified in a similar manner for the kinematic model variants. Only rows that correspond to the n coordinates are included. For example, for the planar 2 joint ($n = 2$) model, only rows 1 and 19, which correspond to C0-C1 pitch and C6-C7 pitch are included, resulting in a $(2 \times m)$ moment arm matrix.

Two equivalent methods for grouping muscles are described here. First, we construct an $(m \times g)$ binary muscle grouping matrix, S_m. Here, m is the number of muscles elements in the full model and g is the number of muscle groups. If muscle i is to be in group j, a 1

is placed in location $S_{m(i,j)}$, while all other elements of row i remain 0. In the case where all muscle elements are controlled independently, $S_m = I_{mxm}$. This simple example of a muscle grouping matrix would group 6 muscles (rows) into 3 muscle groups (columns), group 1 containing muscles 2 and 3, group 2 containing muscles 1 and 4, and group 3 containing muscle 5.

$$S_m = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We utilize the muscle grouping matrix in our linear mapping between muscle activations and joint torques.

$$\tau^{n\times 1} = R^{n\times m} \ \ \mathrm{diag}(F)^{m\times m} \ S_m^{m\times g} \ \alpha^{g\times 1}$$

Where the superscripts represent the dimensions of the matrix or vector.

The muscle activations, α, are now reduced to the number of muscle groups g. All muscles within a group are constrained to share the same activation. However, the strength vector (in diagonal form here) and the moment arm matrix are still defined in terms of all $m = 96$ individual muscle elements. We can scale the strength of a muscle group by scaling all elements in that group individually. However, if a lower dimensional representation of the model is desired, strength is defined for each muscle group.

To obtain a representative simplified moment arm matrix $R_g \in \mathbb{R}^{n \times g}$ and strength vector $F_g \in \mathbb{R}^{g \times 1}$, we can use a weighted average approach.

First, we compute the contribution of each group to the joint torques, $\tau_g \in \mathbb{R}^{n \times g}$, using the S_m matrix.

$$\tau_g = R \ \text{diag}(F) \ S_m$$

Each element of the new strength vector, F_g, is computed as the norm of the strengths of the muscles in that group.

$$F_g(i) = \|F \odot S_m(:, i)\|$$

Finally, the equivalent moment arm matrix for the simplified model is computed for each muscle group in the corresponding column, so that the i^{th} column of R_g is constructed by dividing the joint torque contribution of the group, $\tau_g(:, i)$ by the group strength $F_g(i)$.

$$R_g(:, i) = \frac{\tau_g(:, i)}{F_g(i)}$$

The resulting grouped moment arm matrix, R_g, and muscle strength vector, F_g, represent simplified models that have the same joint torque producing capabilities as the $m = 96$ muscle model when individual muscle elements are constrained to be activated together. This approach is useful when it is desirable to reduce the number of parameters.

4.2.2. Neck Strength Computation

We define the neck strength of a model as the set of maximum wrenches the model can produce from the head in 25 directions. This concept is similar to a feasible force

set (FFS), which is a method for representing the set of all feasible forces that can be produced at an end effector using computational geometry to map polytopes from muscle activation space to end effector task space [106]. FFS has been used for lower dimensional systems, such as a finger [104] and a planar leg [62], but our 96 dimensional muscle space and non-invertible Jacobian make the FFS approach unrealistic. Instead, we use linear programming to compute the maximum end effector force that can be produced by the model when constrained to a specific direction.

To insure that the task is isometric, the joint torques produced by the wrench on the head must be balanced by the joint torques produced by the muscle activity.

$$\sum \tau = J^T w - R \text{ diag}(F) \ \alpha = 0$$

If we break down the wrench into a magnitude $w = \|\mathbf{w}\|$ and unit vector $u = \mathbf{w}/\|\mathbf{w}\|$, we can express our constraint in matrix form.

$$\begin{bmatrix} R \text{ diag}(F) & -J^T u \end{bmatrix} \begin{bmatrix} \alpha \\ w \end{bmatrix} = 0$$

The maximum wrench magnitude w^* that can be produced in direction u by a model with muscle strengths F is computed by maximizing the wrench magnitude w over (α, w). The equality constraint is the matrix form of the joint torque balance equations. Additional constraints include that muscle activations α are between 0 and 1, and the wrench magnitude w must be positive.

$$FFS(u_i, F) = (\alpha, w)_i^* \ \forall \ i = 1, ...25$$

$$
\begin{aligned}
\max_{\alpha, w} \quad & w \\
\text{s.t.} \quad & \begin{bmatrix} R\ \mathrm{diag}(F) & -J^T u \end{bmatrix} \begin{bmatrix} \alpha \\ w \end{bmatrix} = 0, \\
& 0 \leq \alpha \leq 1, \\
& 0 \leq w
\end{aligned} \tag{4.1}
$$

4.2.3. Test Directions

Twenty-five locations on the surface of the skull were selected as test points. The points were selected based on the following criteria: (1) directions where a force could reasonably be produced (the crown of the skull would not be a reasonable choice), (2) an attempt was made to provide good coverage of all possible directions, and (3) where a clinician could reasonably provide resistance to head movement. The test directions are force vectors $F = (F_x, F_y, F_z)$ oriented approximately normal to the skull surface at each test point as shown in Figure 4.1. A 3×3 grid, shown in Figure 4.2 displaying the test directions in slices of the sagittal, transverse, and coronal planes was used to visualize the multi-directional neck strength.

4.2.3.1. Coordinate Transformations. The neck strength computation calls for the direction vector u_i to be a unit wrench vector in the skull frame. The skull frame is located at the base of the skull, at the C0-C1 joint, and oriented with the x-axis directed anterior, y-axis superior, and z-axis to the right.

To transform a force $Fa = (Fa_x, Fa_y, Fa_z)$ applied at the location of a test point $p = (p_x, p_y, p_z)$ to a 6-D wrench in the skull frame (w_{skull}), we use the adjoint of the spatial transform adT_p.

$$w_{skull} = adT_p * \ F_a$$

$$adT = \begin{bmatrix} R & 0 \\ \hat{p}\ R & R \end{bmatrix} \quad \text{where} \quad \hat{p} = \begin{bmatrix} 0 & -p_z & p_y \\ p_z & 0 & -p_x \\ -p_y & p_x & 0 \end{bmatrix} \quad \text{and} \quad R = I_{3x3}$$

The unit wrench vector is then $u = \mathbf{w}_{skull} / \left\| textbf w_{skull} \right\|$.

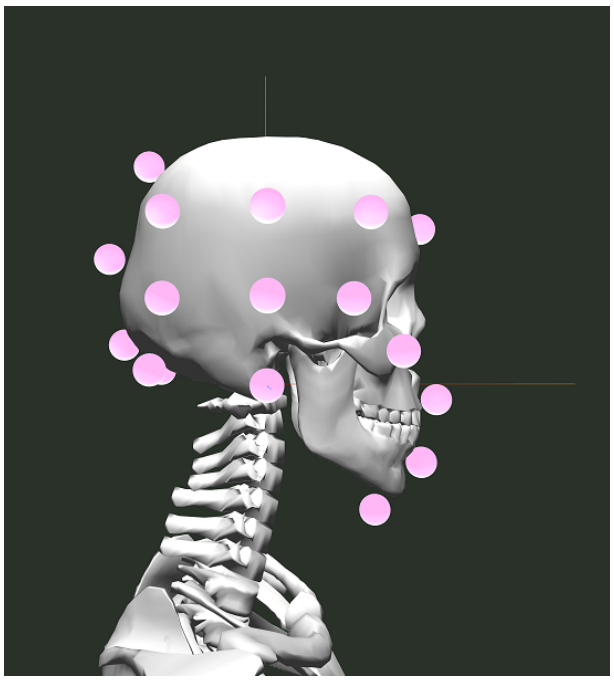


Figure 4.1. The markers represent test locations where a force is applied perpendicular to the surface of the skull. The location and orientation of the skull frame is indicated by the axes: x-axis (red) points anterior, y-axis (green) points superior, and the z-axis (blue) points to the right.

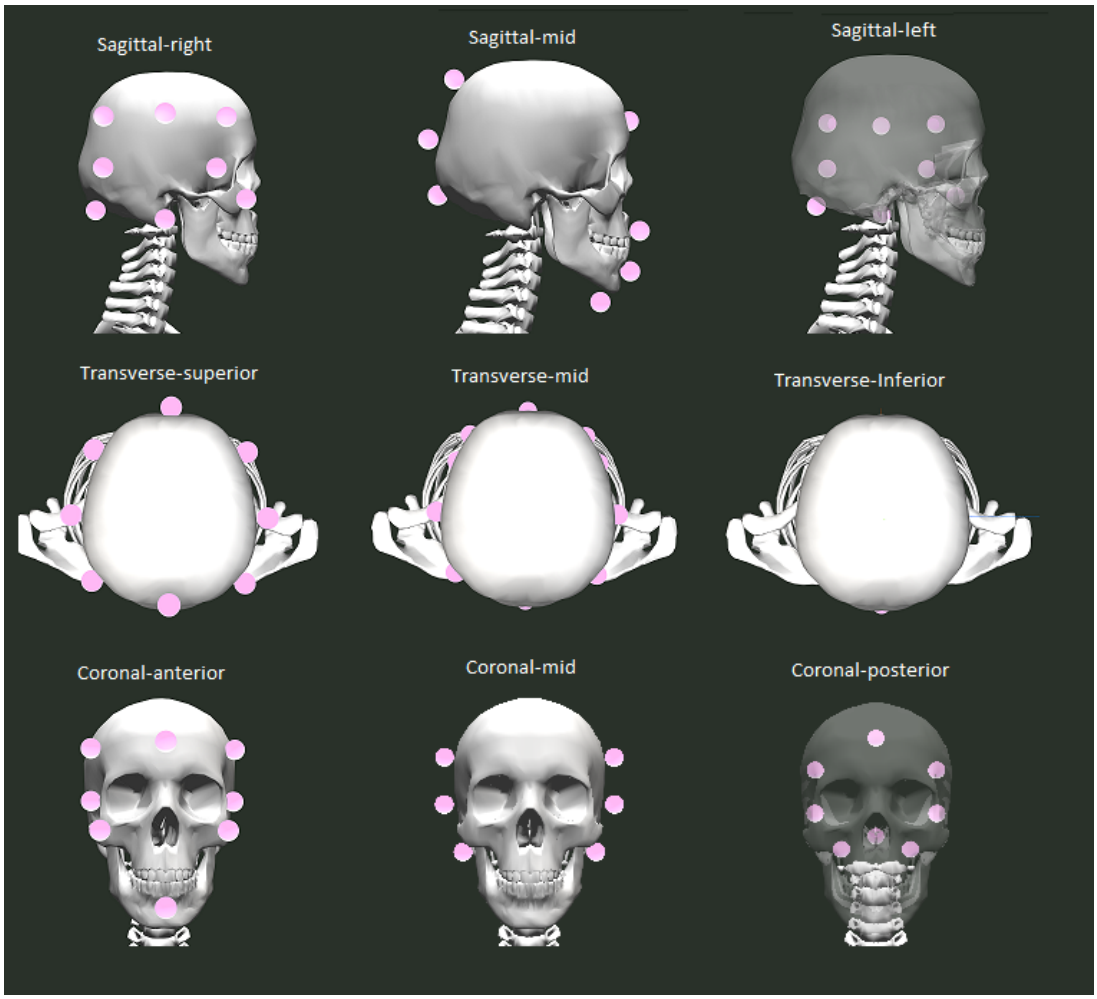


Figure 4.2. A 3x3 grid showing the test directions in multiple planes used to visualize the neck strength and muscle activation patterns. Note that some test positions appear in multiple planes. The top row shows the right-, mid-, and left- sagittal plane test directions. The middle row shows the superior-, mid-, and inferior- transverse plane test directions. The bottom row shows the anterior-, mid-, and posterior- coronal plane test directions.

4.2.4. Muscle Strength

The default muscle strength (maximum isometric muscle force, F) for all the muscles was obtained from the OpenSim Neck Model 1.6 [108] with the addition of the hyoid muscles from [78]. The muscle parameters from the Vasavada neck model [108] were

derived primarily from anatomical studies by Kamibayashi and Richmond [56]. Note that this study was interested only in isometric strength in a neutral posture, therefore other muscle parameters, such as optimal fascicle length and pennation angle were ignored. Muscle strength (peak force) was computed by scaling physiological cross-sectional area (PCSA) by a muscle specific tension coefficient of $35N/cm^2$. While the magnitude of this scaling factor is debated in the literature, the relative strength between muscles remains unaffected by this choice. Noting a relative weakness in cervical flexion strength in the model, previous studies chose to scale the flexors and extensors independently to better match experimental data [78]. This study used the default strength values based on anatomical studies in order to preserve the relative strength between individual muscles based on PCSA.

4.3. Results

4.3.1. Deep extensor muscle activation

The directions in which the right deep extensor muscles are active at its default strength are shown in Figure 4.3. The muscle group is most active in the test directions oriented posterior in the mid-sagittal plane (pure extension), but it is also active in posteriorly oriented directions that combine extension, side-bending, and rotation. In addition, it acts bilaterally, as shown in the coronal-posterior frame of Figure 4.3.

4.3.2. Effect of Deep Cervical Extensor Strength on Neck Strength

The effect of deep cervical extensor strength on multi-directional neck strength is further analyzed using the spatial 3 joint ($n = 9$) model with the logical muscle grouping

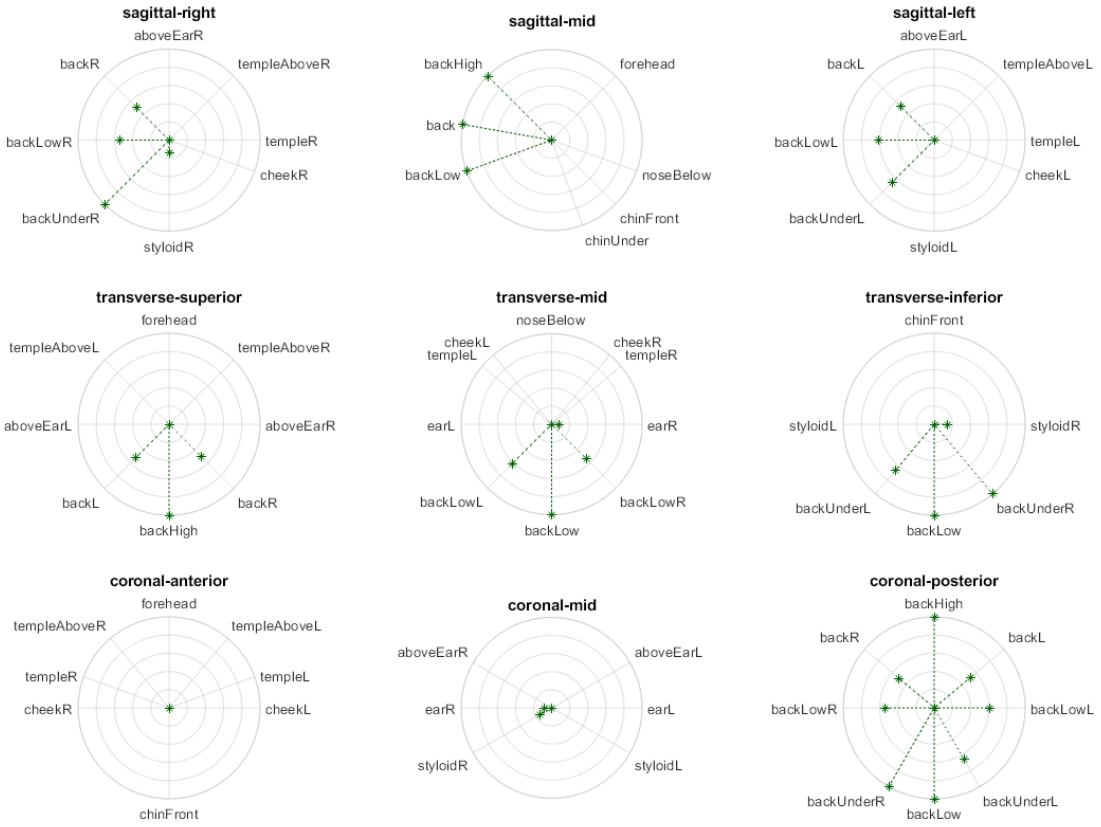


Figure 4.3. Muscle activation of the right deep extensor muscles during simulated maximum isometric strength measurements using the spatial 3 joint (n = 9 DOF) with logical muscle grouping (m = 24) model. The magnitude in the radial axis corresponds to muscle activation between 0 and 1.

($m = 24$). This model has a level of both kinematic and muscle redundancy that creates behaviors that mimic the full model for this isometric strength task, but with muscles grouped in a way that is logical and interpretable.

Neck strength increases as the deep cervical extensor strength increased for test directions with a posterior component, with the greatest effect in the mid sagittal posterior directions. Figure 4.4 shows neck strength in the sagittal plane with varying deep extensor

muscle strength. The strength of the deep extensors did not influence neck strength in directions anterior to the mid coronal line.

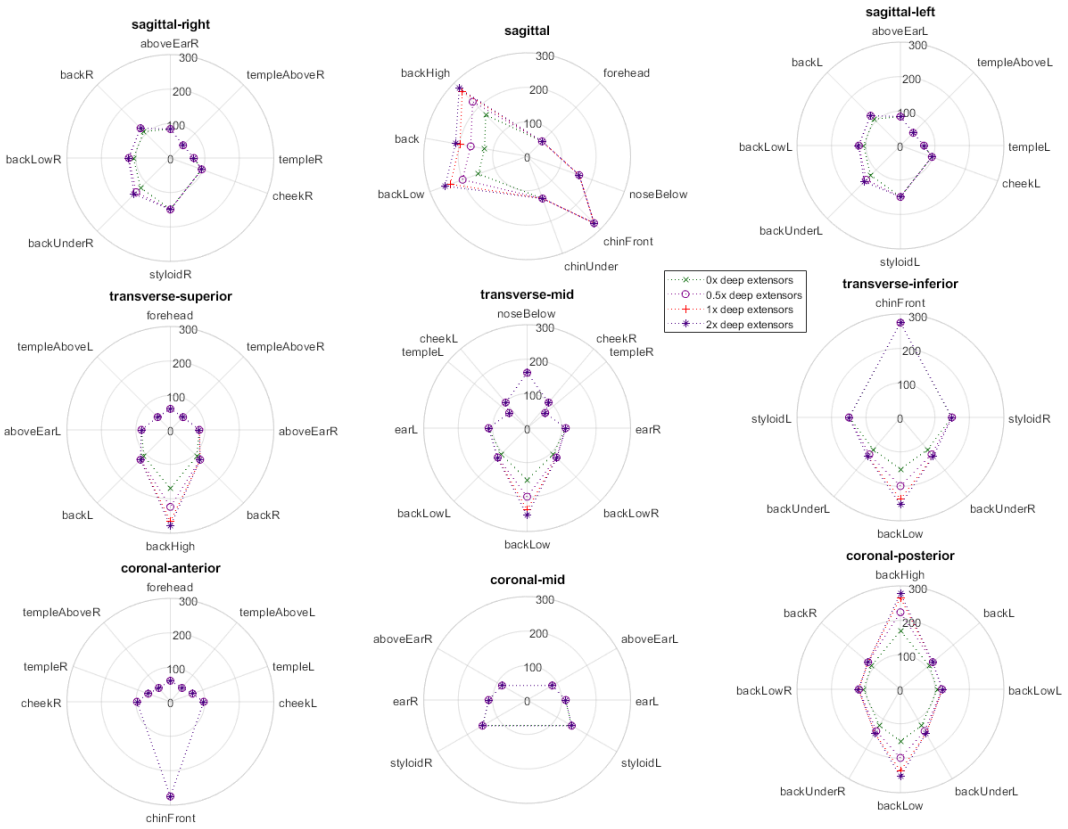


Figure 4.4. Neck strength with varying strength of the bilateral deep extensor muscles (0x, 0.5x, 1x, and 2x) using the spatial 3 joint (n = 9 DOF) and logical muscle grouping (m = 24 muscles) model. The magnitude in the radial axis represents the maximum force the model can produce in the indicated direction in Newtons.

4.3.3. Effect of Deep Cervical Extensor Strength on Muscle Activation Pattern

The relationship between deep cervical extensor strength and the activation pattern of all muscles can be visualized in Figure 4.5. The 24 muscle by 25 direction grid gives a heat map representation of the slope of the regression line for bilateral deep extensor

muscle strength vs. muscle activation level. Muscles that increased their activation with a decrease in deep extensor strength are represented in blue color (negative slope). It is apparent that activation of the semispinalis capitis and longus capitis was increased in response to deep extensor muscle weakness, while suboccipitals and longus colli activation level increased when the deep extensors were stronger. Directions in which the deep extensors were not normally active did not produce altered activation patterns with changes in strength.

Another way to visualize how individual muscle activation responds to deep extensor weakness relies on polar plots in a single plane, where each axis contains the activation levels for an individual muscle. Figure 4.6 shows the sagittal-mid, sagittal-right, and coronal-post planes, respectively. The mid-sagittal view (Figure 4.6a) reveals that, in the absence of the deep extensors, the deep capital extensors (suboccipitals) were no longer active in the posterior (extension) directions, while the deep capital flexors (longus capitis) and superficial capital extensors (splenius capitis) became more active. The impact of deep extensor weakness becomes more complex in the directions that involve axial rotation and side-bending components. The coronal-posterior plane views (Figures 4.6b) reveals that deep extensor weakness led to a reduction in SCM and hyoid activation in some directions. The other superficial capital extensor (semispinalis capitis) became more active in the absence of the deep extensors for directions that were not mid-sagittal.

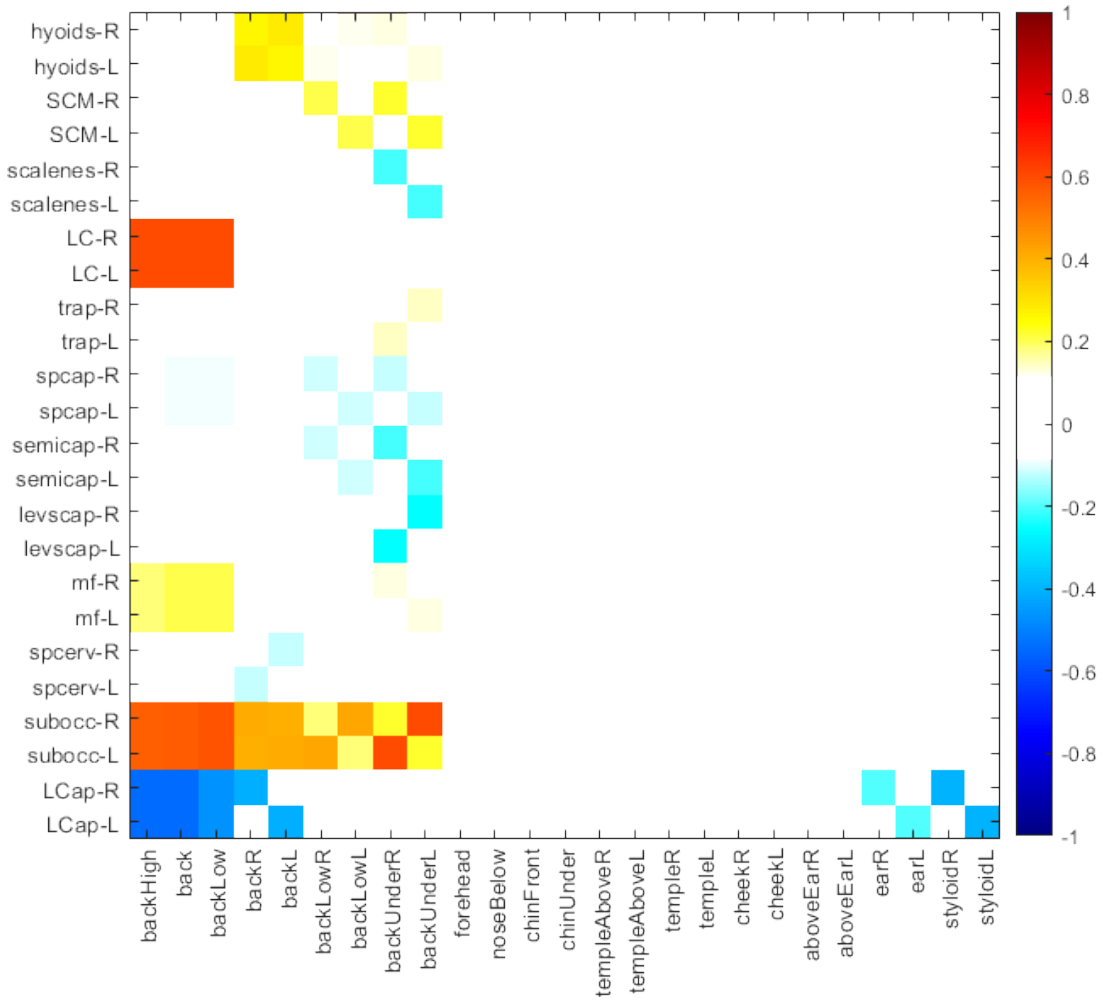
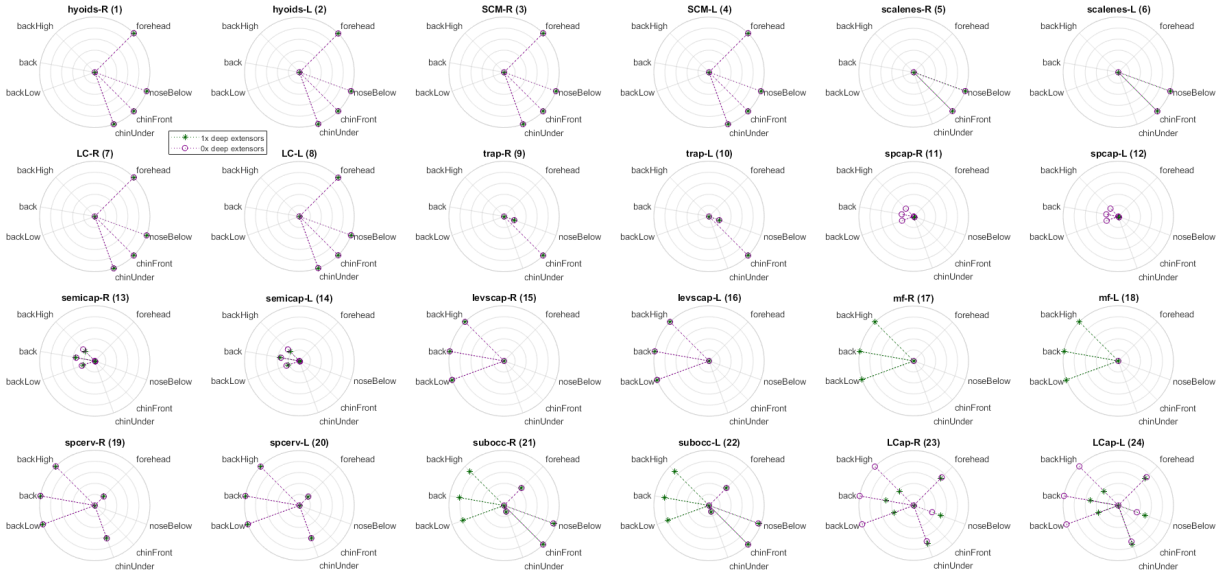


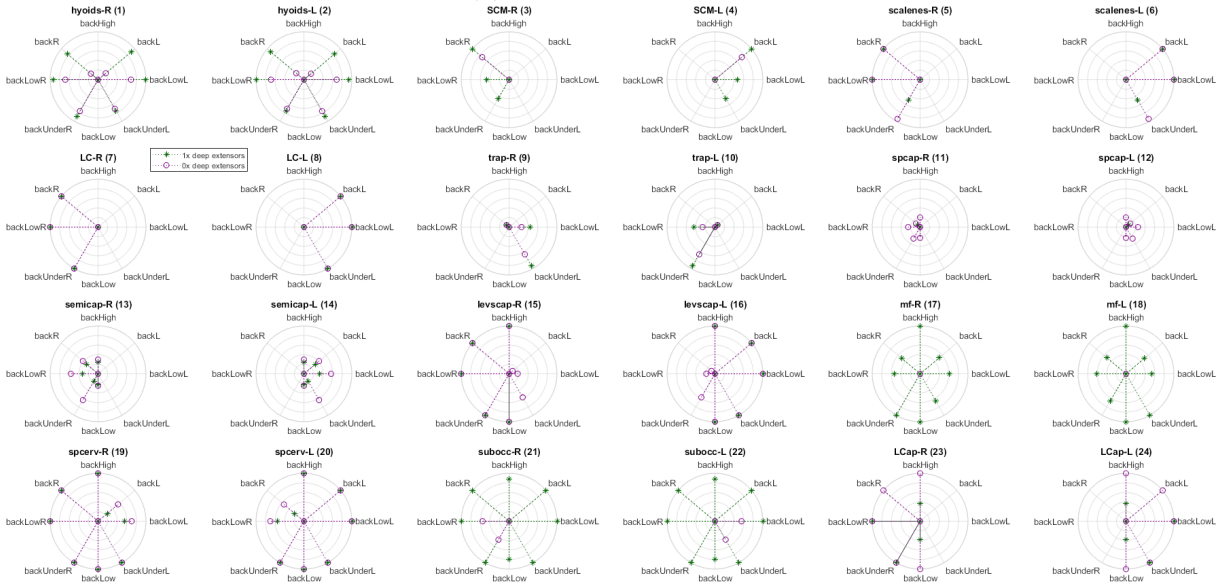
Figure 4.5. Slope of the linear regression of the activation of muscle on the y-axis vs. bilateral deep extensor strength. The model used in this simulation in the 3 joint (n = 9 DOF) with the logical muscle groupings (24 muscle groups). The color represents the slope, with warmer colors depicting more positive slope, and cooler colors representing a negative slope.

4.3.4. Effect of Model Complexity on Neck Strength and the Role of the Deep Extensor Muscles

Figure 4.7 shows that increasing the kinematic degrees of freedom while maintaining the same muscle set generally decreases neck strength in all directions. Model variants with



(a) Sagittal-Mid Plane



(b) Coronal-Posterior Plane

Figure 4.6. Muscle activation [0 1] for individual muscles in the logical muscle group (m=24) for the full strength model (1x deep extensors, green) and with the deep extensors (mf-R and mf-L) removed from the model (0x deep extensors, purple). Note the particular muscle groups that change their activation with the loss of the deep extensors: suboccipitals, SCM, and hyoids reveal decreased activation while the longus capitis (LCap), splenius capitis (spcap), semispinalis capitis (semicap), levator scapulae (levscap), and splenius cervicis (spcerv) increase activation in some directions.

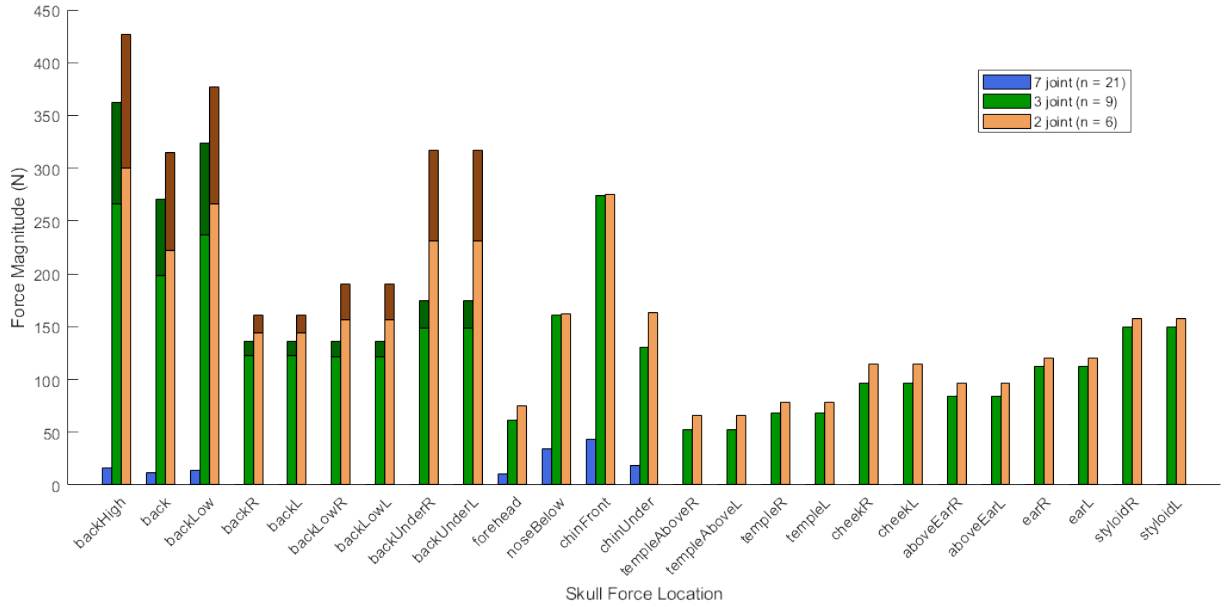
a high number of degrees of freedom and low number of muscles were unable to produce forces in some test directions.

Increasing the number of muscle groups while maintaining the same kinematic degrees of freedom increased neck strength in most directions. This relationship is seen in Figure 4.8. However, the choice of which muscles to group plays a role, as seen when comparing the logical (m = 24) and sag (m = 24) cases, where both have 24 independent muscles. The logical grouping variant produced greater neck strength in most directions compared to the sagittal variant. There does appear to be a ceiling effect, as seen when comparing muscle grouping variants applied to the 2 and 3 joint planar models.

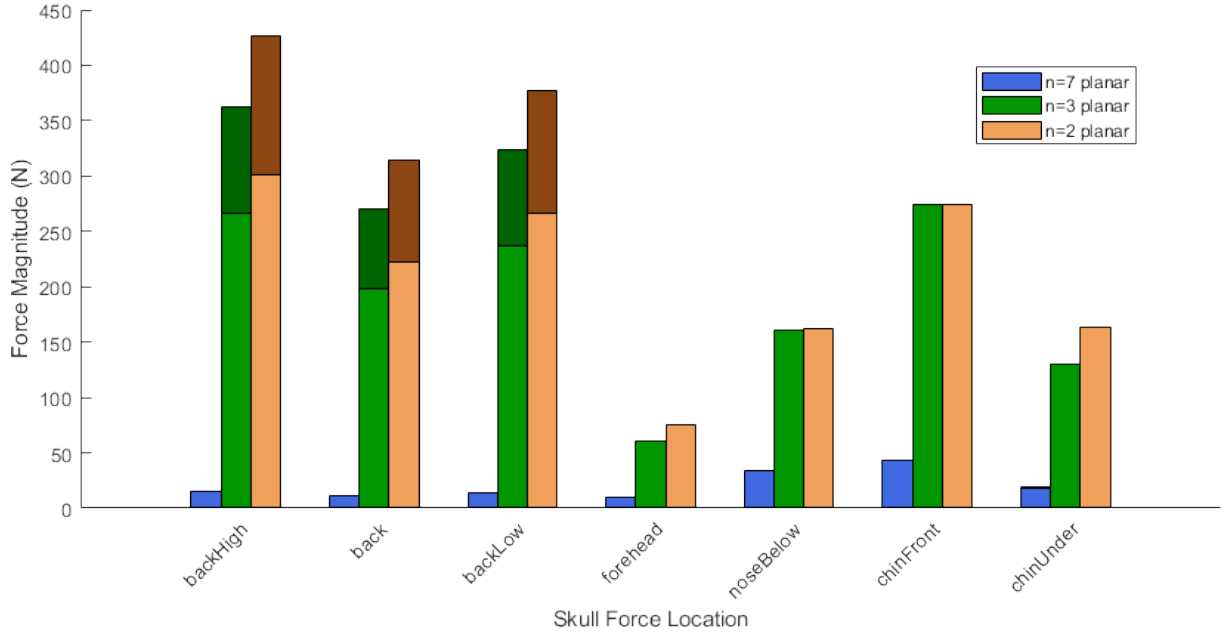
The difference in neck strength with and without the deep cervical extensor muscles is shown in both figures 4.7 and 4.8. The shaded area is the contribution of the deep extensors. To be clear, the deep extensors could not produce this force alone, but instead work in concert with other muscles.

4.4. Discussion

This study utilized computational modeling of the neck musculoskeletal system to examine the role of the deep cervical extensor muscles in multi-directional neck strength. The spatial 3 joint (n=9 DOF) musculoskeletal model with muscles grouped anatomically into 12 pairs (m = 24 muscles groups) was used to simulate neck strength in 25 directions with varying deep extensor muscle strength. As would be expected based on the line of action, the strength of the deep extensor muscles impacted neck strength in the posterior directions (extension), with the greatest contribution in the mid-sagittal plane. Even in the case of complete loss of the bilateral deep extensor muscles, force could be generated in

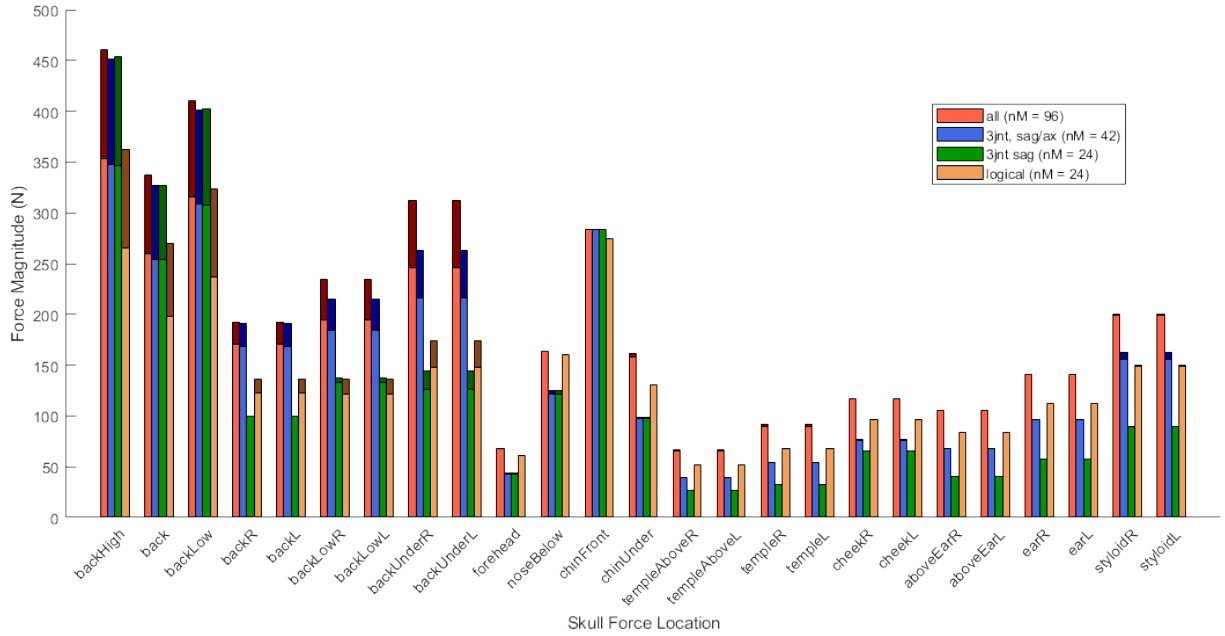


(a) Spatial model

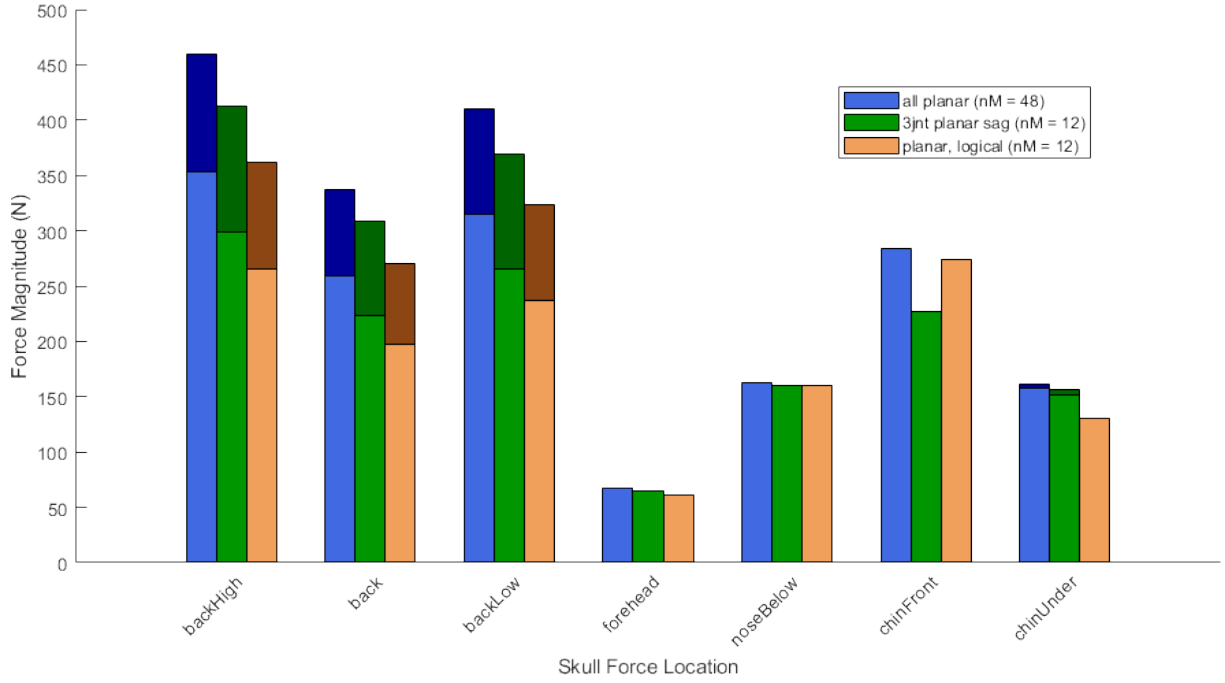


(b) Planar Model

Figure 4.7. Effect of Kinematic Complexity on neck strength and the role of the deep cervical extensor muscles. The logical muscle grouping, where muscles are grouped by anatomical name, is used for each model. The deeper shade portion of some bars represents the additional contribution of the deep extensor muscles to the force output.



(a) Spatial model



(b) Planar Model

Figure 4.8. Effect of muscle complexity on neck strength and the role of the deep extensors using the 3 joint models (planar: n=3, spatial: n=9). The darker shade portion of some bars represents the additional contribution of the deep extensor muscles to the force output.

all 25 test directions. Two primary synergistic strategies for neck extension were observed in response to varying the strength of the bilateral deep extensor muscles. As the deep extensor strength increased, the activation of the suboccipitals also increased. In the presence of a weak deep extensors, activation of the superficial capital extensors (splenius and semispinalis capitis) and deep capital flexors (longus capitis) was increased.

Model complexity had an influence on both the simulated neck strength and the role of the deep extensor muscles in generating endpoint forces. Generally speaking, as kinematic degrees of freedom increase, neck strength decreases. As the number of muscle groups increases, neck strength also increases. This relationship is not linear in all directions, affecting some directions more than others. The ratio of kinematic DOF to number of independent muscle groups must be low enough to allow sufficient control of all degrees of freedom. The role of the deep extensors expands in model variants with more independent muscle groups, where the right and left deep extensors activate independently from each other and some elements become active in directions that do not involve a posteriorly directed force.

The functional role of the deep extensor muscles has been described as a local stabilizer, acting to stiffen and stabilize the vertebral segments independent of direction [8, 16]. Studies that agree with the local stabilizer concept tend be open chain, low load, with an anticipatory stabilization task [31, 46]. These are arguably more realistic for everyday function and may be more representative of normal behavior. Studies that have investigated spatial tuning of muscles with isometric resistance have generally found specific preferred directions more representative of phasic muscles, even for the deep cervical extensors [11, 110]. The task simulated in this study is a maximal isometric force in a

specific direction, with the head and torso fixed to ground. For diagnostic and muscle strengthening purposes, the role of the deep extensors for the isometric, statically supported task is appropriate. However, it is important to keep in mind that as the task changes, the neural control of the muscles will also change.

The simulations revealed two synergistic strategies for producing forces in posterior directions (extension), as shown in Figures 4.5 and 4.6. The first strategy, which we will call the "local" strategy, pairs the deep cervical extensors and suboccipital muscles to produce independent torques about the lower and upper cervical spines. The second strategy combines the actions of the superficial extensors (splenius capitis and semispinalis capitis) with the deep capital flexors (longus capitis). The longus capitis is needed to counteract the excessive extension torque that the superficial extensors apply to the upper cervical spine. This pattern is consistent with a common observation of individuals with chronic neck pain, where the more superficial muscles exhibit high tone. Interestingly, the deep cervical flexors (longus capitis and colli) have been shown to exhibit dysfunctional feedforward activation in individuals with chronic WAD [30]. These muscles appear to be particular important when the deep extensors are weak, adding to the complex clinical picture.

The musculoskeletal model used in our simulations had 3 joints (9 DOF) and 12 muscle pairs (24 independent muscle groups). This simplification from 96 muscle elements to 24 allows a more manageable number of parameters to interpret. As shown in our analysis, kinematic and muscle complexity in model design affect both multi-directional neck strength and the role of the individual muscle groups. Note that design decisions pertaining to muscle grouping make assumptions regarding the neural control of muscles,

which is poorly understood for the cervical spine [11]. Additionally, recruitment patterns are known to change in response to neck pain and WAD, therefore, assumptions based on "healthy" control patterns may not apply to these conditions.

There are several limitations and assumptions that have an impact on the results and interpretations from this study. First, an important assumption in the formulation of the neck strength computation (Equation 4.1) is that the model is producing a maximum effort in each specified direction. The only constraints are that joint torque equilibrium be maintained (no intervertebral movement) and muscles can only pull (not push). In reality, humans may have additional objectives or constraints besides force balance, such as avoiding pain or maintaining a margin of stability, that would affect neck strength and activation patterns. Second, the muscle strengths are derived from the generic OpenSim neck model [108, 78], which combined several anatomical studies to obtain PCSA and geometrical data for the neck muscles. We chose to use the unscaled strength values to maintain relative strength magnitudes between muscles rather than scaling arbitrary groups to match experimental neck strength values. While we investigated the effect of varying the strength of the deep extensors, we did not vary the strengths the other muscles. Third, the model design choice of muscle groupings has a significant effect on both neck strength and muscle activation. While we believe that the muscle grouping used in this model provides appropriate redundancy and constraint for the task, other muscle grouping choices based on innervation, experimental EMG studies, or muscle action could be easily defended. Finally, the computational models analyzed in this study ignore passive components such as joint stiffness or passive muscle forces. However, joint stiffness is very low in a neutral spine posture according to the neutral zone hypothesis [80]. The

passive and active length tension curve for the neck muscles was also ignored because our analysis was restricted to isometric task in a neutral posture. In reality, no task is truly static and these components could play a small role.

The computational approach in this study allows us to examine the role of the deep cervical extensor muscles in multi-directional neck strength and determine how weakness of the group would affect neck strength and muscle activation patterns. The test directions were selected relative to specific landmarks on the skull such that a clinician could replicate the test by applying a normal force at the location of the marker (shown in Figure 4.1) to target specific muscles. While others have characterized the directional preference and spatial tuning of individual neck muscles using EMG [110, 11, 33, 58], this study is the first to demonstrate the effect of an individual muscle weakness on neck strength and activation patterns. Such results could support the delivery of tailored exercise interventions to target specific neck muscles.

The increased activation of the superficial capital extensors (splenius and semispinalis capitis) and deep capital flexors (longus capitis) in response to deep cervical extensor weakness may be clinically relevant and merits further investigation. The fact that deep cervical flexor dysfunction has been reported in WAD [30] makes this particular compensatory pattern more interesting and concerning. The general idea of deep spinal stabilizers becoming weak or less active while the more superficial, global muscles are over-active in pain conditions is not new [16], but a biomechanical explanation following from weakness of a particular muscle group had not previously been reported. The longus capitis is a particularly difficult muscle to access for EMG measurement, but verification of dominance

of this muscle pattern in WAD compared to controls would provide further evidence of deep cervical extensor muscle dysfunction or weakness.

CHAPTER 5

Model-Based Bayesian Inference for Individual Muscle Strength Estimation

In this chapter, we present a novel framework for individual muscle strength estimation that involves (1) collecting neck strength in 25 test directions and surface EMG profiles of a subset of neck muscles, and (2) utilizing a model-based Bayesian estimator to infer specific individual muscle strengths from the collected data. The feasibility of the approach is demonstrated through experimental data collection with $n = 5$ healthy participants. We show that the algorithm can identify weakness of the deep cervical extensor muscles when provided a rich enough set of measurement data.

5.1. Introduction

Imaging evidence suggests that individuals with chronic whiplash-associated disorders (WAD) develop muscle fat infiltration (MFI) in the deep cervical extensor muscles in the months after their whiplash injuries [28, 29]. An obvious assumption is that increased MFI leads to a decrease in functional muscle strength, but there are significant barriers to testing this hypothesis. Neck strength is grossly reduced in WAD [82, 86, 76] in the cardinal directions (flexion/extension, lateral flexion, and rotation), but the strength of individual muscles has not been assessed.

The problem of identifying individual muscle strength is difficult for several reasons. The core issue is that muscle forces cannot be measured directly. In the case of the neck,

wrench (force and torque) measurements are taken from the end effector of the neck, the head. Clinically, this involves the patient resisting a force applied to a location on the head by the clinician, and the clinician grading the strength on a 0-5 scale or using a linear force sensor [44]. The human neck is a complex musculoskeletal structure with 7 intervertebral joints and over 20 pairs of named muscles, most having multiple attachments and crossing multiple joints. The kinematic and muscle redundancy sets up an under-constrained and ill-conditioned inverse problem. An additional complication is that we are interested in the muscle strength, the maximum force that a muscle can produce, not just the contribution of the muscle for a particular end effector measurement.

Estimates of muscle strength used in musculoskeletal modeling are typically derived from physiological cross sectional area (PCSA) measurements from anatomic cadaver studies, which is considered to be proportional to the maximum isometric force that the muscle can produce [108, 17]. This method produces a generic set of muscle strengths based on geometry and muscle size that can be scaled uniformly so that the model produces outputs similar to experimental measurements [78, 110]. The uniform scaling method does not address the individual muscle strengths in a clinical population, where the relative strengths of muscles may change significantly. Subject specific measures of PCSA have been attempted using Magnetic Resonance Imaging (MRI), but this method is time consuming and costly, neck muscles are particularly difficult to segment, and it may not differentiate between functional contractile tissue and injured tissue.

Efforts to determine subject-specific model parameters, such as individual muscle strength, have been made in the hand and wrist [41, 59, 49, 15]. These studies utilize least squares optimization methods to minimize the difference between model predicted

and measured data sets. A bounded, non-linear least squares approach similar to these studies could be a viable approach to the strength estimation problem in the neck, the data available is often noisy and scarce, setting up an ill-posed optimization problem that is known to give poor results for these standard methods [5, 112]. Statistical Bayesian approaches are well-suited for dealing with non-linear models and naturally include uncertainties such as measurement and modeling errors. In fact, Approximate Bayesian Computation (ABC) methods are similar to least squares optimizations in some ways. In a least squares optimization, the sum of the squared error between the simulated and observed data is minimized in an attempt to find a point estimate of the parameters. Like a least squares optimization, ABC relies on a metric to compare the simulated data to observed data but with a different goal: to obtain an estimate of the posterior distributions for the unknown parameters [103]. This probabilistic framework provides uncertainties and correlations between parameters, rather than just point estimates, which is beneficial in a system with the potential to have many solutions [19].

The over-arching goal of this work is to indirectly measure the strength of individual neck muscles using non-invasive methods for application in the chronic WAD population. The novel method developed to fulfill this goal involves (1) collecting neck strength in multiple directions and surface EMG profiles of a subset of neck muscles in vivo, and (2) utilizing a simulation-based Bayesian estimator based on Approximate Bayesian Computation (ABC) with rejection and importance sampling to infer subject specific individual muscle strength from the collected data. The primary contribution of this work is a model-based algorithm that provides a framework for parameter estimation in a highly

redundant and partially observable musculoskeletal system. The ability to identify individual muscle strength deficits would enable clinicians and researchers to understand the biomechanical implications of muscle changes seen in imaging in WAD and to guide diagnosis and treatment.

5.2. Methods

From a high level view, this framework involves collecting a rich set of experimental data that contains information about the parameters of interest and utilizing a model-based ABC algorithm to estimate the parameter values. Because the parameter of interest is the maximum isometric force of individual muscles, the models presented here compute the maximum task space wrench (forces and moments measured at the head) given a set of muscle strengths and a target direction. The experimental protocol to create experimental data sets, likewise, requires that the subject produce a maximum effort in a set of target directions. The algorithm attempts to find individual muscle strengths that would produce simulated wrench measurements matching the experimental data set.

5.2.1. Model

The geometrical and muscle properties of the musculoskeletal model of the neck is derived from the generic OpenSim Neck Model [108] with several modifications, as described in Chapter 4. The cervical spine is configured in a neutral, upright posture and is constrained to maintain static equilibrium through all joints while isometrically resisting forces applied to the head. The moment arm matrix, $R = R(q)$, and skull frame Jacobian, $J = J(q)$, for the neutral neck posture with joint angles, q, are obtained through the OpenSim Matlab

API. All further computations are made in Matlab (Mathworks, Inc., Natick, MA). The following description of the model equations applies to any of the model variants presented in the applications section of the chapter.

To insure that the task is isometric, the joint torques produced by the wrench measured in the skull frame, τ_w must be balanced by the joint torques produced by the muscle activity τ_m. Muscle force is the product of a diagonal matrix of the muscle strength, $f \in \mathbb{R}^m$, and a vector of muscle activations $\alpha \in \mathbb{R}^m$. Each muscle activation is a scalar that can vary between 0 to 1.

$$\begin{aligned} \sum \tau &= \tau_w + \tau_m = 0 \\ &= J^T \mathbf{w} - R \ \mathrm{diag}(f) \ \alpha = 0 \end{aligned} \tag{5.1}$$

Taking a closer look at Equation 5.1, there are two known constants, $J \in \mathbb{R}^{d \times n}$ and $R \in \mathbb{R}^{n \times m}$, the measured quantity $\mathbf{w} \in \mathbb{R}^d$, a independent variable $\alpha \in \mathbb{R}^m$, and the unknown model parameter $f \in \mathbb{R}^m$. The quantities d, n, and m, represent the dimensions of the task (wrench), joint (torque), and muscle (force) spaces respectively. Both the forward and inverse solutions (from muscle activations to wrench or wrench to muscle activations) are ill-posed because both the Jacobian, J, and moment arm matrix R, are not invertible. The inverse problem is under-constrained, meaning that for a given sub-maximal wrench there may be many solutions for α.

We avoid the need to invert the Jacobian or moment arm matrix by setting up the joint torque constraint equations in the torque space. In order to put the joint torque

constraint in matrix form, the wrench is broken down into a scalar magnitude, $w = \|\mathbf{w}\|$, and unit vector, $u = \mathbf{w}/\|\mathbf{w}\|$.

Algorithm 2 Maximum Wrench Model

function M(f,u,[α_{lb} α_{ub}])

$$\begin{aligned} \max_{\alpha,\, w} \quad & w \\ \text{s.t.} \quad & \begin{bmatrix} R\ \mathrm{diag}(f) & -J^T u \end{bmatrix} \begin{bmatrix} \alpha \\ w \end{bmatrix} = 0, \\ & \alpha_{lb} \leq \alpha \leq \alpha_{ub} \end{aligned} \tag{5.2}$$

return $\{w, \alpha\}$
end function

The problem is now set up as a parameter estimation problem, with the muscle strengths f as the unknown parameter of interest, and the maximum wrench magnitude w as the measurement variable. The model equation $M(f, u, [\alpha_{lb}\ \alpha_{ub}]) = \{w, \alpha\}$ shown in Algorithm 2 provides a mapping between muscle strength f and maximum wrench magnitude w in direction u.

5.2.2. Algorithm

The purpose of the algorithm is to estimate the m individual neck muscle strengths, $f \in \mathbb{R}^m$, given the experimental neck strength measurements and subset of m' muscle activations measured by surface electromyography (EMG), $\{w_{meas} \in \mathbb{R}^{N_U}, \alpha_{meas} \in \mathbb{R}^{N_U \times m'}\}$, where N_U is the number of test directions. In constructing the algorithmic pipeline, inspiration is taken from Approximate Bayesian Computation (ABC) algorithms [70, 7, 66]. ABC algorithms are used instead of traditional Bayesian Inference when the likelihood function is intractable [101]. They approximate the posterior distribution by producing simulated data with a model and updating the estimate based on the discrepancy between

simulated and experimental data. The algorithm leverages rejection- and importance-based sampling, in conjunction with a muscle specific weighting function to generate muscle strength estimates. Clinical domain knowledge is encoded in the prior distribution and muscle model specification.

The algorithm for neck muscle strength estimation is described in pseudocode in Algorithm 3. A visual representation of the full algorithm is shown in Figure 5.1. Each component of the algorithm and associated functions are described in detail in the following subsections. The technical algorithm, Algorithm 4, is presented at the end of the section.

5.2.2.1. Constructing an Informative Prior. In Bayesian statistics, the prior is a probability distribution that represents beliefs about the parameter before data is available. Particularly for a system with significant redundancy and potential for multiple solutions, thoughtful selection of a prior can improve performance.

A baseline muscle strength estimate, $f_{base} \in \mathbb{R}^{96}$, is extracted from the maximum isometric force parameter in the OpenSim Vasavada neck model [108] and more recent model updates [78, 110]. This and other model parameters are reported in Appendix 7.2 [add ref to specific appendix section]. Muscle parameters, including maximum isometric muscle force, in musculoskeletal models are derived from published anatomical cadaver studies that report physiological cross-sectional area (PCSA) of the cervical spine muscles or cross-section area from imaging studies.

To create a computationally tractable and interpretable model, the neck muscle model was simplified from its original 96 muscle elements to m muscle groups. The detailed procedure for computing the muscle strength and moment arm for each grouped muscle

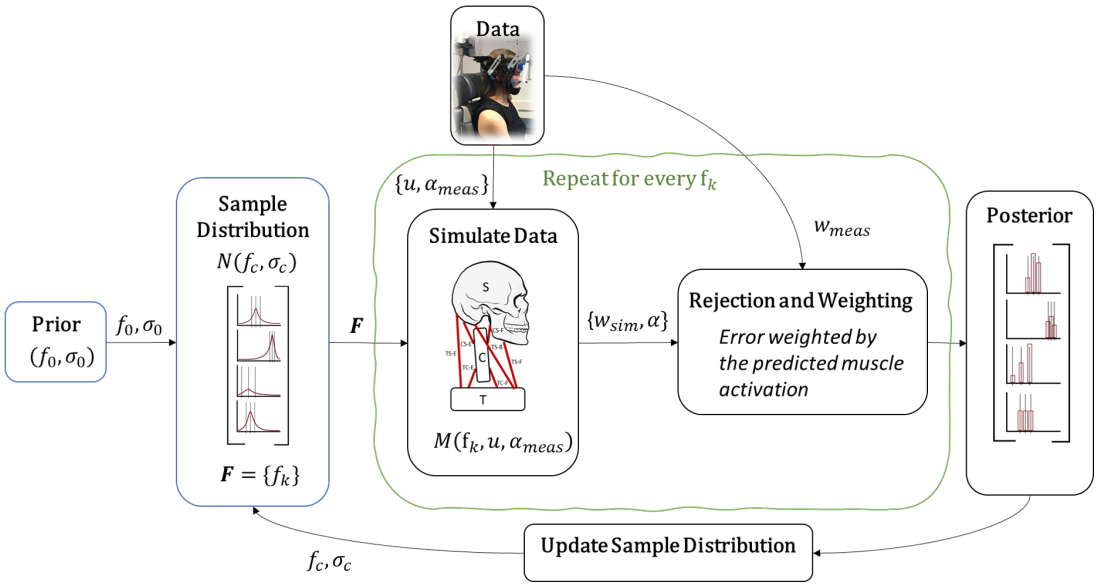


Figure 5.1. **Full Bayesian Estimator for Strength Estimation**. A high level view of the full algorithm. The pseudocode Algorithm 3 provides a step by step explanation of blocks. It starts with constructing the prior. The Sample Distribution is a multivariate Gaussian taking it's parameters from the current point estimate f_c and spread parameter σ_c. A set of candidate strength vectors is sampled from the Sample Distribution. Inside the green box is the likelihood function approximation, where compare data simulated with the model with the measured data through the rejection and weighting filter.

is presented in the previous chapter, in Section 4.2.1.3. In brief, the strength of each muscle group is computed as the norm of individual muscle elements within the group. The new moment arm matrix is computed to maintain the torque producing potential of the muscles from the original model. The resulting $f'_{base} \in \mathbb{R}^m$ is the baseline strength estimate for the associated reduced model, serving as a reasonable starting point with realistic relative strengths between muscle groups.

The final step for computing the initial point estimate is to bootstrap the baseline strength estimate to a single, subject-specific measured wrench from the data set. The

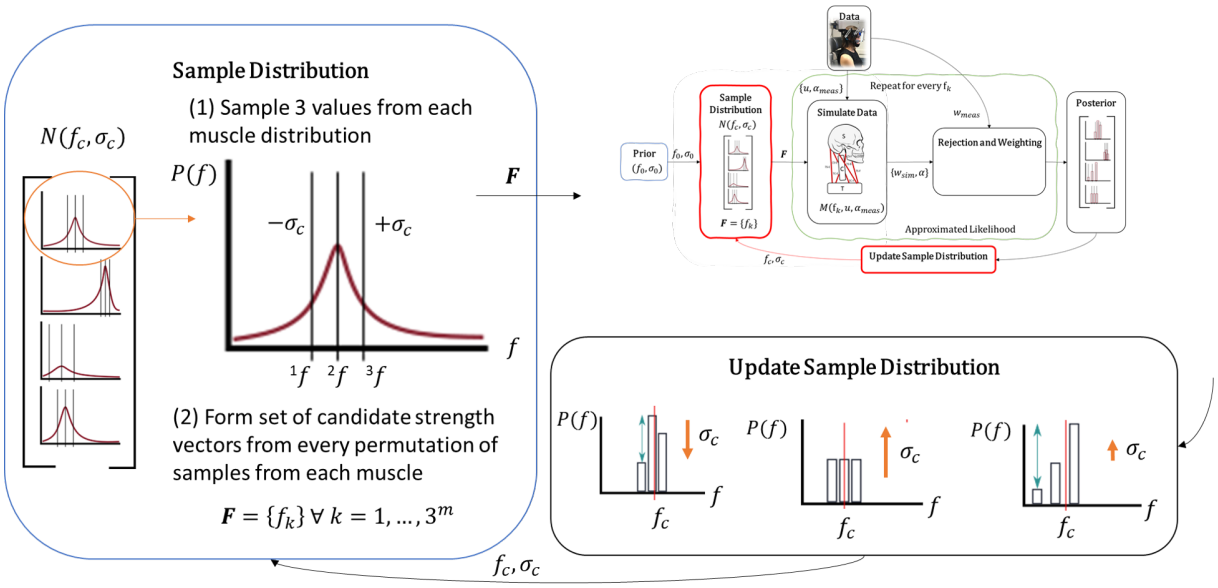


Figure 5.2. **Importance Sampling**. The importance sampling component of the of the algorithm iteratively updates the sampling distribution. The Update Sample Distribution box shows a visual representation of equations 5.10 and 5.11, which compute the new central point estimate f_c and spread parameter σ_c from the posterior distribution. The updated sample distribution is a multivariate Gaussian centered at f_c with a standard deviation of σ_c. Three values are sampled from each muscle's sample distribution according to Equation 5.4. Finally, the set of candidate strength vectors $\mathbf{F}$ is computed from every unique permutation of the 3 samples from each muscle.

resulting point estimate maintains the relative strengths of individual muscle groups based on PCSA, but is scaled to produce a maximum isometric wrench matching an experimental data point. This generic muscle strength scaling procedure is common in musculoskeletal modeling to create subject specific models [110, 78].

The scaled strength estimate, $f_0 = af'_{base}$, is computed by finding the scalar, a, that minimizes the difference between the measured wrench magnitude, w_{meas} and simulated wrench magnitude, w_{sim}, in a single direction, u. We compute $w_{sim} = M(af'_{base}, u)$ from the model function in Algorithm 2.

Algorithm 3 Individual Muscle Strength Estimation Pseudocode

Construct Prior
(1) Derive initial point estimate $f_{base} \in \mathbb{R}^{96}$ from literature
(2) Compute equivalent strengths for simplified model $f'_{base} \in \mathbb{R}^{m}$
(3) Bootstrap initial strength estimate to single measurement direction by scalar a: $f_0 = af'_{base}$
Current iterate, $c = 0$
repeat
(1) **Sample**: Construct set of candidate strength vectors $\mathbf{F} = \{f_k \ \forall \ k = 1, .., N_F\}$ from the current point estimate f_c and spread parameter σ_c.
(2) **Simulate Data**: Compute simulated wrench magnitude and muscle activations $M(f_k, u_i, \alpha_{meas}) = \{w_{sim}, \alpha\}_{i,k}$ for each candidate strength vector $f_k \in \mathbf{F}$ and test wrench direction $u_i \in \mathbf{U}$.
(3) **Data Discrepancy Score**: Compute the data discrepancy score $d(w_{sim,i,k}, w_{meas,i}) \ \forall \ f_k, u_i$ based on the error between the measured and simulated data
(4) **Rejection**: Reject the f_k strength candidates data discrepancy below the q percentile
(5) **Update Posterior**: Compute the posterior probability as the normalized weighted sum of the muscle-specific composite data discrepancy scores $d_{j,k}$ for each remaining candidate strength vector
(6) **Update Sample Distribution**: Compute the updated point estimate, f_c, as the expected value of the posterior, and the updated spread variable, σ_c, based on the shape of the posterior.
Update Current Iterate, $c = c + 1$;
until convergence criterion is met
return Muscle Strength Estimate and Spread $\{f_c, \sigma_c\}$

$$\begin{aligned} \min_{a} \quad & |w_{meas} - w_{sim}| \\ \text{s.t.} \quad & 0 \leq a \leq \infty \end{aligned} \tag{5.3}$$

$$\text{where } w_{sim} = M(af'_{base}, u)$$

The initial point estimate, f_0, is now computed by simply scaling the baseline strength, f'_{base} by a.

The final component of the prior is the initial spread parameter, σ_0. The spread parameter is the standard deviation of the sampling distribution, therefore dictating the size of the parameter space that is sampled on the first iteration. A good starting point for our data is $\sigma_0 = 0.2$. Note that a small σ_0 may lead to slower convergence if the initial point estimate is far from the true value.

5.2.2.2. Sample. The sampling function produces a set of candidate strength estimates, $\mathbf{F} = \{f_k \in \mathbb{R}^m; k = 1, .., N_F\}$, based on the current strength estimate, f_c, and spread, σ_c, parameters. In a more traditional ABC framework, strength estimates would be proposed by randomly sampling from a prior distribution. However, the inherent high dimensionality of the neck model, even with the dimension reduction achieved by grouping similar muscles (a reduction from $m = 96$ to $m = 24$ muscles), necessitates a different sampling strategy. The magnitude of the sampling problem can be illustrated by considering that in order to sample every permutation of just 3 discrete values for each muscle in a $m = 24$ muscle model would require over $N_F = 3^{24} = 2.8 \times 10^{11}$ samples. That is over 2 times the estimated number of neurons in the human brain, but would still not be nearly enough to adequately sample the parameter space. Instead of taking a single pass at sampling the full parameter space, which would be computationally intractable, we employ an iterative approach to adequately sample a region of the parameter space in each iteration, and adjust the sampled region for the next iteration based on the updated estimate. The diagram in Figure 5.2 provides a visual representation of the importance sampling scheme.

The sampling distribution is an m-dimensional multivariate Gaussian distribution centered on the current muscle strength point estimate, $f_c \in \mathbb{R}^m$, with individual standard deviations of $\sigma_c \in \mathbb{R}^m$. Rather than randomly sampling from this Gaussian distribution, three values are sampled explicitly from each muscle, essentially discretizing the multivariate distribution to a set of 3 values for each muscle to insure that representative samples are obtained. For convenience and to capture the full distribution, the three sampled values are the mean (f_c), one standard deviation below the mean ($f_c\ (1 - \sigma_c)$) and one standard deviation above the mean ($f_c\ (1 + \sigma_c)$), as shown in Equation 5.4.

We use the superscript p on ${}^p\mathbf{f}$ to indicate a discrete sampled value, where $p = 1, 2,$ or 3 for the low, mid, and high values respectively.

$$\mathbf{f}_c = \begin{bmatrix} f_{low} \\ f_{mid} \\ f_{high} \end{bmatrix} = \begin{bmatrix} {}^1\mathbf{f}_c \\ {}^2\mathbf{f}_c \\ {}^3\mathbf{f}_c \end{bmatrix} = \begin{bmatrix} f_c \cdot (1 - \sigma_c) \\ f_c \\ f_c \cdot (1 + \sigma_c) \end{bmatrix} \tag{5.4}$$

where each $f_c \in \mathbb{R}^m$ and $\mathbf{f}_c \in \mathbb{R}^{p \times m}$.

The set of candidate strength estimates, $\mathbf{F}$, is composed of every unique permutation of $\mathbf{f}_c$ for each muscle. This combinatorial problem can also be described in terms of m-tuples of 3-sets, where every unique m-dimension ordered list drawn from the 3 values of each muscle makes up the set $\mathbf{F}$. Because there are three sample values for each muscle, the set contains $N_F = 3^m$ f_k candidate strength vectors. This number of samples is much less than would be required to adequately randomly sample from the full parameter space.

5.2.2.3. Simulate Data. A simulated maximum wrench and associated muscle activation pattern, $M(f_k, u_i, \alpha_{meas,i}) = \{w_{sim}, \alpha\}_{i,k}$, is produced for each sample muscle strength candidate, f_k, and unit wrench direction, u_i, at each iteration of the algorithm. The model is represented by Equation 5.2, taking as its inputs f_k, u_i, and any available upper and lower bounds on muscle activations $\alpha_{meas,i} = [\alpha_{lb}, \alpha_{ub}]$. Including the lower and upper bounds on activation of individual muscles allows for optional inclusion of electromyography (EMG) data from accessible muscles during an experiment. If the muscle activation range is not available, the default muscle activation range is $0 \leq \alpha \leq 1$.

The data simulation step is by far the most computationally expensive part of the algorithm. Because the model is evaluated independently for each candidate muscle strength and wrench direction, parallel computing is employed to speed up this step in the algorithm.

5.2.2.4. Data Discrepancy Score. The data discrepancy function assigns an error score to each candidate strength sample, f_k, in each wrench direction, u_i, based on a Gaussian radial basis function with a tuning parameter, $\epsilon = 3$.

$$d_{i,k} = e^{-(\epsilon\delta_{i,k})^2} \tag{5.5}$$

$$\text{where } \delta_{i,k} = \| \frac{w_{sim,i,k} - w_{meas,i}}{w_{meas,i}} \|$$

The choice of a Gaussian radial basis function as the metric for data discrepancy gives us the benefit of an exponential function where the result is strictly positive and a small error is highly rewarded (for a steeper function, the shape parameter ϵ can be increased).

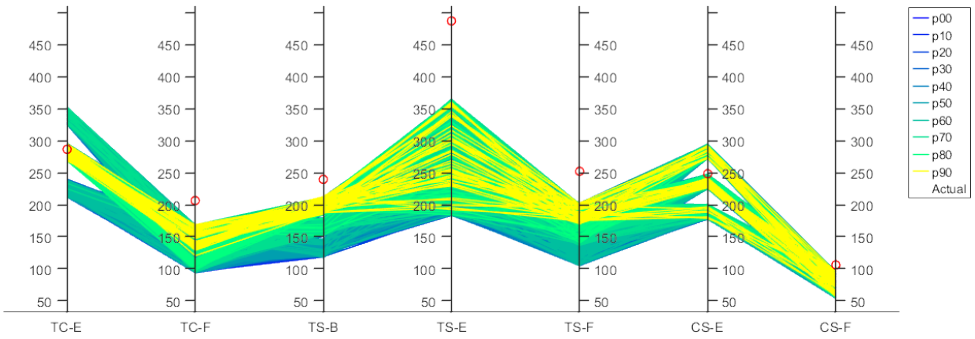
5.2.2.5. Rejection Step. The rejection step removes the lowest scoring candidate strength vectors. First, a composite score for each f_k is computed as the mean of the N_U data discrepancy scores (one for each direction).

$$D_k = \frac{\sum_{i=1}^{N_U} d_{i,k}}{N_U} \tag{5.6}$$

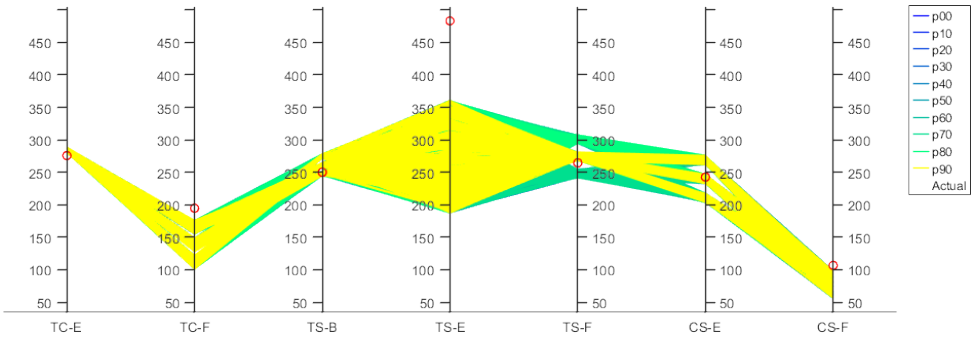
The candidate strength vectors with composite scores below the q percentile are discarded, while the remaining f_k candidates are included in the posterior update. The remaining set of f_k strength candidates is denoted $\mathbf{F}'$.

By considering the scores on candidate strength vectors in this step, we maintain the multivariate property. In the remaining steps of the iteration, the weighting and distributions are considered on a muscle by muscle basis, essentially decoupling the strength vectors for the posterior updates. It can be informative to visualize this step using a parallel coordinate plot as shown in Figure 5.3. Each vertical axis represents a muscle strength, and each trace is an individual candidate strength vector, f_k. The yellow traces are the candidate strength samples f_k that scored in the 90th percentile and above.

5.2.2.6. Update Posterior. Recall that the set of candidate strength vectors, $\mathbf{F}$, is formed by all unique permutations of 3 sampled values for each muscle. The domain of the posterior PDF of muscle j is, therefore, the three discrete sampled values $\mathbf{f}_{samples,j} = \{{}^1\mathbf{f},\ {}^2\mathbf{f},\ {}^3\mathbf{f}\}_j$. The posterior PDF for muscle j, denoted π_j, is computed as the discretized, weighted sum of the data discrepancy scores across test directions and candidate strength vectors and normalized to a sum of 1. Equation 5.7 shows the weighted sum computation for muscle j and sample p. Each score is weighted by the muscle activation level, $\alpha_{j,i,k}$,



(a) Iteration 1



(b) Iteration 5

Figure 5.3. Parallel coordinate plots can be be informative for visualizing the multivariate parameter space. These plots show all $N_F = 3^7$ f_k muscle strength candidate vectors at the (a) first and (b) fifth iterations using the p2m7 model (planar $n = 2$, $m = 7$). There are seven vertical axes, one for each muscle (see Figure 5.6b). The colormap represents the percentile of composite data discrepancy score D_k, where a higher percentile means less error (better score). The red circles are the true strengths.

for muscle j, in direction u_i, simulated with strength, f_k. Recall that the superscript p in ${}^{p}f$ indicates one of 3 sampled values from a particular muscle. The indicator function in Equation 5.8 essentially sorts the candidate strength vectors, subscripted by k, into the appropriate domain value.

$$\phi_{j,p} = \sum_{k}^{N_{F'}} \sum_{i=1}^{N_U} \alpha_{j,i,k} d_{i,k} \mathbb{1}_p(f_{k,j}) \tag{5.7}$$

$$\mathbb{1}_p(f_{k,j}) = \begin{cases} 1 & f_{k,j} = {}^{p}\mathbf{f}_j \\ 0 & \text{else} \end{cases} \tag{5.8}$$

The posterior for muscle j is then computed in Equation 5.9 by normalizing the elements of the weight function by the sum across the domain, resulting in a discrete PDF that sums to 1.

$$\pi_j = \frac{\Phi_j}{\sum_{p=1}^{3} \phi_{j,p}} \text{ where:} \tag{5.9}$$

$$\Phi_j = [\phi_{j,1} \ \phi_{j,2} \ \phi_{j,3}]$$

By utilizing the simulated muscle activation level in the weighted sum for Equation 5.7, only data discrepancy scores that are relevant to each individual muscle will affect the posterior update for that muscle. This point is illustrated by observing the distribution of $d_{i,k}$ values across the test directions, as shown in Figure 5.4. Recall that a data discrepancy score, $d_{i,k}$, is computed for each candidate strength vector, f_k, in each of the test directions, u_i. Muscles are not active in every test direction, in fact, most muscles are predicted to have an activation level, α of 0, in approximately half of the test directions in the simulations. The histograms show that all candidate strength vectors produce wrenches in the "back", "backLow", and "backHigh" directions that very closely match the measured

wrenches (very low error), but there is more variation in score in the other test directions. This weighted sum approach helps minimize the loss of information that occurs when combining the scores across test directions, only averaging scores that are relevant to the particular muscle based on activation level.

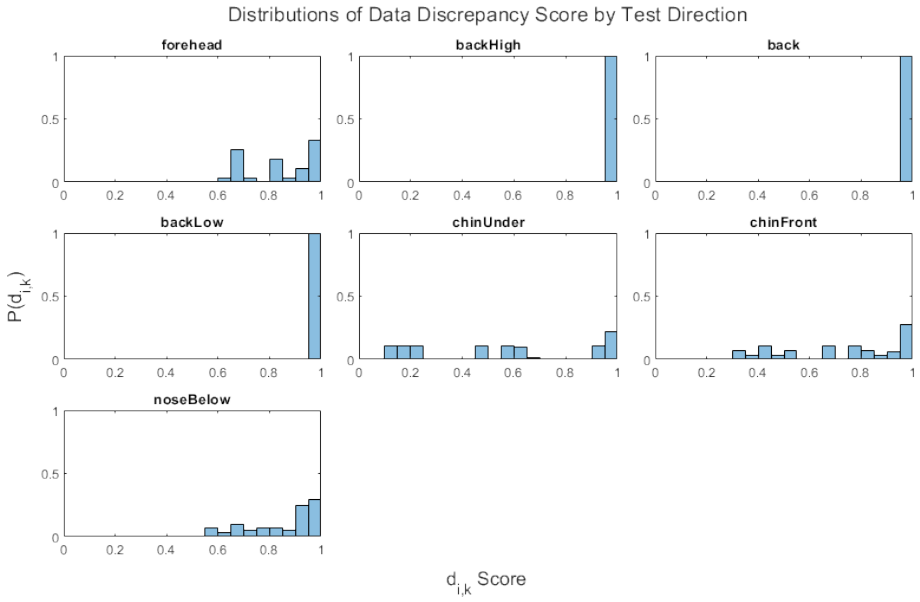


Figure 5.4. Distributions of data discrepancy score $d_{i,k}$ by test direction for a single iteration of the algorithm. Scores range from 0 to 1, with a score of 1 indicating no error between the measured wrench and simulated wrench magnitude in that direction. All data discrepancy scores in the "back" directions (backHigh, back, and backLow) are above 0.95, indicating very low error. The scores in the other 4 directions vary quite significantly. By utilizing the simulated muscle activation level in the weighted sum in Equation 5.7, only data discrepancy scores that are relevant to each individual muscle will affect the posterior update for that muscle.

5.2.2.7. Update Point Estimate. The updated point estimate, f_c, is computed as the expected value of the posterior taken over $\mathbf{f}_{samples}$ for each muscle, Equation 5.10. The spread parameter, σ_c, is updated to reflect confidence in the point estimate based on the relative probabilities of the three discrete samples for each muscle. In the case where

the probability of the middle valued sample ($\pi({}^2\mathbf{f})$) is greatest, σ_c is decreased relative to the difference in magnitude in order to reflect increased confidence in the posterior point estimate. When the middle value is not greatest, σ_c is increased to sample a larger region of the parameter space for that muscle for the next iteration. The spread parameter update function, $K_\sigma(\pi(\mathbf{f}))$, is computed according to Equation 5.12. We limit the value for σ_c to a range of [0.01 0.3] to avoid too narrow or wide of a sampling region.

$$f_c = \mathbb{E}_p(\pi(\mathbf{f}))) = \sum_{p=1}^{3} {}^p\mathbf{f}\ \pi({}^p\mathbf{f}) \tag{5.10}$$

$$\sigma_c = K_\sigma(\pi(\mathbf{f}))\sigma_{c-1} \tag{5.11}$$

$$K_\sigma(\pi(f)) = \begin{cases} 1-\Delta & \text{if } \max(\pi(f)) = \pi({}^2\mathbf{f}) \\ 1+e^{(-5\Delta-1)} & \text{else} \end{cases} \tag{5.12}$$

$$\text{where } \Delta = \max(\pi(f)) - \min(\pi(f))$$

5.2.2.8. Convergence Criteria. The simulation continues until the convergence criteria is met or the number of iterations exceeds $maxIter$. The simulation is considered to have converged when all candidate strength vectors f_k in $\mathbf{F}$ have a composite data discrepancy score D_k above the cutoff ϵ_D, which we set at $\epsilon_D = 0.95$. A composite data discrepancy score D_k above 0.95 indicates that the candidate strength vector f_k produces

simulated maximum wrenches very close to the measured wrenches. A perfect match would result in $D_k = 1$. Figure 5.5 shows the distribution of D_k scores by iteration, with convergence being met at iteration $c = 10$.

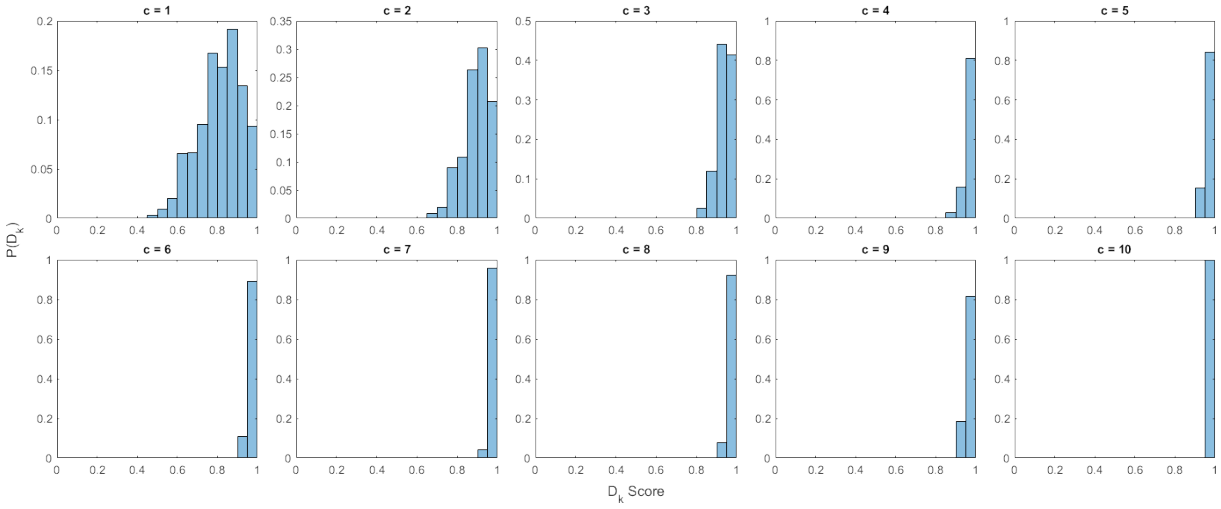


Figure 5.5. Histogram plots of composite data discrepancy scores, D_k, for all candidate strength vectors over 10 iterations. Plotting the distribution of D_k scores by iteration can be informative for visualizing convergence. By iteration 10, all scores are above 0.95, indicating very low discrepancy between the measured data set and simulated wrenches.

5.2.3. Experiment

Maximum neck strength measurements and muscle activation of a subset of superficial neck muscles were collected from 5 healthy participants, as detailed in Appendix A. In brief, participants were seated and attached to a custom built neck strength measurement system that couples the head to a 6-DOF load cell. Real-time feedback of the 6-DOF wrench produced by the participant and the target direction were visible on a screen placed in front of the participant. Participants were instructed to produce a maximum effort while matching the direction of the wrench in each of 25 directions in a randomized order.

Algorithm 4 Individual Muscle Strength Algorithm

Require: initial point estimate f_0 and spread parameter σ_0

$c = 0$

repeat

$\mathbf{f}_c = \begin{bmatrix} f_c \cdot (1 - \sigma_c) \\ f_c \\ f_c \cdot (1 + \sigma_c) \end{bmatrix}$ ▷ Sample

$\mathbf{F} = \{f_k \in \mathbb{R}^m; k = 1, .., 3^m\}$ ▷ every unique permutation of $\mathbf{f}_c$

for all $f_k \in \mathbf{F}$ **do**

for all $u_i \in \mathbf{U}$ **do**

$\{w_{sim}, \alpha\}_{i,k} = M(f_k, u_i, \alpha_{meas,i})$ ▷ Simulate Data: Equation 5.2

$\delta_{i,k} = \| \frac{w_{sim,i,k} - w_{meas,i}}{w_{meas,i}} \|$

$d_{i,k} = e^{-(\epsilon \delta_{i,k})^2}$ ▷ Data Discrepancy Score

end for

$D_k = \frac{\sum_{i=1}^{N_U} d_{i,k}}{N_U}$ ▷ Composite Data Discrepancy Score

end for

$\mathbf{F}' = \{f_k \ \forall \ k \ s.t. \ D_k > q \text{ percentile}\}$ ▷ Rejection Step

for all $j = 1, ..., m$ **do**

for all $p = 1, 2, 3$ **do**

$$\mathbb{1}_p(f_{k,j}) = \begin{cases} 1 & f_{k,j} = {}^p\mathbf{f}_j \\ 0 & \text{else} \end{cases}$$

$\phi_{j,p} = \sum_k^{N_{F'}} \sum_i^{N_U} \alpha_{j,i,k} d_{i,k} \mathbb{1}_p(f_{k,j})$ ▷ Posterior Update

end for

$\pi_j = \frac{\Phi_j}{\sum_{p=1}^3 \phi_{j,p}}$

$f_c = \mathbb{E}_p(\pi(\mathbf{f}))) = \sum_{p=1}^3 {}^p\mathbf{f} \, \pi({}^p\mathbf{f})$ ▷ Point Estimate Update

$\Delta = \max(\pi(\mathbf{f})) - \min(\pi(\mathbf{f}))$

$$K_\sigma(\pi(\mathbf{f})) = \begin{cases} 1 - \Delta & \text{if } \max(\pi(\mathbf{f})) = \pi({}^2\mathbf{f}) \\ 1 + e^{(-5\Delta - 1)} & \text{else} \end{cases}$$

$\sigma_c = K_\sigma(\pi(\mathbf{f}))\sigma_{c-1}$ ▷ Spread Parameter Update

end for

$c++$

ConvCrit = $(\min D_k > \epsilon_D)$ **or** $(c > maxIter)$

until ConvCrit

Each direction was repeated 2-3 times, with the maximum value taken as the measured wrench w_{meas} for that direction. EMG activity of 5 pairs of muscles was simultaneously

recorded, including the right and left sternocleidomastoid (SCM), upper trapezius (trap), semispinalis capitis (SSCap), splenius capitis (SpCap), and levator scapulae (LS). The low pass filtered full wave rectified EMG data were normalized to the maximum over all trials for each muscle as a measure of activation. The value at 250 ms prior to the maximum wrench across trials in each direction was taken as the muscle activation for that direction to account for muscle fiber activation delays. The data set from each participant includes measured wrenches for the six cardinal directions, flexion, extension, left and right sidebending, and left and right rotation, as well as the 25 unit wrench directions that make up the test direction set $\mathbf{U}$. For more details regarding the experiment and neck strength results, refer to Appendix A.

5.3. Applications

The algorithm was applied to several variations of the neck model to demonstrate the utility of the method and highlight particular limitations, categorized based on computational shortcomings or intrinsic biomechanical properties. First, two planar neck models that reduce the biomechanics to the sagittal plane (flexion/ extension) are presented. The lower dimensionality allows for a clear demonstration of the process and easy interpretation of results. Application of the algorithm was then expanded to the full spatial model, with a parameter space of $m = 24$ muscles and $N_U = 25$ test wrench directions. For each variation of the model, we present results for both simulated and experimental data sets (collected from n=5 healthy controls).

The advantage of exploring the simulated data sets is that the true muscle strengths are known, and therefore, accuracy of the final estimate can be assessed. Additionally,

simulated data sets can be designed to mimic particular deficits that we are interested in, such as weak deep extensor muscles. The experimental data sets offer the opportunity to demonstrate feasibility of the framework estimating individual neck muscle strength in a laboratory environment, and the potential to translate to clinical settings.

Two simulated experimental data sets were created for each model variant, one healthy and one weak, where the strength of the muscle corresponding to the deep extensors is reduced. First, the 'actual' muscle strength vector $f_a \in \mathbb{R}^m$ for the 'healthy' model was computed by pseudo-randomly perturbing the f_{base} baseline strength estimate for our model by scaling by a factor $a = 1.5$ and adding random Gaussian noise with a variance of $v = 0.4 f_a$. The simulated experimental data set was then computed using the model Equation 5.2: $\{w_{meas}, \alpha_{meas}\}_i = M(f_a, u_i) \; \forall \; u_i \in \mathbf{U}$. We limit α_{meas} to the superficial muscles for which we recorded surface EMGs in the experiments with participants and that apply to the corresponding model variant. To produce the 'weak' data set, the strength of the deep extensor muscle(s) of the 'healthy' muscle strength vector f_a was scaled by 25%.

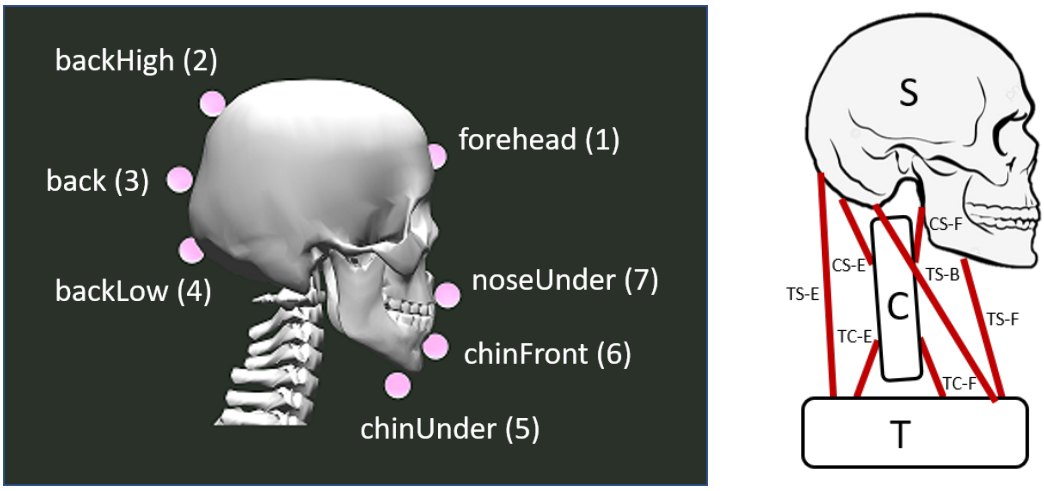
5.3.1. Planar Neck Models

Muscle geometry is symmetric about the sagittal plane, allowing the left and right muscles to be combined, reducing the muscle space dimensionality, m, in half for a planar model compared to the spatial model. Additionally, the kinematic degrees of freedom at each joint are reduced from three in the spatial model to one in a planar model. Overall, the ratio of number of muscles to kinematic DOFs increases, which tends to indicate increased muscle redundancy.

Models that are constrained to the sagittal plane have a task (wrench) space dimension of $d = 3$, so the test direction unit wrench has 3 components, $u = [M_z, F_x, F_y]$. The task space wrench is applied to a frame located at the base of the skull (coincident with the upper cervical spine joint). The subset of sagittal plane wrench directions, $\mathbf{U}_p$ is a set of $N_{Up} = 7$ directions, applied at the markers shown in Figure 5.6a.

5.3.1.1. Models. The most simplified neck model, the *p2m7* model (Figure 5.6b), has $n = 2$ kinematic DOFs and $m = 7$ muscles. The muscles are grouped according to their torque contribution about each joint, with single vs. multi-joint muscles grouped separately. For example, muscles that cross only the lower cervical joint and produce an extension torque are grouped together, separately from muscles that cross both the lower and upper cervical joints. The other planar model, the p3m12 model is reduced to $n = 3$ kinematic DOFs and $m = 12$ muscles. The muscles are grouped logically based on anatomical name, as opposed to the systematic groupings of the *p2m7* model that were based on moment arms.

5.3.1.2. Results with Simulated Data Sets. The simulation results of the individual muscle strength estimation algorithm with simulated healthy and weak data sets for the *p2m7* and *p3m12* models are shown in Figures 5.7 and 5.8 respectively. The initial strength estimate from each model, f_{base}, was bootstrapped to the wrench measurement taken in the "back" direction. The initial spread parameter was set at $\sigma_0 = 0.3$. The simulations for the *p2m7* model are shown for 30 iterations, and the simulations for the *p3m12* model are shown for 20 iterations, but all 4 simulations met the convergence criteria at earlier iterations. The final point estimate, f_c, and spread parameter, σ_c, and

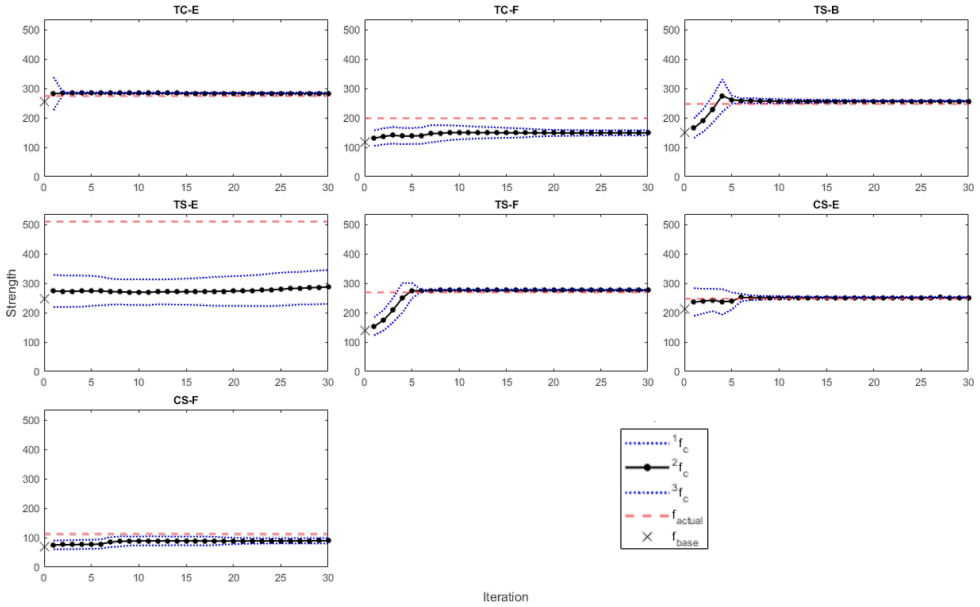


(a) (b)

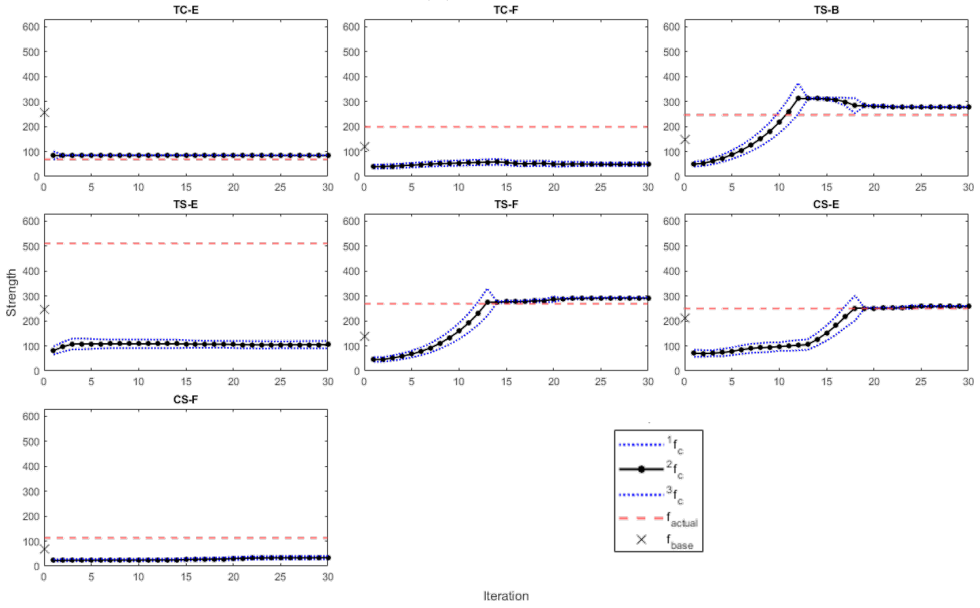
Figure 5.6. (a) Sagittal Plane Test Locations: Each circle marks the location of force application for a test direction in the subset of 7 test directions for the sagittal plane. The unit wrench direction for a test direction is computed by transforming the force applied normal to the surface of the skull to the skull frame at the base of the skull, and then normalized to a unit vector. (b) The *p2m7* model. The naming scheme for the p2m7 muscles is based on the muscle attachment and action, where T = trunk, C = C-spine, S = Skull, E = Extension, F = Flexion, and B = Both (the SCM produces a flexion torque about the lower c-spine and an extension torque about the upper c-spine).

the final composite data discrepancy score D_{f_c} are tabulated in Table 5.1 and 5.2 for models *p2m7* and *p3m12*, respectively.

All 4 simulations for the planar models with simulated data sets met the convergence criteria of $\min(D_k) > 0.95$, indicating that even the lowest scoring candidate strength estimate f_k produced maximum wrenches that closely fit the simulated data set. Convergence criteria was met at iteration $c = 6$ for the *p2m7* healthy model, $c = 15$ for the *p2m7* weak model, $c = 6$ for the *p3m12* healthy model, and $c = 12$ for the *p3m12* weak model.

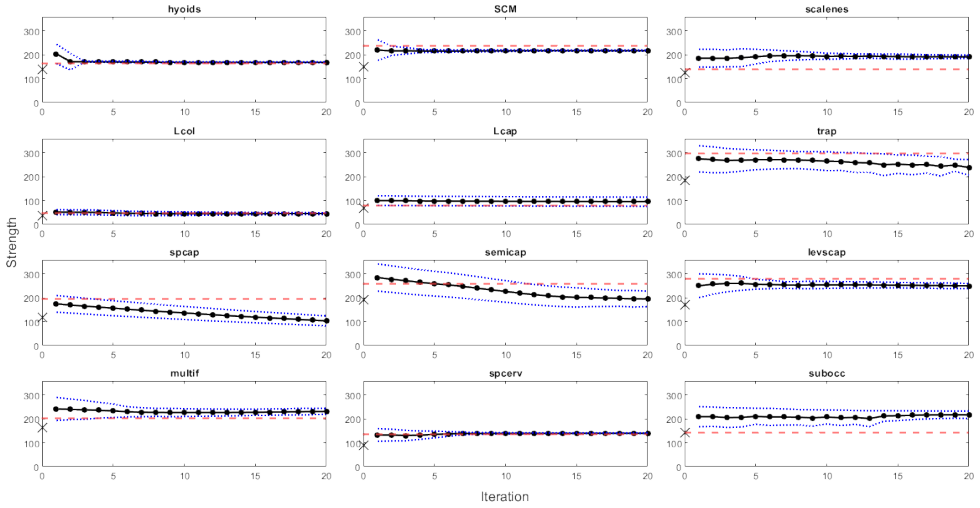


(a) healthy

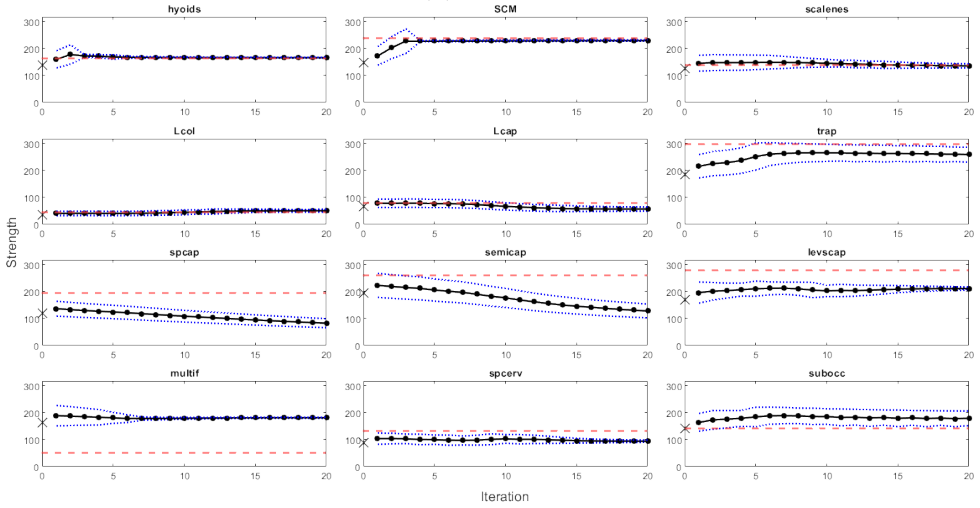


(b) weak deep extensors (TC-E)

Figure 5.7. **Parameter Estimation Results for Simulated Data Sets with *p2m7* model**. Each subplot represents a muscle group (defined in Figure 5.6b). Accuracy and convergence vary between muscles, with similar patterns emerging between the two data sets for which muscles are able to produce accurate and confident estimates. The baseline strength estimate, f'_{base}, prior to bootstrapping is marked by an x on the y-axis. The "actual" strength is the dashed red trace. The current point estimate, ${}^2\mathbf{f}_c = f_c$ is the black trace. The low and high samples, ${}^1\mathbf{f}_c$ and ${}^3\mathbf{f}_c$ are the dashed blue traces.



(a) healthy



(b) weak deep extensors (multif)

Figure 5.8. **Parameter Estimation Results for Simulated Data Sets with *p3m12* model**. Each subplot represents a muscle group. Accuracy and convergence vary between muscles, with similar patterns emerging between the two data sets for which muscles are able to produce accurate and confident estimates. The baseline strength estimate, f'_{base}, prior to bootstrapping is marked by an x on the y-axis. The "actual" strength is the dashed red trace. The current point estimate, ${}^2\mathbf{f}_c = f_c$ is the black trace. And the low and high samples, ${}^1\mathbf{f}_c$ and ${}^3\mathbf{f}_c$ are the dashed blue traces.

		TC-E	TC-F	TS-B	TS-E	TS-F	CS-E	CS-F	D(f_c)	Iters
norm	f_a	275	197	247	510	270	249	112	1.00	6
	f_c	284	146	258	272	276	251	85		
	σ_c	0.01	0.20	0.04	0.16	0.01	0.03	0.20		
weak	f_a	69	197	247	510	270	249	112	0.98	15
	f_c	85	52	307	107	278	182	27		
	σ_c	0.01	0.20	0.03	0.13	0.01	0.20	0.20		
S003	f_c	184	12	259	207	128	60	338	0.78	30+
	σ_c	0.01	0.30	0.02	0.03	0.01	0.05	0.05		
S004	f_c	85	2	98	2	12	45	126	0.65	30+
	σ_c	0.01	0.30	0.01	0.30	0.08	0.23	0.04		
S005	f_c	160	54	189	169	180	26	40	0.99	11
	σ_c	0.03	0.25	0.02	0.29	0.04	0.15	0.30		
S006	f_c	199	93	183	163	344	59	178	0.86	30+
	σ_c	0.01	0.13	0.01	0.01	0.15	0.05	0.05		
S007	f_c	293	139	178	283	47	41	69	0.84	30+
	σ_c	0.01	0.04	0.01	0.04	0.01	0.30	0.04		

Table 5.1. **Parameter Estimation Results using the *p2m7* model**. Results are tabulated for the two simulated data sets (norm and weak) and five experimental data sets (S003 - S007). The algorithm was able to identify weakness of the deep extensor muscle group (TC-E) in the "weak" data set, with high confidence. Similar patterns emerge among the simulations regarding the relative uncertainty in the estimates (represented by the value of the spread parameter σ_c). Note the large error in the strength estimate for muscle group TS-E for the two simulated data sets (comparing f_a with f_c). Despite this error, the fit of the data was excellent for these two data sets, with a data discrepancy score of $D(f_c) = 1.00$ and 0.98, indicating that an alternative solution was found that fit the data and there was not enough information in the data set or model to discriminate between solutions. [*KEY*: f_a: "actual" muscle strength; f_c: final strength point estimate; σ_c: final spread parameter value; $D(f_c)$: data discrepancy score for final point estimate; Iters: number of iterations to convergence; TC-E, TC-F, TS-B, TS-E, TS-F, CS-E, CS-F: 7 muscles groups of the *p2m7* model (as defined in Figure 5.6b)]

The accuracy of the final point estimates varied by muscle. Taking the *p2m7* weak data set simulation (Figure 5.7b) for example, the final point estimates for muscles TC-E,

		hy	SCM	scal	Lc	LC	trap	spC	ssC	LS	mf	spc	SO	D(f_c)	Iters
	f_a	164	237	139	47	79	297	195	261	279	203	134	141		
norm	f_c	170	216	196	46	98	271	148	248	254	228	139	208	1.00	6
	σ_c	0.02	0.02	0.11	0.20	0.20	0.14	0.20	0.20	0.05	0.08	0.04	0.16		
	f_a	164	237	139	47	79	297	195	261	279	51	134	141		
weak	f_c	167	229	141	49	60	264	100	156	205	180	100	181	1.00	12
	σ_c	0.01	0.01	0.09	0.16	0.20	0.11	0.20	0.20	0.09	0.02	0.13	0.17		
S003	f_c	278	206	122	64	80	204	170	305	230	248	87	220	0.57	20+
	σ_c	0.02	0.03	0.20	0.15	0.20	0.20	0.04	0.06	0.02	0.03	0.01	0.02		
S004	f_c	69	90	44	18	34	185	59	99	40	118	68	68	0.56	20+
	σ_c	0.03	0.20	0.20	0.19	0.20	0.20	0.06	0.06	0.20	0.06	0.06	0.12		
S005	f_c	289	137	49	27	45	52	30	77	64	157	120	229	0.97	20+
	σ_c	0.05	0.03	0.14	0.20	0.20	0.20	0.14	0.20	0.20	0.04	0.08	0.14		
S006	f_c	304	224	80	33	71	59	37	95	27	223	131	109	0.82	20+
	σ_c	0.09	0.03	0.20	0.20	0.20	0.20	0.20	0.16	0.20	0.06	0.04	0.03		
S007	f_c	223	157	101	54	96	130	31	199	287	237	225	161	0.69	20+
	σ_c	0.18	0.15	0.20	0.20	0.20	0.20	0.20	0.14	0.06	0.03	0.03	0.14		

Table 5.2. **Parameter Estimation Results using the *p3m12* model** : Results are tabulated for the two simulated data sets (norm and weak) and the five experimental data sets (S003-S007). The algorithm was not able to identify the weakness of the deep extensor muscle group (mf) using this model (note the error between f_a and f_c for the weak data set in the mf column). Despite this error, the fit of the parameter estimate to the data set was excellent, with a data discrepancy score of $D(f_c) = 1.00$. This indicates that an alternative solution was found. The data available combined with the planar models may not have enough information to discriminate between solutions due to the level of redundancy in the system. [*KEY*: f_a: "actual" muscle strength; f_c: final strength point estimate; σ_c: final spread parameter value; $D(f_c)$: data discrepancy score for final point estimate; Iters: number of iterations to convergence]

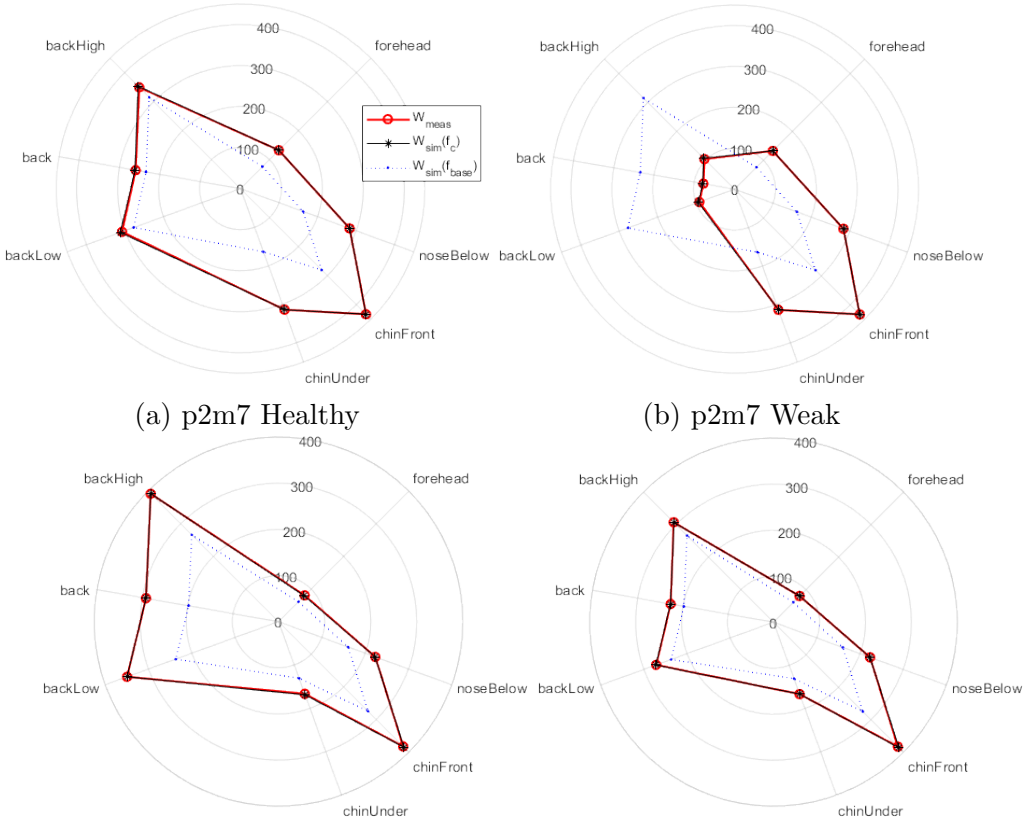
TS-B, TS-F, and CS-E are within 5% of the true value. However, the algorithm produces an estimate for muscle TS-E of 107 N, when the muscle's f_{actual} value was 510 N, for an error of 79%. The *p3m12* weak data set simulation over-estimates the true strength of the multifidus muscle by more than double, for an error of 255%, and underestimates the strength of the two more superficial extensors, the splenius capitis and semispinalis capitis by 48% and 40%, respectively.

Despite the significant error in the strength estimate of several muscles when compared to the actual strengths that produced the data sets, the final point estimates produce nearly identical maximum wrench measurements in each test direction when compared to the simulated data sets, as shown in Figure 5.9 and indicated by the composite data discrepancy score for the final point estimate, f_c, of $D_c = 1.00$. This indicates that the algorithm found an alternate solution - another set of muscle strengths that produce the same set of maximum wrench measurements.

5.3.1.3. Distinguishing between healthy vs. weak deep extensor data sets. A primary motivation of this work is to identify weak deep extensor muscles from neck strength measurements. While it is clear from the radar plots in Figure 5.9 that measured wrenches in the extension direction (test directions "back", "backlow", and "backHigh") are reduced, muscle redundancy leads to multiple muscle strength solutions that are not disambiguated by the test directions. For example, the *p3m12* weak data set simulation (Figure 5.8b) ends with an estimate that overestimates the strength of the multifidus (the deep extensor muscle) and underestimates the strength of the semispinalis capitis and splenius capitis (the superficial extensors). In this case, the weak deep extensors would be missed by the algorithm. These results lead to the conclusion that there is not enough information in the planar data sets to distinguish between the superficial and deep extensor muscles with the models constrained to the sagittal plane with 7 test directions.

5.3.2. Spatial Neck Model

With a transition to a spatial neck model and data set, the task (wrench) space dimension increases to $d = 6$, so that a test direction unit wrench has 6 components,



(a) p2m7 Healthy (b) p2m7 Weak

(c) p3m12 Healthy (d) p3m12 Weak

Figure 5.9. **Radar Plots Showing Parameter Estimate Fit to Data Set**: These plots compare the maximum wrench measurements w_{meas} from the simulated data set (red trace) vs. maximum wrenches produced by the final point estimate $w_{sim}(f_c)$ for simulations (black trace) with the two planar models (a & b) *p2m7* and (c & d) *p3m12*. The dashed blue line shows the maximum wrenches produced by the baseline strength estimate for each model. The black traces overlay the red traces nearly identically, indicating an excellent fit of the parameter estimates to the to the measured data. This is confirmed by the near perfect data discrepancy scores $D(f_c) = 1.00, 0.98, 1.00,$ and 1.00 for (a-d) respectively.

$u = [M_x, M_y, M_z, F_x, F_y, F_z]$. The set of test directions, $\mathbf{U}$, for the spatial data sets is now $N_U = 25$, applied at the markers shown in Figure 5.10. Each joint in the model is associated with 3 kinematic DOFs: pitch, roll, and yaw.

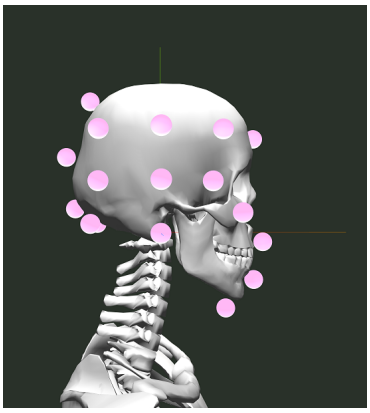


Figure 5.10. **Spatial test locations**: Each of the 25 circles marks the location of force application for a strength test. The unit wrench direction for a test direction is computed by transforming the force applied normal to the surface of the skull to the skull frame at the base of the skull, and then normalized to a unit vector.

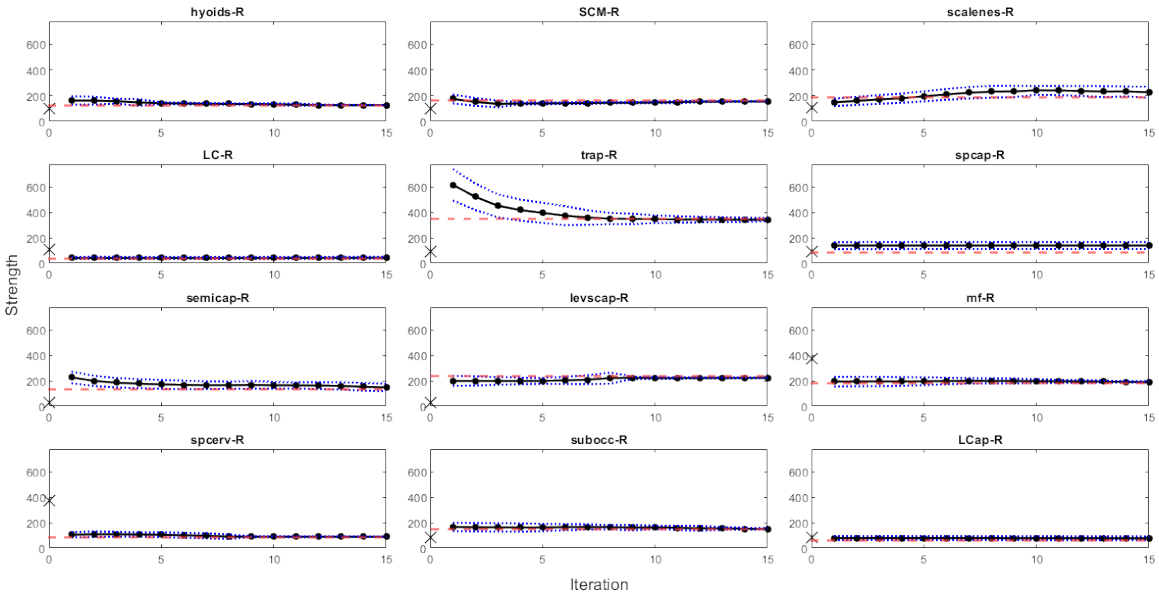
5.3.2.1. Model. The spatial neck model, *s3m24*, has 3 joints ($n = 9$ kinematic DOFs) and $m = 24$ muscles. The muscles are grouped in the same manner as the p3m12 model, but with right and left muscle pairs, resulting in twice the number of muscles. In order to make the simulations computationally feasible, we make an assumption during our sampling step that the right and left muscle pairs have equal strength. This assumption is reasonable for our population of interest, chronic WAD, where muscle atrophy is distributed throughout the deep extensor muscles bilaterally [1]. This reduces the set of candidate muscle strengths $\mathbf{F}$ to $N_F = 3^{12}$ from $N_F = 3^{24}$.

5.3.2.2. Results with Simulated Data Sets. The results of the muscle strength estimation algorithm with simulated healthy and weak deep extensor data sets and the spatial *s3m24* model are shown in Figures 5.11a and 5.11b respectively. The results are also tabulated in Table 5.3. As with the planar model simulations, the initial strength

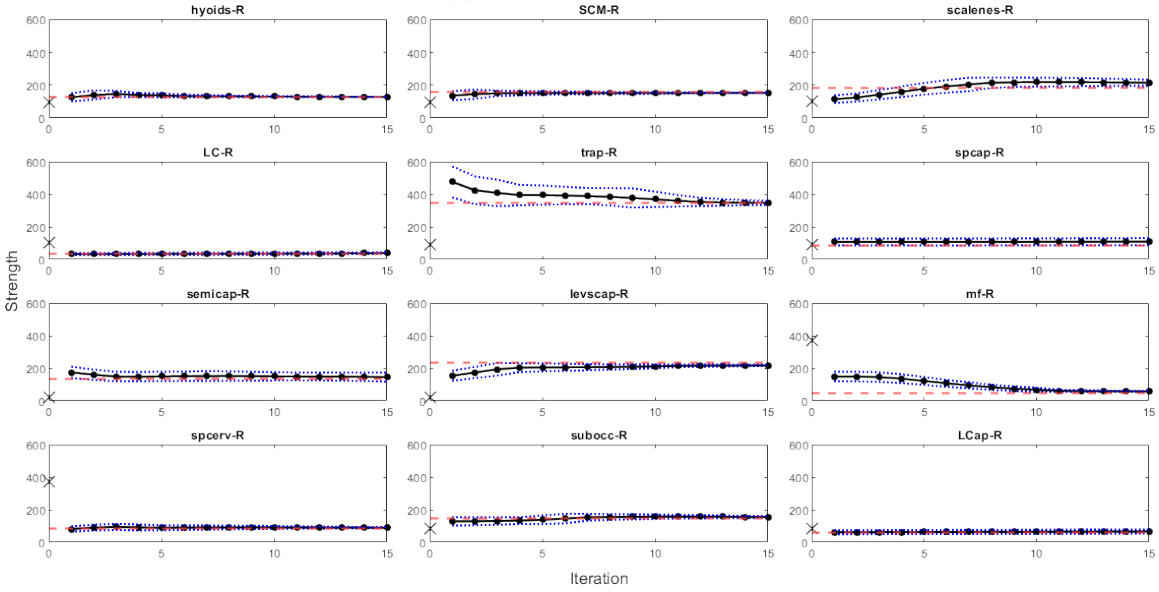
estimate, f'_{base} was bootstrapped to the wrench measurement taken in the "back" direction. The initial spread parameter was set at $\sigma_0 = 0.2$. The simulations were run for 15 iterations, with both meeting the convergence criteria of $\min(D_k) < 0.95$ by iteration $c = 9$.

The final point estimates for the individual muscle strength estimation simulations for the 'healthy' and 'weak mf' data sets both produced excellent fits to the simulated experimental data set, with composite data discrepancy scores of $D_c = 1.00$ for each. The algorithm converged to the set of wrench measurements produced by the f_a strength vector representing the true strength for all muscles. To further confirm that the final strength estimates produce a wrenches that closely match the data set, the distribution of data discrepancy scores for each candidate strength vector in the final iteration is shown in Figure 5.12 for the s3log12 weak deep extensor simulation. As shown in Figure 5.12b, nearly all individual data discrepancy scores are above $d_{i,k} = 0.95$ in each test direction, indicating an excellent fit of all candidate strength vectors by iteration $c = 9$ to the simulated data set. The error of the point estimate , f_c, at convergence compared with the actual strength vector, f_a, is below 10% error with the exception of the scalenes, splenius capitis, semispinalis capitis, and longus capitis.

5.3.2.3. Results with Experimental Data Sets. The results of the muscle strength estimation simulations for two of the $n = 5$ participants is shown in Figures 5.13a and 5.13b. All simulation parameters were set the same as for the simulated data sets in the previous section. The algorithm was run for 20 iterations, and the convergence criteria of $\min(D_k) > 0.95$ was not achieved for any of the simulations.

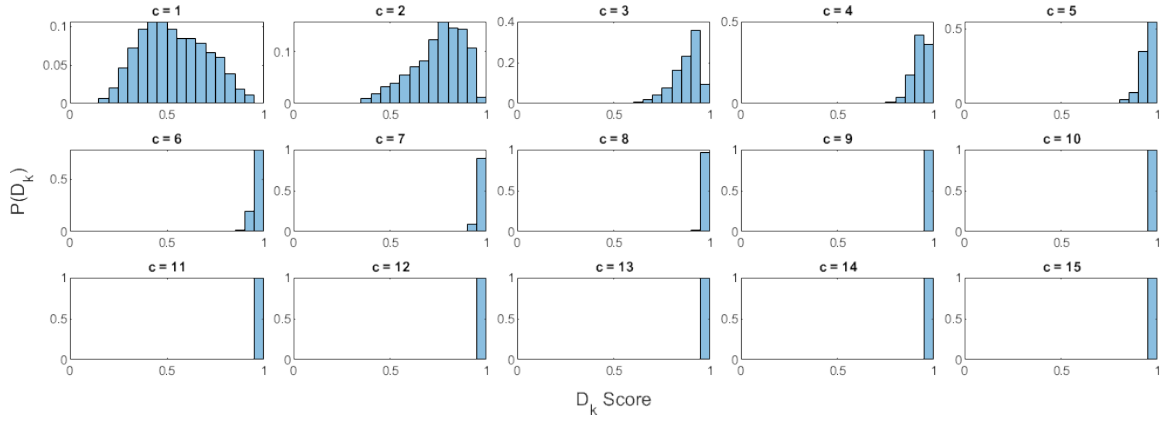


(a) Simulated Normal

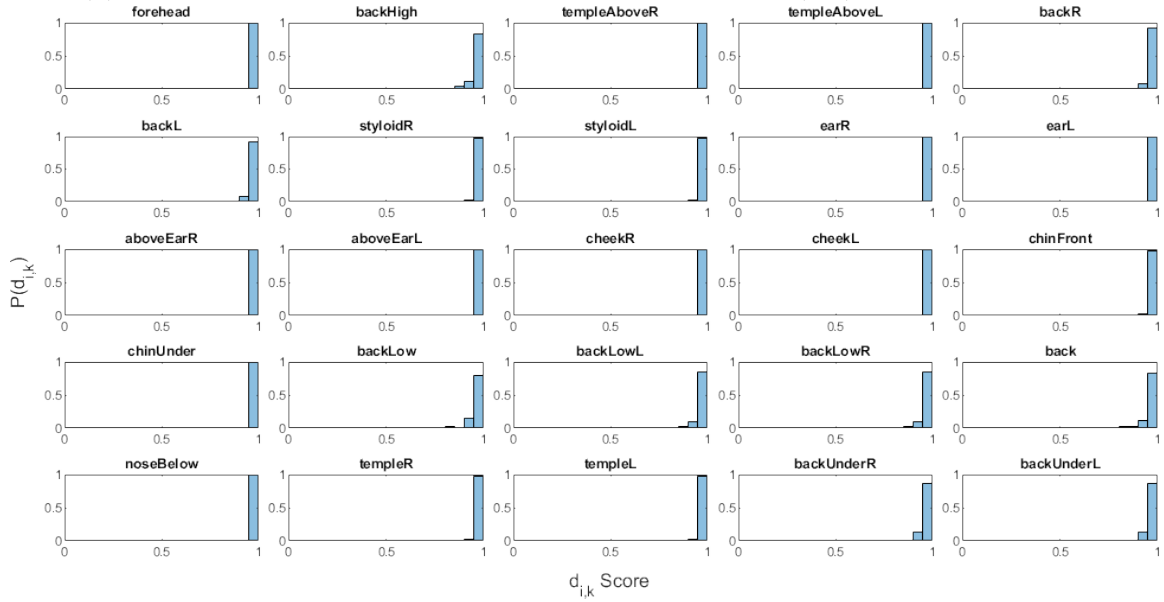


(b) Simulated Weak (mf-R)

Figure 5.11. **Parameter Estimation Results for Simulated Data Sets with *s3m24* model**. Each subplot represents a muscle group. The baseline strength estimate, f'_{base}, prior to bootstrapping is marked by an x on the y-axis. The "actual" strength is the dashed red trace. The current point estimate, ${}^2\mathbf{f}_c = f_c$ is the black trace. And the low and high samples, ${}^1\mathbf{f}_c$ and ${}^3\mathbf{f}_c$ are the dashed blue traces. The algorithm was able to identify isolated weakness of the deep extensor muscle group (mf) and the final parameter estimates achieved an excellent fit to the data, with data discrepancy scores of $D(f_c) = 1.00$.

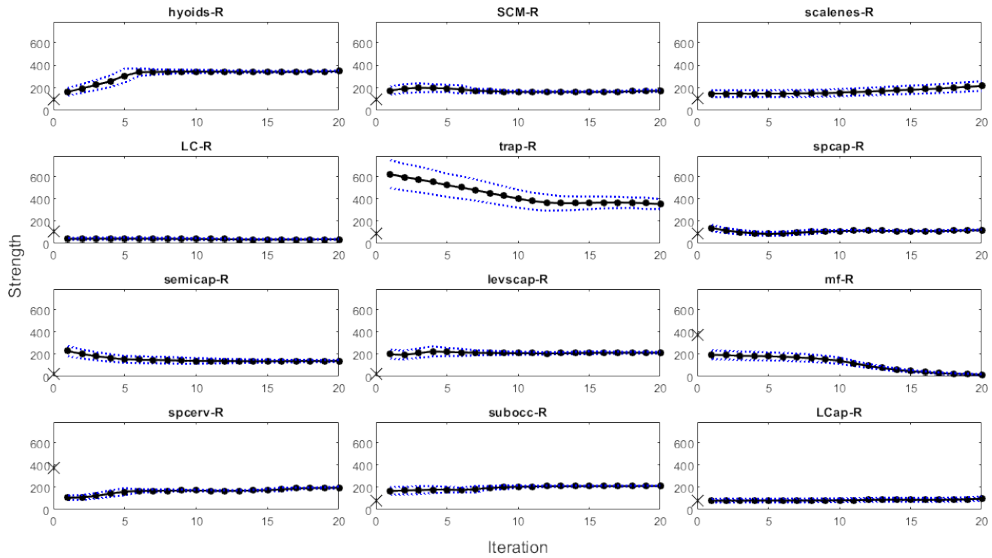


(a) Distribution of Composite Data Discrepancy Scores (D_k) over 20 Iterations

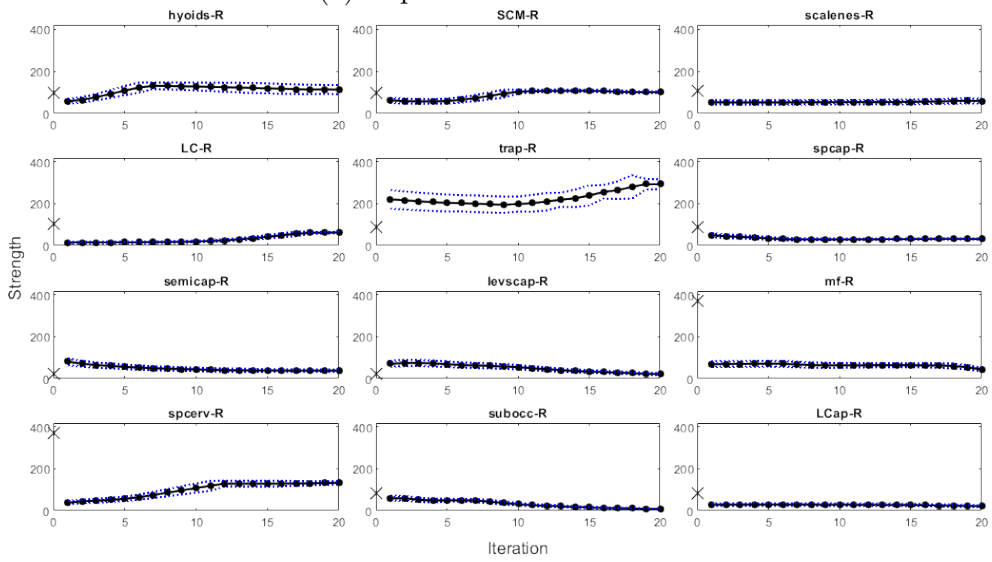


(b) Distribution of Data Discrepancy Scores by Test Direction at Iteration $c = 9$

Figure 5.12. **Distributions of Data Discrepancy Scores for the Weak Data Set and *s3m24* Model**. (a) The distribution of all N_F composite data discrepancy scores over 20 iterations. By the 9^{th} iteration, all D_k scores were above 0.95, indicating that the algorithm had met the convergence criteria. (b) The data discrepancy scores ($d_{i,k}$) for all 25 test directions at iteration $c = 9$ shows that all f_k strength candidates produce wrench magnitudes that closely match the measured data set in all directions (where a $d_{i,k} = 1$ would indicate no error).



(a) Experimental Data S003



(b) Experimental Data S004

Figure 5.13. **Parameter Estimation Results for two Experimental Data Sets with *s3m24* model**. The left and right muscles of each muscle pair are assumed to have equal strengths, so only the right side muscles are shown here. The baseline strength estimate, f'_{base}, prior to bootstrapping is marked by an x on the y-axis. The "actual" strength is the dashed red trace. The current point estimate, ${}^2\mathbf{f}_c = f_c$ is the black trace. And the low and high samples, ${}^1\mathbf{f}_c$ and ${}^3\mathbf{f}_c$ are the dashed blue traces. The algorithm was run for 20 iterations, and the convergence criteria of $\min(D_k) > 0.95$ was not achieved for any of the simulations. The data discrepancy of the final point estimates of these two simulates was $D(f_{c=20}) = 0.66$ and 0.54, indicating a moderate fit of the final parameter estimates to the data.

		hy	SCM	scal	Lc	trap	spC	ssC	LS	mf	spc	SO	LC	D(f_c)	Iter
	f_a	125	160	184	37	350	84	132	233	184	84	149	62		
norm	f_c	133	147	241	41	346	140	165	221	197	89	163	78	1.00	9
	σ_c	0.05	0.04	0.14	0.19	0.09	0.20	0.17	0.03	0.08	0.05	0.09	0.20		
	f_a	125	160	184	37	350	84	132	233	46	84	149	62		
weak	f_c	131	152	219	35	371	108	151	213	68	91	158	64	1.00	9
	σ_c	0.03	0.03	0.12	0.20	0.13	0.20	0.17	0.05	0.20	0.10	0.08	0.20		
S003	f_c	346	174	216	38	353	120	140	214	19	197	211	98	0.66	20+
	σ_c	0.02	0.03	0.20	0.10	0.12	0.05	0.06	0.03	0.20	0.05	0.02	0.20		
S004	f_c	113	101	60	63	294	31	38	21	44	133	8	21	0.54	20+
	σ_c	0.20	0.03	0.20	0.05	0.08	0.08	0.15	0.20	0.20	0.06	0.20	0.20		
S005	f_c	263	198	66	46	128	131	76	47	31	299	138	54	0.56	20+
	σ_c	0.07	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.10	0.20		
S006	f_c	314	131	43	46	274	176	164	50	32	276	105	60	0.50	20+
	σ_c	0.09	0.05	0.20	0.10	0.20	0.04	0.20	0.20	0.11	0.04	0.05	0.20		
S007	f_c	70	251	141	40	1096	41	140	396	165	156	111	94	0.22	20+
	σ_c	0.06	0.20	0.20	0.18	0.14	0.20	0.20	0.20	0.20	0.01	0.20	0.20		

Table 5.3. **Parameter Estimation Results using the *s3m24* model** : Results are tabulated for the two simulated data sets (norm and weak) and the five experimental data sets (S003-S007) utilizing all 25 test directions. The algorithm was able to identify the weakness of the deep extensor muscle group (mf) using this model (note the error between f_a and f_c for the norm and weak data sets in the mf column). The parameter estimates for the simulated data sets were also able to achieve an excellent fit to the data, with a data discrepancy scores of $D(f_c) = 1.00$. The results for the experimental data sets reveal some difficulty with convergence and achieving confidence in estimates, with spread parameter values remaining relatively high for most muscles. [*KEY*: f_a: "actual" muscle strength; f_c: final strength point estimate; σ_c: final spread parameter value; $D(f_c)$: data discrepancy score for final point estimate; Iters: number of iterations to convergence]

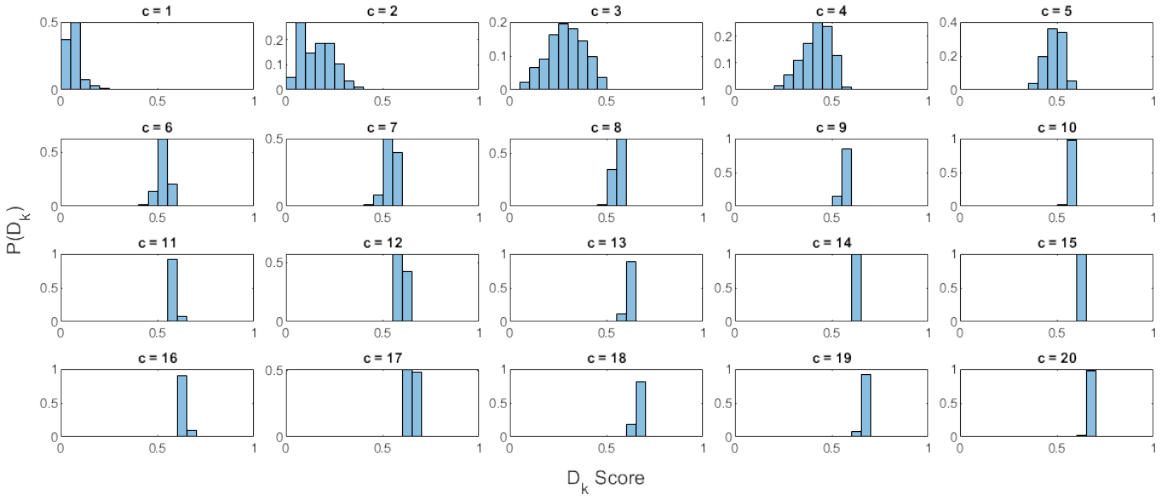
While it is not possible to verify the accuracy of the final point estimate in the simulations for the experimental data sets, a closer look at the data discrepancy scores for Participant 1 (Figure 5.14a) reveals that at the final iteration ($c = 20$), the composite scores were all between 0.65-0.70. This indicates that there was a moderate fit to the data, but error still remains. Investigating the distribution of data discrepancy scores in each direction at iteration $c = 20$ in Figure 5.14b, it is clear that there was a better fit

for some directions than others. The composite data discrepancy score of the model for the final point estimate, f_c, for 4 of the 5 participants was between $0.50 - 0.66$.

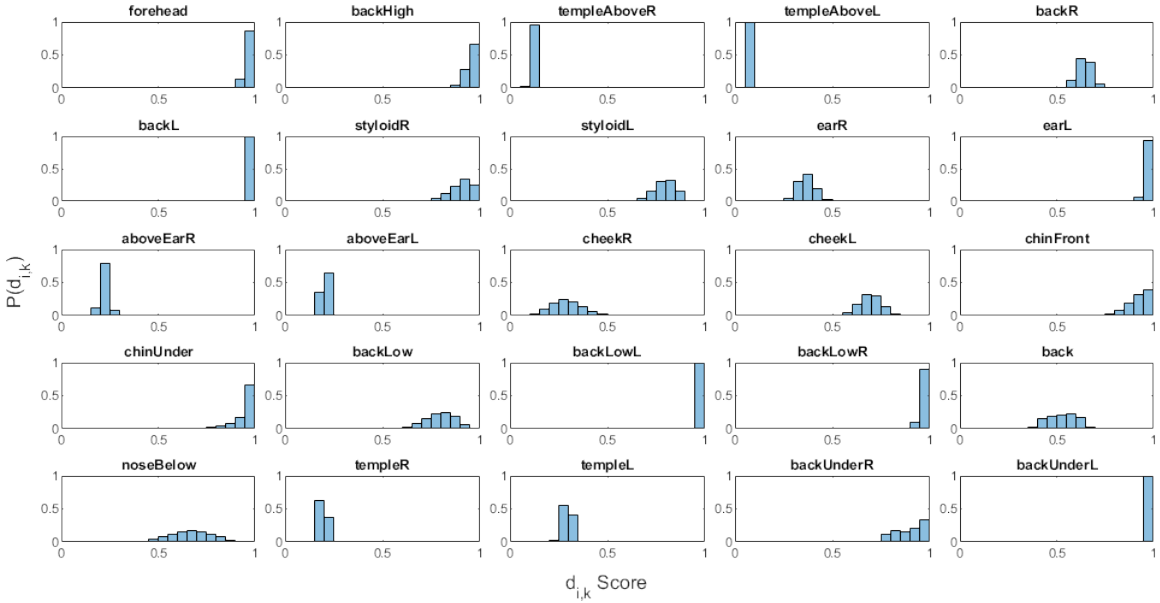
5.4. Discussion

This chapter provides a framework for parameter estimation in musculoskeletal systems and demonstrates its use in the particularly complex and redundant human neck. The model-based algorithm draws from Approximate Bayesian Inference with rejection and importance sampling and flexibly allows for inclusion of any number of muscle activation measurements that can be adjusted based on available information. The algorithm can handle moderately high dimensionality due to the incorporation of iterative sampling from a sample distribution, requiring 3^m samples for each iteration. The feasibility of the approach is demonstrated with both simulated data sets and neck strength and EMG measurements collected from 5 healthy participants.

There are several limitations to the framework. The algorithm itself does not solve the issue of multiple solutions that is inherent in a redundant musculoskeletal system like the neck. The model and data set must contain enough information to differentiate between muscles, as demonstrated in the example applications where the simulations with the planar models and data sets were prone to finding alternate solutions. The algorithm also becomes prohibitively computationally expensive for models with more than 12 independent muscles. While utilizing parallel processing will improve the speed, the computational cost increases exponentially with each additional muscle. The models presented in this chapter compute a maximum task space wrench in a specified direction based purely on torque balance, not considering other factors that may affect the maximum effort in



(a) Distribution of Composite Data Discrepancy Scores (D_k) over 15 Iterations



(b) Distribution of Data Discrepancy Scores by Test Direction at Iteration $c = 20$

Figure 5.14. **Distributions of Data Discrepancy Scores for the Experimental S003 Data Set and *s3m24* Model** (a) The distribution of all N_F composite data discrepancy scores over 20 iterations. The algorithm had not reached the convergence criteria of all $D_k > 0.95$ by the 20^{th} iteration. (B) The data discrepancy scores ($d_{i,k}$) for all 25 test directions at iteration $c = 20$ shows a wide degree of variance between error in the different test directions. Excellent fit to the data was achieved in some directions (forehead, backL, earL, backLowL, backLowR, and backUnderL), a high degree of error persists in other directions (templeAboveR, templeAboveL, aboveEarR). The algorithm was unable to find a better fit to the data given the constraints of the model.

a human experiment like coordination, understanding the task, and pain. The level of cocontraction of the superficial muscles that was measured in the experimental data sets was higher than what would be predicted by the models, indicating the models did not capture some factors contributing to muscle activation. This difference makes inclusion of the surface EMG measurements into the data set even more important, because it constrains the muscle activations to more closely match those in the experiment.

This framework requires that all measurements are collected prior using the algorithm. Particularly in a clinical setting, it would be advantageous to limit the number of measurements required. The spatial model data set included 25 test directions, which was fatiguing even for healthy participants without neck pain. A real-time iterative approach to selecting the next measurement direction based on information measures would be a practical extension of this approach. Another extension that would likely improve performance of the algorithm for individual muscle estimation would be including different postures to take advantage of muscle length-tension curves and altered moment arms to further differentiate muscles, in addition to the multiple test directions.

This chapter presented a framework for parameter estimation in musculoskeletal systems and demonstrated its use on the difficult problem of estimating individual neck muscle strengths from non-invasive wrench and EMG measurements. When provided with a rich enough data set and representative model, including spatial wrench measurements and surface EMG of a subset of muscles, the algorithm can identify weakness of the deep cervical extensor muscles. This method can be extended to other musculoskeletal systems and parameters estimation problems that are challenging to solve due to redundancy.

CHAPTER 6

Automated Test Sequence Selection for Musculoskeletal Parameter Estimation

This chapter expands on the individual strength estimator from the previous chapter. Clinicians have limited time for their evaluations and individuals with chronic WAD would not tolerate the number of neck strength measurements included in the experiment due to pain and fatigue. The next step in making this framework practical for clinical application is to reduce the number of measurements to obtain an informative estimate of individual muscle strengths. In this chapter, we propose a novel framework for musculoskeletal parameter estimation that sequentially selects the next measurement that maximizes expected information gain. We demonstrate the utility of the approach by showing with simulated data sets and 5 experimental data sets that the algorithm would provide information allowing the clinician to take fewer measurements. While limited by computational cost in its current form, this algorithmic framework is a step towards a clinically translatable test for clinician in the loop, efficient musculoskeletal parameter estimation.

6.1. Introduction

Individuals with chronic whiplash-associated disorders (WAD) present with a wide range of symptoms, which can include persistent pain, weakness, motor dysfunction, dizziness, and psychological distress [100, 99, 25]. Despite the severity of pain and disability

in WAD, no salient pathoanatomical lesion has been identified, hindering diagnosis and treatment. Imaging studies have identified increased muscle fat infiltration in the deep cervical extensor muscles, which could indicate muscle atrophy and weakness [1, 29]. The complex structure of the neck, with its kinematic and muscle redundancy, is a barrier for determining individual muscle strengths. There is currently no available method to test the hypothesis that these individuals have functional weakness of the deep extensor muscles, which would be valuable for understanding mechanisms, diagnosis, and guiding treatment decisions.

To address this need, a model-based Bayesian estimator was presented in the previous chapter to estimate individual muscle strengths from maximum isometric wrench (moment and force) measurements taken at the head. The framework involved collecting a full set of data that included 2-3 trials of maximum efforts in 25 test directions. The data was collected in full, then processed by the algorithm post-hoc to find the set of individual strength estimates that best fit the data. The framework was successful at identifying muscle strength deficits in the deep cervical extensors with the simulated data sets and the feasibility of the approach was demonstrated with experimental data sets from 5 healthy participants. *The strength testing protocol was found to be fatiguing and uncomfortable, even for some healthy individuals, making it unlikely that someone with chronic pain would tolerate the full protocol.* The next step in making this framework practical for a clinical application is to reduce the number of measurements required to obtain an informative estimate of individual muscle strengths.

There is an opportunity for clinician-in-the-loop algorithms to automate aspects of clinical examination and evaluation. These algorithms are not intended to replace the

clinicians intuition, in fact the clinician is a critical part of the framework. The complex architecture of the human neck, with its kinematic and muscular redundancy, is an ideal setting for musculoskeletal modeling and computational algorithms for parameter estimation. The complexity makes it difficult for a clinician to have good intuition about how to isolate individual muscles. A Bayesian framework allows for a priori knowledge of the clinician to be incorporated through construction of the prior and interpretation of the parameter estimate, which is presented as a probability distribution rather than a single solution. Clinicians are accustomed to working with uncertainty as they consider differential diagnoses and responses to treatment. Building on the intuition of the clinician who formulates the prior belief and hypotheses, we propose an extension of the model-based Bayesian estimator of individual neck muscle strengths that keeps the clinician in the loop, proposes the next measurement, and updates the parameter estimate based on the result of that measurement.

The primary contribution of this chapter is a method for automated sequential experimental design for individual muscle strength estimation. This implementation is the next step towards developing a framework that can be translated to the clinic. The framework utilizes a model-based likelihood-free estimator based on Approximate Bayesian Computation (ABC) algorithms [103] for parameter estimation and an entropy-reducing procedure based on infotaxis [111], a technique that seeks to maximize information gain with each new measurement. The clinician is an integral part of the process, exchanging information with the algorithm at each iteration. The algorithm optimizes the strength testing sequence, ultimately reducing the number of measurements that the clinician would need to perform to achieve a useful estimate of individual neck muscle strength. Challenges

with computational efficiency and translation of measurement methods from the lab to the clinic will need to be address before implementation in a clinical setting. For the remainder of the chapter, we will refer to the individual interacting with the algorithm as the "user". We envision the "user" being a physical therapist or rehabilitation physician for a clinical implementation of the framework in the future. The clinical implications for a tool that assists with selection of diagnostic testing is improved efficiency and accuracy in diagnosis and therapeutics. For the specific application of interest in this thesis, estimation of individual neck muscle strength in individuals with chronic WAD, reducing the number of tests is crucial because patients with WAD are not able to tolerate extensive or invasive tests due to high levels of pain.

6.2. Background

The goal of a parameter estimation problem is to produce an estimate of the set of the unknown parameters, θ, given observations X. In the previous chapter, the data set (observations) were collected independently and the parameter estimation algorithm was performed post-hoc on the full data set. In this chapter, the user is in the loop, exchanging information with the algorithm at each iteration. The algorithm proposes the next measurement for the user to take and updates the muscle strength estimate after the measurement is taken. The next sections introduce necessary background information on Bayesian inference, experimental design, and information theoretic measures that are utilized in framework.

6.2.1. Bayesian Inference

The set of unknown parameters is θ. The observations (or measurements) are represented by the random variable X, where a single observation is denoted $X = x$.

Baye's rule, shown in Equation 6.1, is used to update the belief about the parameter after data has been collected.

$$p(\theta|x) = \frac{p(x|\theta)\ p(\theta)}{p(x)} \quad \text{where} \quad p(x) = \int p(x|\theta)p(\theta)d\theta \tag{6.1}$$

The set of prior distributions, $p(\theta)$, for parameter set θ, reflects the prior beliefs about the parameter values before data is available. The likelihood function $p(x|\theta)$ relates the variables to the model and indicates how likely it is to observe the data x under the model with parameters θ. The posterior distribution $p(\theta|x)$ is the probability density function (PDF) over θ after having observed the data x. In other words, it expresses the uncertainty about the parameter set θ after the prior and data have been taken into account. The denominator, $p(x)$, is the marginal likelihood of the data and describes what the data should look like according to the model, prior to any observations. This value computes to a scalar, so that the posterior is proportional to the product of the prior and likelihood, $p(\theta|x) \propto p(x|\theta)\ p(\theta)$. Direct computation of the evidence $p(x)$, which can be computationally expensive, can be avoided by simply normalizing the numerator to sum to 1.

The algorithms in this and the previous chapter (Chapter 5) are modeled from Approximate Bayesian Computation (ABC), a family of numerical approximation techniques

for situations where the likelihood function is either computationally intractable or not available [103]. These methods approximate the likelihood function through point-wise model simulations. The likelihood function, $p(x|\theta)$, gives the probability of observing the data x given the fixed parameter values θ. The approximation of the likelihood function is achieved by creating a simulated data set through repeated model simulations with parameter values sampled from the prior, $p(\theta)$, and utilizing some metric to compare the simulated and measured data. Many specific extensions of the basic ABC framework have been developed to address challenges with sampling the parameter space and convergence, including MCMC, sequential Monte-Carlo methods importance sampling, and regression adjustments [103, 66, 101, 70]. The implementation in this chapter utilizes a relatively low parameter space ($m = 7$ muscles), so the sampling scheme simply randomly samples from the initial prior distribution.

6.2.2. Statistical Information Theoretic Approaches

Information theoretic measures have been used quite extensively in the field of optimal experimental design (OED) [95]. An objective of OED in the context of parameter estimation is to select an experiment that will produce the best statistical quality of the parameter estimates [77]. One strategy is to maximize the information gain, from prior to posterior. The field of information theory defines self-information of an event x as $I(x) = -\log(p(x))$. $I(x)$ can be thought of as the information content of the random variable X. Shannon entropy [90], the average amount of information in a probability

distribution, is computed using Equation 6.3 by taking the expected value of the self-information over all possible values of x. Shannon entropy can also be thought of as the measure of uncertainty of a probability distribution [48].

The expected value of a distribution of a random variable is the weighted sum of the probabilities over the domain of the random variable.

$$\mathbb{E}[X] = \sum_{x} x \; p(x) \tag{6.2}$$

So the average entropy over x is

$$S(x) = \mathbb{E}(I(x)) = \sum_{i=1}^{N} -p(x_i) \log(p(x_i)) \tag{6.3}$$

where X takes N possible values x_i, $i = 1, ..., N$. Note that $S \geq 0$ and $S = 0$ only if there is no uncertainty in X. The maximum value that S can have is $\ln N$, where all possible outcomes of X have equal probability (a uniform distribution over the full range of X).

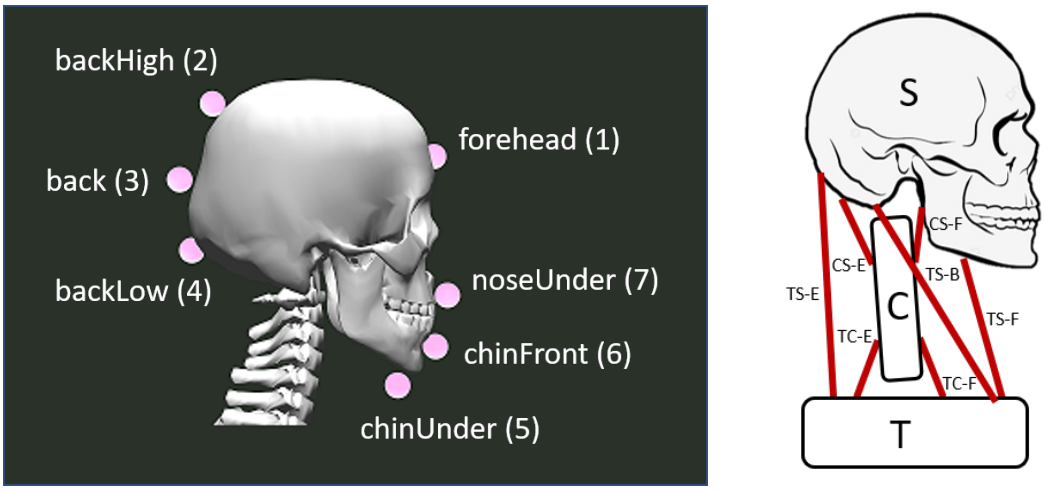
The strategy of utilizing information theoretic measures for reducing uncertainty is used in applications involving search processes and source localization, where the unknown parameter is the location of an object or chemical source. A particular entropy reducing controller designed for locating a chemical source with sparse information is infotaxis [111]. As new measurements are acquired, the algorithm updates the posterior $p(\theta|x)$, to reflect the updated belief about parameter θ based on measurement x.

In a search problem, the control authority is over where to move to collect data for the next observation. For example, if the search area is divided into a grid, the decision that the controller needs to make is which neighboring grid location to search next. The infotaxis algorithm makes this decision at each iteration based on maximizing expected entropy reduction. The pseudocode algorithm for infotaxis: entropy reducing control for source localization is presented in Algorithm 5. In this search application, the entropy measure S should be reduced at each step as more information is collected and will equal zero when the source is located.

Algorithm 5 Infotaxis - Entropy Reducing Control for Source Localization

Input: prior $p(\theta)$
Compute initial entropy: $S_0 = -\sum_x p(\theta)\log(p(\theta))$
repeat
(1) **Take Measurement** x
(2) **Compute Posterior** $p(\theta|x) = \frac{p(x|\theta)\ p(\theta)}{p(x)}$
(3) **Update Prior**: $p(\theta) = p(\theta|x)$
(4) **Compute Entropy**: $S = -\sum_x p(\theta)\log(p(\theta))$
(5) **Compute Control** $u_i = \arg\max_u \mathbb{E}[\Delta S(u)]$
(6) **Apply Control**: move to the next area according to u_i
until $S = 0$

Several important modifications to the infotaxis algorithm are required for implementation in the musculoskeletal parameter estimation application. First, the problem is higher dimensional in both the parameter and measurement space, meaning that a metric will be required to make the control decision based on the set of differential entropies. Second, the posterior update will be computed using a likelihood-free, Approximate Bayesian Computation (ABC) method due to the intractable likelihood function. Third, instead of a local optimization considering just the neighboring grid squares in the search, the next test direction will be selected from the full set of test directions with the condition that



(a) Test Locations (b) *p2m7* Model

Figure 6.1. **Labeled test locations and a schematic of the *p2m7* model**. (a) The 7 test directions in the set $\mathbf{U}_P$ are created by a force applied perpendicular to the skull surface at the location of each of the pink markers. The test directions are referred to by labelled name or direction index (in parentheses). (b) The *p2m7* model with $n = 2$ DOF and $m = 7$ muscles. The muscle names represent the attachment points (T: trunk, C: c-spine, S: skull) and action (E: extensor, F: flexor, B: both).

each direction can only be tested once. Finally, entropy is not expected to go to zero, so the stop criteria will be based on a metric on the expected gain in information.

6.3. Model

The musculoskeletal parameter estimation problem addressed in this chapter is to estimate the strength of $m = 7$ muscles of an $n = 2$ DOF neck model limited to the sagittal plane based on wrench (moment and force) measurements taken at the head and EMG measurements from a subset of muscles. The wrench measurements can be taken in 7 predefined test directions, represented by the set of sagittal plane test directions $\mathbf{U}_P = \{u_i \in \mathbb{R}^3 \; \forall \; i = 1, ..., 7\}$.

Algorithm 6 Data Generating Model

function M(f, u, $[\alpha_{lb}, \alpha_{ub}]$)

$$\max_{\alpha,\, w} \quad w$$

$$\text{s.t.} \quad \begin{bmatrix} R \ \text{diag}(f) & -J^T u \end{bmatrix} \begin{bmatrix} \alpha \\ w \end{bmatrix} = 0,$$

$$\alpha_{lb} \leq \alpha \leq \alpha_{ub}$$

return $\{w, \alpha\}$

end function

The baseline geometric and muscle parameters for the model, including the moment arm matrix $R \in \mathbb{R}^{n \times m}$ and Jacobian $J \in \mathbb{R}^{3 \times n}$, are taken from the OpenSim Neck Model [108], and simplified for the *p2m7* model as described in detail in Chapter 5. The wrench is the set of moments and forces measured at the end effector (the head), which for a planar model is $\mathbf{w} = [M_z, F_x, F_y]$. For the remainder of the chapter, we will denote a scalar magnitude of the wrench as $w = \|\mathbf{w}\|$ with unit direction $u = \mathbf{w}/\|\mathbf{w}\|$, where a bold-face $\mathbf{w}$ is the wrench vector in $\mathbb{R}^3$.

The data generating model function (Algorithm 6) computes the maximum isometric end effector wrench magnitude and set of corresponding muscle activations, $\{w, \alpha\}$, that can be produced given the set of muscle strengths, f, and wrench direction u_i. Lower and upper bounds on individual muscle activations can also be specified, $[\alpha_{lb}, \alpha_{ub}]$ if available. This replicates the physical experiment, where participants produce a maximum isometric wrench matching a target wrench direction while observing feedback on a screen. During the experiment, the wrench magnitude and a subset of muscle activations are collected. For the full experimental protocol, see Appendix A.

In the source localization problem, the unknown parameter, θ, was the location of the chemical plume, the observations, x, were the number of odor encounters, and the

algorithm had control authority over which direction to move for the next search, u_i. In the musculoskeletal parameter estimation application of interest in this chapter, the unknown parameters are individual neck muscle strengths, $\theta = f \in \mathbb{R}^m$, where m is the number of muscles in the model. The observations are neck strength (wrench) measurements and a subset of muscle activation measurements, $x = \{w_{meas}, \alpha_{meas}\}$. The control authority is over which unit wrench test direction, $u_i \in \mathbf{U}$, to measure next. The discrete set of test directions, $\mathbf{U}$, is a set of unit wrenches mapped from locations on the skull where a user can provide resist force to test neck strength (Figure 6.1b).

6.4. Algorithms

6.4.1. Sequential Muscle Strength Estimation Algorithm

The framework for musculoskeletal parameter estimation presented in Chapter 5 was centered on an algorithm drawn from ABC algorithms with rejection and importance sampling. While there are similarities in this implementation, there are some important differences that are clarified here. (1) The parameter estimate is updated sequentially as each new measurement is collected, rather than a post-hoc analysis of a full set of collected data. This allows the user to be in the loop, exchanging information with the algorithm at each iteration. (2) This algorithm does not utilize importance sampling, where a sampling function is updated to sample a new subset of the parameter space each iteration, eventually converging to a solution. Instead, candidate strength samples used across iterations are drawn from the initial prior, which is selected to encompass the full set of possible solutions.

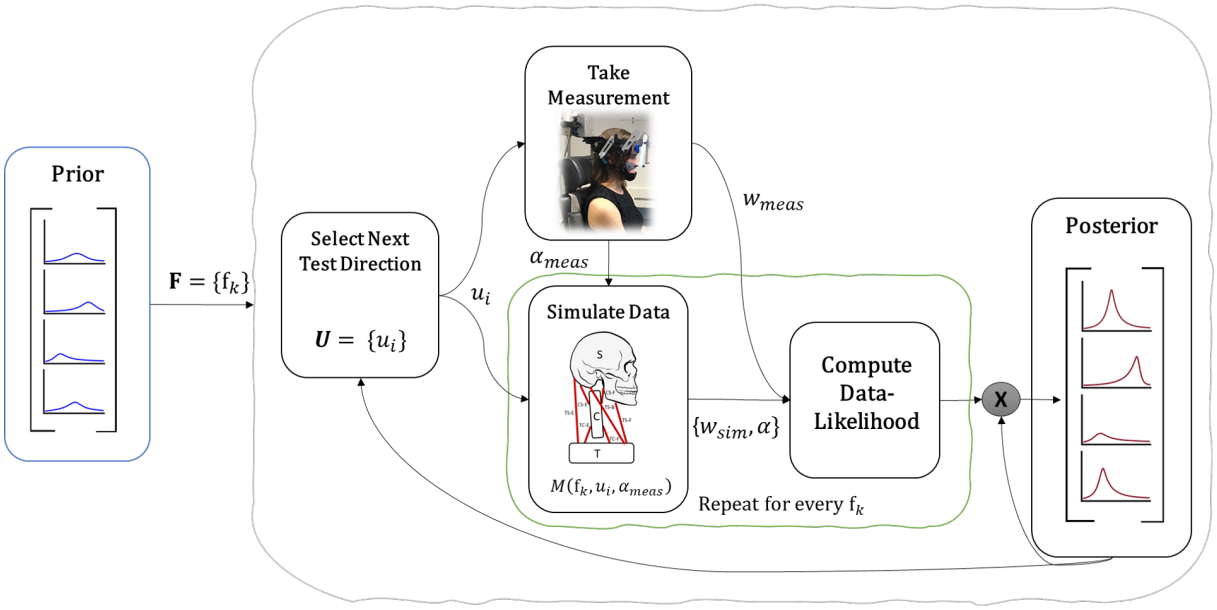


Figure 6.2. **Active Sequential Strength Estimation Algorithm**. Algorithm 7 details the steps of the algorithm visualized here. The set of candidate strength vectors **F** is sampled a single time from the initial prior distribution. At each iteration of the algorithm, a test direction is selected, a measurement is taken, data is simulated with the model, the data-likelihood function approximates the likelihood, and the posterior is updated based on the single measurement.

The Sequential Muscle Strength Estimation Algorithm, Algorithm 7, begins with defining the initial prior, $p(f) = \mathcal{N}(f_0, \sigma_0^2)$, which is a multivariate Gaussian distribution centered on the initial point estimate and variance, $f_0, \sigma_0^2 \in \mathbb{R}^m$. A Gaussian was selected to increase sampling in the region of the initial point estimate, f_c, which is an informed estimate based on anatomic studies and subject specific uniform scaling. In a context where less is known about possible parameter values, the prior could be a uniform distribution, centered on the point estimate and with lower and upper bounds representing the range of possible parameter values. A set of N_F candidate strength vectors, f_k, are randomly drawn from the prior to form the set **F**.

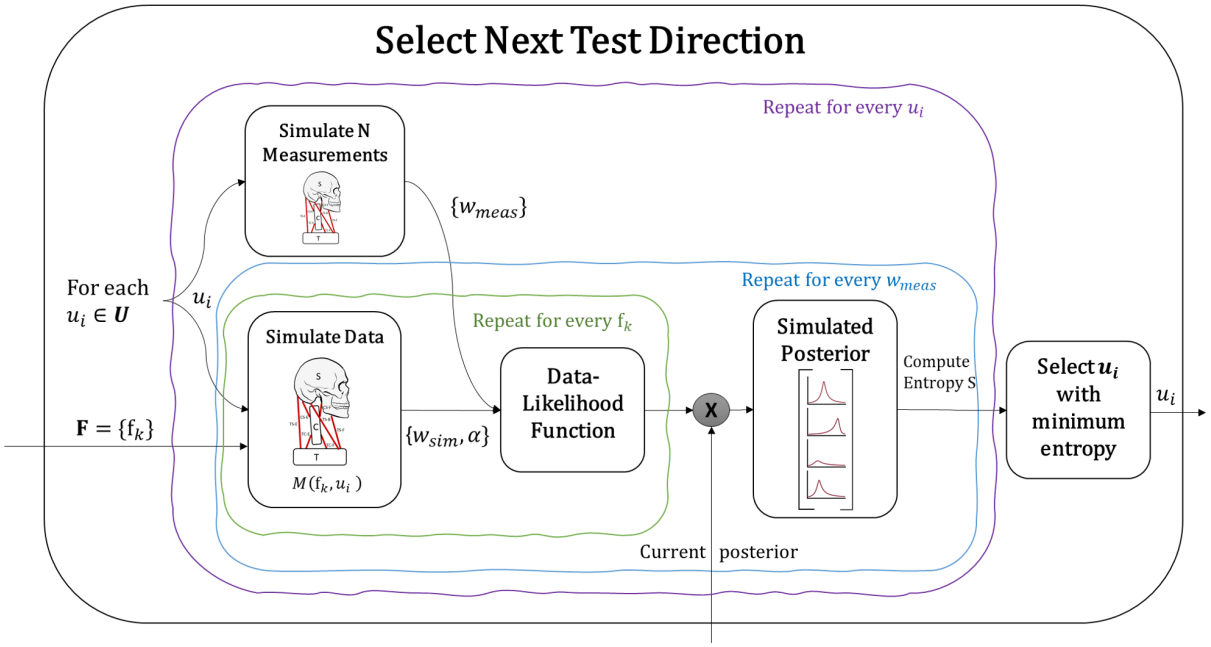


Figure 6.3. **Diagram for Select Next Test Direction Algorithm**. This algorithm is executed at each iteration of the full Active Sequential Strength Estimation Algorithm 7 shown in the *Select Next Test Direction* block in Figure 6.2. For each remaining test direction, u_i, the update to the posterior distribution is simulated for N_S simulated wrench measurements in that direction. The test direction with the minimum average entropy of simulated posterior updates is selected as the next test direction.

Each iteration begins with selection of the test direction, $u_i \in \mathbf{U}$, that maximally reduces the expected entropy of the prior using the entropy reducing controller detailed in Algorithm 9. For this application, the heuristic decision is made that each test direction can only be tested a single time to ensure coverage, so once test direction u_i has been tested, it is removed from set $\mathbf{U}$. The details of the test direction selection algorithm are explained in detail in the next section, Section 6.4.2. Once a test direction is selected, a measurement is taken by the user, $\{w_{meas}, \alpha_{meas}\}_i$, in the selected direction u_i. This measurement is the maximum wrench magnitude that the subject can produce in the specified direction, u_i. The option of including surface electromyography (EMG) measurements of

muscle activation, α_{meas}, adds more information to the data set, increasing the chance that the algorithm will converge to a single solution.

Steps 3-4 of Algorithm 7 are what classify the approach as an Approximate Bayesian Computation algorithm. Because the likelihood function is intractable, the conditional probability of the observed measurement, w_{meas}, on the unknown parameter, f, is approximated using simulated data. First, the maximum wrench and muscle activations, $\{w_{sim}, \alpha_{sim}\}_{i,k}$ in direction u_i, are simulated for each candidate strength vector, f_k, in set $\mathbf{F}$ with the model (Algorithm 6). The data-likelihood function (Algorithm 8) serves as an approximation to a likelihood function, producing $p(w_{meas}|\mathbf{F})$, the conditional probability of the observed measurement, w_{meas}, given the set of candidate strength vectors $\mathbf{F}$. The function computes the data discrepancy scores (Equation 6.4), a measure of error between the observed and simulated wrenches, to reject samples with poor performance (low data discrepancy score) and then approximates the likelihood as the sum of data discrepancy scores of the remaining samples weighted by muscle specific predicted activation levels (Equation 6.5).

The posterior update is computed as the product of the data-likelihood probability and the current prior, $p(f|w_{meas,i}) \propto p(w_{meas,i}|\mathbf{F})\ p(f)$, and then normalized to a sum of 1. For the start of the next iteration, the prior is updated as the current posterior. This process is repeated until convergence is achieved or the set of test directions, $\mathbf{U}$, is exhausted.

Note that, because the set of candidate strength vectors, $\mathbf{F}$, is sampled from the initial prior and is not updated each iteration, the simulated data and data likelihood functions are not dependent on previous iterations. This decoupling means that the order of the

tests does not change the final estimate as long as the same set of directions are measured. This design choice makes sense for our application, because we would not want to remove areas of the parameter space due to a single noisy measurement early in the experiment. However, there could be applications where it would be advantageous to resample from the updated posterior at each iteration.

The next section further details step (1) of Algorithm 7, the selection of the test direction $u_i \in \mathbf{U}$ that maximizes entropy reduction of the current estimate.

6.4.2. Entropy-Reducing Sampler for Test Sequence Selection

The motivation behind using an information theoretic metric such as maximum entropy reduction to select the next measurement is to avoid performing unnecessary measurements. In the human subject experiment presented in Appendix A, participants were asked to provide maximum efforts in 25 test directions, with multiple trials in each direction. The experiment proved to be fatiguing, sometimes uncomfortable, and time-consuming, even for healthy control participants. This algorithm, applied to that experiment, could potentially identify the most informative test directions and reduce the number of test directions required to obtain an estimate of individual muscle strength.

The Entropy-Reducing Test Selection procedure, Algorithm 9, is run at step (1) of each iteration of the Sequential Muscle Strength Estimation Algorithm. The question it answers is: which of the available test directions in set $\mathbf{U}$ will produce the greatest reduction the entropy of the posterior? In other words, which test direction will provide the most informative update? To answer this question, the *NextTest* function compares

Algorithm 7 Sequential Individual Muscle Strength Estimation

Require: $f_0 \in \mathbb{R}^m$: initial point estimate of muscle strengths
$v_0 \in \mathbb{R}^m$: variance
m: number of muscles (dimension of parameter space)
$\mathbf{U}$: set of test directions
N_F: number of f_k samples to draw from prior
N_S: number of simulated w_{meas} for entropy reduction algorithm 9

Define Initial Prior:

$$p(f) = \mathcal{N}(f_0, \sigma_0^2) \text{ where } f_0, \sigma_0 \in \mathbb{R}^m$$

Sample from Prior:

$$\mathbf{F} = \{f_k \sim p(f) \ \forall \ k = 1, ..., N_F\}$$

Precompute General Simulated Data Set:

$$\{w_{sim}, \alpha_{sim}\}_{i,k}^{g} = M(f_k, u_i) \ \forall \ f_k \in \mathbf{F} \text{ and } u_i \in \mathbf{U}$$

repeat

(1) **Select Test Direction**: $u_i \in \mathbf{U}$ ▷ Algorithm 9

$$u_i = NextTest(p(f), \mathbf{U}, N_S, \{w_{sim}, \alpha_{sim}\}^g)$$

(2) **Take Measurement**: Collect $\{w_{meas}, \alpha_{meas}\}_i$ in direction u_i

(3) **Simulate Data**: ▷ Algorithm 6

$$\{w_{sim}, \alpha_{sim}\}_{i,k} = M(f_k, u_i, \alpha_{meas,i}) \ \forall \ f_k \in \mathbf{F}$$

(6) **Compute Data-Likelihood:** ▷ Algorithm 8

$$p(w_{meas,i}|\mathbf{F}) = DataLikelihood(\{w_{sim}, \alpha_{sim}\}_{i,k}, w_{meas,i}) \ \forall \ f_k \in \mathbf{F}$$

(7) **Update Posterior**:

$$p(f|w_{meas,i}) \propto p(w_{meas,i}|\mathbf{F}) \ p(f)$$

(8) **Update Prior**: to current posterior $p(f) = p(f|w_{meas,i})$

(9) Remove u_i from set $\mathbf{U}$

until $\mathbf{U} = \emptyset$ **or** convergence criteria met

possible outcomes of the next posterior update for each remaining test direction. Specifically, the expected entropy change, the difference between the entropy of the current prior

Algorithm 8 Data-Likelihood Function

Require: $\{w_{sim}, \alpha_{sim}\}_k \; \forall \; f_k \in \mathbf{F}$: Set of simulated data
w_{meas}: wrench measurement

function DATALIKELIHOOD($\{w_{sim}, \alpha_{sim}\}$,$w_{meas}$)
(1) Compute Data Discrepancy scores: $d_k \; \forall \; f_k \in \mathbf{F}$

$$d_k = e^{-(\epsilon \delta_k)^2} \text{ where } \delta_k = \| \frac{w_{sim,k} - w_{meas}}{w_{meas}} \| \tag{6.4}$$

(2) Rejection Step: keep subset $\mathbf{F'} \subseteq \mathbf{F}$ where $d_k \geq q$ percentile
(3) Compute Data-Likelihood: over domain f, for each muscle j

$$p(w_{meas}|\mathbf{F'}) \propto \sum_{k}^{N_{F'}} \alpha_{j,k} \; d_k \tag{6.5}$$

return $p(w_{meas}|\mathbf{F'})$
end function

and the simulated posterior updates, is compared across test directions, and the direction with the maximum predicted reduction in entropy is selected as the next test direction.

Computing the expected entropy change through forward simulation can be prohibitively computationally expensive, particularly in a high dimensional parameter space. Two important modifications are made to reduce the computational cost. First, a set of baseline simulated data is precomputed prior to the start of the repeat loop in Algorithm 7. This set contains a simulated wrench magnitude and activation pattern, $\{w_{sim}, \alpha_{sim}\}_{i,k} = M(f_k, u_i)$ for each $f_k \in \mathbf{F}$ and $u_i \in \mathbf{U}$. Because it is computed prior to any experimental data collection, measured muscle activations are not input to the model, so it defaults to a range of [0 1]. This data set is used in the Entropy-Reducing Test Sequence Selection function for the forward simulation step, avoiding the need to repeat the costly simulations at each iteration. The second modification that decreases computation cost is to forward simulate the posterior update with only a subset of N_S

randomly selected wrenches from the simulated data set, $w^s_{meas,i} \sim \{w_{sim}\}_i$. This subset can be orders of magnitude lower than N_F, the number of candidate muscle strength samples used to approximate the likelihood function.

6.5. Application

The implementation of the framework is demonstrated through our example problem initially introduced in Section 6.3. The objective is to estimate the strength of neck muscles from wrench (moment and force) measurements taken at the head and EMG measurements from a subset of muscles using planar neck model with $m = 7$ muscles and $n = 2$ DOF. The wrench measurements can be taken in 7 predefined test directions, represented by the set of sagittal plane test directions $\mathbf{U}_P = \{u_i \in \mathbb{R}^3 \ \forall \ i = 1, ..., 7\}$ and visualized in Figure 6.1. The parameter estimation results incorporating all 7 test direction measurements are presented first in Section 6.5.2. Then, the utility of the entropy-reduction sampler for automated test sequence selection is presented in Section 6.5.3. We show that the framework gives users the opportunity to take fewer measurements by selecting the most informative tests first.

6.5.1. Implementation

The algorithm was implemented for a simulated data set, where the "actual" muscle strengths are known, and for 5 experimental data sets. To produce the simulated data set, the "actual" muscle strength vector $f_a \in \mathbb{R}^m$ was formed by pseudo-randomly perturbing the f'_{base} baseline strength estimate for the *p2m7* model by scaling by a factor 1.5 and adding random Gaussian noise with a variance of $v = 0.4 f_a$. The simulated data set

Algorithm 9 Entropy-Reducing Next Test Selection

Require: $p(f)$: current prior
$\mathbf{U}$: set of test directions
N_S: number of simulated samples
$\{w_{sim}, \alpha_{sim}\}_{i,k} = M(f_k, u_i)\ \forall\ u_i \in \mathbf{U}$ and $f_k \in \mathbf{F}$: set of simulated data

function NEXTTEST($p(f)$, $\mathbf{U}$, N_S, $\{w_{sim}, \alpha_{sim}\}$)

Compute Current Entropy:

$$S_c = -\sum p(f) \log p(f)$$

for all $u_i \in \mathbf{U}$ **do**

Randomly sample N_S wrenches $w^s_{meas,i} \sim \{w_{sim}\}_i$ in direction u_i

for all $s = 1, ..., N_S$ **do**

Compute Data-Likelihood: ▷ Algorithm 8

$$p(w^s_{meas,i}|\mathbf{F}) = \mathit{DataLikelihood}(\{w_{sim}, \alpha_{sim}\}_i, w^s_{meas,i})$$

Compute Simulated Posterior:

$$p(f|w^s_{meas,i}) \propto p(w^s_{meas,i}|\mathbf{F})\ p(f)$$

Compute Entropy of Simulated Posterior:

$$S^s_{sim,i} = -\sum p(f|w^s_{meas,i}) \log p(f|w^s_{meas,i})$$

end for

Compute Average Entropy Change for direction u_i:

$$\Delta S_i = S_c - \frac{\sum_s^{N_S} S^s_{sim,i}}{N_S} \quad \text{where} \quad \Delta S_i \in \mathbb{R}^m \tag{6.6}$$

end for

Find test direction with maximum entropy reduction (across m muscles):

$$\max_i \ \frac{\sum_j^m \Delta S_{i,j}}{m} \text{ for } j = 1, .., m \tag{6.7}$$

return test direction u_i

end function

was computed using the model function in Algorithm 6: $\{w_{meas}, \alpha_{meas}\}_i = M(f_a, u_i)$ for all 7 test directions in the sagittal plane, $u_i \in \mathbf{U}_P$. The experimental data sets were

collected from 5 healthy participants as described in Chapter A. Only the subset of data from the 7 sagittal plane test directions was used. Muscle activation measurements for the sternocleidomastoid (SCM) and semispinalis capitis (SSCap) muscles were mapped to the TS-B and TS-E muscles of the *p2m7* model (See Figure 6.1).

The initial prior was selected using the process described in the previous chapter (section 5.2.2.1), reiterated briefly here. The baseline point estimates, $f'_{base} \in \mathbb{R}^7$, for the *p2m7* model are the 2-norm of the individual muscle elements that make up the muscle groups (derived from the OpenSim model). For each data set, the baseline point estimate was bootstrapped to the data by uniformly scaling by the average scalar value, a, that minimized the difference between the measured wrench magnitude, w_{meas} and simulated wrench magnitude, w_{sim}, across the 7 test directions. The simulated wrench measurements were computed from the model function in Algorithm 6, $w_{sim} = M(af'_{base}, u)$. The prior was then set as the multi-variate Gaussian $p(f) = \mathcal{N}(f_0, v_0)$, with $f_0 = af'_{base}$ and $v_0 = 0.3f_0$.

The prior was randomly sampled to create a set of candidate strength vectors $\mathbf{F}$ of size $N_F = 500,000$. The number of simulated wrench measurements for the Entropy-Reducing Next Test Selection Algorithm (Algorithm 9) was set at $N_S = 100$.

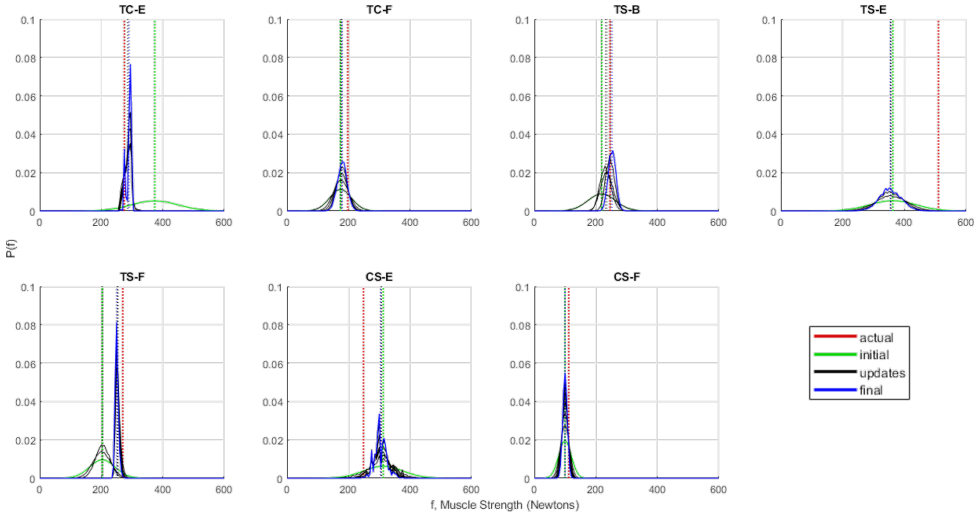
6.5.2. Parameter Estimation Results

Muscle strength estimation results for the simulated and experimental data sets are tabulated in Table 6.1. Figures 6.4 and 6.5 show the (a) posterior updates and (b) iterative point value estimates and variance for the simulated data set and experimental data set for participant S005, respectively. The algorithm produced estimates with low variance

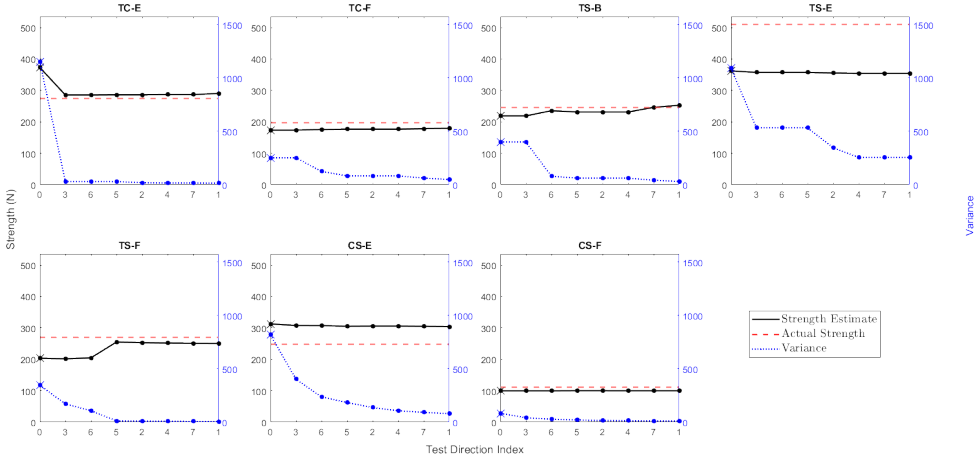
		TC-E	TC-F	TS-B	TS-E	TS-F	CS-E	CS-F	D(f_c)
Sim	f_a	275	197	247	510	270	249	112	0.97
	f_c	290	180	253	354	251	304	100	
	v_c	14	49	30	255	6	80	10	
S003	f_c	284	104	146	191	139	161	59	0.23
	v_c	20	35	35	118	30	118	10	
S004	f_c	125	39	56	62	59	24	21	0.41
	v_c	9	29	31	107	13	20	10	
S005	f_c	160	86	131	165	119	117	48	0.58
	v_c	2	33	22	149	13	133	10	
S006	f_c	297	97	142	196	131	149	57	0.23
	v_c	17	34	31	148	24	157	11	
S007	f_c	303	104	142	226	125	177	63	0.44
	v_c	14	29	38	136	26	117	10	

Table 6.1. **Results of the simulations for the simulated data set (Sim) and the 5 experimental data sets (S003-S007) with the *p2m7* model**. The 7 muscle groups and the data discrepancy value for the final point estimate, D(f_c), are the column headers. f_a is the "actual" strength of the muscles in the simulated data set. f_c is the final individual muscle strength point estimate, computed as the expected value of the posterior distribution after all 7 measurements have been considered. v_c is the variance of the same final posterior distribution. The data discrepancy score, D(f_c), is a measure of fit of the final point estimate to the measured data, where D(f_c) = 1 would be a perfect fit (no error). The data discrepancy score for the simulated data set indicates an excellent fit, but the fit was much lower for the experimental sets, implying that the model may not capture the full biomechanics of the human subjects. A similar pattern of variance emerges among the data sets, where the estimates for particular muscles have a relatively low variance (TC-E) and others have a relatively high variance (TS-E).

for a subset of muscles but indicates less confidence in the estimates of others. Comparing the subplots for muscle group TC-E vs. TS-E in Figure 6.4a, the final posterior PDF (in blue) for muscle TC-E has a low variance of 13.9 N^2 and is closely aligned with the "actual" strength (red vertical line). In contrast, the final posterior PDF for muscle

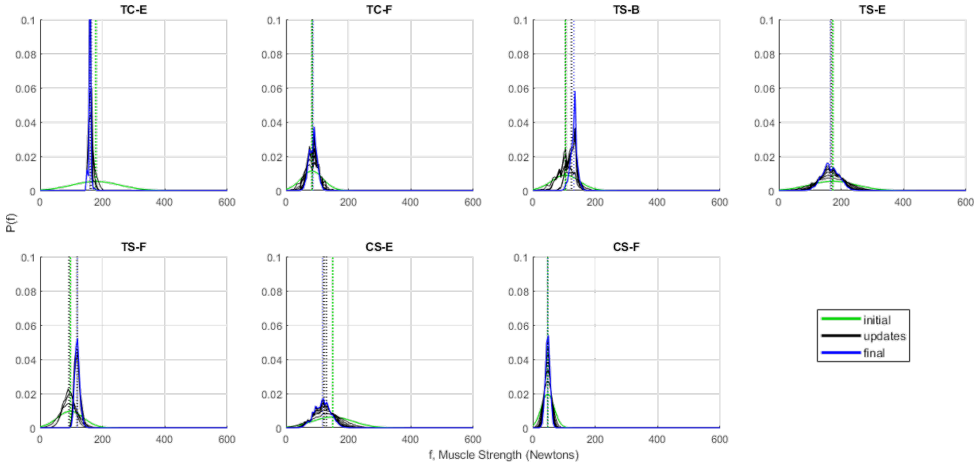


(a) Posterior Updates

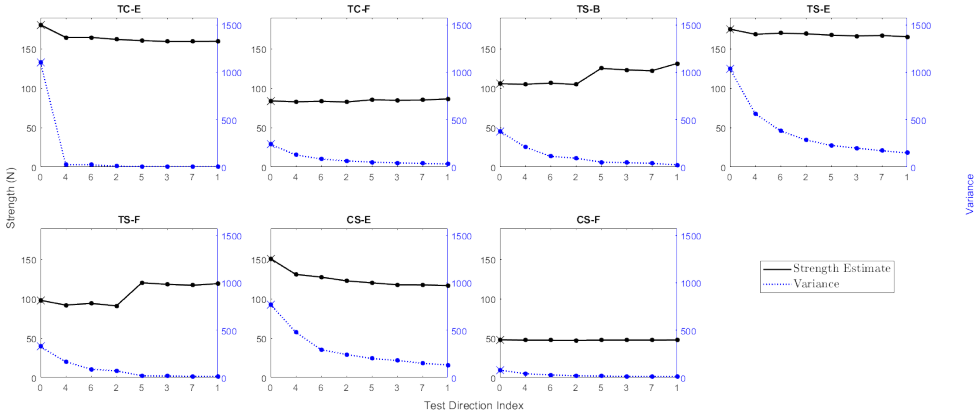


(b) Expected Value and Variance by Iteration

Figure 6.4. **Parameter estimation results for the simulated data set**. Because this is a simulated data set, the "actual" strength is known and shown in red. (a) The posterior PDFs for each of the 7 muscles over all 7 iterations. The expected value is represented by the dashed vertical lines. (b) The expected value of the posterior PDF (black) and variance (blue) over the 7 iterations. The values on the x-axis refer to the test direction index (see Figure 6.1). Note that the algorithm was able to achieve low variance for some muscles (TC-E, TS-F, CS-F) but not others (TS-E, CS-E) and that those with greater error in the point estimate also have a larger variance.



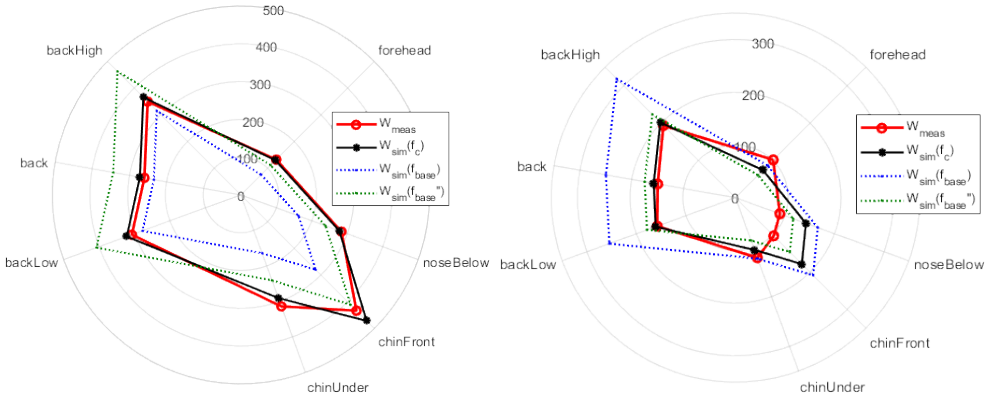
(a) Posterior PDF



(b) Expected Value and Variance

Figure 6.5. **Parameter estimation results for the experimental data set of Participant S005.** (a) The posterior PDFs for each of the 7 muscles over all 7 iterations. The expected value is represented by the dashed vertical line. (b) The expected value of the posterior PDF (black) and variance (blue) over the 7 iterations. The values on the x-axis refer to the test direction index (see Figure 6.1). Note that the algorithm was able to achieve low variance for some muscles (TC-E, TS-F, CS-F) but not others (TS-E, CS-E).

TS-E is close to the initial estimate (in green) with a variance of 255.2 N^2 and an error of 155.6 N between the expected value of the final posterior and the "actual" strength. Although the ideal result would be a confident estimate in all muscle strengths, this result



(a) Simulated Data Set (b) Experimental Data Set

Figure 6.6. **Radar plots showing the fit of the final parameter estimate to the (a) simulated data set and (b) experimental data set for subject S005.** The point estimate, f_c, is the expected value of the final posterior PDF. The black trace is the set of maximum wrenches computed by the model with strengths f_c. The blue and green traces are the wrenches computed with model for the baseline strength and scaled baseline strengths respectively. The algorithm was able to find a set of muscle strengths that fit the simulated data set (a) accurately. The fit of the parameter estimates for (b) was good in some directions (backHigh, back, and backLow), but less accurate for others (particularly noseBelow and chinFront).

is informative, revealing that the data did not provide enough information to distinguish between all possible solutions in this redundant system.

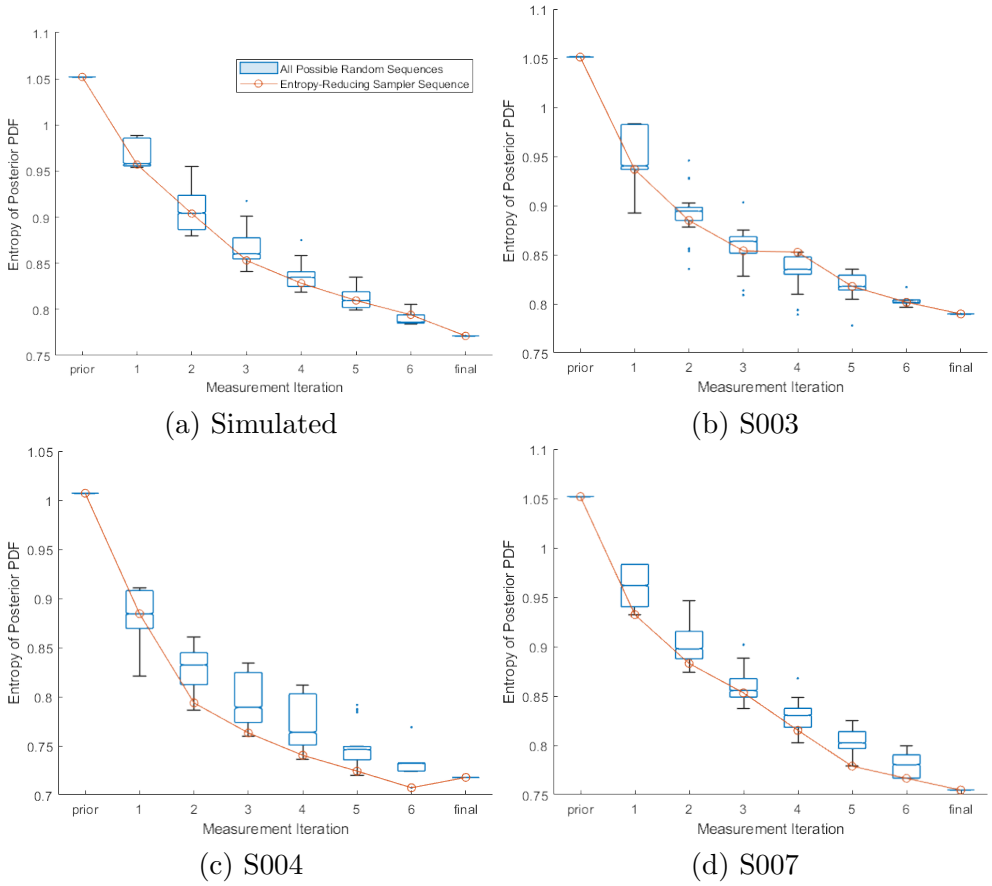
A point estimate for muscle strength is computed as the expected value of the posterior PDF of each muscle. The fit of the final point estimate to the data set was analyzed by computing the set of maximum wrenches that the model could produce with final point estimate, f_c, and comparing these values to the measured wrenches in the data set. Figure 6.6 shows the fit of the final point estimate for (a) the simulated data set and (b) the experimental data set for subject S005. The final point estimates received data discrepancy scores of 0.97 and 0.58 for the simulated and experimental data sets respectively, where a score of 1 indicates no error.

The variance in the final posterior PDFs for the simulated and experimental data sets exhibit a similar pattern across muscles. Although the "actual" strength of the muscles in the experimental data set is unknown, the higher variance for muscles TS-E and CS-E (Figure 6.5b) give an indication that there is less confidence in the estimate for these muscles. Also note that there is very little change in both the expected value and variance in the final 3 iterations for all muscles in Figure 6.4b. The implications of this plateau, related to the automated test sequence selection, are discussed in the next section.

6.5.3. Performance of the Automated Test Sequence Selection

The utility of the entropy-reducing sampler for automated test sequence selection is demonstrated by (1) comparing iterative results of the algorithm utilizing the entropy-reducing sampler vs. using a random test sequence and (2) showing that the number of tests can be reduced without significant change in outcome when using the entropy-reducing sampler to select the most informative tests.

Figure 6.7 shows the average entropy of the posterior PDFs at each iteration for the entropy-reducing sampler sequence (red trace) and all possible random test sequences (blue box plots). For 7 test directions, there are $7! - 1 = 3049$ permutations to represent all possible sequences (excluding the sequence selected by the entropy-reducing sampler). The entropy-reducing sampler selected test directions performed well on the experimental data sets for participants (c) S004 and (d) S007, reducing entropy of the posterior PDFs more rapidly than 75% of random sequences in all but one iteration. It performed less well for the (a) simulated data set and (b) experimental data set of participant S003, producing posterior PDFs with entropy higher than the median of random sequences



(a) Simulated (b) S003

(c) S004 (d) S007

Figure 6.7. **Entropy of the posterior PDF by iteration for (a) the simulated data set and (b-d) experimental data sets for participants S003, S004, and S007.** The box plots show the distributions of the results for all other possible test sequences (permutations of 1 through 7, producing $(7! - 1) = 5039$ random test sequences). The red trace represents the results for the sequence selected by the entropy-reducing sampler. The ideal result would show the entropy-reducing test sequence producing posteriors with the lowest entropy at each iteration. The sampler performed well for (c) participant S004 and (d) participant S007, producing posterior updates with entropy in the lowest quartile for all but one iteration. The sampler performed less well for (a) the simulated data set and (b) participant S003, with entropy values closer to the median value in most iterations. The sampler selects the sequence based on average predicted change in entropy across simulated measurements. The actual change in entropy results from the value of the measurement and may not match the prediction.

in several iterations. An optimal result would be if the entropy of the posterior PDFs in the sequence selected by the sampler was less than all random sequences at each iteration. However, the entropy-reducing technique is optimal on average, not for every measurement result. The average entropy change (Equation 6.6) averages across the entropy of the forward simulated posterior distributions created by all $N_S = 50$ randomly sampled simulated wrench measurements. The posterior update that results from an individual measurements is not guaranteed to follow the prediction. Note that the entropy of the prior and final iterations do not depend on the test sequence. If all 7 test directions are measured, the final estimate will be the same regardless of test sequence when using this algorithm. However, if only a subset of test directions are measured, the test directions included in that subset will affect the parameter estimate. The entropy-reducing sampler prioritizes the test directions that it predicts will produce the greatest decrease in average entropy of the posterior PDFs, resulting in a more rapid decrease in entropy (decrease in uncertainty) in the estimates.

We have demonstrated that the automated test sequence selection algorithm produces a sequence of tests that reduces the average uncertainty in the individual strength estimates more rapidly than most random test sequences. This result is beneficial if it enables us to predict when further testing would not improve the parameter estimate so that fewer tests can be performed.

Further investigation into the expected entropy change provides insights to aid decisions as to when to stop taking measurements. The expected change in entropy, ΔS from Equation 6.6, is an $N_U \times m$ matrix (in this case, $N_U = 7$ test directions by $m = 7$ muscles) that represents the predicted change in entropy of the posterior for each individual muscle

in each remaining test direction. The expected entropy change over all 7 iterations for the experimental data set of subject S005 is shown in Figure 6.8. It is difficult to define a cutoff for expected entropy reduction that justifies taking an additional measurement, as this is context dependent, with factors to weigh such as time, cost, level of acceptable confidence, and patient comfort, among others. However, a relative comparison between the expected entropy reduction from a measurement in any test direction selected in Iteration 1 and those remaining in Iterations 5-7, reveals there is likely little additional benefit in continuing to take measurements.

Figure 6.9 confirms what was identified in the analysis of expected entropy reduction by iteration in Figure 6.8 and the actual entropy reduction in Figure **??**, that minimal change occurs in the posterior PDF after the 4^{th} measurement. The posterior PDF after 4 measurements is traced in blue and the posterior PDF after the full 7 measurements is traced in black. The point estimates, computed by the expected values, are indicated by the vertical dashed lines. The blue and black traces nearly completely overlay each other, indicating that there was little change produced by the 3 final measurements. This indicates that the number of tests can be reduced without a significant change in outcome when using the automated test sequence algorithm.

The entropy-reducing sampler for automated test sequence selection provides a method for selecting the most informative tests and determining at what point to stop taking measurements. Clinical tests, such as the neck strength measurements in our application, can be time consuming and uncomfortable for the patient. This framework allows for the number of tests to be minimized, while providing information for the user to make informed decisions.

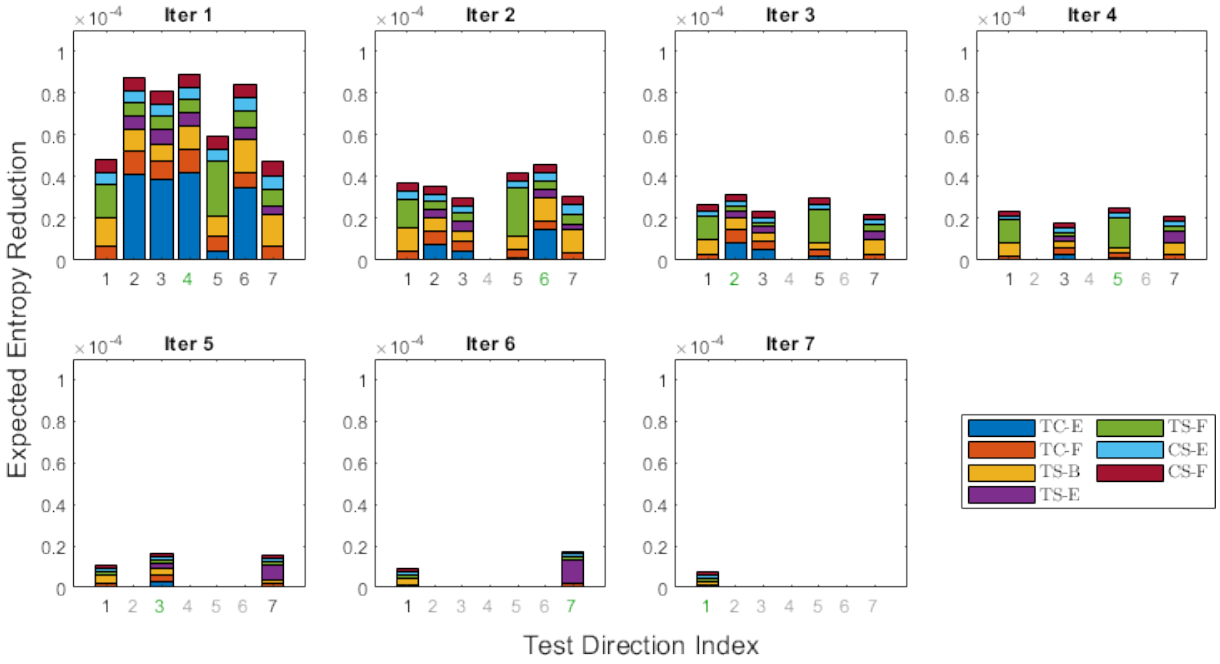


Figure 6.8. **Expected entropy reduction of the posterior PDF** (ΔS) from Equation 6.6. Each subplot presents ΔS for a single call to the function *NextTest* in Algorithm 9 over the course of 7 iterations. The x-axis shows the test direction indices, with the selected index in green, and the unavailable (already previously selected) indices in grey. The y-axis is the total expected entropy reduction of the posterior PDFs from forward simulations. Each element in the stacked bars represent the contribution of an individual muscle group, defined by color in the legend. This information could be utilized by the user to weigh the benefit vs. costs of taking an additional measurement. For example, the total reduction in entropy predicted for Iteration 5, in the direction selected by the entropy-reducing sampler (test direction 3), is small when compared to the scale of the changes predicted in the previous iterations. By the 5^{th} iteration, there is little benefit to taking additional measurements.

6.6. Discussion

6.6.1. Challenges and Limitations

There are several limitations to this experimental design and parameter estimation framework in its present form. First, the scalability to higher dimensional parameter spaces is

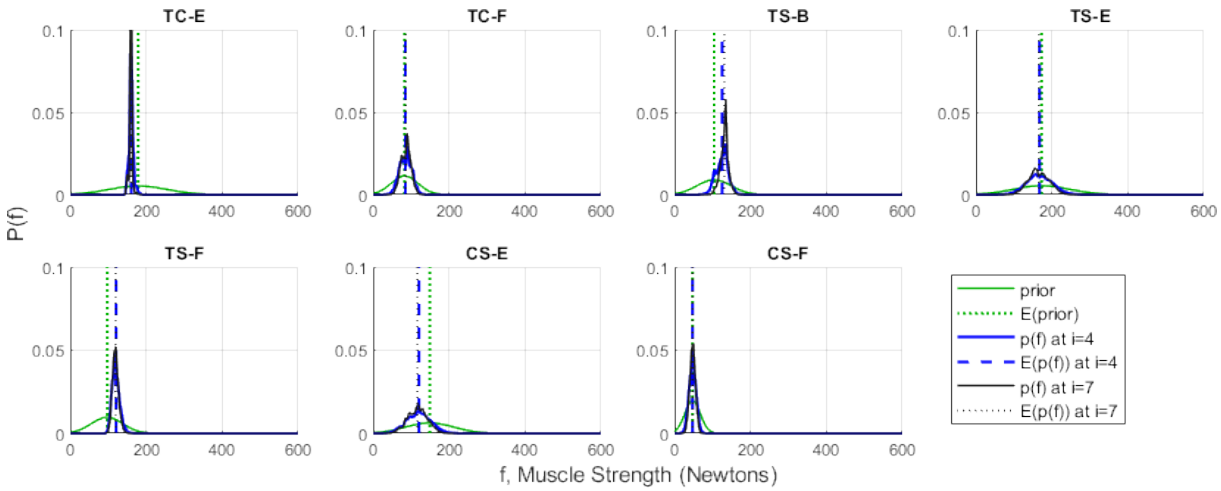


Figure 6.9. **Posterior PDFs for each of the 7 muscles after 4 measurements** (blue traces) **and after 7 measurements** (black traces) **for the experimental data set of Participant S005**. The vertical dashed lines represent the expected value of the posterior PDF . The prior is shown in green. The black and blue traces almost entirely overlay one another, indicating that there is very little change in the posterior PDFs between measurement 4 and measurement 7. This confirms the conclusion that the final 3 measurements add little additional information to the parameter estimate.

limited due to computational and memory cost associated with sampling and computing expected entropy on forward simulations in addition to the parameter estimation step. Implementation of the importance sampling techniques from Chapter 5, as described in the previous section, would partially address this issue. However, the algorithm needs to run sufficiently fast to provide the next test direction to measure during clinical testing. Additional solutions for improving efficiency also need to be considered.

Another limitation of this, and any other framework that would attempt to reduce the number of measurements, is that the algorithm cannot predict a particularly unexpected measurement that would dramatically change the strength estimate. Figure 6.10 shows an example of a late change in strength estimate for muscle group TC-E after the 5^{th}

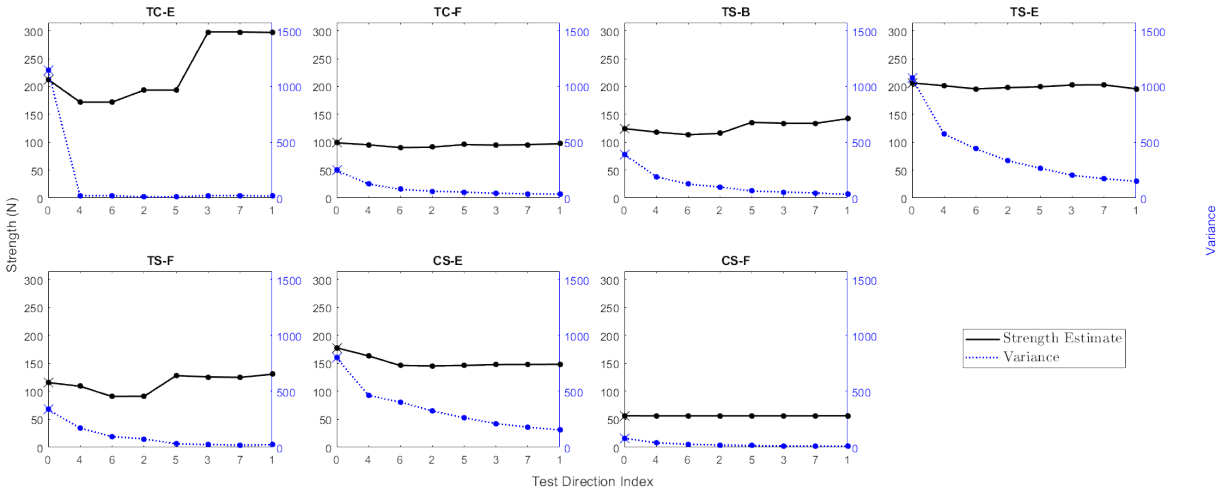


Figure 6.10. **Expected value** (black) **and variance** (blue) **of the strength estimate over 7 iterations with the experimental data set from Participant S006 using model *p2m7***. The values on the x-axis refer to the test direction index - see Figure 6.1. Note the large jump in the point estimate of the strength for muscle group TC-E at the 5^{th} estimate (test direction index 3). If only 4 measurements had been taken, the strength estimate for this muscle would be more than 33% lower. Despite a low variance, a single unexpected measurement can change an estimate dramatically.

measurement. If the user had decided to stop testing after 4 measurements due to low expected entropy change with the remaining test directions, the point estimate for muscle TC-E would be 33% lower.

A strength of the Bayesian approach is that it is able to represent the uncertainty in the parameter estimates, which is particularly relevant in the muscularly redundant human neck. This framework does not "solve" the problem of redundancy, but it does present user with the full space of solutions so they can use their domain knowledge to make informed decisions based on the information available. It is important to note that the point estimate of the muscle strength estimate, which is computed by taking the expected value of the posterior PDF, is not guaranteed to represent the best solution

or even the most probable solution. The parameter estimate in this framework is a set of probability distributions that can be interpreted by a clinician in the context of the uncertainty of the estimates and the presentation of the patient in front of them.

6.6.2. Extensions and Future Directions

There are opportunities to extend this framework to allow for integration of other sources of information or parameter estimation applications. One example is the situation where domain knowledge from a clinician is available regarding which muscles are priorities for improving strength estimates. For example, the clinician could have interest in a particular muscle due to a known mechanism of injury or nerve injury. In this case, Algorithm 9 can be modified to include an input, $\gamma \in \mathbb{R}^m$ that weights the importance of each muscle in the maximum entropy computation in Equation 6.7. The equation would then be rewritten as:

$$\max_i \ \frac{\sum_j^m \gamma_j \ \Delta S_{i,j}}{m} \tag{6.8}$$

This simple modification allows input from the user to prioritize particular muscles over others.

Other physiological based insights can be incorporated into the algorithm to improve performance in a context-specific manner. The selection of the prior distribution is one way that the clinician can add insight to the framework. The candidate strength vectors are sampled from the prior distribution, so it is important to make sure that the prior covers the possible range of strength values. One way to improve the algorithm's ability to distinguish between muscles would be to add different neck postures in addition to the

set of test directions. The force-length relationship affects maximum muscle force [87] and could help further reduce the space of solutions by selectively biasing particular muscle groups.

The sequential approach presented in this chapter, in which the next test direction is proposed by the algorithm and a measurement is taken at each iteration, can be combined with the importance sampling techniques employed in the previous chapter. This would create a hierarchical scheme, where within each measurement iteration, the estimate is fit to the data through iterative sampling from the sampling distribution. This modification to the sampling would allow the algorithm to be used with higher dimensional parameter spaces, such as the 24 muscle model in Chapter 5.

6.6.3. Clinician-Algorithm Interface

The interface between the clinician and the algorithm in the proposed framework requires active input and interpretation on the part of the clinician. The role of the algorithm is to compute and update probability distributions representing the parameter estimates and suggest the next test. The clinician is responsible for interpreting these outputs in the context of the patient in front of them, while being aware of the limitations of any model-based simulation of a complex musculoskeletal system. This would require training, but as described in the introduction of this chapter, Bayesian inference should be intuitive to a clinician as it mimics the clinical decision-making process. Cooperation between the algorithm and clinician can be facilitated by ensuring sufficient control of the procedure by the clinician, who is the domain expert using the algorithm as a tool.

6.7. Conclusion

This chapter provides a framework for automated test sequence selection for model-based estimation of individual muscle strengths with the clinician in the loop. The method is applied to estimate the strength of individual muscles based on neck strength measurements taken at 7 test locations on the head. It was demonstrated that the entropy-reducing sampler for test sequence selection optimizes the testing sequence, ultimately reducing the number of tests that the clinician would need to perform. Before this framework can be used in a clinical context, the computational efficiency of the algorithm would need to be improved. The clinical implications of a tool that optimizes selection of clinical tests is the potential to improve efficiency and accuracy in diagnosis and therapeutics. For the specific application of interest in this thesis, estimation of individual neck muscle strength in individuals with chronic Whiplash Associated Disorders (WAD), reducing the number of tests is crucial because patients with WAD are not able to tolerate extensive or invasive tests due to high levels of pain.

CHAPTER 7

Conclusions and Future Directions

This work began with a motivation to expand our understanding of muscle changes and motor dysfunction in chronic whiplash-associated disorders (WAD) in order to improve clinical outcomes and advance diagnostic and treatment regimes. We began with a relatively straightforward, clinically motivated question: Are the deep cervical extensor muscles functionally weak in chronic WAD? Analyzing the functional impact of compositional muscle changes requires the ability to isolate the strength of individual neck muscles. The lack of an existing framework for in vivo measurement of individual muscles strengths presented both a barrier to our progress and an opportunity for a wider contribution to the field of musculoskeletal rehabilitation. A multi-disciplinary approach was employed, drawing from the fields of physical therapy, quantitative magnetic resonance imaging, musculoskeletal modeling, Bayesian inference, optimal experimental design, and information theory. In the end, the major contribution of this thesis is a framework for clinician-in-the-loop individual muscle strength estimation utilizing non-invasive measurements in a complex musculoskeletal system.

The first step towards unravelling the muscle changes and motor dysfunction in chronic WAD was to confirm and expand on the evidence of muscle composition changes that appears to be unique in individual with chronic WAD. In **Chapter 3**, we mapped the magnitude and spatial distribution of muscle fat infiltration (MFI) within the deep cervical extensor muscles in individuals with chronic WAD, recovered WAD, and healthy

controls. Results confirmed an overall increase in MFI in these muscles in the chronic WAD group, further motivating investigations into the functional implications of these compositional changes. In **Chapter 4**, we investigated the functional role of the deep cervical extensor muscles in producing multi-directional neck strength and examined how weakness of this muscle group affects overall strength and muscle activation patterns utilizing a musculoskeletal model of the neck. We demonstrated how kinematic and muscle model complexity impacts the predicted neck strength and muscle activations. The results contribute to our understanding of the function of the deep cervical extensor muscles, but no clear set of test directions for isolating the strength of the deep cervical extensor muscles emerged.

The next two chapters develop frameworks for determining individual muscle strengths from neck strength measurements. In **Chapter 5**, a parameter estimation algorithm with an Approximate Bayesian Computation (ABC) with rejection and importance sampling approach was developed to infer the strength of individual neck muscles based on a pre-collected data set of multi-directional neck strength and electromyography (EMG) measurements. The algorithm was able to identify deep extensor weakness when utilizing the full spatial neck model and 25 test directions with simulated data sets. Feasibility of the approach was demonstrated with experimental measurements from 5 healthy participants. **Chapter 6** presented a sequential framework to include the clinician in the loop and an entropy-reducing sampler to select the next test direction that provides the most informative update to the individual muscle strength estimates. Feedback regarding expected information gain is provided to the clinician to make informed decisions about

when to stop collecting data, which is particularly valuable when testing individuals with chronic neck pain.

7.1. Technical Challenges

Several technical challenges remain to be addressed before the frameworks presented in the final chapters can firmly answer the motivating question: are the deep cervical extensor muscles weak in WAD? These challenges are detailed here, and a road map for addressing the challenges is presented.

7.1.1. Scalability and Efficiency

The framework for clinician-in-the-loop sequential individual muscle strength estimation presented in **Chapter 6** was implemented on the 7 muscle *p2m7* model limited to the sagittal plane due to the computational cost of increasing dimensionality of the parameter space. The *s3m24* model, with 25 test directions and 24 muscle groups performed better in **Chapter 5** to identify weakness of the deep extensor muscles. However, even with the simplified *p2m7* model, the computation time is unrealistic for real-time use in the clinic in its current state. The following extensions to the framework are proposed to address the scalability and efficiency limitations.

7.1.1.1. Sampling Methods. A major drawback of ABC approaches is their immense computational cost, requiring repeated simulation of the model over sampled parameter values. The additional entropy-reducing test sequence selection component of the sequential individual muscle strength estimation framework adds additional computational cost.

It requires averaging over forward simulated data sets in order to estimate expected information content. For this reason, Bayesian experimental design approaches have mainly focused on linear models with conjugate priors and tractable likelihood functions [38]. There is an inherent trade-off between computational and statistical efficiency. Reducing the number of samples reduces the computational cost, but can also decrease the accuracy and confidence of the estimates. Sampling strategies have been developed to overcome this limitation by intelligently sampling the parameter space in areas of higher likelihood.

The sampling strategy used in Chapter 6 consisted of simply sampling parameter values randomly from the prior distribution. This standard prior sampling scheme, utilized in the original ABC-Rejection framework [85], samples parameter values from the prior distribution. This approach may be reasonable for a low dimensional parameter space and with informative priors, but the number of samples required for coverage increases exponentially with each additional parameter. One might consider this a brute force sampling method.

The sampling strategy used in the algorithm in Chapter 5 was based on importance sampling (IS) methods. Instead of sampling from the prior distribution, samples are drawn from a proposal distribution that is iteratively updated to move to a region of higher likelihood. This modification enabled us to use the *s3m24* model with 24 muscles and data sets with 25 test directions. However, the framework in this chapter relied on all data being collected prior to the parameter estimation step. A fusion of the Chapter 5 and 6 frameworks would require a multi-level iterative algorithm with sequential importance sampling occurring after each measurement is taken to approximate the likelihood function, and within the entropy-reducing sampler algorithm to forward simulate and

determine the most informative next test. While this strategy may produce a tractable computation, additional steps will be required to improve efficiency enough to be clinically usable.

A number of extensions to the standard ABC algorithm have been developed and expanded upon in the literature with the specific motivation to improve efficiency through more intelligent sampling [6, 66]. The Markov Chain Monte Carlo (MCMC) ABC algorithm uses the Metropolis-Hastings MCMC kernel that is common in Bayesian statistics. Instead of sampling from the prior distribution, candidate parameters are iteratively proposed from the Metropolis-Hastings kernel [72]. One important downside to MCMC is that the simulations are no longer independent, so the code is not easily parallelized. Ease of parallelization is a practical advantage of traditional ABC and crucial when speed of computation is important. Sequential Monte Carlo (SMC) ABC methods perform several steps of ABC to iteratively concentrate the sampling effort on areas of the parameter space with a high likelihood [7]. The preliminary simulations are used to identify a set of particles that are then repeatedly resampled according to a weighting scheme, perturbed via a transition kernel, and filtered [92, 102]. An advantage of this approach over the importance sampling scheme used in Chapter 5 and MCMC methods is that the algorithm avoids getting caught in local minima by initially sampling coarsely over the full prior. While the SMC ABC approach is computationally more efficient than the classic ABC rejection algorithm, the overall computational burden relies not only on the complexity of the model and amount of data, but also on the details of the specific sequential scheme. There is a trade-off between convergence and exhaustive exploration of the parameter space in regards to selection of the particular perturbation kernel and rejection cutoff

schedule. Strategies for selecting optimal kernels have been proposed in several studies, with promising results [35, 65]. Optimization of the sampling scheme for improving efficiency of the algorithm will be a crucial step for feasible for real-time, clinician-in-the-loop implementation.

7.1.1.2. Computational Efficiency. The computational power of standard computers continues to increase rapidly, with a basic office computer now able to perform computing tasks that required a super computer a decade ago. As technology progresses, the computational cost of these algorithms will be less of a barrier. In the meantime, thoughtful parallelization and use of approximations can decrease the computational cost. The costliest parts of the ABC framework is the data generating step, where the sampled parameter values are simulated through the model to produce a simulated data set. This step occurs in our algorithm in Chapter 6 in both the strength estimate update and the entropy-reducing test sequence selection algorithm. The set of sampled candidate strength vectors are independent from one another, so these simulations can be split to multiple independent workers. All computations in this thesis were run in Matlab. Translation to a platform with less overhead would improve efficiency. In fact, there are libraries available in python and R for ABC that my provide a starting point for these implementations [50, 67].

7.1.2. Redundancy

The level of biomechanical kinematic and muscle redundancy in the neck underlies the challenge of obtaining estimates of individual neck muscle strength. The interactions between model complexity and the resulting neck strength and muscle activation patterns

were explored in Chapter 4, revealing the importance of balancing the number of kinematic degrees of freedom with the number and structure of independent muscle groups. Redundancy is also related to task specification [62]. The more complex the task (the more constraints), the less redundancy the system will appear to have.

7.1.2.1. Model Selection. Model-based parameter estimation methods attempt to tune the model parameters to best fit the experimental data collected. In this thesis, the model computes the maximum end effector wrench and muscle activations given a set of muscle strengths and a target direction (Algorithm 6). This computation is dependent on several important assumptions, including geometric properties such as number of joints, joint angles, and subject specific dimensions (encoded in the Jacobian and Moment Arm Matrix) and independent muscle groupings (encoded in the muscle strength vector and moment arm matrix). The OpenSim neck model [108] is equipped with $m = 96$ muscle elements, each with independent activation control. While research regarding how humans control neck muscles and the level of independent control over muscles is scarce [34, 11], an independent muscle set in the range of 24 muscles is likely more realistic. The level of muscle complexity should be considered in the context of kinematic complexity. As seen in Chapter 4, the ratio between kinematic degrees of freedom and number of independent muscle groups has an effect on the wrench producing capability of the model (note: the distribution of the muscle groups is also important, not just the quantity).

Techniques for model selection, between a discrete set of candidate models, are available in the Bayesian parameter estimation framework. The ABC framework provides a straightforward way to investigate the goodness-of-fit of a model. This is demonstrated in Chapters 5 and 6 with the data discrepancy scores and radar plots which show how

close the simulated data are to the measured data. A low final data discrepancy score may indicative of a deficient model. An interesting extension of this work could include an analysis of model fit to healthy control data sets comparing candidate models with various muscle groupings and model complexities. These results would reveal the levels of kinematic and muscle redundancy inherent in a multi-directional isometric neck strength task.

7.1.2.2. Informative Measurements. The information content in the observed data is another important component in resolving redundancy and reducing parameter estimate uncertainty. Observing the intermediate steps in the Bayesian algorithm can provide insight into how well the observed data discriminates between parameter values. Specifically, analyzing the shape of the approximated likelihood function, the data-likelihood function Algorithm 8, can indicate whether the data is providing sufficient information to identify the model parameters. In both Chapters 5 and 6, the parameter estimation results using the *p2m7* model (planar, 2 DOF, 7 muscle) produced muscles strength estimates with high confidence for some muscles, but consistently low confidence for others (TS-E and CS-E specifically). The data-likelihood function produces an essentially flat distribution for these muscles for measurements in each of the 7 planar test directions, indicating that the model output is not sensitive to the sampled muscle strength values of those particular muscles in those test directions. More informative measurements that can discriminate between parameters would improve performance.

One strategy to produce more informative measurements is to take advantage of the muscle length-tension curve and state dependent Jacobian and moment arm matrix by including various neck postures in addition to test directions. The relative contribution

of a muscle group to the end effector wrench can be manipulated by altering joint angles. This is a common procedure for clinicians when performing strength testing. To test the strength of the brachialis muscle, for example, a clinician will resist elbow flexion with the forearm pronated to reduce the contribution of the biceps. Biceps strength is tested with the forearm in supination. The OpenSim model is equipped with the muscle parameters, based on anatomic studies, required for providing the Jacobian, moment arm matrix, and force-length scalar, given a particular neck posture. The automated test sequence selection framework in Chapter 6 would be valuable for selecting the subset of postures and test directions to test, as increasing the number of measurements would be burdensome on the clinician and patient.

7.1.2.3. Summary Statistics. The optimal choice of summary statistics, a mapping of the raw observations to a lower dimensional and simpler set of metrics, can improve performance ABC algorithms. The composite data discrepancy score is the analogue to the summary statistic in the algorithms presented in this thesis. Equation 5.5 computes a data discrepancy score for each candidate strength vector in each test direction using a Gaussian radial basis function on the error between the simulated and measured wrench. The performance of a candidate strength vector is determined by the composite data discrepancy score, which is the average data discrepancy score across all test directions. Other statistics could be used to score performance, such as metrics on the 3-D shape of the wrenches (ellipsoid axes or PCA), the minimum error or maximum error, or some combination of these measures. Methods have been developed to analyze the sufficiency of the summary statistics and to select an optimal subset given a set of candidate statistics

[84, 32, 12]. Improving the summary statistics has the potential to improve both efficiency and accuracy, as it shapes the comparison between simulated and observed data.

7.1.3. Validation

While validation with respect to the accuracy of the muscle strength estimates is not possible because there is no other method to determine individual muscle strengths, there are several steps that can be taken to verify parts of the framework. First, a more extensive experimental study including intra-muscular EMG of additional neck muscles would help to verify the muscle activation patterns being proposed by the algorithm. There is no direct mapping from EMG activity to muscle force when the muscle strength is unknown, but the additional EMG data would provide a check that the muscle activation pattern producing the maximum wrench magnitude corresponds between the model and a human participant. Implementation of the framework for a different area of the body, such as the shoulder, would be a good test of generalizability and would make it possible to obtain a full set of EMG measurements. As described in the model selection section above (Section 7.1.2.1), the goodness of fit of the model can be evaluated based on the data discrepancy scores. Other tests of the ABC algorithm performance have been suggested that involve testing the final posterior distribution to assess whether the spread is appropriate and check for coverage [114, 83].

The neck is a particularly complex musculoskeletal system, making the muscle strength estimation problem challenging due to the high dimensionality of the parameter space, level of redundancy, and measurement difficulties. A clinical tool for individual muscle

strength estimation is a valuable contribution to entire field of musculoskeletal rehabilitation, as the problem of isolating individual muscle strength is not limited to the neck alone. The technical challenges and potential solutions may be less relevant for other simpler body systems with fewer muscles and less redundancy. However, implementation of the framework for the neck is an excellent test of its capabilities and potential.

7.2. Clinician in the Loop Design

Diagnostic and therapeutic tools utilizing the latest advances in computational parameter estimation and optimization of experiments have the potential to complement and enhance the knowledge and skills of clinicians. In the Bayesian frameworks presented in this thesis, domain knowledge from the expert (clinician) is an integral part of every step of the process. Any semi-automated framework faces challenges related to the human-algorithm interface. There is an inherent trade-off between flexibility and usability. We could compare this to the difference between running Linux vs. a mac OS. With Linux, the user has the power to customize and optimize the operating system to suit their specific needs. However, they also have the power to make mistakes and crash the system. On the other hand, a mac OS gives little power to the user to either optimize the system or to mess it up. This trade-off is important to consider in the context of implementation of this proposed framework and the benefits vs. risks associated. It is my opinion that if a user is given the appropriate initial training and orientation, the power and flexibility of the Linux OS is worthwhile and risks can be mitigated. The clinician-in-the-loop Bayesian framework for musculoskeletal diagnostics builds upon the clinical expertise of the clinician with the computational advances in musculoskeletal modeling

and computational methods for parameter estimation to integrate information to enhance clinical decision making.

Clinical decision making is a natural fit for Bayesian inference. Whether or not they use the language, clinicians formulate hypotheses and differential diagnoses based on the interview and screen (prior), perform clinical tests (observations) to update their beliefs (posterior). An experienced clinician even optimizes the subset and sequence of clinical tests to provide the most information that discriminates between differential diagnoses, minimizes uncertainty, and minimizes the number of tests that need to be performed (entropy-reducing test sequence selection). When framed in this way, clinicians should have an intuitive understanding of the framework and the ways in the framework can complement the diagnostic work flow. While the user does not need to understand the technical details of the computational algorithms, training should include the limitations of musculoskeletal modeling and model based parameter estimation. It is crucial that users know that the algorithm is fitting the model to the data, and is limited by the assumptions inherent in the model. Clinicians should be encouraged to trust their clinical judgement and use the tool to add to the information available to make a clinical decision, considering the uncertainty of the parameter estimates. Example scenarios are now presented to demonstrated possible uses of this framework for aiding in clinical diagnosis and therapeutics.

***Scenario 1*: Targeted muscle weakness in WAD** A clinician is evaluating a patient with chronic neck pain and disability after a whiplash injury. From the interview and screen, general neck weakness and poor postural endurance are noted. It is unclear from available clinical tests whether the patient has global neck weakness or weakness

of a specific subset of muscles. To refine the PT diagnosis and target a treatment plan, the clinician utilizes the individual muscle strength estimation framework with entropy-reducing test sequence selection. The algorithm iteratively suggests the sequence of test directions in which to take maximum isometric strength measurements from the patient. It also provides a visual representation of the expected information gain (reduction in uncertainty) for the next measurement. Weighing the comfort and fatigue of the patient, the confidence of the specific muscle strength estimates, and the expected information gain for the next measurement, the clinician decides to stop testing after the 5^{th} measurement. Results of the parameter estimation reveals a high confidence of bilateral weakness of the deep cervical extensors and flexors relative to the superficial muscles, but with low confidence in the estimates of strength estimate of the suboccipital and scalene muscles. The clinician targets the treatment plan on strengthening the deep cervical extensors and flexors, while continuing to monitor and consider the possibility of specific dysfunction of the muscle groups with higher uncertainty in their estimates.

***Scenario 2*: Automated Rotator Cuff Strength Testing** Extension of this framework to individual muscle strength estimation of muscles at the shoulder would be relatively straight-forward. Musculoskeletal models of the shoulder, with well-defined, state dependent muscle parameters are freely available through Opensim [18]. A testing setup that allows for multiple joint configurations and 6-DOF force/torque measurements would provide the necessary data for discriminating between muscles. If necessary, surface EMG measurements of the deltoid, biceps, and triceps would provide additional information for parameter discrimination. The entropy-reducing sampler remains important in

this context, allowing the clinician to prioritize the strength estimates of particular muscles of interest and reducing the number of measurements to gain the desired information. A clinician who suspects a rotator cuff tear has a series of clinical tests that can be performed to attempt to identify the muscles involved, with varying levels of sensitivity, specificity, and repeatability [47]. This framework would provide the clinician with quantitative estimates of individual muscle strengths, which can be selectively biased towards reducing uncertainty in the 4 rotator cuff muscles.

Much of the work of this thesis was motivated by the lack of an available method for measuring the strength of individual neck muscles to test the hypothesis that individuals with chronic WAD have weak deep extensor muscles. The framework for individual muscle strength estimation with entropy-reducing test sequence selection proposed in the final chapter represents a step towards answering this and other open parameter estimation questions in musculoskeletal rehabilitation. The proposed solution integrates clinical expertise, musculoskeletal modeling, Bayesian inference, and information theory. While limited to lower dimensional models in its current form, the framework has the potential to have a clinical impact if some of the extensions proposed in this chapter are implemented. Clinician-in-the-loop, musculoskeletal model-based Bayesian inference for individual muscle strength estimation offers a flexible, quantitative method for determining muscle strength.

APPENDIX A

Neck Strength Experiment

A.0.1. Methods

Isometric neck strength measurements were taken using a custom made device. Participants were seated in a comfortable, neutral posture with the torso restrained to the chair using two crossed straps. The head was rigidly attached to a 6-axis load cell with 10 pads and a chin strap firmly tightened around the head, as shown in Figure A.1.

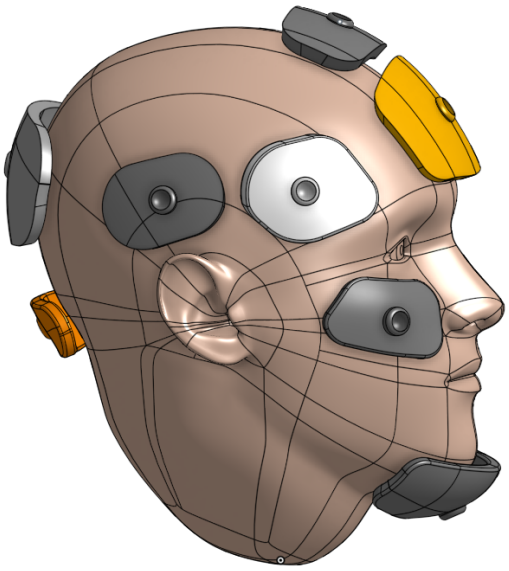


Figure A.1. Custom pads to interface with the head for strength measurement device.

Load cell and EMG measurements were sampled at 1000 Hz by a data acquisition device (NI-DAQ 6218). Load cell measurements were averaged over a 100 ms window for feedback and direction matching.

The moments and forces measured by the load cell are resolved to a wrench $w_s = (M_x, M_y, M_z, F_x, F_y, F_z)$ at the skull frame, with its origin at the occipito-axial joint, the x-axis directed anteriorly, y-axis superiorly, and the z-axis to the right, horizontal to the ground. The lower posterior pad is adjusted to be centered just under the occipital protuberance. The location of the occipito-axial joint (C0-C1) is then estimated as 8 cm anterior and 2 cm inferior to the to the center of the lower posterior pad.

The target direction and real-time feedback was provided in 6-DOF on a screen in front of the participant. The wrench direction was represented by the posture of an anthropometric figure of the head and neck, with a pyramidal head where apex of the pyramid is the nose facing into the screen, as shown in Figure A.2. The participant was instructed to match the direction of the wrench while maximizing the magnitude, indicated by the height of bars on each side of the screen. If unable to match the direction for at least 2 seconds, the trial was repeated. A maximum effort was measured in each target direction at least 2 times. A third measurement was taken if the second measurement magnitude was more than 10% greater than the first (indicating that learning may still be occurring). The largest magnitude of the measurements was taken as the true value of the set.

Maximum wrenches in each of the 6 cardinal directions (extension, flexion, right sidebending, left sidebending, right rotation, left rotation) were measured at the beginning and end of the experiment to test for fatigue and allow for comparison to other neck strength literature. For these directions, feedback was limited to the relevant components

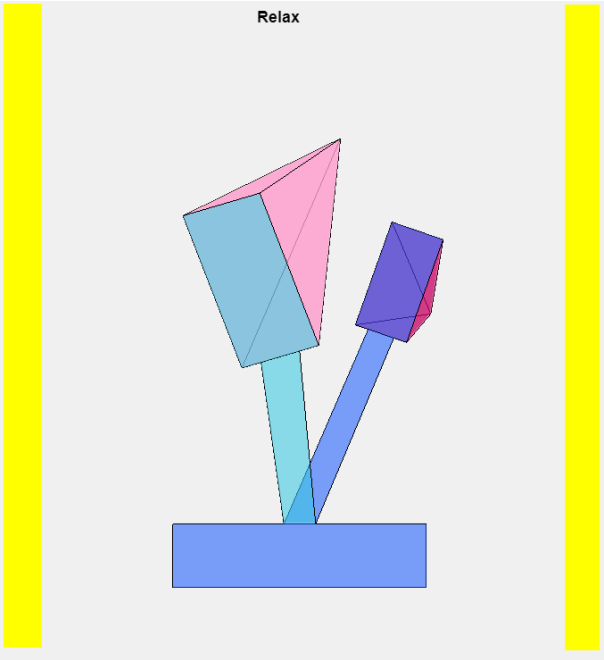


Figure A.2. Visual target and feedback for 6-DOF wrench direction provided to the participant during the experiment. The apex of the pyramid represents the participant's nose. The bars on the either side of the screen change color to green when the direction is matched within the allowable error.

(for flexion/extension, the M_z, F_x, and F_y), and error window for direction matching was increased to allow variation in the center of rotation (ratio of moments to forces).

Maximum wrenches in each of the 25 target directions were measured in randomized order. The target directions (shown in Figure A.3) were selected as locations and normal forces on the surface of the skull where resistance could be applied manually by a clinician. These locations and force directions were mapped to 6-D wrenches in the skull frame. To transform a force $Fa = (Fa_x, Fa_y, Fa_z)$ applied at the location of a test point $p = (p_x, p_y, p_z)$ to a 6-D wrench in the skull frame (w_{skull}), we use the adjoint of the spatial transform adT_p.

$$w_{skull} = adT_p * \ F_a$$

$$adT = \begin{bmatrix} R & 0 \\ \hat{p}\ R & R \end{bmatrix} \quad \text{where} \quad \hat{p} = \begin{bmatrix} 0 & -p_z & p_y \\ p_z & 0 & -p_x \\ -p_y & p_x & 0 \end{bmatrix} \quad \text{and} \quad R = I_{3x3}$$

Verbal directions to the participant included the location of the point on the skull to push into (for example, 'push into a point located over your R ear').

An error of 20 degrees was allowed in the unit vector direction of the moment and force components and an error of 10% was allowed in the force-moment ratio.

The 6-DOF wrench $w_s = (M_x, M_y, M_z, F_x, F_y, F_z)$ is a vector with components that have different units (moments in Newton-meters, forces in Newtons). Representation of a wrench as magnitude and direction components is not strictly proper, as it would involve summing different units. In order to fairly assess wrench direction matching without biasing the force or moment components, wrench directions are represented as a moment unit vector (M_u), force unit vector (F_u), and ratio between force and moment magnitudes (R_{fm}).

(A.1)
$$w = \begin{bmatrix} M \\ F \end{bmatrix} \qquad M = \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} \qquad F = \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix}$$

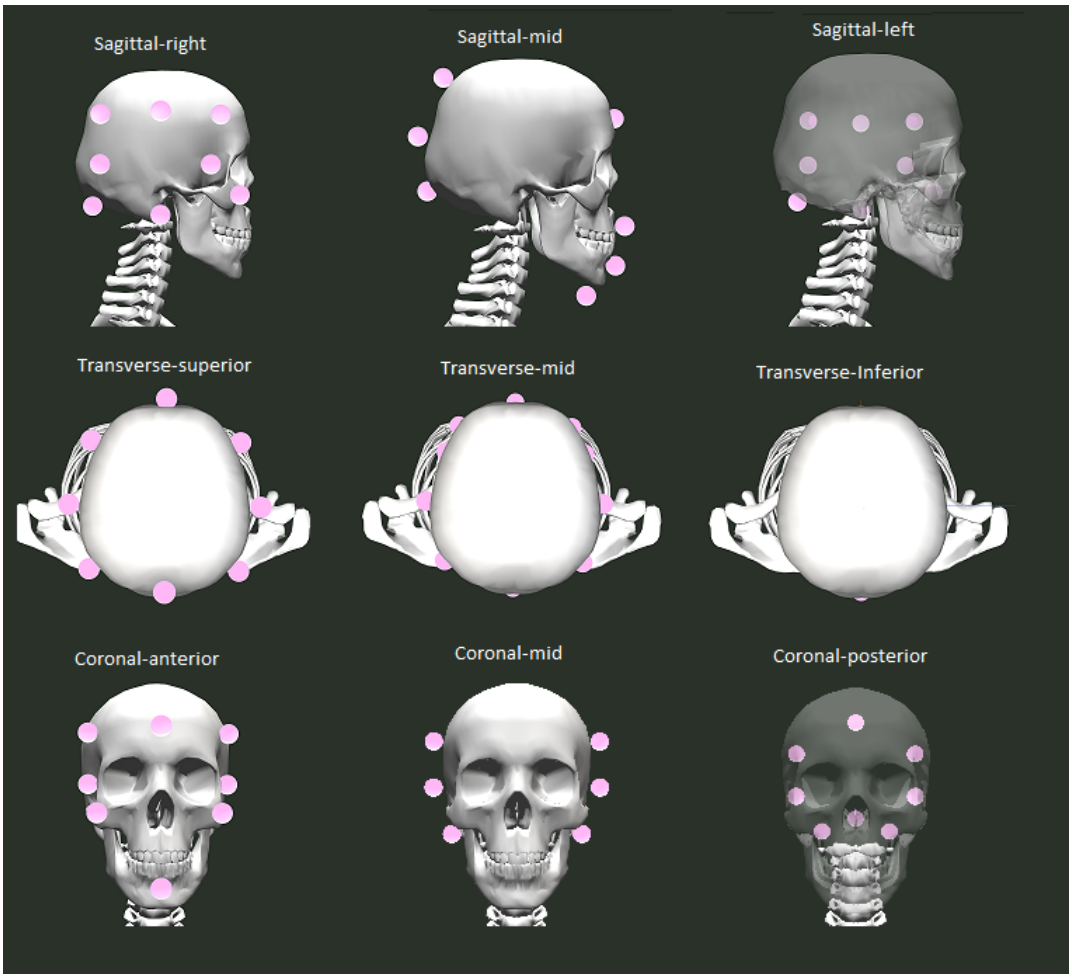


Figure A.3. A 3x3 grid showing the test directions in multiple planes used to visualize the neck strength and muscle activation patterns. Note that some test positions appear in multiple planes. The top row shows the right-, mid-, and left- sagittal plane test directions. The middle row shows the superior-, mid-, and inferior- transverse plane test directions. The bottom row shows the anterior-, mid-, and posterior- coronal plane test directions.

(A.2) $$M_u = \frac{M}{|M|} \qquad F_u = \frac{F}{|F|} \qquad R_{fm} = \frac{|F|}{|M|+|F|}$$

	Electrode Placement
Sternocleidomastoid	Over muscle belly, $\sim 1/3$ of its length rostral to sternal attachment
Trapezius	Over muscle belly, at the level of C6-C7
Semispinalis Capitis	At level of C1/C2 spinous process, 2 cm lateral to midline
Splenius Capitis	Between trapezius and sternocleidomastoid, 3 cm below mastoid process
Levator Scapulae	Between trapezius and sternocleidomastoid at C4/C5 level

Table A.1. **EMG Electrode Placement**. These locations are based on [58] and [3]

Note that if a target direction $R_{fm} = 0$, then it is a pure moment, and if $R_{fm} = 1$, it is a pure force. In that case, the unit direction for the other component is meaningless therefore ignored. Instead, the error window on R_{fm} captures the component that should maintain a zero magnitude.

Surface electrodes were placed bilaterally on the sternocleidomastoid (SCM), upper trapezius (Trap), splenius capitis (SpCap), semispinalis capitis (SSCap), and levator scapulae (LS). Electrode placement is described in Table A.1, and verified by palpation. The ground electrode was placed over the right acromion.

EMG records were rectified and filtered using a moving average filter centered over a 250 ms window. EMG levels were normalized with respect to the maximum value obtained over all trials for each muscle. The normalized EMG level at the timepoint 250 ms prior to the maximum wrench magnitude for each trial was taken as the muscle activation level for that trial.

A.1. Results

Five healthy participants with no history of neck pain were tested according to the protocol. The maximum strength measurements in 9 planes for all 5 subjects are visualized in Figure A.4. Overlaid on the strength measurements is the strength prediction from the

s3log24 model, with 3 joint ($n = 9$) and $m = 24$ muscle model with f'_{base} baseline strength values.

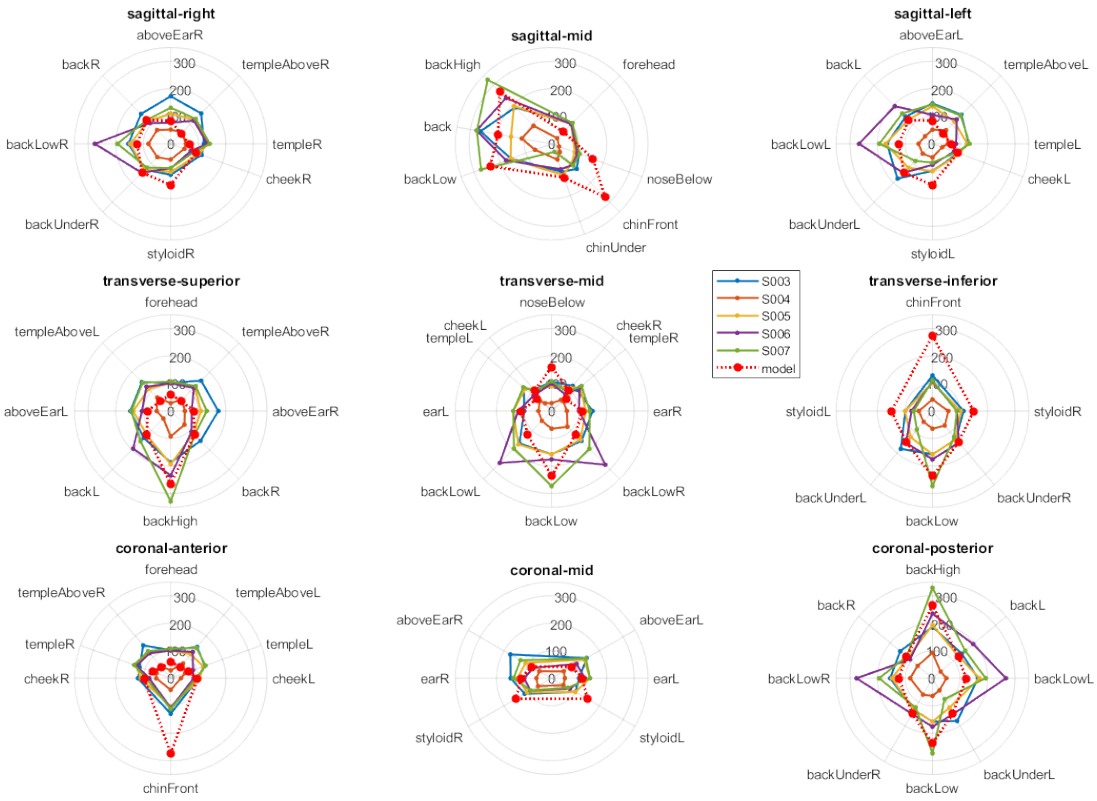
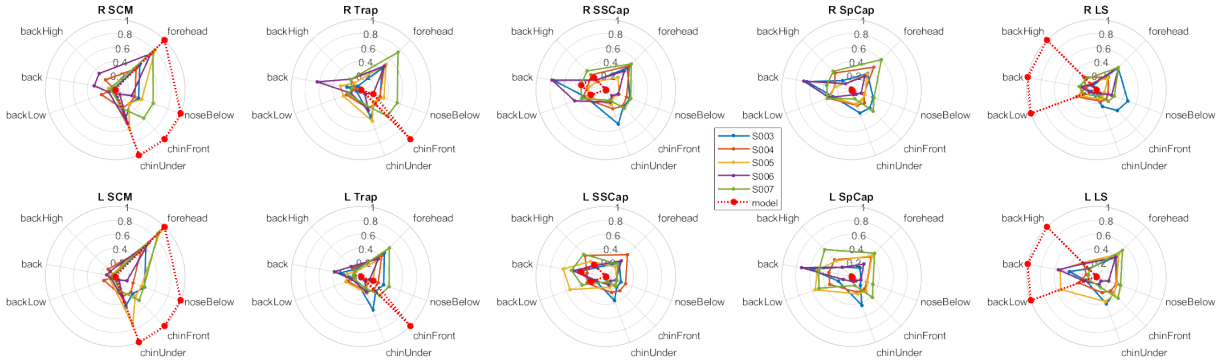


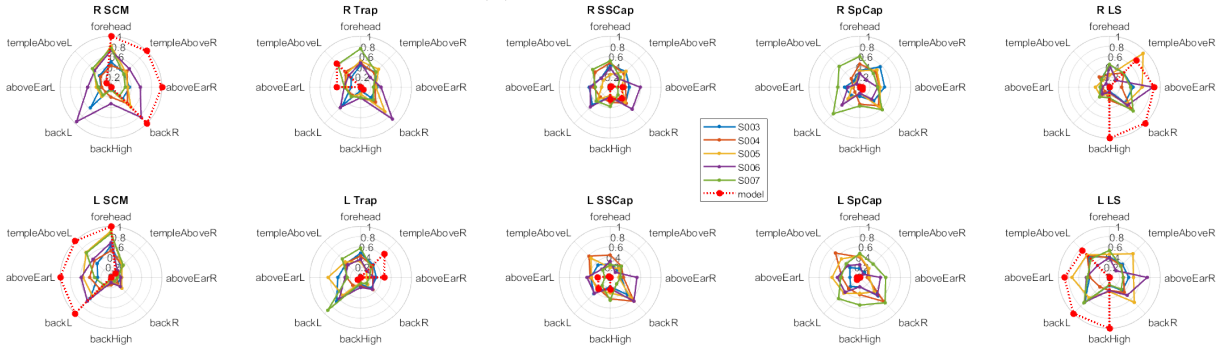
Figure A.4. Neck strength measurements from (n=5) participants (dashed lines). The maximum wrench magnitude from the 's3log24' neck model (spatial 3 joint ($n = 9$) and $m = 24$ muscle model) is the red solid line.

The muscle activations measured from the 5 pairs of muscles with surface EMG during production of the maximum wrench are presented for the mid-sagittal and mid-transverse planes in Figure A.5. The predicted activation from the from the s3log24 model is overlaid in red.

The muscle activations of each of the 10 muscles is illustrated in Figures A.6 - A.13. In each radar plot, the red dashed trace is the muscle activation level predicted by the *s3m24* model with the baseline muscle strengths. The experimentally measured EMG



(a) mid-sagittal view



(b) mid-transverse view

Figure A.5. Muscle activation measured by surface EMG for $n = 5$ healthy participants during maximum effort (dashed lines). The predicted muscle activation using the 's3log24' model (spatial 3 joint ($n = 9$) and $m = 24$ muscle model) is the red solid line.

values for all muscles exhibit more co-activation and less directional preference than the model prediction.

	Maximum Force (N)						
	S003	S004	S005	S006	S007	Avg	SD
forehead	102.7	28.7	101.2	100.8	107.3	88.16	33.32
backHigh	186.9	92.5	193.4	234.8	329.0	207.32	85.66
templeAboveR	156.0	58.3	122.5	118.9	128.3	116.81	35.84
templeAboveL	149.0	69.6	113.7	125.0	145.0	120.45	31.90
backR	153.1	70.4	115.9	106.9	122.8	113.84	29.84
backL	138.8	38.3	123.6	193.1	155.5	129.89	57.35
styloidR	113.1	57.3	102.6	88.9	88.3	90.05	21.03
styloidL	98.9	49.7	99.6	76.3	68.1	78.53	21.27
earR	148.8	55.4	124.1	99.8	138.6	113.33	37.27
earL	99.8	48.1	139.4	121.3	139.6	109.66	38.09
aboveEarR	173.0	51.3	110.0	76.0	130.8	108.23	47.43
aboveEarL	146.9	51.0	136.3	104.9	143.5	116.51	40.22
cheekR	119.4	52.9	109.0	77.9	89.6	89.74	26.21
cheekL	100.2	37.1	105.0	83.4	92.1	83.56	27.22
chinFront	129.1	42.1	102.5	107.8	111.3	98.57	33.12
chinUnder	108.6	61.1	121.3	98.6	31.5	84.23	37.06
backLow	157.6	64.2	158.1	176.2	273.0	165.85	74.25
backLowL	165.8	50.5	171.3	267.1	193.4	169.63	77.90
backLowR	154.0	81.2	147.6	276.1	194.0	170.57	71.57
back	262.6	110.7	149.8	277.2	276.0	215.28	79.03
noseBelow	107.3	28.8	89.5	100.1	108.4	86.82	33.30
templeR	120.8	55.1	135.0	132.4	142.0	117.07	35.46
templeL	131.8	56.0	125.6	85.9	134.8	106.82	34.59
backUnderR	133.1	69.0	133.4	149.2	122.6	121.46	30.84
backUnderL	179.4	50.6	123.5	150.1	87.9	118.29	50.69

Table A.2. The maximum force magnitude (in Newtons) measured in each test direction for each participant (Denoted S00_).

	Normalized Activation at Max Force						
	S003	S004	S005	S006	S007	Avg	SD
forehead	0.50	0.42	0.81	0.72	0.77	0.64	0.17
backHigh	0.02	0.19	0.06	0.32	0.06	0.13	0.13
templeAboveR	0.37	0.44	0.41	0.51	0.36	0.42	0.06
templeAboveL	0.32	0.29	0.47	0.46	0.52	0.41	0.10
backR	0.32	0.46	0.55	0.84	0.24	0.48	0.23
backL	0.57	0.14	0.26	0.96	0.22	0.43	0.34
styloidR	0.36	0.40	0.29	0.69	0.25	0.40	0.17
styloidL	0.21	0.19	0.25	0.61	0.17	0.28	0.18
earR	0.32	0.45	0.31	0.56	0.38	0.40	0.11
earL	0.17	0.25	0.28	0.69	0.24	0.33	0.21
aboveEarR	0.36	0.32	0.33	0.58	0.28	0.37	0.12
aboveEarL	0.20	0.18	0.29	0.46	0.25	0.28	0.11
cheekR	0.41	0.45	0.87	0.47	0.46	0.53	0.19
cheekL	0.25	0.29	0.58	0.33	0.47	0.38	0.14
chinFront	0.29	0.28	0.28	0.09	0.58	0.30	0.17
chinUnder	0.44	0.28	0.59	0.51	0.29	0.42	0.14
backLow	0.02	0.21	0.09	0.10	0.06	0.10	0.07
backLowL	0.15	0.08	0.17	0.66	0.15	0.24	0.24
backLowR	0.23	0.25	0.15	0.50	0.23	0.27	0.13
back	0.07	0.02	0.10	0.31	0.05	0.11	0.11
noseBelow	0.36	0.30	0.41	0.25	0.58	0.38	0.13
templeR	0.24	0.31	0.45	0.66	0.27	0.39	0.17
templeL	0.23	0.30	0.31	0.58	0.35	0.35	0.13
backUnderR	0.06	0.14	0.13	0.18	0.04	0.11	0.06
backUnderL	0.09	0.16	0.16	0.35	0.03	0.16	0.12

Table A.3. The normalized muscle activation level measured at the time of maximum effort in each test direction for each participant (Denoted S00_).

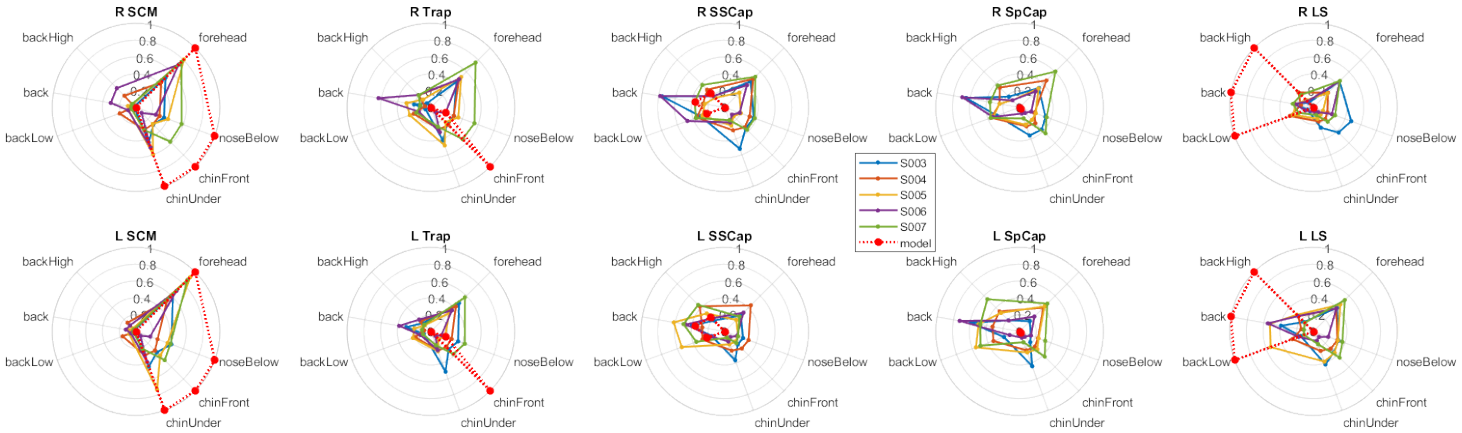


Figure A.6. Mid-Sagittal Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

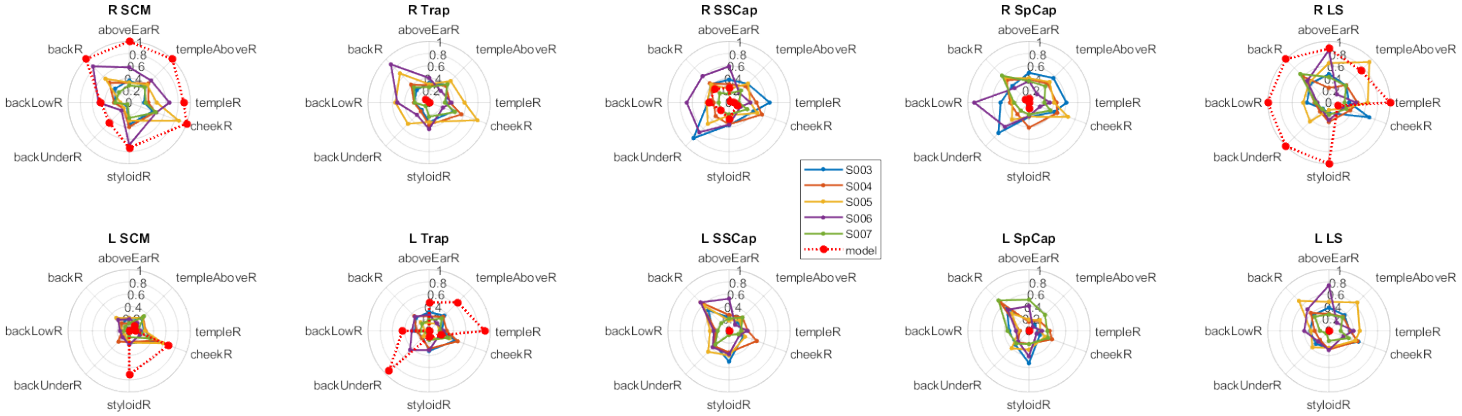


Figure A.7. Right-Sagittal Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

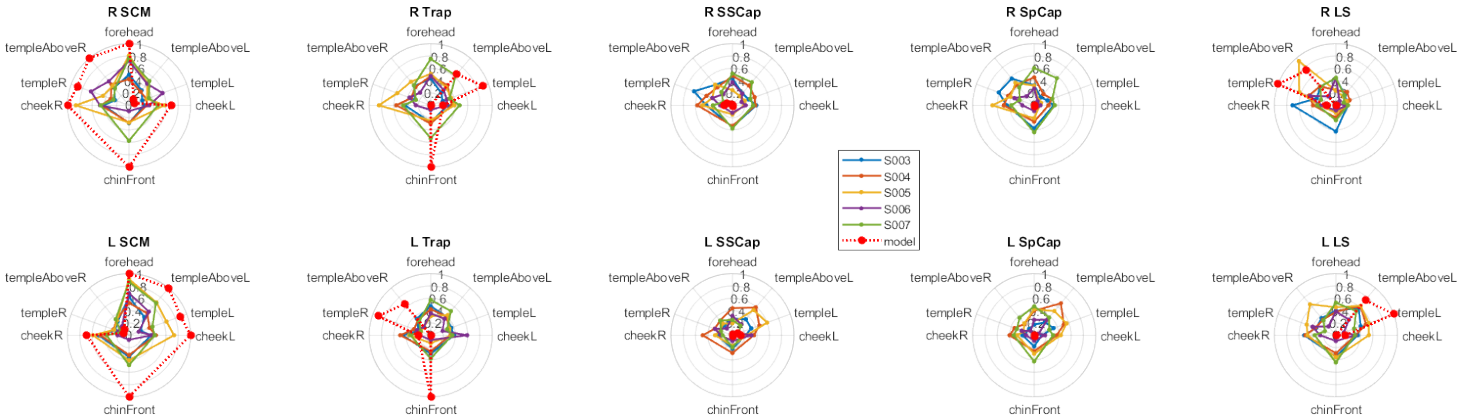


Figure A.8. Anterior Coronal Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

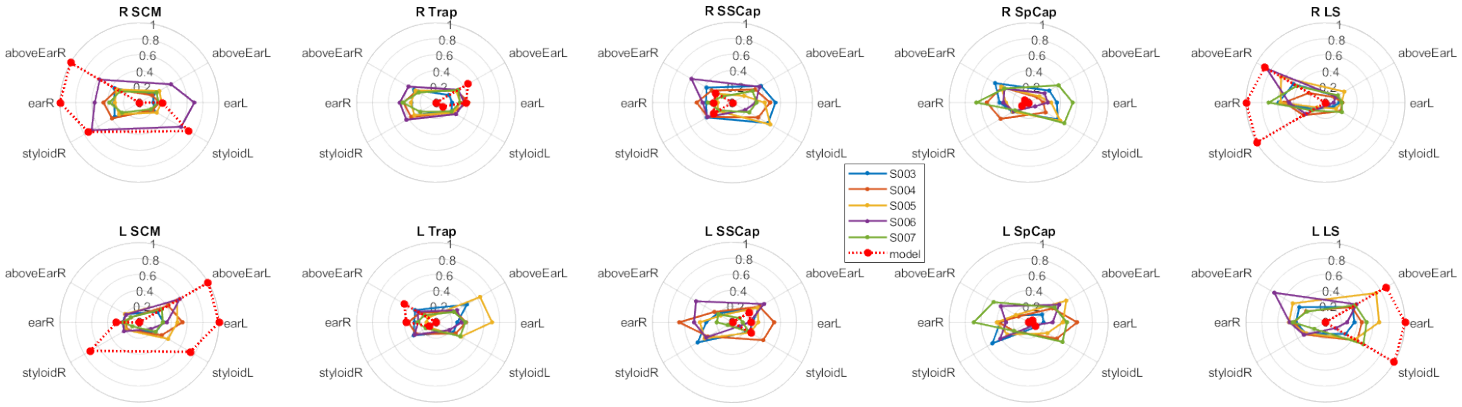


Figure A.9. Mid-Coronal Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

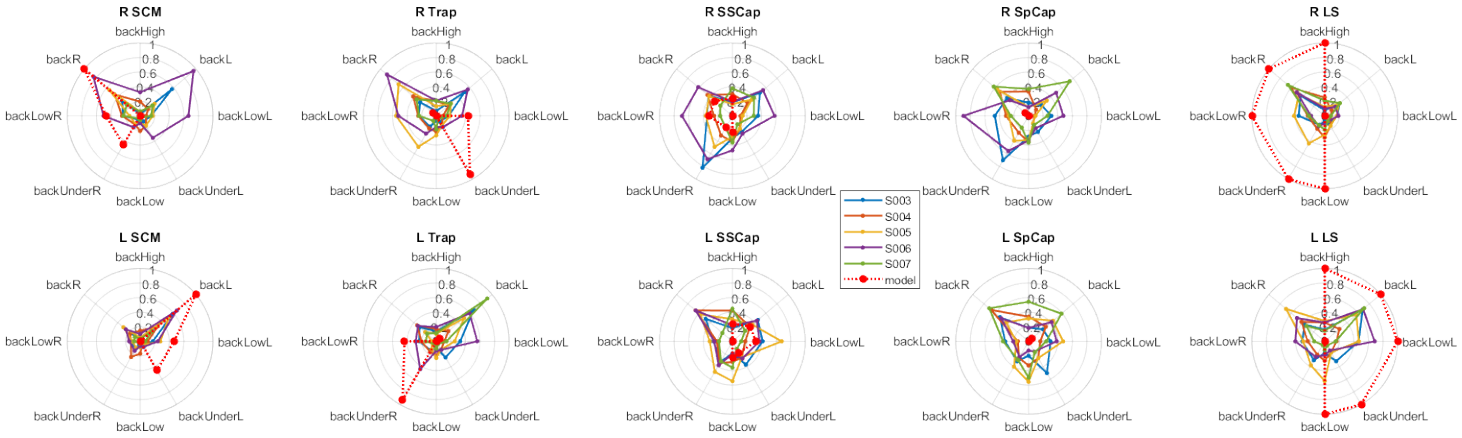


Figure A.10. Posterior Coronal Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

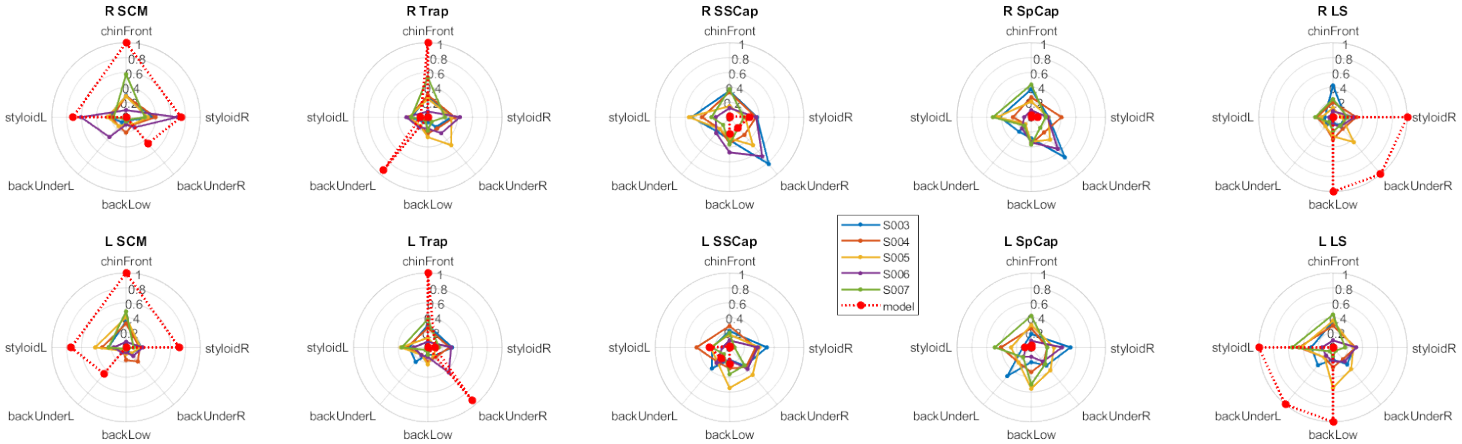


Figure A.11. Inferior Transverse Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

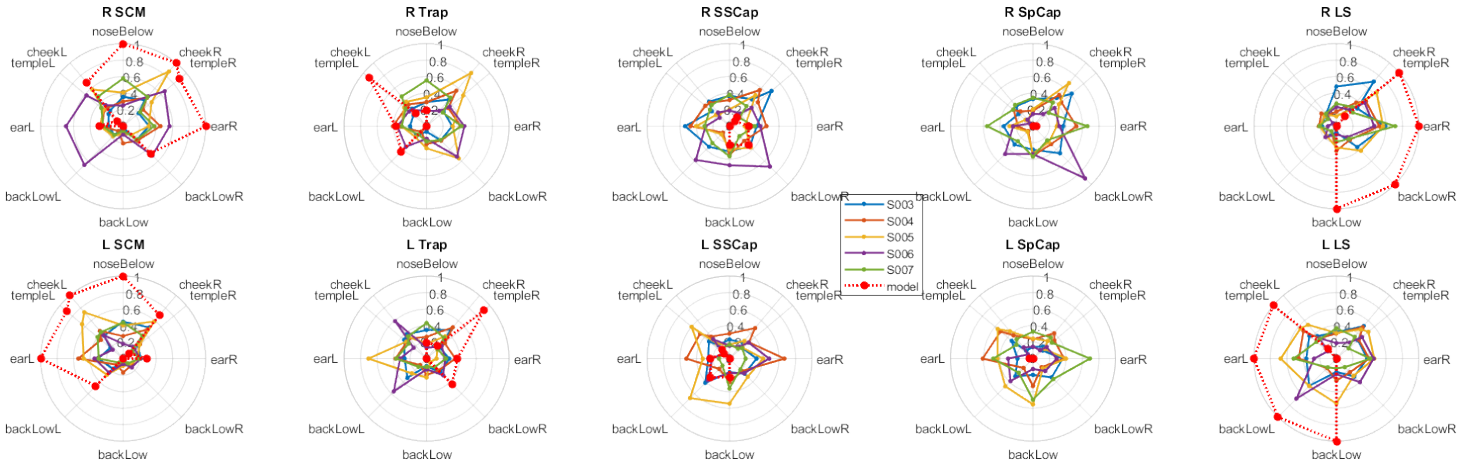


Figure A.12. Mid-Transverse Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

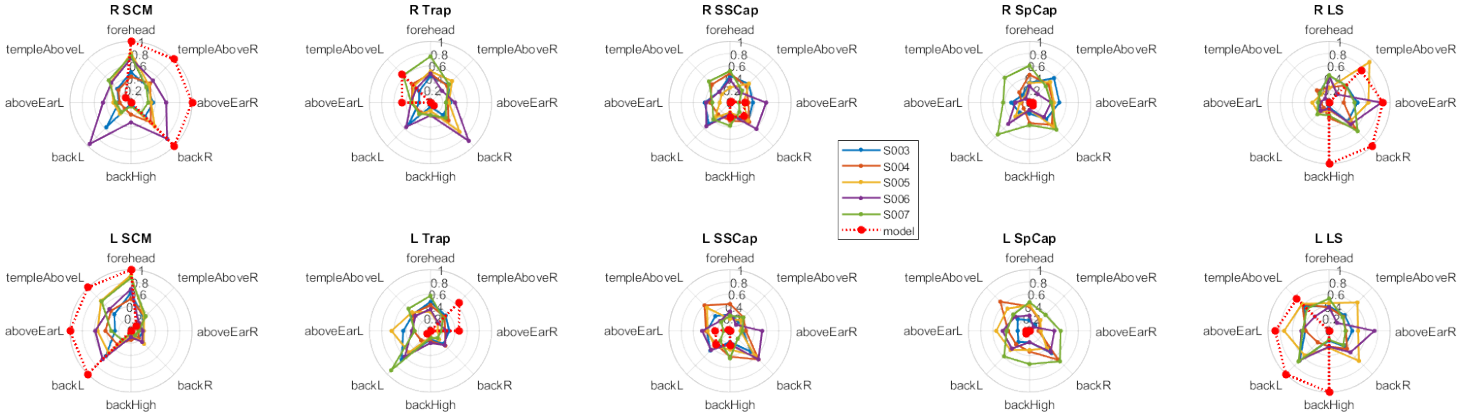


Figure A.13. Superior Transverse Plane View: Muscle activations measured by EMG in $n = 5$ healthy participants.

APPENDIX B

Model Parameters

Table B.1. Muscle parameters for all $m = 96$ muscle elements in the full musculoskeletal neck model. These values are derived from the OpenSim Neck Models [108, 78]. Pitch: (+) Extension, (-) Flexion; Roll: (+) R Sidebend, (-) L Sidebend; Yaw: (+) L Rotation, (-) R Rotation

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
1	digastric post	40.6							0.0038	0.0393	0.0168
2	digastric post L	40.6							0.0038	-0.0393	-0.0168
3	digastric ant	40.6							-0.0209	0.0624	0.0238
4	digastric ant L	40.6							-0.0209	-0.0624	-0.0238
5	mylohyoid ant	55.7							-0.0410	0.0579	0.0453
6	mylohyoid ant L	55.7							-0.0410	-0.0579	-0.0453
7	stylohyoid lat	6.8							-0.0104	0.0354	0.0095
8	stylohyoid lat L	6.8							-0.0104	-0.0354	-0.0095
9	stylohyoid med	6.8							-0.0107	0.0356	0.0094
10	stylohyoid med L	6.8							-0.0107	-0.0356	-0.0094
11	sternohyoid	20.3	-0.0453	0.0102	-0.0093	-0.0509	0.0065	-0.0089			
12	sternohyoid L	20.3	-0.0453	-0.0102	0.0093	-0.0509	-0.0065	0.0089			
13	geniohyoid	34.0							-0.0681	-0.0069	-0.0087
14	geniohyoid L	34.0							-0.0681	0.0069	0.0087

Table B.1 continued from previous page

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
15	mylohyoid post	18.6							-0.0399	0.0459	0.0190
16	mylohyoid post L	18.6							-0.0399	-0.0459	-0.0190
17	sternothyroid	22.8	-0.0384	0.0134	-0.0032	-0.0401	0.0127	-0.0031			
18	sternothyroid L	22.8	-0.0384	-0.0134	0.0032	-0.0401	-0.0127	0.0031			
19	omohyoid	26.3	-0.0199	0.0263	-0.0264	-0.0322	0.0152	-0.0252			
20	omohyoid L	26.3	-0.0199	-0.0263	0.0264	-0.0322	-0.0152	0.0252			
21	stern mast	86.1	-0.0265	0.0251	0.0144	-0.0174	0.0341	0.0135	0.0033	0.0575	0.0119
22	stern mast L	86.1	-0.0265	-0.0251	-0.0144	-0.0174	-0.0341	-0.0135	0.0033	-0.0575	-0.0119
23	cleid mast	43.1	-0.0284	0.0568	0.0146	-0.0187	0.0583	0.0144	0.0034	0.0622	0.0141
24	cleid mast L	43.1	-0.0284	-0.0568	-0.0146	-0.0187	-0.0583	-0.0144	0.0034	-0.0622	-0.0141
25	cleid occ	43.1	-0.0215	0.0546	0.0224	-0.0079	0.0563	0.0222	0.0243	0.0605	0.0219
26	cleid occ L	43.1	-0.0215	-0.0546	-0.0224	-0.0079	-0.0563	-0.0222	0.0243	-0.0605	-0.0219
27	scalenus ant	65.8	-0.0120	0.0363	0.0017	-0.0048	0.0277	0.0027			
28	scalenus ant L	65.8	-0.0120	-0.0363	-0.0017	-0.0048	-0.0277	-0.0027			
29	scalenus med	47.6	-0.0016	0.0441	0.0060	0.0054	0.0332	0.0072			
30	scalenus med L	47.6	-0.0016	-0.0441	-0.0060	0.0054	-0.0332	-0.0072			

Table B.1 continued from previous page

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
31	scalenus post	36.8	0.0030	0.0420	0.0078						
32	scalenus post L	36.8	0.0030	-0.0420	-0.0078						
33	long cap sklc4	48.0							-0.0120	0.0078	-0.0021
34	long cap sklc4 L	48.0							-0.0120	-0.0078	0.0021
35	long col c1thx	14.4	-0.0121	0.0039	-0.0008	-0.0145	0.0038	-0.0007			
36	long col c1thx L	14.4	-0.0121	-0.0039	0.0008	-0.0145	-0.0038	0.0007			
37	long col c1c5	14.4				0.0021	0.0217	-0.0093			
38	long col c1c5 L	14.4				0.0021	-0.0217	0.0093			
39	long col c5thx	14.4	0.0014	0.0201	-0.0031						
40	long col c5thx L	14.4	0.0014	-0.0201	0.0031						
41	trap cl	132.0	-0.0006	0.0751	0.0411	0.0142	0.0605	0.0427	0.0502	0.0229	0.0452
42	trap cl L	132.0	-0.0006	-0.0751	-0.0411	0.0142	-0.0605	-0.0427	0.0502	-0.0229	-0.0452
43	trap acr	348.6									
44	trap acr L	348.6									
45	splen cap sklc6	55.0				0.0565	0.0061	-0.0225	0.0581	0.0339	-0.0244
46	splen cap sklc6 L	55.0				0.0565	-0.0061	0.0225	0.0581	-0.0339	0.0244

Table B.1 continued from previous page

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
47	splen cap sklthx	53.2	0.0512	0.0151	-0.0197	0.0487	0.0241	-0.0207	0.0395	0.0473	-0.0223
48	splen cap sklthx L	53.2	0.0512	-0.0151	0.0197	0.0487	-0.0241	0.0207	0.0395	-0.0473	0.0223
49	semi cap sklc5	91.7				0.0135	0.0207	0.0124	0.0551	0.0167	0.0127
50	semi cap sklc5 L	91.7				0.0135	-0.0207	-0.0124	0.0551	-0.0167	-0.0127
51	semi cap sklthx	101.5	0.0432	0.0332	0.0142	0.0516	0.0279	0.0148	0.0705	0.0143	0.0157
52	semi cap sklthx L	101.5	0.0432	-0.0332	-0.0142	0.0516	-0.0279	-0.0148	0.0705	-0.0143	-0.0157
53	splen cerv c3thx	50.1	0.0330	0.0151	-0.0101	0.0242	0.0180	-0.0104			
54	splen cerv c3thx L	50.1	0.0330	-0.0151	0.0101	0.0242	-0.0180	0.0104			
55	semi cerv c3thx	107.1	0.0431	0.0169	0.0107	0.0434	0.0088	0.0116			
56	semi cerv c3thx L	107.1	0.0431	-0.0169	-0.0107	0.0434	-0.0088	-0.0116			
57	levator scap	109.2	0.0461	0.0505	-0.0051	0.0348	0.0380	-0.0038			
58	levator scap L	109.2	0.0461	-0.0505	0.0051	0.0348	-0.0380	0.0038			
59	longissi cap sklc6	34.3				0.0207	0.0292	-0.0035	0.0324	0.0510	-0.0050
60	longissi cap sklc6 L	34.3				0.0207	-0.0292	0.0035	0.0324	-0.0510	0.0050
61	longissi cerv c4thx	52.2	0.0326	0.0245	-0.0120	0.0216	0.0237	-0.0119			
62	longissi cerv c4thx L	52.2	0.0326	-0.0245	0.0120	0.0216	-0.0237	0.0119			

Table B.1 continued from previous page

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
63	iliocost cerv c5rib	36.4	0.0310	0.0300	-0.0106						
64	iliocost cerv c5rib L	36.4	0.0310	-0.0300	0.0106						
65	rectcap post maj	58.8							0.0504	0.0168	-0.0159
66	rectcap post maj L	58.8							0.0504	-0.0168	0.0159
67	rectcap post min	32.2							0.0391	0.0076	-0.0065
68	rectcap post min L	32.2							0.0391	-0.0076	0.0065
69	obl cap sup	30.8							0.0136	0.0183	0.0379
70	obl cap sup L	30.8							0.0136	-0.0183	-0.0379
71	obl cap inf	68.3									
72	obl cap inf L	68.3									
73	mult sup C45C2	14.7				0.0085	0.0225	0.0237			
74	mult sup C45C2 L	14.7				0.0085	-0.0225	-0.0237			
75	mult sup C56C2	19.3				0.0217	0.0191	0.0181			
76	mult sup C56C2 L	19.3				0.0217	-0.0191	-0.0181			
77	mult sup C67C2	15.8	0.0221	0.0277	0.0132	0.0290	0.0167	0.0144			
78	mult sup C67C2 L	15.8	0.0221	-0.0277	-0.0132	0.0290	-0.0167	-0.0144			

Table B.1 continued from previous page

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
79	mult sup T1C4	16.3	0.0353	0.0243	0.0149	0.0325	0.0100	0.0165			
80	mult sup T1C4 L	16.3	0.0353	-0.0243	-0.0149	0.0325	-0.0100	-0.0165			
81	mult sup T1C5	11.7	0.0351	0.0211	0.0215						
82	mult sup T1C5 L	11.7	0.0351	-0.0211	-0.0215						
83	mult sup T2C6	6.5	0.0415	0.0114	0.0215						
84	mult sup T2C6 L	6.5	0.0415	-0.0114	-0.0215						
85	mult deep C45C2	7.4				0.0178	0.0245	0.0122			
86	mult deep C45C2 L	7.4				0.0178	-0.0245	-0.0122			
87	mult deep C56C3	12.3				0.0205	0.0178	0.0124			
88	mult deep C56C3 L	12.3				0.0205	-0.0178	-0.0124			
89	mult deep C67C4	16.1	0.0231	0.0260	0.0112	0.0238	0.0114	0.0128			
90	mult deep C67C4 L	16.1	0.0231	-0.0260	-0.0112	0.0238	-0.0114	-0.0128			
91	mult deep T1C5	12.3	0.0332	0.0222	0.0169						
92	mult deep T1C5 L	12.3	0.0332	-0.0222	-0.0169						
93	mult deep T1C6	8.3	0.0278	0.0145	0.0252						
94	mult deep T1C6 L	8.3	0.0278	-0.0145	-0.0252						

Table B.1 continued from previous page

	Muscle Name	Strength (N)	C6-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
			pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
95	mult deep T2C7	14.0									
96	mult deep T2C7 L	14.0									

Table B.2. Muscle groupings for the *s3m24* neck model. Only 12 muscle groups are listed here but each has a left and right component, resulting in 24 muscle groups.

hyoids	SCM	scalenes	LC	Lcap	mf
digastric post	stern mast	scalenus ant	long col c1thx	long cap sklC4	semi cerv C3thx
digastric ant	cleid mast	scalenus med	long col c1c5		mult sup C45C2
mylohyoid ant	cleid occ	scalenus post	long col c5thx		mult sup C56C2
mylohyoid post					mult sup C67C2
stylohyoid lat					mult sup T1C4
stylohyoid med					mult sup T1C5
sternohyoid					mult sup T2C6
geniohyoid					mult deep C45C2
sternothyroid					mult deep C56C3
omohyoid					mult deep C67C4
					mult deep T1C5
					mult deep T1C6
					mult deep T2C7

trap	spcap	semicap	levscap	spcerv	subocc
trap cl	splen cap sklc6	semi cap sklc5	splen cerv c3thx	longissi cerv c4thx	rectcap post maj
trap acr	splen cap sklthx	semi cap sklthx	levator scap	iliocost cerv c5rib	rectcap post min
	longissi cap sklc6				obl cap sup
					obl cap inf

Table B.3. Muscle parameters for the *s3m24* neck model. The muscle groupings are listed in Table B.2. Pitch: (+) Extension, (-) Flexion; Roll: (+) R Sidebend, (-) L Sidebend; Yaw: (+) L Rotation, (-) R Rotation.

	Name	Strength	T1-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
		(N)	pitch	roll	yaw	pitch	roll	yaw	pitch	roll	yaw
1	hyoids-R	98.0	-0.0236	0.0123	-0.0097	-0.0285	0.0084	-0.0093	-0.0630	0.0863	0.0445
2	hyoids-L	98.0	-0.0236	-0.0123	0.0097	-0.0285	-0.0084	0.0093	-0.0630	-0.0863	-0.0445
3	SCM-R	105.5	-0.0420	0.0659	0.0269	-0.0251	0.0746	0.0259	0.0140	0.0971	0.0244
4	SCM-L	105.5	-0.0420	-0.0659	-0.0269	-0.0251	-0.0746	-0.0259	0.0140	-0.0971	-0.0244
5	scalenes-R	89.1	-0.0084	0.0677	0.0077	-0.0006	0.0382	0.0058			
6	scalenes-L	89.1	-0.0084	-0.0677	-0.0077	-0.0006	-0.0382	-0.0058			
7	LC-R	24.9	-0.0062	0.0139	-0.0023	-0.0071	0.0147	-0.0058			
8	LC-L	24.9	-0.0062	-0.0139	0.0023	-0.0071	-0.0147	0.0058			
9	trap-R	372.7	-0.0002	0.0266	0.0145	0.0050	0.0214	0.0151	0.0178	0.0081	0.0160
10	trap-L	372.7	-0.0002	-0.0266	-0.0145	0.0050	-0.0214	-0.0151	0.0178	-0.0081	-0.0160
11	spcap-R	83.8	0.0325	0.0096	-0.0125	0.0765	0.0312	-0.0293	0.0765	0.0731	-0.0322
12	spcap-L	83.8	0.0325	-0.0096	0.0125	0.0765	-0.0312	0.0293	0.0765	-0.0731	0.0322
13	semicap-R	136.8	0.0321	0.0246	0.0106	0.0473	0.0346	0.0193	0.0893	0.0218	0.0202
14	semicap-L	136.8	0.0321	-0.0246	-0.0106	0.0473	-0.0346	-0.0193	0.0893	-0.0218	-0.0202

Table B.3 continued from previous page

	Name	Strength	T1-C7 Moment Arm (m)			C4-C5 Moment Arm (m)			C0-C1 Moment Arm (m)		
15	levscap-R	120.1	0.0557	0.0522	-0.0089	0.0418	0.0420	-0.0078			
16	levscap-L	120.1	0.0557	-0.0522	0.0089	0.0418	-0.0420	0.0078			
17	mf-R	116.7	0.0619	0.0323	0.0221	0.0595	0.0227	0.0247			
18	mf-L	116.7	0.0619	-0.0323	-0.0221	0.0595	-0.0227	-0.0247			
19	spcerv-R	63.6	0.0444	0.0373	-0.0159	0.0177	0.0194	-0.0098			
20	spcerv-L	63.6	0.0444	-0.0373	0.0159	0.0177	-0.0194	0.0098			
21	subocc-R	100.5							0.0462	0.0179	0.0002
22	subocc-L	100.5							0.0462	-0.0179	-0.0002
23	Lcap-R	48.0							-0.0120	0.0078	-0.0021
24	Lcap-L	48.0							-0.0120	-0.0078	0.0021

Table B.4. Muscle groupings for the *p3m12* planar neck model. The right and left muscles are combined due to the symmetry in the sagittal plane.

hyoids	SCM	scalenes	Lcol	Lcap	multif
digastric post	stern mast	scalenus ant	long col c1thx	long cap sklC4	semi cerv C3thx
digastric ant	cleid mast	scalenus med	long col c1c5		mult sup C45C2
mylohyoid ant	cleid occ	scalenus post	long col c5thx		mult sup C56C2
mylohyoid post					mult sup C67C2
stylohyoid lat					mult sup T1C4
stylohyoid med					mult sup T1C5
sternohyoid					mult sup T2C6
geniohyoid					mult deep C45C2
sternothyroid					mult deep C56C3
omohyoid					mult deep C67C4
					mult deep T1C5
					mult deep T1C6
					mult deep T2C7

trap	spcap	semicap	levscap	spcerv	subocc
trap cl	splen cap sklc6	semi cap sklc5	splen cerv c3thx	longissi cerv c4thx	rectcap post maj
trap acr	splen cap sklthx	semi cap sklthx	levator scap	iliocost cerv c5rib	rectcap post min
	longissi cap sklc6				obl cap sup
					obl cap inf

Table B.5. Muscle parameters for the *p3m12* planar neck model. The muscle groupings are listed in Table B.4. The moment arm is in the sagittal plane where: (+) Extension, (-) Flexion

	Name	Strength (N)	Moment Arm (m) C6-C7	C4-C5	C0-C1
1	hyoids	138.6	-0.0334	-0.0403	-0.0891
2	SCM	149.1	-0.0594	-0.0355	0.0198
3	scalenes	126.1	-0.0119	-0.0009	
4	Lcol	35.3	-0.0087	-0.0101	
5	Lcap	67.8			-0.0170
6	trap	186.6	-0.0008	0.0201	0.0710
7	spcap	118.5	0.0459	0.1082	0.1081
8	semicap	193.4	0.0453	0.0670	0.1262
9	levscap	169.9	0.0787	0.0591	
10	multif	163.9	0.0882	0.0848	
11	spcerv	89.9	0.0628	0.0250	
12	subocc	142.1			0.0653

Table B.6. Muscle groupings for the *p2m7* planar neck model. The right and left muscles are combined due to the symmetry in the sagittal plane.

TC-E	TC-F	TS-B	TS-E	TS-F	CS-E	CS-F
scalenus post	scalenus ant	stern mast	trap cl	digastric post	splen cap sklc6	long cap sklC4
splen cerv c3thx	scalenus med	cleid mast	splen cap sklthx	digastric ant		
semi cerv c3thx	long col c1thx	cleid occ	semi cap sklthx	mylohyoid ant		
levator scap	long col c5thx			mylohyoid post		
longissi cerv c4thx				stylohyoid lat		
iliocost cerv c5rib				stylohyoid med		
mult sup C67C2				sternohyoid		
mult sup T1C4				geniohyoid		
mult sup T1C5				sternothyroid		
mult sup T2C6				omohyoid		
mult deep C67C4						
mult deep T1C5						
mult deep T1C6						

Table B.7. Muscle parameters for the *p2m7* planar neck model. The muscle groupings are listed in Table B.6. The moment arm is in the sagittal plane where: (+) Extension, (-) Flexion

	Name	Strength (N)	Moment Arm (m) T1-C7	C0-C1
1	TC-E	254.8	0.1323	
2	TC-F	118.4	-0.0172	
3	TS-B	149.1	-0.0594	0.0198
4	TS-E	247.2	0.0569	0.1285
5	TS-F	138.6	-0.0334	-0.0891
6	CS-E	213.1		0.1314
7	CS-F	67.8		-0.0170

References

[1] Rebecca Abbott, Ashley Pedler, Michele Sterling, Julie Hides, Todd Murphey, Mark Hoggarth, and James Elliott. The geography of fatty infiltrates within the cervical multifidus and semispinalis cervicis in individuals with chronic whiplash-associated disorders. *The Journal of orthopaedic and sports physical therapy*, 45(4):281–288, 4 2015.

[2] Rebecca E. Abbott, Todd B. Parrish, Mark A. Hoggarth, Andrew C. Smith, and James M. Elliott. Letter to the editor regarding Smuck M, Cristostomo RA, Demirjian R, et al. Morphologic change in the lumbar spine after lumbar medial branch radiofrequency neurotomy: A quantitative radiological study., 2014.

[3] Mina Alizadeh, Gregory G. Knapik, Prasath Mageswaran, Ehud Mendel, Eric Bourekas, and William S. Marras. Biomechanical musculoskeletal models of the cervical spine: A systematic literature review, 2020.

[4] J S Anderson, A W Hsu, and A N Vasavada. Morphology, architecture, and biomechanics of human cervical multifidus. *Spine (Phila Pa 1976)*, 30(4):86–91, 2005.

[5] Richard C. Aster, Brian Borchers, and Clifford H. Thurber. *Parameter Estimation and Inverse Problems*. Elsevier, 2019.

[6] Mark A. Beaumont. Approximate Bayesian Computation. *Annual Review of Statistics and Its Application*, 6(1):379–403, 3 2019.

[7] Mark A Beaumont, J.-M. Cornuet, J.-M. Marin, and Christian P Robert. Adaptive approximate Bayesian computation. *Biometrika*, 96(4):983–990, 12 2009.

[8] Anders Bergmark. Stability of the Lumbar Spine: A study in mechanical engineering. *Acta Orthop Scand*, 60(230), 1989.

[9] N.A. Bernstein. *The Co-ordination and Regulation of Movement*. Pergamon Press, Oxford, 1967.

[10] C S Bexander and P W Hodges. Cervico-ocular coordination during neck rotation is distorted in people with whiplash-associated disorders. *Exp Brain Res*, 217(1):67–77, 2012.

[11] J S Blouin, G P Siegmund, M G Carpenter, and J T Inglis. Neural control of superficial and deep neck muscles in humans. *J Neurophysiol*, 98(2):920–928, 2007.

[12] M. G. B. Blum, M. A. Nunes, D. Prangle, and S. A. Sisson. A Comparative Review of Dimension Reduction Methods in Approximate Bayesian Computation. *Statistical Science*, 28(2):189–208, 5 2013.

[13] Grethe Borchgrevink, Olaug Smevik, Irma Haave, Olav Haraldseth, Asbjørn Nordby, and Inggard Lereim. MRI of cerebrum and cervical columna within two days after whiplash neck sprain injury. *Injury*, 28(5-6):331–335, 6 1997.

[14] Linda J. Carroll, Lena W. Holm, Sheilah Hogg-Johnson, Pierre Côtè, J. David Cassidy, Scott Haldeman, Margareta Nordin, Eric L. Hurwitz, Eugene J. Carragee, Gabrielle van der Velde, Paul M. Peloso, and Jaime Guzman. Course and Prognostic Factors for Neck Pain in Whiplash-Associated Disorders (WAD). *Journal of Manipulative and Physiological Therapeutics*, 32(2):S97–S107, 2 2009.

[15] Francesco M. Colacino, Rustighi Emiliano, and Brian R. Mace. Subject-specific musculoskeletal parameters of wrist flexors and extensors estimated by an EMG-driven musculoskeletal model. *Medical Engineering and Physics*, 34(5), 2012.

[16] M J Comerford and S L Mottram. Movement and stability dysfunction–contemporary developments. *Manual therapy*, 6(1):15–26, 2 2001.

[17] E. de Bruijn, F. C.T. van der Helm, and R. Happee. Analysis of isometric cervical strength with a nonlinear musculoskeletal model with 48 degrees of freedom. *Multibody System Dynamics*, 2016.

[18] Scott L. Delp, Frank C. Anderson, Allison S. Arnold, Peter Loan, Ayman Habib, Chand T. John, Eran Guendelman, and Darryl G. Thelen. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering*, 54(11), 11 2007.

[19] Wolfgang von der Linden, Volker Dose, and Udo von Toussaint. *Bayesian Probability Theory*. Cambridge University Press, Cambridge, 2014.

[20] J P Dulor, B Cambon, P Vigneron, Y Reyne, J Nouguas, L Casteilla, and F Bacou. Expression of specific white adipose tissue genes in denervation-induced skeletal muscle fatty degeneration. *FEBS Lett.*, 439(1-2):89–92, 1998.

[21] J C Eck, S D Hodges, and S C Humphreys. Whiplash: a review of a commonly misunderstood injury. *Am. J. Med.*, 110(8):651–656, 2001.

[22] J Elliott, G Jull, J T Noteboom, R Darnell, G Galloway, and W W Gibbon. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders: a magnetic resonance imaging analysis. *Spine (Phila Pa 1976)*, 31(22):847–55, 2006.

[23] James Elliott, Gwendolen Jull, Jon Timothy Noteboom, Ross Darnell, Graham Galloway, and Wayne W Gibbon. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders. *Spine*, 31(22):847–55, 2006.

[24] James Elliott, Gwendolen Jull, Jon Timothy Noteboom, Ross Darnell, Graham Galloway, and Wayne W. Gibbon. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders: A magnetic resonance imaging analysis. *Spine*, 2006.

[25] James Elliott, Michele Sterling, Jon Timothy Noteboom, Julia Treleaven, Graham Galloway, and Gwendolen Jull. The clinical presentation of chronic whiplash and the relationship to findings of MRI fatty infiltrates in the cervical extensor musculature:

a preliminary investigation. *European Spine Journal*, 18(9), 9 2009.

[26] James M Elliott. Are There Implications for Morphological Changes in Neck Muscles After Whiplash Injury? *Spine*, 36(25 Suppl):S205–S210, 12 2011.

[27] James M Elliott, Jon Timothy Noteboom, Timothy W Flynn, and Michele Sterling. Characterization of acute and chronic whiplash-associated disorders. *J Orthop Sports Phys Ther*, 39(5):312–323, 5 2009.

[28] James M Elliott, Shaun O'Leary, Michele Sterling, Joan Hendrikz, Ashley Pedler, and Gwen Jull. Magnetic resonance imaging findings of fatty infiltrate in the cervical flexors in chronic whiplash. *Spine*, 35(9):948–954, 2010.

[29] James M Elliott, Ashley Pedler, Gwendolen Jull, Luke Van Wyk, Graham Galloway, and Shaun P OLeary. Differential changes in muscle composition exist in traumatic and nontraumatic neck pain. *Spine*, 39(1):39–47, 2014.

[30] D. Falla, G. Jull, and P. W. Hodges. Feedforward activity of the cervical flexor muscles during voluntary arm movements is delayed in chronic neck pain. *Experimental Brain Research*, 157(1):43–48, 7 2004.

[31] Deborah Falla, Gina Bilenkij, and Gwendolen Jull. Patients with chronic neck pain demonstrate altered patterns of muscle activation during performance of a functional upper limb task. *Spine*, 2004.

[32] Paul Fearnhead and Dennis Prangle. Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474, 6 2012.

[33] Jason B. Fice, Gunter P. Siegmund, and Jean Sébastien Blouin. Prediction of three dimensional maximum isometric neck strength. *Annals of Biomedical Engineering*, 2014.

[34] Jason B. Fice, Gunter P. Siegmund, and Jean-Sébastien Blouin. Neck muscle biomechanics and neural control. *Journal of Neurophysiology*, 120(1):361–371, 7 2018.

[35] Sarah Filippi, Chris P. Barnes, Julien Cornebise, and Michael P.H. Stumpf. On optimality of kernels for approximate Bayesian computation using sequential Monte Carlo. *Statistical Applications in Genetics and Molecular Biology*, 12(1), 1 2013.

[36] J. L. Fleckenstein, D. Watumull, K. E. Conner, M. Ezaki, R. G. Greenlee, W. W. Bryan, D. P. Chason, R. W. Parkey, R. M. Peshock, and P. D. Purdy. Denervated human skeletal muscle: MR imaging evaluation. *Radiology*, 187(1):213–218, 4 1993.

[37] Michele Gaeta, Emanuele Scribano, Achille Mileto, Silvio Mazziotti, Carmelo Rodolico, Antonio Toscano, Nicola Settineri, Giorgio Ascenti, and Alfredo Blandino. Muscle fat fraction in neuromuscular disorders: dual-echo dual-flip-angle spoiled gradient-recalled MR imaging technique for quantification–a feasibility study. *Radiology*, 259(2):487–494, 2011.

[38] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 11 2013.

[39] C. Gerber, D. C. Meyer, A. G. Schneeberger, H. Hoppeler, and B. Von Rechenberg. Effect of tendon release and delayed repair on the structure of the muscles of the rotator cuff: An experimental study in sheep. *Journal of Bone and Joint Surgery - Series A*, 2004.

[40] Christian Gerber, Alberto G. Schneeberger, Hans Hoppeler, and Dominik C. Meyer. Correlation of atrophy and fatty infiltration on strength and integrity of rotator cuff repairs: A study in thirteen patients. *Journal of Shoulder and Elbow Surgery*, 2007.

[41] Benjamin Goislard de Monsabert, G. Rao, A. Gay, E. Berton, and L. Vigouroux. A scaling method to individualise muscle force capacities in musculoskeletal models of the hand and wrist using isometric strength measurements. *Medical & Biological Engineering & Computing*, 55(12):2227–2244, 12 2017.

[42] Karen Grimmer. Measuring the endurance capacity of the cervical short flexor muscle group. *Australian Journal of Physiotherapy*, 1994.

[43] H Grip, G Sundelin, B Gerdle, and J Stefan Karlsson. Cervical helical axis characteristics and its center of rotation during active head and upper arm movements-comparisons of whiplash-associated disorders, non-specific neck pain and asymptomatic individuals. *J Biomech*, 41(13):2799–2805, 2008.

[44] H J Hislop and Jacqueline Montgomery. *Daniels and Worthingham's Muscle testing*. 2014.

[45] P Hodges, A K Holm, T Hansson, and S Holm. Rapid atrophy of the lumbar multifidus follows experimental disc or nerve root injury. *Spine (Phila Pa 1976)*, 31(25):2926–2933, 2006.

[46] P W Hodges and C A Richardson. Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. *Spine (Phila Pa 1976)*, 21(22):2640–2650, 1996.

[47] Phillip C. Hughes, Nicholas F. Taylor, and Rod A. Green. Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. *Australian Journal of Physiotherapy*, 54(3), 2008.

[48] Aaron Courville Ian Goodfellow, Yoshua Bengio. *Deep Learning*. MIT press, Cambridge, vol. 1, no edition, 2016.

[49] Benjamin W. Infantolino and John H. Challis. Measuring Subject Specific Muscle Model Parameters of the First Dorsal Interosseous In Vivo. *Annals of Biomedical Engineering*, 42(6):1331–1339, 6 2014.

[50] Franck Jabot, Thierry Faure, and Nicolas Dumoulin. EasyABC: performing efficient approximate Bayesian computation sampling schemes using R. *Methods in Ecology and Evolution*, 4(7):684–687, 7 2013.

[51] G. Jull, E. Kristjansson, and P. Dall'Alba. Impairment in the cervical flexors: A comparison of whiplash and insidious onset neck pain patients. *Manual Therapy*, 2004.

[52] Gwendolen a Jull, Shaun P O'Leary, and Deborah L Falla. Clinical assessment of the deep cervical flexor muscles: the craniocervical flexion test. *Journal of manipulative*

and physiological therapeutics, 31(7):525–33, 9 2008.

[53] Birgit Juul-Kristensen, Brian Clausen, Inge Ris, Rikke Vikær Jensen, Rasmus Fischer Steffensen, Shadi Samir Chreiteh, Marie Birk Jørgensen, and Karen Søgaard. Increased neck muscle activity and impaired balance among females with whiplash-related chronic neck pain: A crosssectional study. *Journal of Rehabilitation Medicine*, 2013.

[54] Christopher A. Kahn and Catherine S. Gotschall. The Economic and Societal Impact of Motor Vehicle Crashes, 2010 (Revised). *Annals of Emergency Medicine*, 66(2):194–196, 8 2015.

[55] S. Kamath, N. Venkatanarasimha, M. A. Walsh, and P. M. Hughes. MRI appearance of muscle denervation. *Skeletal Radiology*, 37(5):397–404, 5 2008.

[56] Lynne K. Kamibayashi and Frances J.R. Richmond. Morphometry of Human Neck Muscles. *Spine*, 23(12):1314–1323, 6 1998.

[57] Henry Kendall and Florence Kendall. *Muscles. Testing and Function*. Williams & Wilkins Company, Baltimore, 1949.

[58] E. A. Keshner, D. Campbell, R. T. Katz, and B. W. Peterson. Neck muscle activation patterns in humans during isometric head stabilization. *Experimental Brain Research*, 1989.

[59] Terry K.K. Koo, Arthur F.T. Mak, and L.K Hung. In vivo determination of subject-specific musculotendon parameters: applications to the prime elbow flexors in normal and hemiparetic subjects. *Clinical Biomechanics*, 17(5):390–399, 6 2002.

[60] Jostein Krakenes, B. Kaale, G. Moen, H. Nordli, N. Gilhus, and J. Rorvik. MRI assessment of the alar ligaments in the late stage of whiplash injury – a study of structural abnormalities and observer agreement. *Neuroradiology*, 44(7):617–624, 7 2002.

[61] Søren Krogh and Helge Kasch. Whiplash injury results in sustained impairments of cervical muscle function: A one-year prospective, controlled study. *Journal of Rehabilitation Medicine*, 2018.

[62] Jason J. Kutch and Francisco J. Valero-Cuevas. Muscle redundancy does not imply robustness to muscle dysfunction. *Journal of Biomechanics*, 2011.

[63] Haejung Lee, Leslie L. Nicholson, and Roger D. Adams. Neck Muscle Endurance, Self-Report, and Range of Motion Data From Subjects With Treated and Untreated Neck Pain. *Journal of Manipulative and Physiological Therapeutics*, 28(1):25–32, 1 2005.

[64] J P Lefaucheur, B Gjata, H Lafont, and A Sebille. Angiogenic and inflammatory responses following skeletal muscle injury are altered by immune neutralization of endogenous basic fibroblast growth factor, insulin-like growth factor-1 and transforming growth factor-beta 1. *J Neuroimmunol*, 70(1):37–44, 1996.

[65] Maxime Lenormand, Franck Jabot, and Guillaume Deffuant. Adaptive approximate Bayesian computation for complex models. *Computational Statistics*, 28(6):2777–2796, 12 2013.

[66] Jarno Lintusaari, Michael U. Gutmann, Ritabrata Dutta, Samuel Kaski, and Jukka Corander. Fundamentals and Recent Developments in Approximate Bayesian Computation. *Systematic Biology*, 66(1):syw077, 10 2016.

[67] Jarno Lintusaari, Henri Vuollekoski, Antti Kangasrääsiö, Kusti Skytén, Marko Järvenpää, Pekka Marttinen, Michael U. Gutmann, Aki Vehtari, Jukka Corander, and Samuel Kaski. ELFI: Engine for Likelihood-Free Inference. *Journal of Machine Learning Research*, 8 2017.

[68] S M Lord, L Barnsley, B J Wallis, and N Bogduk. Chronic cervical zygapophysial joint pain after whiplash. A placebo-controlled prevalence study. *Spine (Phila Pa 1976)*, 21(15):1735–1737, 1996.

[69] D A MacDonald, G L Moseley, and P W Hodges. The lumbar multifidus: does the evidence support clinical beliefs? *Man Ther*, 11(4):254–263, 2006.

[70] Jean Michel Marin, Pierre Pudlo, Christian P. Robert, and Robin J. Ryder. Approximate Bayesian computational methods. *Statistics and Computing*, 22(6), 2012.

[71] Ali Marjaninejad and Francisco J. Valero-Cuevas. Should Anthropomorphic Systems be "Redundant"? In *Springer Tracts in Advanced Robotics*, pages 7–34. 2019.

[72] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavare. Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324–15328, 12 2003.

[73] J. Martin Bland and DouglasG. Altman. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, 327(8476):307–310, 2 1986.

[74] J-B Masson, M. Bailly Bechet, and M. Vergassola. Chasing information to search in random environments. *Journal of Physics A: Mathematical and Theoretical*, 42(43):434009, 10 2009.

[75] Dominik C. Meyer, Christoph Pirkl, Christian W.A. Pfirrmann, Marco Zanetti, and Christian Gerber. Asymmetric atrophy of the supraspinatus muscle following tendon tear. *Journal of Orthopaedic Research*, 2005.

[76] Q. C. Miranda Moore and Simon Iveson. Maximal Voluntary Isometric Neck Strength Deficits in Adults With Whiplash. *Digital Evidence and Electronic Signature Law Review*, 2012.

[77] Douglas C Montgomery. Design and Analysis of Experiments. *Technometrics*, 48(1):158–158, 2 2006.

[78] Jonathan D. Mortensen, Anita N. Vasavada, and Andrew S. Merryweather. The inclusion of hyoid muscles improve moment generating capacity and dynamic simulations in musculoskeletal models of the head and neck. *PLoS ONE*, 2018.

[79] Marc J. Nederhand, Maarten J. IJzerman, Hermie J. Hermens, Chris T. M. Baten, and Gerrit Zilvold. Cervical Muscle Dysfunction in the Chronic Whiplash Associated Disorder Grade II (WAD-II). *Spine*, 25(15):1938–1943, 8 2000.

[80] M M Panjabi, J Cholewicki, K Nibu, L B Babat, and J Dvorak. Simulation of whiplash trauma using whole cervical spine specimens. *Spine (Phila Pa 1976)*,

23(1):17–24, 1998.

[81] M M Panjabi, J Cholewicki, K Nibu, J Grauer, L B Babat, and J Dvorak. Critical load of the human cervical spine: an in vitro experimental study. *Clin Biomech (Bristol, Avon)*, 13(1):11–17, 1998.

[82] Isabelle Pearson, Alison Reichert, Sophie J. De Serres, Jean-Pierre Dumas, and Julie N. Côté. Maximal Voluntary Isometric Neck Strength Deficits in Adults With Whiplash-Associated Disorders and Association With Pain and Fear of Movement. *Journal of Orthopaedic & Sports Physical Therapy*, 39(3):179–187, 3 2009.

[83] D. Prangle, M. G. B. Blum, G. Popovic, and S. A. Sisson. Diagnostic tools for approximate Bayesian computation using the coverage property. *Australian & New Zealand Journal of Statistics*, 56(4):309–329, 12 2014.

[84] Dennis Prangle. Adapting the ABC Distance Function. *Bayesian Analysis*, 12(1):289–309, 3 2017.

[85] Jonathan K. Pritchard, Mark T. Seielstad, Anna Perez-Lezaun, and Marcus W. Feldman. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, 16(12):1791–1798, 12 1999.

[86] Tamara Prushansky, Reuven Gepstein, Carlos Gordon, and Zeevi Dvir. Cervical muscles weakness in chronic whiplash patients. *Clinical Biomechanics*, 20(8):794–798, 10 2005.

[87] D. E. Rassier, B. R. MacIntosh, and W. Herzog. Length dependence of active force production in skeletal muscle, 1999.

[88] Hans R. Ronnen, P J de Korte, Peter R.G. Brink, H J van der Bijl, Alphons J. Tonino, and Cees L. Franke. Acute whiplash injury: is there a role for MR imaging?–a prospective study of 100 patients. *Radiology*, 201(1):93–96, 10 1996.

[89] Jochen Schomacher and Deborah Falla. Function and structure of the deep cervical extensor muscles in patients with neck pain. *Manual Therapy*, 18(5):360–366, 10 2013.

[90] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(4):623–656, 10 1948.

[91] G P Siegmund, B S Myers, M B Davis, H F Bohnet, and B A Winkelstein. Mechanical evidence of cervical facet capsule injury during whiplash: a cadaveric study using combined shear, compression, and extension loading. *Spine (Phila Pa 1976)*, 26(19):2095–2101, 2001.

[92] S. A. Sisson, Y. Fan, and Mark M. Tanaka. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765, 2 2007.

[93] AD Smith, Gwendolen Jull, and Geoff Schneider. Cervical radiofrequency neurotomy reduces central hyperexcitability and improves neck movement in individuals with chronic whiplash. *Pain Medicine*, 15:128–141, 2014.

[94] Mathew Smuck, Ralph a. Crisostomo, Ryan Demirjian, David S. Fitch, David J. Kennedy, and Michael E. Geisser. Morphologic changes in the lumbar spine after lumbar medial branch radiofrequency neurotomy: a quantitative radiological study.

The Spine Journal, 11 2013.

[95] Byran Smucker, Martin Krzywinski, and Naomi Altman. Optimal experimental design. *Nature Methods*, 15(8):559–560, 8 2018.

[96] M. Hongchul Sohn, J. Lucas McKay, and Lena H. Ting. Defining feasible bounds on muscle activation in a redundant biomechanical task: Practical implications of redundancy. *Journal of Biomechanics*, 2013.

[97] M. Hongchul Sohn, Daniel M. Smith, and Lena H. Ting. Effects of kinematic complexity and number of muscles on musculoskeletal model robustness to muscle dysfunction. *PLoS ONE*, 2019.

[98] Michele Sterling, James M Elliott, and Peter J Cabot. The course of serum inflammatory biomarkers following whiplash injury and their relationship to sensory and muscle measures: a longitudinal cohort study. *PloS one*, 8(10):e77903, 1 2013.

[99] Michele Sterling, Gwendolen Jull, and Justin Kenardy. Physical and psychological factors maintain long-term predictive capacity post-whiplash injury. *Pain*, 122(1):102–108, 5 2006.

[100] Michele Sterling, Gwendolen Jull, Bill Vicenzino, Justin Kenardy, and Ross Darnell. Development of motor system dysfunction following whiplash injury. *Pain*, 2003.

[101] Mikael Sunnåker, Alberto Giovanni Busetto, Elina Numminen, Jukka Corander, Matthieu Foll, and Christophe Dessimoz. Approximate Bayesian Computation. *PLoS Computational Biology*, 9(1):e1002803, 1 2013.

[102] Tina Toni, David Welch, Natalja Strelkowa, Andreas Ipsen, and Michael P.H. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, 6(31):187–202, 2 2009.

[103] Brandon M. Turner and Trisha Van Zandt. A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2):69–85, 4 2012.

[104] Francisco J. Valero-Cuevas. A mathematical approach to the mechanical capabilities of limbs and fingers. *Advances in Experimental Medicine and Biology*, 2009.

[105] Francisco J. Valero-Cuevas. *Fundamentals of Neuromechanics*, volume 8 of *Biosystems & Biorobotics*. Springer London, London, 2016.

[106] Francisco J. Valero-Cuevas. *Fundamentals of Neuromechanics*, volume 8 of *Biosystems & Biorobotics*. Springer London, London, 2016.

[107] Francisco J. Valero-Cuevas, Felix E. Zajac, and Charles G. Burgar. Large index-fingertip forces are produced by subject-independent patterns of muscle excitation. *Journal of Biomechanics*, 31(8):693–703, 8 1998.

[108] Anita N. Vasavada, Siping Li, and Scott L. Delp. Influence of Muscle Morphometry and Moment Arms on the Moment-Generating Capacity of Human Neck Muscles. *Spine*, 23(4):412–422, 2 1998.

[109] Anita N. Vasavada, Siping Li, and Scott L. Delp. Three-dimensional isometric strength of neck muscles in humans. *Spine*, 2001.

[110] Anita N. Vasavada, Barry W. Peterson, and Scott L. Delp. Three-dimensional spatial tuning of neck muscle activation in humans. *Experimental Brain Research*, 2002.

[111] Massimo Vergassola, Emmanuel Villermaux, and Boris I. Shraiman. 'Infotaxis' as a strategy for searching without gradients. *Nature*, 445(7126):406–409, 1 2007.

[112] Curtis R. Vogel. *Computational Methods for Inverse Problems*. Society for Industrial and Applied Mathematics, 1 2002.

[113] S. D. Walter, M. Eliasziw, and A. Donner. Sample size and optimal designs for reliability studies. *Statistics in Medicine*, 1998.

[114] Daniel Wegmann, Christoph Leuenberger, and Laurent Excoffier. Efficient Approximate Bayesian Computation Coupled With Markov Chain Monte Carlo Without Likelihood. *Genetics*, 182(4):1207–1218, 8 2009.

[115] B A Winkelstein, R W Nightingale, W J Richardson, and B S Myers. The cervical facet capsule and its role in whiplash injury: a biomechanical investigation. *Spine (Phila Pa 1976)*, 25(10):1238–1246, 2000.

[116] Astrid Woodhouse, øyvind Stavdahl, and Ottar Vasseljen. Irregular head movement patterns in whiplash patients during a trajectory task. *Experimental Brain Research*, 2010.

[117] Astrid Woodhouse and Ottar Vasseljen. Altered motor control patterns in whiplash and chronic neck pain. *BMC Musculoskeletal Disorders*, 2008.