Image registration and statistical analysis for quantitative in vivo spin-lock magnetic resonance imaging of the intervertebral disc response to compression

Joey Ann Kimdon

Electrical Engineering and Computer Sciences
University of California at Berkeley

August 2, 2007

Image registration and statistical analysis for quantitative *in vivo* spin-lock magnetic resonance imaging of the intervertebral disc response to compression

by

Joey Ann Kimdon

B.S. (Harvey Mudd College) 2000

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Thomas Budinger, Chair
Professor Sharmila Majumdar
Professor Jitendra Malik

Fall 2007

The dissertation of Joey Ann Kimdon is approved:

_____

Chair   Date

_____

Date

_____

Date

University of California, Berkeley

Fall 2007

Image registration and statistical analysis for quantitative *in vivo* spin-lock magnetic

resonance imaging of the intervertebral disc response to compression

Abstract

Image registration and statistical analysis for quantitative *in vivo* spin-lock magnetic

resonance imaging of the intervertebral disc response to compression

by

Joey Ann Kimdon

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Thomas Budinger, Chair

Low back pain decreases the quality of life and affects almost 80% of the population. Quantitative spin-lock magnetic resonance imaging of the $T_{1\rho}$ relaxation parameter is a promising marker of early degeneration in intervertebral discs. However, $T_{1\rho}$ measurements may be affected by disc compression, and understanding this effect is necessary for studying abnormal dynamics of disc response to compression or correlations between degeneration and $T_{1\rho}$ values.

We developed an algorithm to register images of intervertebral discs in different compression states and to calculate the statistical significance of local changes in $T_{1\rho}$. Our procedure includes automatic registration during image acquisition to view the same location across exams, segmentation of intervertebral discs with minimal user intervention,

automatic registration of discs using a non-rigid transformation guided by rigid transformations of vertebrae, and a method of investigating the statistical significance of changes in local neighborhoods using resampling hypothesis testing with variable confidence levels.

Validation tests on phantoms and volunteer data indicate sub-pixel registration accuracy and precision. Registration results agree with a manual gold standard. Hypothesis testing is sensitive to registration accuracy, indicating the need for registration even with patient movements of less than a millimeter. Tests confirm expected trade-offs between type I and type II statistical significance errors depending on neighborhood size, significance level, and confidence level.

We demonstrate the algorithm on varying compression states caused by lying down after 30 minutes of standing with a 20 lb backpack or by supporting 55 pounds while lying using a leg-press-like device. $T_{1\rho}$ varies locally depending on the distribution of compression within the discs. The coefficient of variation was 7.6% while relaxing and 13.9% when applying and releasing the compression device, indicating an effect of the device. Variation was 48% less in degenerated versus healthy discs ($p < 0.04$). $T_{1\rho}$ in the lower discs increased by 20% ($p < 0.05$) when the compression device pressure was released. Local comparisons showed that $T_{1\rho}$ tended to decrease under compression and increase after compression was released. Changes were not significant during an hour of supine resting, suggesting that variable pre-scan compression from normal daily activities does not considerably complicate $T_{1\rho}$ measurements.

To my family

Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my co-advisers, Prof. Thomas Budinger and Prof. Sharmila Majumdar for their guidance and support throughout my graduate studies. Their advice helped me develop an interesting and useful project, and working in their labs taught me many things.

I am grateful to the rest of my dissertation and qualifying exam committee members, Prof. Jitendra Malik and Prof. Stephen Derenzo, for their discussion and suggestions.

My gratitude goes to Dr. Thomas Link who helped me identify which aspects of the spine study would be clinically interesting and gave me a clinical perspective on the $T_{1\rho}$ results.

Thanks to Janet Blumenfeld, a fellow grad student, for assisting me with the prospective registration. We worked together extensively to implement it for the spine and for the axial/sagittal switch.

Xiaojuan Li developed and tested the $T_{1\rho}$-FSE sequence and helped me to keep it working throughout the project, for which I am much indebted.

I extend my thanks to Ronald Huesman, Gregory Klein, and Chaincy Kuo, whose work I built upon for the registration design.

Thanks to David Newitt and Kshama Agrawal for designing and building the compression device.

Many people helped with data acquisition, including Xiaojuan Li, Janet Blumenfeld, David Newitt, Radu Bolbos, Gabby Blumenkrantz, Ben Hyun, Radu Bolbos, Niles Bruce,

Glossary

Anterior: Body position toward the front/belly of a person. Opposite of posterior.

Axial: Plane normal to the long axis of the body. Splits the body into superior and inferior sections.

Coronal: Plane parallel to the long axis of the body, splitting the body into anterior and posterior sections.

Global registration: Alignment of images using the same transformation for the entire image.

Glycosaminoglycan (GAG) chains: Carbohydrate chains that provide osmotic pressure to maintain disc hydration. GAG chains and proteins make up proteoglycans.

Hypothesis testing: A test to determine the likelihood that differences in groups of measurements are due to random noise rather than representing true differences in the underlying distributions.

Inferior: Body position toward the feet of a person. Opposite of superior.

Magnetic resonance imaging (MRI): A non-ionizing method of non-invasively imaging inside the body using the interaction between a magnetic field and the nuclei of atoms in a sample.

Neighborhood: Pixels around a point of interest. Neighborhoods come in different sizes and shapes.

Non-rigid registration: Alignment of images using a transformation that does not necessarily preserve shapes.

p-value: The probability of getting a test statistic as or more extreme than the observed statistic due to random noise. Used in hypothesis testing to determine the significance of differences in a test statistic.

Pixels: Square individual values that make up an image. A pixel is a common choice of basis function to represent an image.

Posterior: Body position toward the back of a person. Opposite of anterior.

Prospective registration: Alignment of a pair of images before or during acquisition rather than after.

Proteoglycans: Negatively charged molecules made of protein and glycosaminoglycan chains. Proteoglycans make up the extracellular matrix of the intervertebral discs and are responsible for maintaining osmotic pressure and regulating the passage of molecules through the matrix. Proteoglycan loss is the most significant biochemical change in disc degeneration.

Registration: Alignment of two or more images such that pixels correspond to approximately the same anatomical region in both images.

Resampling statistics: A method of using samples of a population to represent the underlying distribution rather than assuming a particular distribution.

Rigid registration: Alignment of images using only translation and rotation. This type of transformation preserves shapes.

Sagittal: Plane parallel to the long axis of the body, splitting the body into right and left sections.

Segmentation: Classification of pixels in an image, often according to the type of tissue the pixels represent.

Superior: Body position toward the head of a person. Opposite of inferior.

$T_{1\rho}$: The spin-lattice MRI relaxation parameter in the rotating frame. It is emerging as a marker for early degeneration of intervertebral discs and articular cartilage.

Type I error: Incorrectly rejecting the null hypothesis in hypothesis testing. With this type of error, one asserts that there is a significant effect where there is none.

Type II error: Failing to reject the null hypothesis when it is false. With this type of er-

ror, one does not assert that there is an effect even though there is one.

Chapter 1

Introduction

Low back pain is a common problem causing a significant decrease in quality of life and increase in medical costs and lost workdays. It is the most common reason for filing worker's compensation claims in the United States, accounting for one third of all compensation costs and 40% of absences from work [1]. It is the most frequent cause for limitation in activity for Americans under age 45 [2]. An estimated 25% of Australians experience back pain at any given time, with over 80% experiencing back pain at some point in their lives [2]. Back pain affects youth as well, with 50% of Danish youth under age 20 reporting at least one back pain episode [3]. Degeneration and acute injury of intervertebral discs are common causes of back pain[1].

Magnetic resonance imaging (MRI) is a non-invasive method of imaging the human body without ionizing radiation. Anatomic MRI creates grayscale images that show inter-

---

[1]Back pain symptoms are from pressure on nerve roots and the spinal cord as well as facet joint inflammation. These causes are often the result of disc degeneration.

nal structures. Quantitative MRI determines absolute parameter values such as relaxation parameters T_1, T_2, and $T_{1\rho}$, water diffusion, and intervertebral disc dimensions and volume. $T_{1\rho}$ mapping [4, 5, 6, 7, 8], T_2 relaxometry [9, 10, 11, 12], and apparent tissue water diffusion [13, 9, 14] appear to identify degeneration in intervertebral discs [9, 14, 5, 7, 8, 12] and articular cartilage [13, 10, 4, 6] early in the disease progression, before the degeneration is visible in anatomic MRI. In particular, $T_{1\rho}$ mapping is emerging as a sensitive marker of early disc degeneration. It has been studied in knee articular cartilage [15, 4, 16], and is now being investigated in the intervertebral disc [17, 5, 18, 7]. Studies of disc degeneration seek to correlate *in vivo* $T_{1\rho}$ values with clinical diagnoses and to understand the relationship between proteoglycan content and $T_{1\rho}$ relaxation maps.

Compression is known to alter intervertebral disc properties such as T_2, diffusion, sodium concentration, and volume [19, 20, 21, 13, 22, 9], so it is reasonable to hypothesize that $T_{1\rho}$ measurements are also sensitive to pressure changes. As further motivation to study pressure changes, $T_{1\rho}$ is likely a marker for proteoglycan and water concentration, and water content in the disc depends on external loads and osmotic pressure from the concentration of negatively-charged glycosaminoglycan (GAG) chains [22]. Long-term compression has been shown to alter proteoglycan content of intervertebral discs in dogs [23, 24] and correlations between static loading and disc degeneration has been demonstrated in mouse tails [25].

Time of day, activity preceding the exam, and scanning position affects the amount and distribution of pressure in the intervertebral disc. If $T_{1\rho}$ depends on these factors, care must

be taken to control for them when using $T_{1\rho}$ to study disc degeneration. Disc height and volume measured with MRI differs in the evening, morning, and after 8 hours of walking [26]. Significant decreases have been shown in T_2 [27, 28], T_1, and proton number [19] between morning and evening measurements in sedentary workers. There may be similar changes in $T_{1\rho}$. Extra care must be taken to understand the effects of pressure difference when comparing healthy and degenerate discs since degenerate discs are expected to lose fluid more quickly under load and to have a lower fluid content than healthy discs. Hydraulic permeability increases with the decrease in proteoglycans [22] associated with disc degeneration, thereby affecting relaxation parameter values. For example, the diurnal change was significantly less pronounced in degenerative discs than in healthy discs for T_1 [19] and T_2 [27, 28]. The pressure in the disc at the time of the scan depends on the level of compression-inducing activity preceding the scan. Standing and muscular movements exert up to 20 times the pressure on the L4-L5 disc compared with lying prone [21, 22]. Decompression occurs while lying down [21], so the amount of time spent in the scanner before the $T_{1\rho}$ mapping sequence is run may affect results. Amount of curvature in the back (e.g. caused by varying amount of support placed under the knees during the scan) may cause uneven pressure on the discs. Any of these causes of pressure difference can be minimized during study design or at least considered during analysis if investigators are aware of the degree of influence.

Since the nucleus and annulus of intervertebral discs have different $T_{1\rho}$ values and compression will not be consistent over the entire disc depending on the amount of curvature in

the spine, we do not expect $T_{1\rho}$ values to change uniformly throughout the disc. Thus we consider changes in local neighborhoods at each pixel in the disc between images in varying states of compression. Since different compression states involve changes in position or re-entry into the scanner, registration is required to identify anatomically corresponding regions between images. A local statistical technique is then needed to analyze the registered images.

Dissertation overview   In this work, we develop an algorithm to analyze differences in $T_{1\rho}$ in lumbar intervertebral discs *in vivo*, regardless of position changes between scans. We emphasize automation throughout the process. Other studies attempting to compare quantitative parameters between scans have compared values within regions of interest (usually manually defined) placed on each of the images. Our approach is novel in that it relies on automatic registration and looks at local differences anywhere within the disc. Our specific novel contributions include:

- A data acquisition method for imaging the same anatomical regions across exams

- An automatic method for classifying intervertebral discs in $T_{1\rho}$-weighted MRI images

- An automatic method for warping intervertebral discs to align them between two images in a manner that is consistent with the rigid motion of the surrounding vertebrae

- A method for analyzing the statistical significance of local changes in $T_{1\rho}$ values and selecting the desired confidence level

- Validation testing for the above registration, segmentation, and statistical analysis methods

- Experiments using the above process to investigate the changes in $T_{1\rho}$ when intervertebral discs are subject to compression and decompression.

The algorithm is designed to be applied to studies investigating pressure effects as described above, but it may be used to analyze any $T_{1\rho}$ study with multiple exams per person. The method is likely applicable to studies of other quantitative MRI parameters such as T_2 mapping, but these parameters are not tested in this work.

Note that this dissertation develops a method of analyzing $T_{1\rho}$ changes but does not attempt to make conclusions about those changes in patients.

This work covers the process from image acquisition through statistical analysis of each image. Chapter 2 gives a background of magnetic resonance imaging, the MRI spin-lock technique for imaging the $T_{1\rho}$ parameter, and an overview of spinal anatomy to orient the reader for the images and terminology in the rest of the document. Figure 1.1 contains a diagram briefly describing the flow of our procedure. Chapters 3, 4, and 5 explain each step in the diagram. Specifically, chapter 3 details our procedure for image acquisition, Chapter 4 describes the registration and segmentation portions of our algorithm, and Chapter 5 explains our approach to statistical analysis. The latter two chapters begin with a background of applicable techniques in the field and then explain our methods.

We demonstrate the performance of the algorithm in chapter 6 with validation tests on volunteers and phantoms. Chapter 7 shows the practical use of the algorithm via prelim-

Step 1: Image acquisition

$T_{1\rho}$ maps are calculated from several $T_{1\rho}$-weighted images. We begin by calculating $T_{1\rho}$ maps at different times or in different positions. Registration guides the acquisition of the same anatomical location. We use the $T_{1\rho}$-weighted images for segmentation and registration because they show clearer anatomical information, and we use the $T_{1\rho}$ maps for the statistical analysis since they contain the information we seek.

First $T_{1\rho}$-weighted image
(i.e. position 1)

Second $T_{1\rho}$-weighted image
(i.e. position 2)
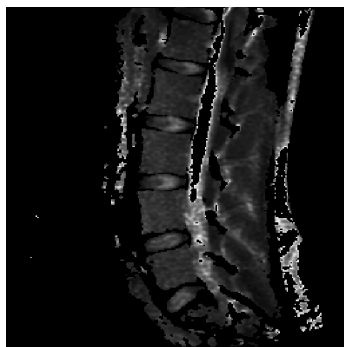
First $T_{1\rho}$ map
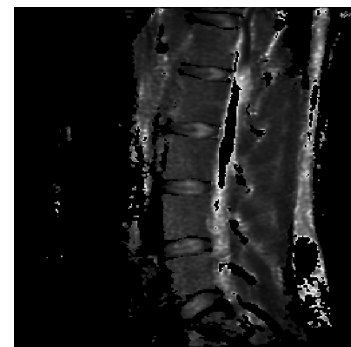
Second $T_{1\rho}$ map

Figure 1.1: The diagrams on this and the following pages outline the process described in this work. See chapter 3 for a more detailed explanation of this step. (Figure continued on the next page.)

Step 2: Segmentation

We use a $T_{1\rho}$-weighted image in the first ("fixed") position to
identify the intervertebral discs, which are the regions of interest
and the approximate vertebral bodies, which are the rigid bodies
that will guide the registration.
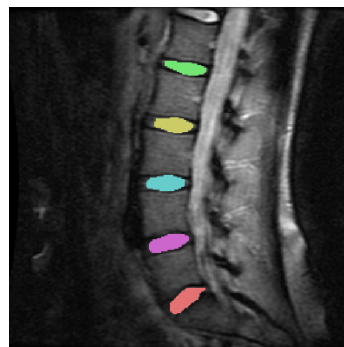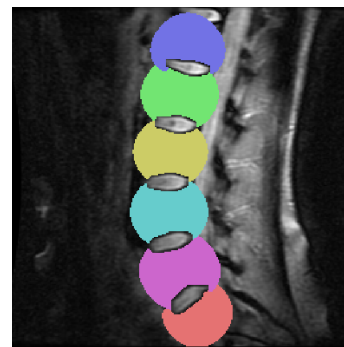
Discs   Rigid bodies

Figure 1.1: Algorithm overview continued. See chapter 4 for a more detailed explanation
of this step.

Step 3: Registration

Each rigid body is individually registered to align the second image with the first image. Points not within a rigid body are transformed according to an elastic transformation guided by the rigid transformations.

Before registration:

| First image | Second image (original position) | Difference between first and second (original position) |
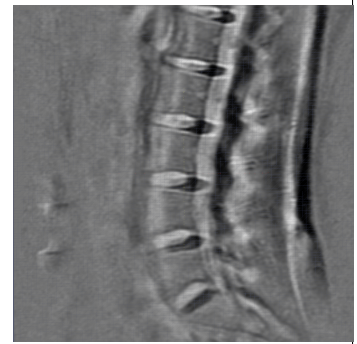
After registration:

| First image | Second image (transformed) | Difference between first and second (transformed) |
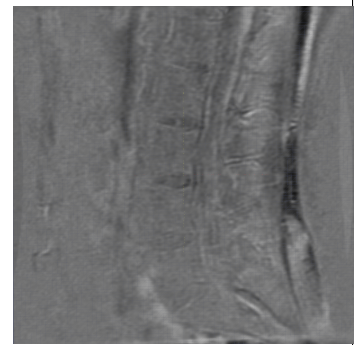
Figure 1.1: Algorithm overview continued. See chapter 4 for a more detailed explanation of this step.

Step 4: Statistical analysis

The transformation found in step 3 is applied to the $T_{1\rho}$ map values within the intervertebral discs. We calculate the statistical significance of the difference in mean $T_{1\rho}$ in a neighborhood around each pixel between the two exams. In the image below right, only pixels with statistically significant differences in mean are colored.
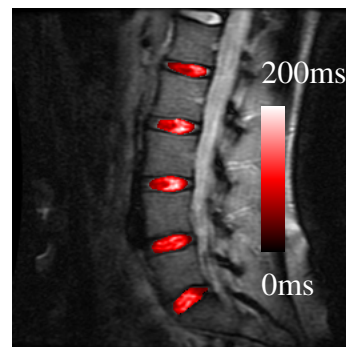
First map
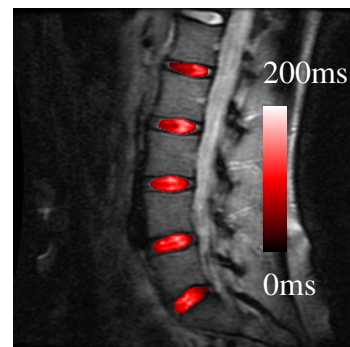
Second map
(transformed)
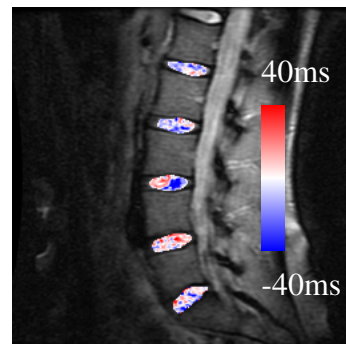
Map difference
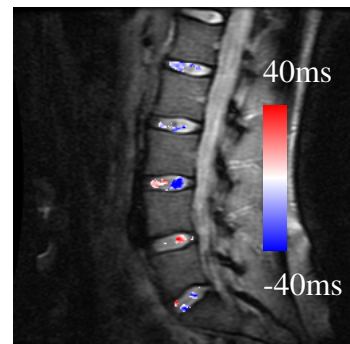
Significant differences.

Figure 1.1: Algorithm overview continued (last page). See chapter 5 for a more detailed explanation of this step.

inary compression experiments. The final chapter contains a summary, an analysis of the strengths and weaknesses of this project, and suggestions for future work.

Chapter 2

Background

2.1 Spinal anatomy

2.1.1 Anatomical terms of location

We first introduce terms to specify relative anatomical locations. Position is identified in three dimensions as right-left, anterior-posterior (front-back), and superior-inferior (top-bottom), illustrated in Figure 2.1. Two-dimensional planes through the body are referred to as sagittal (dividing into right/left sections), coronal (dividing into anterior/posterior sections), and axial or transverse (dividing into superior/inferior sections), as shown in Figure 2.1.
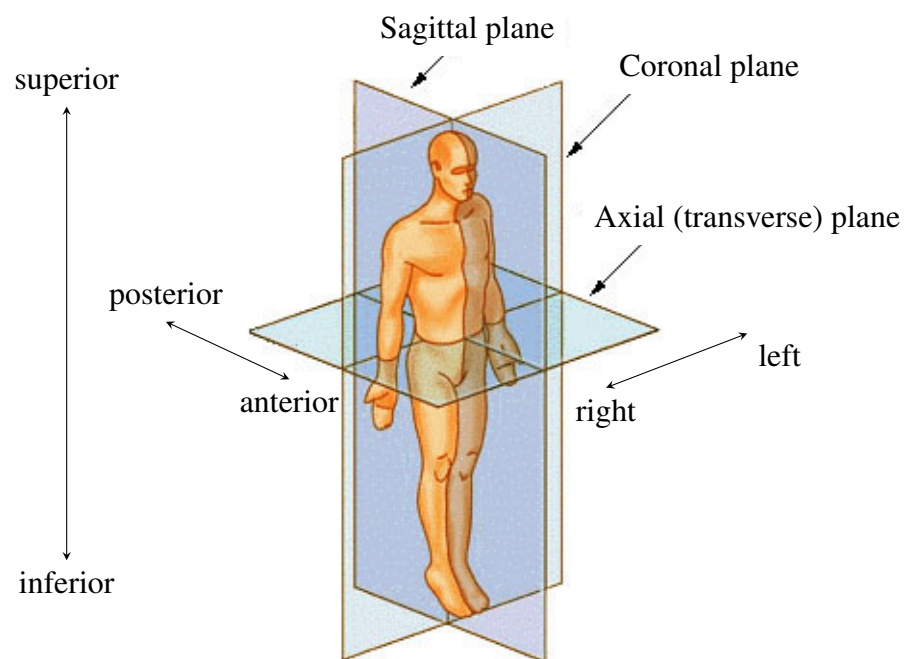
Figure 2.1: Terms of anatomical position. Image is in the public domain, obtained from [29].

2.1.2 Vertebrae

The spinal column consists of seven cervical, twelve thoracic, five lumbar, five fused sacral and four fused coccygeal vertebrae, shown in Figure 2.2. Vertebra shorthand uses the first letter of the location (C,T,L,S) and the number, with numbers increasing from superior to inferior. For example, L1 refers to the first lumbar vertebra, which is immediately inferior to T12, the twelfth and most inferior thoracic vertebra. Between the vertebrae are intervertebral discs, shown in Figure 2.3. We study the discs between L1 and S1, for they take much of the mechanical load of the spine and often exhibit pain and degeneration.
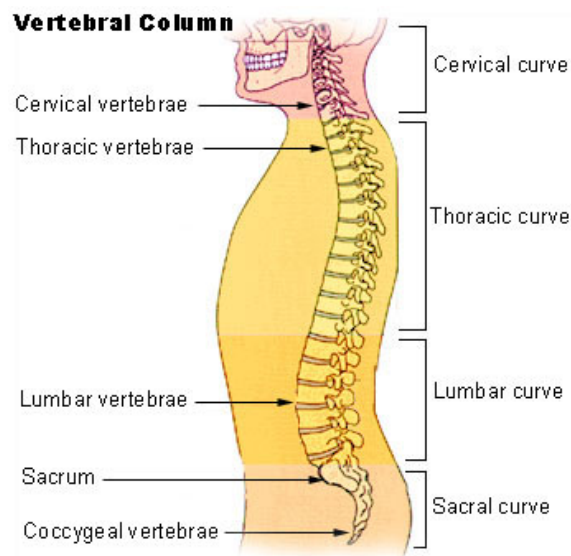
Figure 2.2: The vertebral column. Image is in the public domain, obtained from [30].

To help readers orient themselves in the MR images displayed in this document, we point out some anatomical landmarks in diagrams and sample MR images in the sagittal plane (Figures 2.3 and 2.4) and axial plane (Figures 2.5 and 2.6). The vertebral body is the cylindrical portion joined by intervertebral discs to create the column that bears most of

the compression through the spine. The transverse and spinous processes project outward

as attachment points for muscles and ligaments. The spinal cord runs through the vertebral

canal, the opening made of the vertebral foramen of successive vertebrae.
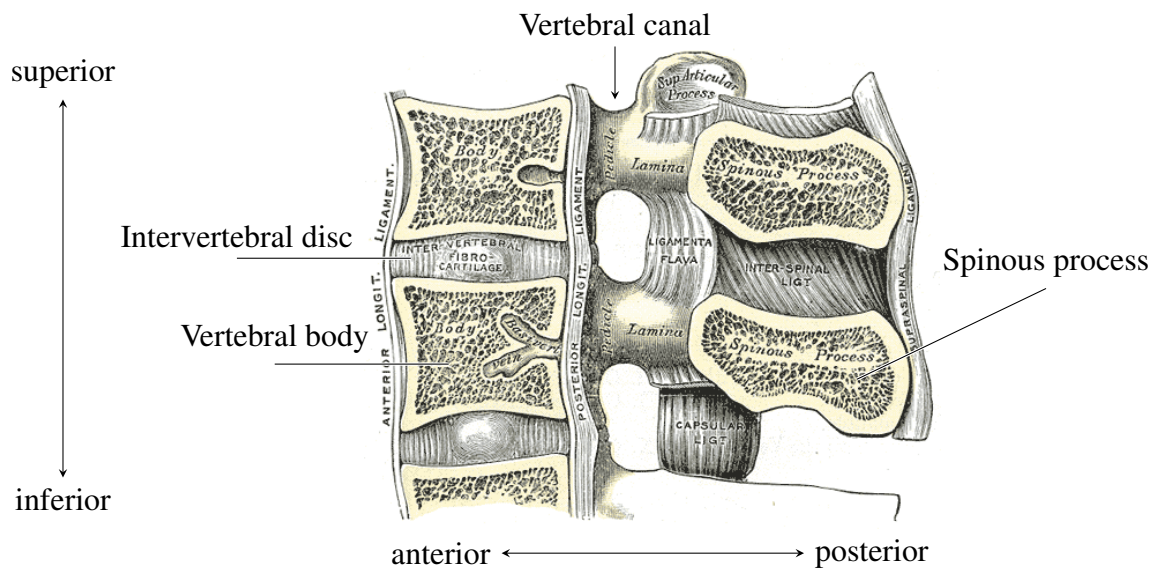
Figure 2.3: Diagram of two lumbar vertebrae, sliced through a sagittal plane. Image is in the public domain, obtained from [31].

2.1.3 Intervertebral discs

Intervertebral discs absorb stresses, protect the vertebral bodies from touching, and

allow the spine to bend. Their inner portion, called the nucleus pulposus, is a jelly-like

substance that cushions the spine from compressive forces. Surrounding the nucleus lat-

erally are concentric rings of fibrocartilage called the annulus fibrosus, which keeps the

nucleus contained and resists both compressive and tensile forces. Figure 2.7 points out

the locations of the nucleus and annulus in the sagittal and axial planes. Fibrocartilage is
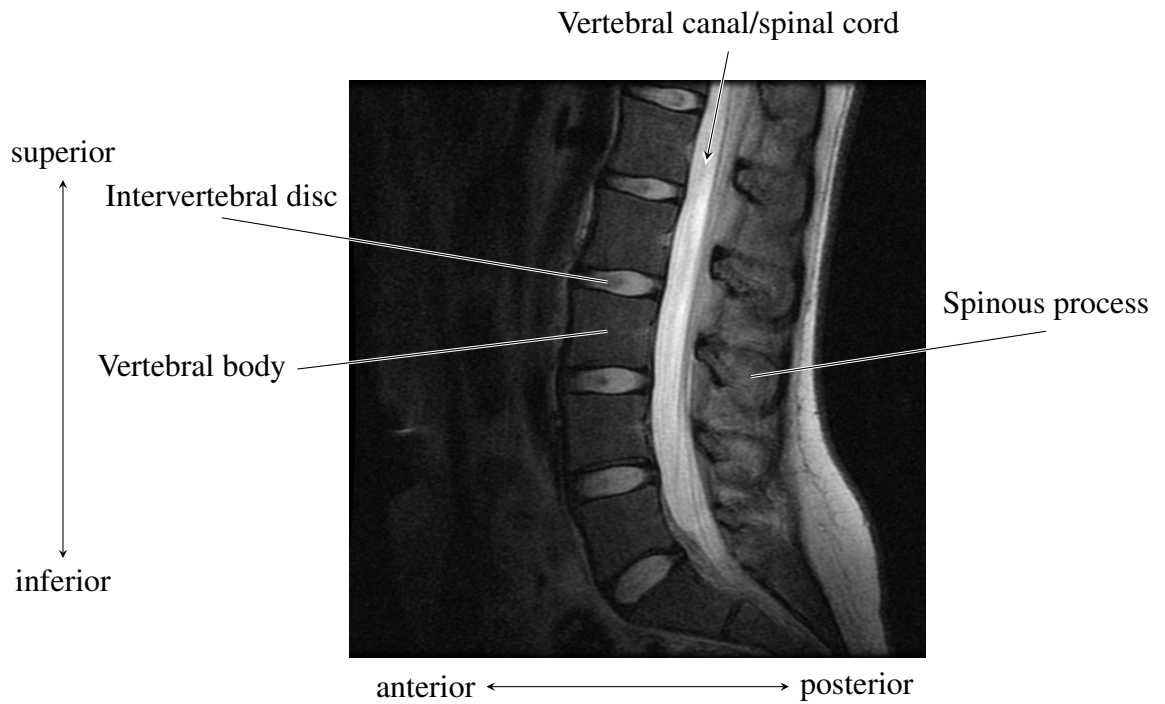
Figure 2.4: MR image of lumbar vertebrae: two-dimensional slice in the sagittal plane.
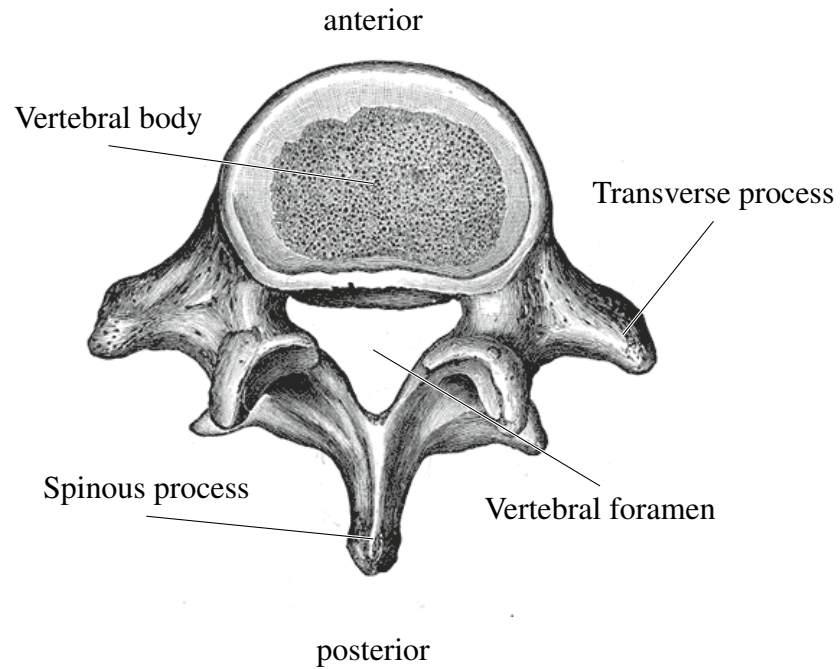
Figure 2.5: Diagram of a lumbar vertebra viewed from above. Image is in the public domain, obtained from [32].

anterior

Vertebral foramen

Vertebral body

Transverse process

Spinous process

posterior
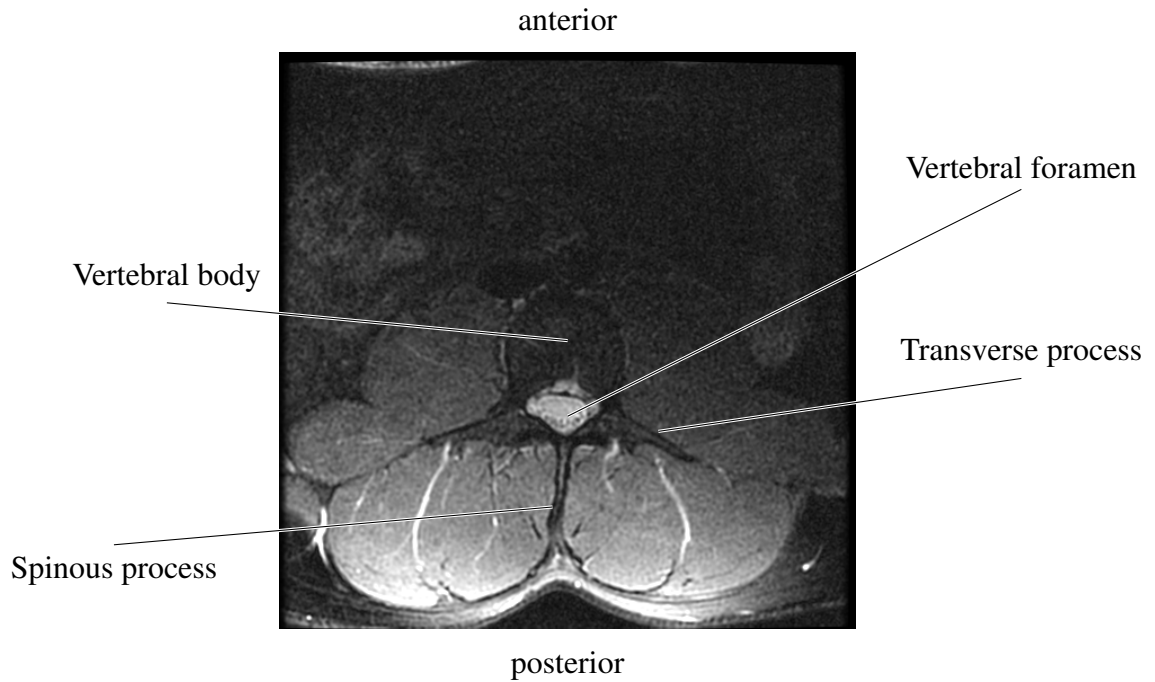
Figure 2.6: MR image of a lumbar vertebra: two-dimensional slice in the axial plane.

comprised of alternating rows of thick collagen fibers and chondrocytes (cartilage cells). The fibers cross through the concentric rings of the annulus to give discs strength in twisting. Among the collagen and chondrocytes is a matrix of proteoglycans, molecules made of protein and glycosaminoglycan (GAG) chains. GAG chains provide osmotic pressure to maintain disc hydration.

Loss of proteoglycan is the most significant biochemical change in disc degeneration, leading to reduced hydration, a more rapid loss of hydration and height under load, and a tendency to bulge [33]. The nucleus supports weight primarily via hydration pressure [34], so loss in hydration related to aging and degeneration reduces the disc's ability to absorb stress and protect the vertebrae. Loss of negatively charged proteoglycans may also permit

abnormal passage of positively charged ions and large uncharged molecules through the matrix, affecting cell behavior. Studies are investigating a direct correlation between the MR parameter $T_{1\rho}$ and proteoglycan content [7, 8, 5] to image early disc degeneration.
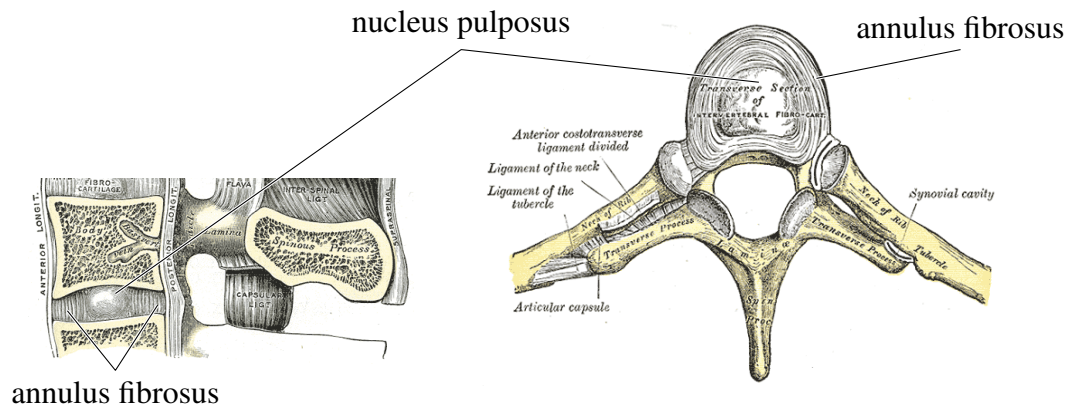
Figure 2.7: Identification of the disc nucleus pulposus and annulus fibrosus in a sagittal plane (left) and axial plane (right).

2.2 Magnetic resonance imaging

Magnetic resonance imaging (MRI) creates images via the interaction between an atom's nucleus and an external magnetic field. Protons and neutrons, nucleons, have a property called nuclear spin, and atoms with unpaired nucleons have a net spin and a magnetic moment. In clinical MRI, we are concerned with hydrogen atoms, which have only one proton and a net spin of 1/2. These magnetic moments are randomly aligned, canceling each other out to create no net magnetization. However, when placed in an external static magnetic field B_0, the magnetic moments align with B_0 in one of their two permitted states, parallel or anti-parallel to the direction of B_0. More protons align in the parallel orientation since

that is the lower energy state. The proportion of proton spins aligned with the imposed

field is slightly larger than those aligned against the field (anti-parallel), and this proportion

increases with the magnetic field strength. Thus the net magnetization is in the direction

of B_0. This net magnetization is represented as a vector M in Figure 2.8a. The magnetic

moments precess around the direction of B_0 at a resonant frequency of

$$\omega = \gamma B_0, \tag{2.1}$$

where ω is called the Larmor frequency and γ is the gyromagnetic ratio, which is specific

to the type of nucleus. γ is 42.6 MHz/T for the hydrogen nucleus (proton), so the resonant

frequency is 127.7 MHz at 3 Tesla.

A radio frequency (RF) pulse is applied to create a magnetic field B_1 perpendicular to

the original field B_0. The RF pulse is at the Larmor frequency so that the energy is absorbed

by the spinning protons, causing more of the magnetic moments to align in the higher en-

ergy, anti-parallel state. At some point, there will be an equal number of moments oriented

in the parallel and anti-parallel states, canceling each other out to yield a net magnetiza-

tion vector of zero in the direction of B_0. This direction is termed z, or the longitudinal

direction, by convention. While the RF pulse is turned on, it also aligns the precessing

magnetic moments in the plane normal to z. That is, the magnetic moments are rotating

around z at the same frequency with a coherent phase. This coherence creates a component

of the precessing magnetization vector M in the xy plane, called the transverse component,

illustrated in Figure 2.8b. For ease in illustration, the magnetization vector is often drawn

in terms of a frame of reference rotating around z at the Larmor frequency so the rotating vector M can be drawn as a static vector. The plane rotating around z is denoted as the $x'y'$ plane to distinguish it from the xy plane in the stationary, laboratory reference frame. Part (c) of Figure 2.8 shows the same vector M as in part (b) but in the rotating frame. When we combine the two effects of the RF pulse on the magnetization vector M, i.e. decreasing the z component and increasing the $x'y'$ component, it appears as if M is tilting from the z axis toward the $x'y'$ plane, as demonstrated in Figure 2.9. The angle of this tilt depends on the amount of time the B_1 field is applied, according to the equation:

$$\theta = \gamma B_1 t, \tag{2.2}$$

where t is time in seconds, θ (radians) is termed the flip angle, γ is the gyromagnetic ratio (rad/s/T), and B_1 is the applied field strength (Tesla). A common flip angle is $\theta = \pi/2$.
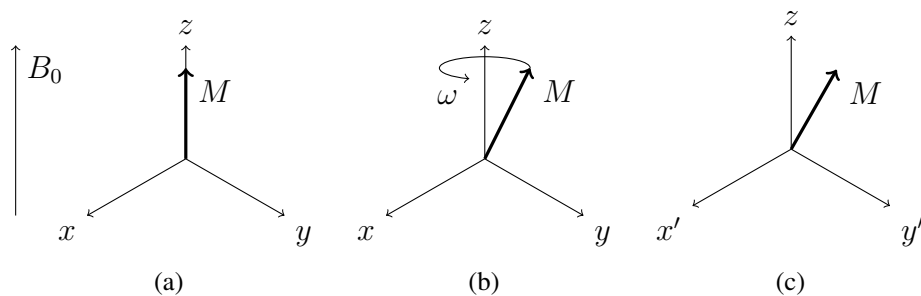
Figure 2.8: We illustrate the rotating frame of reference. In (a), the magnetization vector M begins aligned with B_0. When we add a transverse component to M, it rotates around z at the resonant frequency ω, shown in (b). The rotating frame of reference $x'y'z$ is rotating at the resonant frequency ω, so the precessing magnetization vector M appears stationary, shown in (c).

Once the RF pulse is turned off, the energy it put into the system is released and the
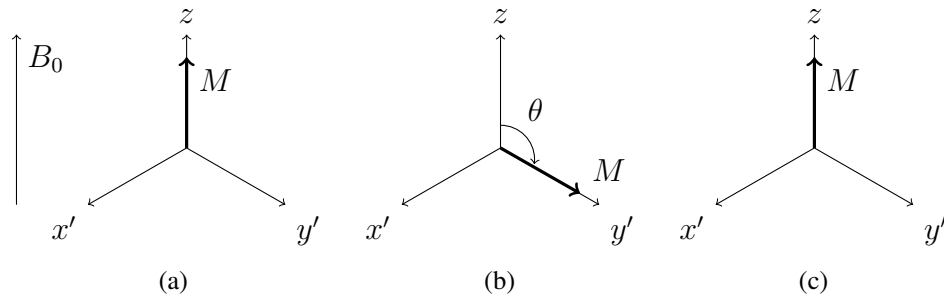
Figure 2.9: The magnetization vector M begins aligned with B_0, shown in (a). In (b), an RF pulse perpendicular to B_0 adds energy, tilting M toward the transverse plane by an angle θ, which depends on the duration of the pulse. In (c), the pulse is turned off and the magnetization relaxes back to z, releasing energy.

magnetization relaxes to its pre-RF state, shown in Figure 2.9, re-aligning along z since the B_0 field is still present. The loss of transverse magnetization is due to several factors. The proton loses energy to the molecular environment, bringing magnetic moments back to the lower energy parallel state, represented by a decrease in the z component of M. This is called spin-lattice or T_1 relaxation. Spin-spin or T_2 relaxation depends on the interaction between protons. They exchange energy, flipping energy states resulting in random phases. Magnetic interaction between protons causes slight differences in precession frequency, also leading to dephasing. Since the proton phases are no longer coherent, they cancel each other out, reducing the $x'y'$ component of M. Loss of phase coherence also occurs because of small local differences in the magnetic field from B_0 inhomogeneities and magnetic susceptibility differences, termed T_2^* relaxation. The rate at which the magnetization returns to its pre-RF state via T_1 and T_2 relaxation is described by the Bloch equations:

$$\frac{dM_{xy}(t)}{dt} = -\frac{M_{xy}(t)}{T_2} \tag{2.3}$$

$$\frac{dM_z(t)}{dt} = \frac{(M_0 - M_z(t))}{T_1}, \tag{2.4}$$

where M_{xy} is the transverse component of the magnetization vector M, M_z is the longitudinal component, M_0 is the steady-state magnitude along z when B_0 is present but B_1 is not, and T_1 and T_2 are the relaxation parameters (units of time) described above. The relaxation after applying a 90° pulse is therefore described by:

$$M_{xy}(t) = M_0 e^{(-t/T_2)} \tag{2.5}$$

$$M_z(t) = M_0 \left(1 - e^{(-t/T_1)}\right). \tag{2.6}$$

As the moments precess in the transverse plane, they induce a current in an RF coil tuned to the Larmor frequency, producing a measurable signal called a free induction decay. The intensity of this measured signal depends on the $M_{xy}(t)$ component of the magnetization.

Since the precession frequency depends on the strength of the magnetic field, by changing the strength along a spatial direction, we can distinguish the location of the protons contributing to the measured signal. This is called applying a magnetic field gradient. We apply gradients in three normal directions to create images in three dimensions. Localizing the source of the signal is done in three different ways. The first is to excite only protons that are within the desired region by applying a RF pulse with select frequency components.

Applying a magnetic field gradient of appropriate strength at the same time as the RF pulse allows us to choose what spatial region will resonate at the range of frequencies in the RF pulse. This is called slice or slab selection. The second method is to apply a "frequency encoding gradient" at the time of signal acquisition so that protons in different locations are precessing at different frequencies. Locations can be extracted from the measured signal using a Fourier Transform since location corresponds to frequency. The third method is to apply a "phase encoding gradient" between slice excitation and signal measurement. The protons precess at different frequencies until the gradient is turned off, at which point they precess at the same frequency but have phase offsets remaining from the gradient application according to their spatial location. Again, the spatial location can be decoded from the measured signal via a Fourier Transform.

There are many types of imaging sequences. In a spin-echo sequence, we apply a slice selection gradient in the z direction during a 90° RF pulse followed by a perpendicular phase encoding gradient in the y direction. Then we apply a 180° pulse at time t_1 along with the slice selection gradient to compensate for the effect of local differences in field strength that caused T_2^* relaxation. The M_{xy} component of the magnetization had decreased due to dephasing from local field variations following the 90° pulse. The compensating 180° pulse inverts the magnetization, causing the the faster protons to be behind the slower ones, so the different precession frequencies now work to realign the phases. We time the signal measurement, called the "echo", for time $2t_1$, the point when the phases are realigned for maximum M_{xy} magnitude. The time $2t_1$ to the center of the echo read-

out is called the echo time, TE. Note that the spin-echo technique removes T_2^* but not T_2 relaxation, which is energy exchanged between protons and therefore not recoverable by inverting the magnetization vector. During the echo measurement, we apply a frequency encoding gradient in the x direction. We want the signal to be highest at the center of the echo readout, so we also apply an earlier dephasing gradient in the x direction so that the frequency encoding gradient causes the protons to rephase. Once all this is done, we have acquired one line in the two-dimensional frequency space (called "k-space") of the selected slice. To acquire the other lines, we repeat the sequence with a slightly different phase encoding gradient strength for each line. The time between the repetitions is called the repetition time, TR. The pulse sequence diagram describing the spin-echo sequence is illustrated in Figure 2.10. In the diagram, the RF pulses are represented by sinusoids with sinc function envelopes since they convey bandpass-filtered, equal strength oscillations at the slice-selection frequencies. The 180° pulse is left on for twice as long as the 90° pulse, which scales the Fourier Transform representation by 1/2, so the 180° pulse is scaled by 2 to compensate.

Tissues vary in proton density, characteristic T_1, and characteristic T_2. We can control the relative intensities between tissues by selecting the TE and TR to emphasize parameters we expect to be different between the tissues of interest according to the signal intensity equation:

$$I(x, y, z) = p(x, y, z)(1 - e^{-TR/T_1(x,y,z)})(e^{-TE/T_2(x,y,z)}), \qquad (2.7)$$
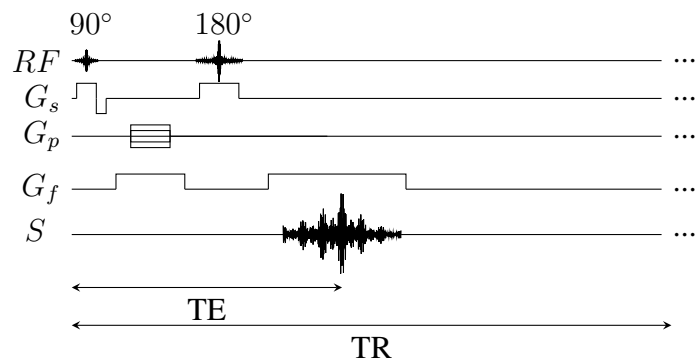
Figure 2.10: Pulse sequence diagram for the spin echo sequence. The RF line shows the timing of the 90° and 180° pulses. The G_s, G_p, and G_f lines show the slice selection, phase encoding, and frequency encoding gradients, respectively. The S line shows the measured signal. The signal is measured after time TE, and the whole sequence is repeated after time TR. The multiple heights of the phase encoding gradient represent the different gradient strengths at each repetition. The negative slice selection gradient and the early frequency encoding gradient are to compensate for dephasing that occurs during the corresponding primary gradients.

where I is the signal intensity, p is the proton density, T_1 is the spin-lattice relaxation and T_2 is the spin-spin relaxation of the tissue (or combination of tissues) at the pixel (x, y, z).

Total image time depends on TR and the size of the k-space imaging matrix, for there needs to be one repetition per matrix line. To speed up acquisition, the fast spin-echo (FSE) technique repeats the 180° refocusing pulse plus echo readout sequence several times before starting a new repetition. With each 180° pulse, the phase encoding gradient changes strength to encode a different line. See Figure 2.11 for the pulse sequence diagram. The number of times we repeat the refocusing pulse is called the echo train length (ETL). An ETL of n reduces imaging time by a factor of n. If n is too high, the later echos will have a very low signal intensity since the magnetization vector has been relaxing during the long time since the 90° pulse.
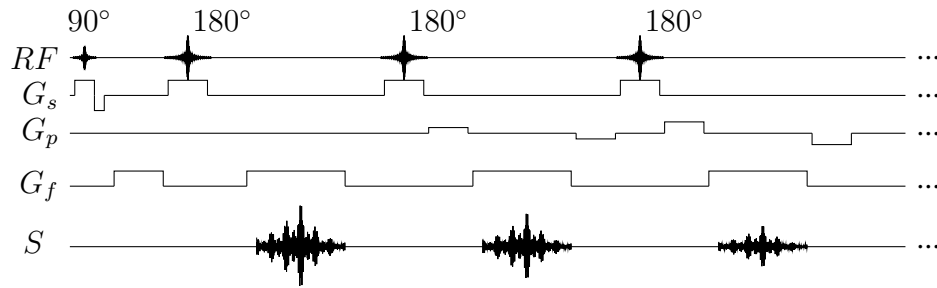
Figure 2.11: Pulse sequence diagram for the fast spin-echo sequence. The diagram is similar to the spin-echo sequence shown in Figure 2.10, but more than one phase encoded line in the imaging matrix is acquired before the 90° pulse is repeated at TR. Each phase encode is rephased following signal acquisition. In this figure, the echo train length is three. Note that the acquired signal S has lower intensity for the later echos due to magnetization relaxation.

2.3 Magnetic resonance imaging of the $\mathbf{T}_{1\rho}$ parameter

The spin-lock technique investigates relaxation that behaves similar to T_1 at very low magnetic field strength but allows imaging at high field strength for a high signal to noise ratio. Imaging begins in a strong static magnetic field B_0. A 90° RF pulse tips the magnetization vector to the transverse plane, as described earlier for the spin echo sequence. The RF coils then generate a very low strength magnetic field B_{SL} that rotates around the z axis at the resonant frequency of the B_0 field, $\omega_0 = \gamma B_0$. The magnetic field is rotating at the same speed as the rotating reference frame, so when viewed in the rotating frame, the magnetic field is static ($\omega'_M = 0$) and pointing in the same direction (y') as the tipped magnetization vector. The parallel alignment of B_{SL} and the tipped magnetization vector in the rotating frame is analogous to the alignment of B_0 and the initial magnetization vector along the z axis in the stationary reference frame. In the rotating frame, precession is governed by the resonant frequency $\omega_{SL} = \gamma B_{SL}$ instead of ω_0. In this situation, trans-

verse and longitudinal relaxation due to B_0 does not occur – the magnetization is said to be "locked" along B_{SL} in the rotating frame. Instead, the magnetization relaxes according to a parameter called $T_{1\rho}$, following the equation:

$$M(t) = M_0 e^{(-t/T_{1\rho})}. \tag{2.8}$$

$T_{1\rho}$ is approximately the value of T_1 that would be determined by imaging with a static magnetic field of B_{SL} and is often called the spin-lattice relaxation in the rotating frame.

$T_{1\rho}$ can be measured by applying a spin-locking preparation sequence followed by a standard imaging sequence. According to equation (2.8), the amount of transverse magnetization remaining after the spin-lock pulse is turned off depends on the amount of time (t_{SL}) during which the pulse was applied. After the spin-lock pulse is finished, a high amplitude -90° pulse brings the magnetization back to the z axis, and a crusher gradient dephases any remaining transverse magnetization. A standard high-field imaging sequence then begins with magnetization $M(t_{SL}) = M_{SL}$ instead of M_0. See Figure 2.12 for a pulse sequence diagram of the spin-lock preparation and Figure 2.13 for an illustration of the progression of the magnetization vector during the preparation. Because of the long duration spin-locking pulse, the examiner must be careful not to exceed specific absorption rate safety limitations set by the Food and Drug Administration.

The spin-lock preparation sequence is followed by a standard imaging sequence. For example, if it is followed by a spin-echo sequence, the signal intensity equation is:
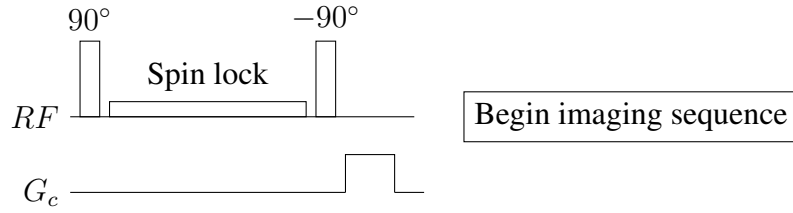
Figure 2.12: Spin-lock preparation sequence diagram. The RF line shows the radio frequency pulses for creating the high amplitude B_1 and low amplitude B_{SL} magnetic fields. The magnitude of the pulse dictates the strength of the magnetic field. We begin with a high amplitude 90° pulse to bring the magnetization into the transverse plane (shown in part (b) of Figure 2.13). Then the low amplitude spin-lock pulse locks the spins into low-field relaxation (shown in parts (c) and (d) of Figure 2.13). A negative 90° pulse brings the remaining unrelaxed portion of the magnetization back to the longitudinal axis (part (e) of Figure 2.13) to serve as the starting magnetization for a standard imaging sequence. The G_c line shows a crusher gradient that dephases any remaining transverse magnetization. Finally the main imaging sequence begins. No slice selection gradient is used in this preparation because the entire volume is excited.

$$I(x, y, z) = p(x, y, z)e^{-t_{SL}/T_{1\rho}(x,y,z)}\left(1 - e^{-TR/T_1(x,y,z)}\right)\left(e^{-TE/T_2(x,y,z)}\right), \tag{2.9}$$

where variables are defined as in equation (2.7).

To estimate the value of $T_{1\rho}$, we acquire several images with spin-lock preparation sequences using different spin-lock durations t_{SL}. Keeping the imaging portion of the sequence consistent means keeping TE and TR constant in equation (2.9). Thus the intensity I at point (x, y, z) in the $T_{1\rho}$-weighted image prepared with spin-lock duration t_{SL} is:

$$I(x, y, z) \propto e^{(-t_{SL}/T_{1\rho}(x,y,z))}$$

or

$$I(x, y, z) = ke^{(-t_{SL}/T_{1\rho}(x,y,z))}, \tag{2.10}$$
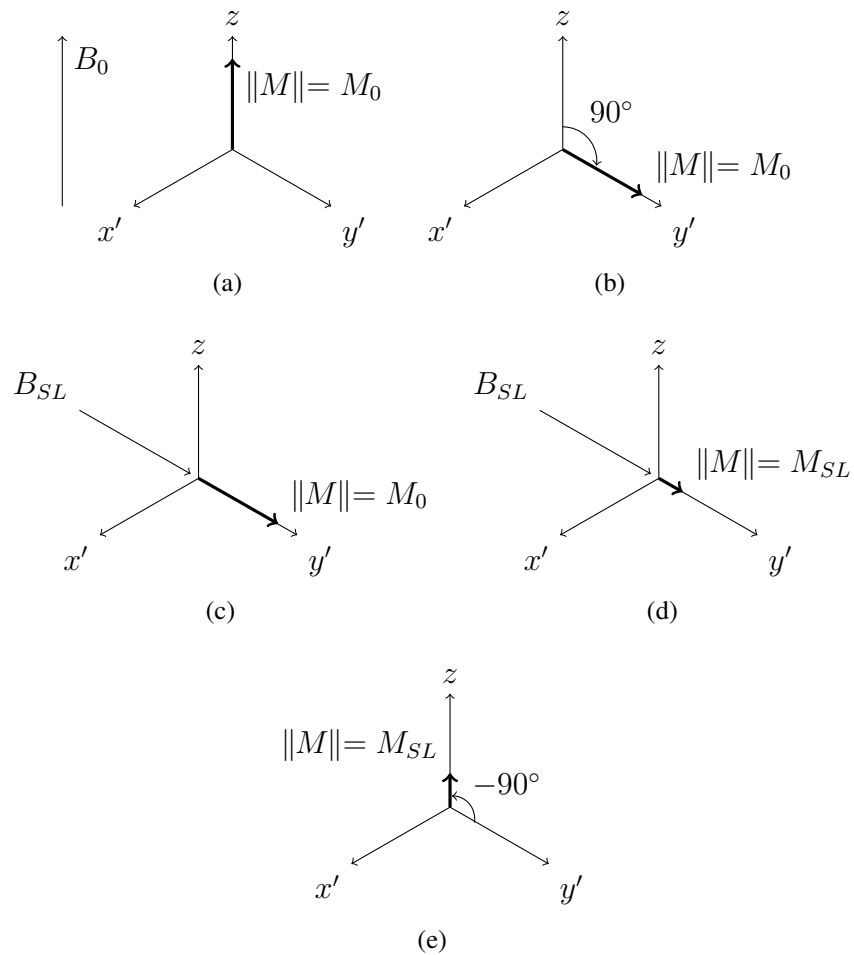
Figure 2.13: We show the progression of the magnetization vector during the spin-lock preparation sequence from Figure 2.12. (a): The magnetization begins with magnitude M_0, aligned with the static magnetic field B_0. (b): A 90° pulse brings M down to the y' axis (in the rotating frame). (c): Next the spin-locking pulse creates a magnetic field B_{SL} along the direction of M. (d): While B_{SL} is applied, M relaxes according to $T_{1\rho}$ but not T_1 or T_2, reducing its magnitude to M_{SL}. (e): After B_{SL} is removed, a -90° RF pulse brings M back to the z axis, and a crusher gradient removes any residual transverse magnetization. A standard high-field imaging sequence then begins starting with magnetization M_{SL}.

where k is an unused proportionality constant. After acquiring several such images, we find the value for $T_{1\rho}$ at each pixel that best fits equation (2.10).

Relaxation mechanisms T_1 relaxation depends on magnetic field strength B_0 and tissue properties that affect the motion and local environment of the protons being imaged. Many of the mechanisms involved in the relaxation can be described in terms of random molecular motion such as vibration, rotation, and diffusion. Correlation time τ_c is a measure of this motion; it is defined as the average time for a molecule to change its orientation by one radian. The effect of molecular motion on T_1 is described with a bi-Lorentzian shape:

$$\frac{1}{T_1} \propto \left[\frac{\tau_c}{1 + \omega_0^2 \tau_c^2} + \frac{4\tau_c}{1 + 4\omega_0^2 \tau_c^2} \right], \tag{2.11}$$

where ω_0 is the resonant frequency for the static magnetic field B_0, and T_1 is the exponential parameter describing the T_1 relaxation. According to this equation, T_1 relaxation is most efficient when there are molecules with correlation times around $1/\omega_0$ and $1/2\omega_0$ [35]. Figure 2.14 illustrates how relaxation rate depends on the relationship between the speed of molecular motion and the resonant frequency of the magnetic field. For frequencies far from resonance, energy exchange is less efficient so T_1 relaxation is longer.

The exponential parameter $T_{1\rho}$ is similar to what the T_1 parameter would be if it were imaged with a static magnetic field at the low spin-lock field strength B_{SL}. We would therefore expect $T_{1\rho}$ relaxation to be most efficient for correlation times related to the lower resonant frequency ω_{SL}. The relaxation is indeed dominated by τ_c around $1/2\omega_{SL}$, as

shown in equation (2.12) and Figure 2.14, though it is not completely independent of the higher resonant frequency ω_0.

$$\frac{1}{T_{1\rho}} \propto \left[\frac{(3/2)\tau_c}{1 + 4\omega_{SL}^2\tau_c^2} + \frac{(5/2)\tau_c}{1 + \omega_0^2\tau_c^2} + \frac{\tau_c}{1 + 4\omega_0^2\tau_c^2} \right], \qquad (2.12)$$

Imaging slow molecular motion using a high field strength is impractical because slow motion relaxation is very inefficient compared with the faster motions near the resonant frequency. Relaxation time differences between tissues will be dominated by the more efficient relaxation mechanisms. Low field imaging focuses better on low molecular tumbling rates since the slower motions are nearer the resonant frequency of the field, where the relaxation is more efficient (demonstrated by the horizontal offset of the curves in Figure 2.14). Molecular motion is slower when there are greater concentrations of macromolecules such as proteoglycans or when water is bound to proteins instead of moving freely.

The two curves in Figure 2.14 are also vertically offset, showing that relaxation rates vary with field strength, termed relaxation dispersion. At magnetic field strengths much larger than 20 mT, thermal movement of water molecules, especially fast-moving free water, is the dominant relaxation mechanism [36]. At lower field strengths, other causes of relaxation become more important, especially water-protein interactions, leading to significantly faster relaxation in tissues containing more proteins. Since T_1 (or $T_{1\rho}$) is shorter for tissues with more protein (e.g. pancreas and liver) at lower magnetic fields as opposed to high fields but does not change much for low-protein tissues (e.g. blood and fat) [36],

the range of relaxation times at lower field strengths is greater than at high field strengths, providing increased T_1 (or $T_{1\rho}$) contrast between tissues.
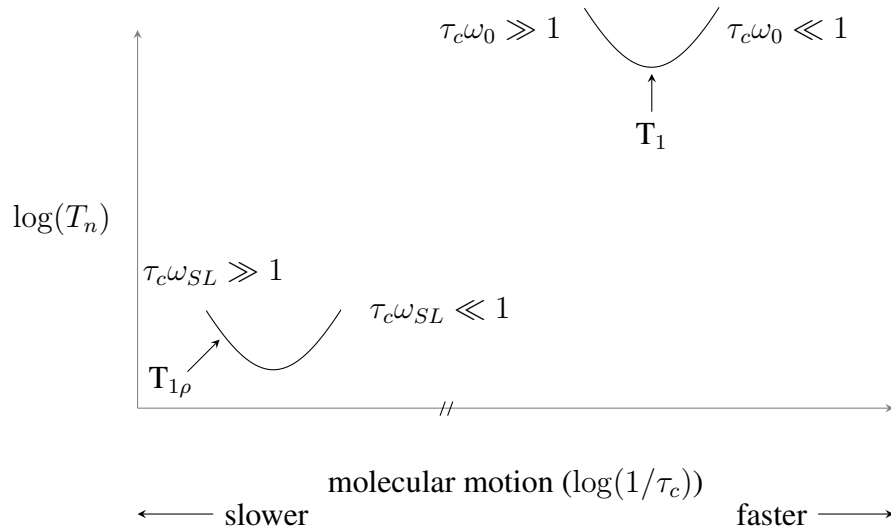
Figure 2.14: Approximate plots of equations (2.12) and (2.11) describe the $T_{1\rho}$ relaxation time (left) and T_1 relaxation time (right), respectively. $T_{1\rho}$ relaxation is more efficient in the presence of slower molecular motion than T_1 relaxation.

At magnetic fields lower than 20 mT, slow thermal movement of medium or large macromolecules, such as proteoglycans, becomes more important in relaxation, as does concentration and movement of water bound to proteins. Cross-relaxation between water protons or between protons in bound water and protein is more likely to occur with slower movement.

At B_{SL} fields lower than 0.2 mT, describing relaxation in terms of τ_c is insufficient [37]. Molecular weight, concentration, and bond structure strongly influence relaxation. Increased molecular weight and concentration decreases $T_{1\rho}$ without affecting τ_c. Breaking and reforming hydrogen and disulfide bonds can either increase or decrease $T_{1\rho}$.

Proton exchange between water and chemical groups NH and OH is an important re-

laxation mechanism at magnetic field strengths less than about 0.1 mT [38]. Through this mechanism, $T_{1\rho}$ depends on chemical structure and pH. More proteoglycans can mean more OH and NH groups available for proton exchange. Higher pH speeds relaxation, possibly by increasing exchange rates of these groups [38].

If B_{SL} is weaker than the local magnetic fields of interacting nuclei, the spin-lock condition is not met and relaxation occurs according to B_0 even while the B_{SL} field is present. In this case, $T_{1\rho}$ relaxation behaves like T_2 relaxation.

$T_{1\rho}$ for imaging degeneration   Disc degeneration involves changes in the extracellular matrix [33, 39]. Proteoglycan loss begins in the early stages of degeneration, so a marker of proteoglycan would be useful for early detection of deteriorating discs. Since low-field $T_{1\rho}$ relaxation is dominated by slower molecular motions associated with macromolecules and protein-water interactions, $T_{1\rho}$ is sensitive to proteoglycan loss.

There are many aspects of the molecular environment of a degenerating disc that affect the $T_{1\rho}$ relaxation. Since movement of large molecules is important in low-field relaxation, proteoglycan motion is a factor. The proteins in proteoglycans provide sites for protein-water interactions, including binding and cross-relaxation. Proteoglycans are negatively charged and affect tissue pH, influencing low-field relaxation. Hydration due to osmotic pressure from the glycosaminoglycan chains in proteoglycans affects relaxation. The change of molecular weight and structure from the breakdown of proteoglycans relates to $T_{1\rho}$. Fewer proteoglycan molecules means there is a different concentration of

molecules in the disc, which affects $T_{1\rho}$ relaxation. In addition, proteoglycans are responsible for regulating the flow of other molecules into and through the extracellular matrix, so there is a change in concentration of other molecules as well due to the loss of regulating proteoglycans.

In the intervertebral disc nucleus pulposus, lower $T_{1\rho}$ (faster relaxation) is associated with degeneration [7, 8, 5]. The reason for the change in relaxation is still under study. $T_{1\rho}$ is strongly correlated with concentration of proteoglycan glycosaminoglycan chains and moderately correlated with water content [7]. Proteoglycan loss leads to dehydration and looser regulation of macromolecular flow into the matrix. This may cause a higher ratio of macromolecules to water, speeding relaxation via reduced mobility (slower molecular tumbling) and more water-protein interactions. Loss of negatively-charged proteoglycans can also lead to higher pH, increasing relaxation rates.

The properties of the disc in compression are greatly influenced by osmotic pressure and hydration [34], which affect $T_{1\rho}$ values. Since proteoglycan loss directly affects osmotic pressure, we expect $T_{1\rho}$ behavior to change with compression and degeneration.

Imaging degeneration in knee articular cartilage with $T_{1\rho}$ preceded the study of $T_{1\rho}$ for intervertebral discs. Articular (hyaline) cartilage is similar in structure to intervertebral discs, containing collagen[1] and chondrocytes in an extracellular matrix dominated by aggrecan proteoglycans. As with intervertebral discs, early degeneration of articular cartilage includes loss of proteoglycans and water. However, $T_{1\rho}$ *increases* with articular cartilage

---

[1]The collagen in hyaline cartilage is of type II, while the fibrocartilage of the intervertebral discs contains both type I and type II collagen.

degeneration [40, 17, 4, 41, 42, 43, 16, 6]. Proteoglycan loss and dehydration associated with degeneration may slow relaxation because of fewer water-protein interactions and reduced molecular concentration. Fewer proteins within the matrix may lead to molecular tumbling rates significantly higher than the low-field resonant frequency, leading to slower relaxation. The opposing effect of $T_{1\rho}$ between disc and cartilage degeneration emphasizes that there are many competing underlying mechanisms in $T_{1\rho}$ relaxation, and understanding $T_{1\rho}$ relaxation in intervertebral discs is not straightforward and requires further study.

2.4   Description of the scanning environment

Some details of the specific scanning environment will help the reader understand this dissertation. A GE Signa MRI scanner is shown in Figure 2.15 with the spine receiver coil on the scanner bed. Knowing the location of the coil is useful in interpreting the images since regions closer to the coil have a higher signal intensity. The subject lies supine and the examiner places a landmark at the anatomy of interest using a laser. This landmark is the reference point for specifying where to aim the image acquisition. The scanner moves the subject into the cylindrical bore headfirst, leaving only their feet outside the bore.

Figure 2.15: MRI scanner

Chapter 3

Data acquisition

We wish to measure $T_{1\rho}$ values in intervertebral discs. As explained in 2.3, we calculate $T_{1\rho}$ maps by finding the exponential parameter for each pixel in the intervertebral discs that best fits the following equation in the least squares sense:

$$I \propto e^{-T_{SL}/T_{1\rho}}, \tag{3.1}$$

where I is the intensity of a pixel in the $T_{1\rho}$-weighted image acquired with spin-lock duration T_{SL}, and $T_{1\rho}$ is the parameter to estimate.

To obtain the values on the left hand side of equation (3.1), one acquires several consecutive $T_{1\rho}$-weighted images with different locking pulse durations in the spin-lock preparation sequence. $T_{1\rho}$-weighted images are acquired by adding a spin-lock preparation to a single-slice Fast Spin Echo (FSE) sequence. The $T_{1\rho}$-FSE sequence used for the data presented in this work was developed and tested by Xiaojuan Li [8]. Table 3.1 lists the

sequence parameters.

FOV (field of view)	20 cm
Imaging matrix	192x128
Slice thickness	8 mm (one central sagittal slice)
In-plane pixel size	0.78 x 0.78 mm
TR (repetition time)	2000 ms
ETL (echo train length)	16
BW (bandwidth)	244.1Hz
NEX (number of excitations)	2
Flip angle	90°
Spin lock frequency	300 Hz (0.007 mT)
Spin lock durations	0/20/50/90 ms
Scanning time	2:11 minutes

Table 3.1: $T_{1\rho}$ Imaging parameters

Since the images cover only a single sagittal slice, we must ensure that we are imaging the same slice even if the two exams we are comparing were acquired during different exam sessions. Before the $T_{1\rho}$-FSE imaging sequence in each session, we acquire an axial T_2-weighted image using a FSE sequence. These T_2-weighted images are automatically aligned using a global rigid transformation during the second imaging session so that we may select the correct right/left offset on the scanner for the $T_{1\rho}$-weighted acquisitions. See section 4.2.1 for registration details. We chose axial for this prospective registration step because the higher in-plane resolution includes the right/left anatomical direction, yielding precise results in the direction we need for positioning a sagittal slice. We selected sagittal for the main imaging sequence to reduce the effects of partial voluming between disc and bone.

We are looking for pressure-related changes in the intervertebral discs, so we created

situations in which we expect differences in disc compression. Measurable differences in MR images have been recorded with subjects positioned with spinal flexion/extension and under axial loading [44]. We tried several different methods, including rounding of the back by flexing the muscles, application of a compression device within the machine, supported position changes (knees bent, propped up with pillows compared with legs extended flat on the scanner bed), and decompression while lying supine in the scanner. For future exams, we suggest taking scans multiple times during the day since discs are expected to be more hydrated and less compressed in the early morning than in the late evening. For convenience in developing and testing the registration algorithm, our data sets were all acquired in single imaging sessions of less than 2 hours. To simulate different imaging sessions, some volunteers exited and re-entered the scanner. Thus the algorithm is prepared to deal with studies spanning multiple imaging sessions.

Chapter 4

Registration

Registration allows us to compare image points that are in anatomically the same position. Prospective registration guides image acquisition, segmentation identifies rigid bodies and regions of interest, and disc registration aligns the intervertebral discs between two images in different positions.

4.1 Background

4.1.1 Registration

Registration of medical images is the process of transforming a "moving" image such that its pixels align anatomically with the pixels in a "fixed" image. See [45, 46, 47, 48] for recent surveys of medical image registration techniques. An image may be registered with another image of a different patient (inter-patient), with an anatomical atlas, or with

another image of the same patient (intra-patient) acquired with different imaging modalities or in different situations such as before and after treatment. In this project, we deal with intra-patient registration.

Transformation type

Registration transformations can either preserve shapes or not. Rigid transformations preserve shapes, whereas affine and curved transformations do not. Rigid transforms allow only rotation and translation. Affine transforms also permit scale and shear. Restrictions on curved transforms depends on the type of transform. For example, some are required to conform to points of known deformations, called control points or landmarks; others are controlled by local smoothness constraints. Usually a single type of transformation is applied over the entire image, called a global transformation, e.g. [49]. However if one knows the locations of image features that are expected to transform differently, different types of transformations can be applied according to tissue type [50, 51, 52]. For example, a rigid transform can align bony regions while a curved transform deforms the surrounding soft tissue.

Section 4.2.3 explains how our registration technique incorporates curved registration using a thin plate spline (TPS) deformation. A two-dimensional thin plate spline is a radial basis function transform using the kernel $\phi(r) = r^2 \ln r$. The radial basis function transform determines the deformation $\mathbf{d}$ at a point $\mathbf{x}$ based on the positions of k landmark points $\mathbf{p_i}$ according to the following equation:

$$\mathbf{d}(\mathbf{x}) = A\mathbf{x} + \sum_{i=1}^{k} \mathbf{w_i}\phi\left(|\mathbf{x} - \mathbf{p_i}|\right), \qquad (4.1)$$

where $|.|$ denotes the Euclidean norm, bold typeface indicates a two-dimensional point, A is an affine transformation matrix, and $\mathbf{w_i}$ is a set of weights determined by solving the above equation after substituting the known control point deformation values for $\mathbf{d}$.

Radial basis functions have been used with success in curved registration of medical images [53, 54, 55, 56]. Basis functions $\phi(r)$ that have been explored for medical image registration include [54]:

- $r^2 \ln r$ (two-dimensional thin plate spline; the basis function we use in this work)
- r (three-dimensional thin plate spline)
- r^3 (volume spline)
- $[\alpha r^2 I - 3xx^T]r$, $\alpha = 12(1 - \nu) - 1$, ν is Poisson's ratio (elastic body spline),

where x is the point in question, r is the Euclidean distance between a landmark and x, and I is the identity matrix. Each function has a physical basis for the resulting deformation. However, it is not yet clear which physical basis is appropriate for different tissues and number of dimensions.

Registration metric

The outcome of the registration depends on the metric chosen to evaluate alignment. Metrics are classified as landmark-based, segmentation-based, or intensity-based. Land-mark and segmentation-based methods bring corresponding points or surfaces identified in

both the fixed and the moving images into alignment. These methods depend on the quality of the landmark placement or segmentation. Intensity-based methods directly compare the image pixel values. Examples include computing the pixel-wise squared difference, cross-correlation, or mutual information.

Mutual information (MI) is defined in terms of the entropies of two images, A and B, as follows:

$$MI(A, B) = H(A) + H(B) - H(A, B), \qquad (4.2)$$

where $H(A)$ and $H(B)$ represent the entropies of A and B, and $H(A, B)$ is the joint entropy of the two images. Shannon's entropy and joint entropy are defined as:

$$H(A) = -\sum_a p_A(a) \log p_A(a) \qquad (4.3)$$

$$H(A, B) = -\sum_{a,b} p_{AB}(a, b) \log p_{AB}(a, b), \qquad (4.4)$$

where $p_A(a)$ is the probability distribution of intensity values of image A, and $p_{AB}(a, b)$ is the joint probability distribution of the intensity values of images A and B. The probability distribution $p_A(a)$ can be estimated with a histogram of pixel intensities in image A. Entropy $H(A)$ provides a measure of the amount of information in an image. If A is mostly uniform, hovering around a value v, $p_A(a)$ will have a large spike around v and will be low elsewhere. Thus $H(A)$ will be low; the image contains little information. We know that if we pick a pixel at random, it will probably be around v, and we don't gain

much information from checking the true value of the pixel because we already knew what to expect. On the other hand, if A contains pixels of many different intensities arranged in a visual representation of a body part, as in an MRI, $p_A(a)$ will have a more uniform distribution of probabilities, leading to a higher entropy value $H(A)$. The image contains a lot of information; it is not clear from the start what the intensity of a randomly drawn pixel will be – when you draw the pixel you learn something about the image. Similarly, joint entropy is a measure of the amount of information contained in two images together. The joint probability distribution $p_{AB}(a, b)$ can be estimated by a normalized two-dimensional histogram of image intensities. The images are placed one on top of the other so that a pixel location now has two values – one from A and one from B. The value of the histogram at point (a, b) is the number of pixels in the intensity range assigned to a in image A and b in image B. If both images contain similar intensity patterns, $p_{AB}(a, b)$ will have clusters of higher intensity leading to a low joint entropy $H(A, B)$. In other words, the information in A and B together is low because once you know the value of A at a randomly drawn pixel, you are likely to be able to guess B at that pixel. B isn't giving you much additional information. If the intensity patterns in A and B are different or if the patterns are similar but misaligned, the joint histogram will be more spread out, meaning $H(A, B)$ is higher and more information is contained in the image pair. In other words, even knowing the value in A of a randomly drawn pixel, you are unlikely to guess the pixel value in B.

Mutual information combines these entropies to give a metric of registration. According to equation (4.2), a high MI value comes from low $H(A, B)$ and high $H(A)$ and $H(B)$.

The low $H(A, B)$ means that the patterns in A and B are similar and aligned, i.e. knowing A provides information about B. Since $H(A)$ and $H(B)$ are calculated only where A and B overlap, a high MI value encourages alignments where the information-filled regions of A and B overlap. That is, $-H(A, B)$ alone is high when A and B only overlap in homogeneous background regions, which is not a desirable registration. Including $H(A)$ and $H(B)$ penalizes this trivial alignment.

Since mutual information depends on histograms rather than directly comparing pixel values, it is useful in aligning patterns that do not necessarily have the same absolute intensity values, such as images acquired using different modalities [57, 58, 48].

Smoothing the joint histogram and moving image makes the mutual information metric differentiable, aiding optimization [59]. The transformed moving image is smoothed using a B-spline interpolator, and the joint histogram is extrapolated by convolving with a smoothing spline kernel, called a Parzen window.

Optimization method

The best registration is found by searching for the transform parameters that yield the maximum or minimum value of the selected metric. Many methods are used successfully with the mutual information metric, including Powell's method, the simplex method, gradient ascent, and simulated annealing [48]. We found the Insight ToolKit [60] implementation of the regular step gradient descent method to yield robust and quick results with the Mattes mutual information metric on our $T_{1\rho}$-weighted MR images.

The regular step gradient descent (or ascent) algorithm moves in the parameter space along the curve representing the metric value. Optimizing all parameters at once, it follows the direction opposite (or along) the gradient of the metric curve. Starting with a maximum step length, it steps along the negative (or positive) gradient by that length to cover most of the distance to the nearby minimum (or maximum). It continues until it hits a point where the gradient direction changes sharply. The step length is then reduced by a set relaxation factor so that the minimum (or maximum) is not overlooked. The search continues in this manner until the gradient magnitude is below a set tolerance, the step size becomes smaller than a set tolerance, or the pre-determined maximum number of iterations is reached. Parameters can be given scaling factors so that steps along that parameter are larger or smaller according to the user's approximation of how much that parameter is expected to change with respect to the other parameters. For example, a rotation parameter usually changes on a different scale than translation parameters. The gradient descent (or ascent) algorithm converges to local minima (or maxima), so the starting parameters are influential in the outcome. Many other optimization methods also have this drawback of tending toward local optima.

A multi-resolution approach to optimization can help steer the solution toward global optima rather than local optima, thereby reducing the influence of the starting parameters [61, 62]. A multi-resolution approach often also speeds up the optimization. This approach begins by optimizing the parameters using smoothed and downsampled versions of the images. The initial coarse level guides the optimization to a global rather than local optimum.

The optimization continues using images that have more detail than the initial coarse step but are still smoothed and downsampled with respect to the original images. Optimizing with higher resolution images yields a more precise result. It is also more susceptible to local optima, but the coarser resolution versions of the image have already placed the higher resolution versions at an initial position closer to a global optimum. This procedure continues until the optimization finishes with a fine-tuning on the original high-resolution images.

Inverse mapping

After finding the optimum registration transformation, we apply it such that the moving image is in the same position as the fixed image. We transform the moving image via an inverse mapping. That is, the transform is defined from the fixed to the moving frames. A blank image is created in the fixed frame, called the morphed image, and the value for each pixel in this new morphed image is found by following the fixed-to-moving transform into the moving frame. The pixel value at the corresponding point in the moving image is placed in the morphed image, interpolating if necessary since the point is likely not an integer pixel position.

An inverse mapping is preferred because with a direct mapping from the moving to fixed images, there may be pixels in the resulting morphed image to which no intensity is assigned or multiple intensities are assigned. Inverse mapping is a straightforward way of approximating the transform along the regularly spaced pixel positions of a digital image.

4.1.2 Segmentation

Segmentation is the process of classifying pixels in an image. In medical images, pixels are typically classified according to tissue type, anatomic organ or object, or functional or chemical properties. As explained in section 4.2.2, our procedure requires anatomical structure segmentation of each intervertebral disc and each vertebra. For a review of segmentation methods for medical images, see [63]. Many different segmentation methods have been developed, though no one method is general enough for all applications. Successful segmentation schemes often combine and tailor methods for a specific application, which is the approach we take. In the following paragraphs, we explain the general segmentation techniques that we combine and modify to meet our needs.

Region growing Region growing classifies pixels according to their intensity and spatial location. The procedure begins with a seed pixel that is known to be within the region of interest. Neighboring pixels are tested to see if they should be included in the region. If so the pixel is marked as included and neighboring pixels of the newly classified pixel are tested. The procedure continues, growing the included region until no more pixels neighboring the enlarged region satisfy the inclusion criteria. Region growing algorithms differ in their inclusion test and in their definitions of neighboring pixels. The inclusion test could be a simple threshold, where pixels with intensities lower than a provided threshold (or higher or within an intensity range) are included. Other common tests are inclusion only if all neighbors satisfy a threshold condition or inclusion only if the pixel intensity

falls within a given number of standard deviations of the mean of the pixels that are already included.

Voting operations Voting operations act on a pixel according to rules that depend on the number of neighboring pixels satisfying a criterion. For example, the mathematical morphological operation of binary dilation can be seen as a voting operation. In binary dilation, a pixel changes classification from background to foreground if the number of neighboring foreground pixels in its neighborhood is greater than or equal to one. The result is a slight increase in the size of the foreground region. The definition of "neighboring pixels" is given via a structuring element, or a kernel, which is a small collection of binary pixels (or voxels for a three-dimensional image) that when overlapped on the image, centered around the pixel in question includes a 1 for pixels in the neighborhood and a 0 for pixels outside the neighborhood. See Figure 4.1 for some examples.

0	0	1	1	1	0	0
0	1	1	1	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0

1	1	1
1	1	1
1	1	1

0	1	0
1	1	1
0	1	0

Figure 4.1: Examples of two-dimensional structuring elements for voting operations.

Another voting operation used in this project is the binary hole filling operation. Pixels are changed from background to foreground if a majority plus n of the neighborhood pixels are foreground. Smaller n will fill larger holes.

Fast marching   The fast marching method is a type of level set method. Level set methods describe the evolution of an interface [64], and if we view the interface as the outline of a segmented region, the level set method becomes a technique for image segmentation [60]. Instead of following the evolution of an interface $\Gamma(t)$ in the N-dimensional space in which it resides, we follow the interface (hypersurface) as if it were the zero level-set at time t of an evolving $N + 1$-dimensional propagating function $\phi(x, t)$. We begin by stating the initial condition

$$\phi(x, t = 0) = \pm d, \tag{4.5}$$

where d is the distance from the N-dimensional point x to the initial interface $\Gamma(t = 0)$, and the sign of d indicates if x is outside (positive) or inside (negative) the hypersurface $\Gamma(t = 0)$. Condition (4.5) means the zero level set of ϕ is equal to the initial interface, i.e. $\Gamma(t = 0) = (x|\phi(x, t = 0) = 0)$. We want to maintain this relationship, i.e.

$$\Gamma(t) = (x|\phi(x, t) = 0), \tag{4.6}$$

so we note that

$$\phi(p(t), t) = 0 \tag{4.7}$$

for all time t, where $p(t)$ is a point on the evolving hypersurface $\Gamma(t)$. We denote the speed of propagation of $p(t)$ as $F(p(t))$, that is $\frac{\partial p}{\partial t} \cdot n = F(p(t))$, where n is the normal to the surface at $p(t)$. Taking the derivative of equation (4.7), we get

$$\frac{\partial \phi}{\partial t} = -F|\nabla \phi| \tag{4.8}$$

as the partial derivative governing the evolution of the function ϕ. To obtain the interface Γ at time t, simply find the zero level set of $\phi(t)$. Advantages of this approach include ease in managing separations, merges, and sharp corners of Γ as it evolves, ease in numerical approximation, ease in determining Γ curvature directly from ϕ, and ease in following interfaces in dimensions higher than two.

The general level set function can evolve in complex ways, including terms that control advection, propagation, and local curvature. When the level set function is in a simple form where the interface only advances in one direction (i.e. shrinks or expands) and the speed of the evolution depends only on position within the image, the propagation can be solved with a fast algorithm called the fast marching method [64]. Without loss of generality, we explain here the case of monotonic expansion. In the fast marching algorithm, we calculate the time $T(x, y)$ at which the interface crosses the point (x, y) by sweeping outward, moving point-by-point from points with lower to higher T according to the speed $F(x, y)$. Once $T(x, y)$ is calculated, the interface at time t is found by thresholding $T(x, y)$ at value t.

To use the fast marching method to segment an image, we give the algorithm an initial guess of the segmentation outline. This guess must be inside the final outline everywhere since the assumption is that the interface always moves outward. Usually the initial guess will be one or more points located inside the anatomical structure of interest. The easy-

merging property of the level set method makes it simple to combine several initial points. We also give the algorithm an image with pixel intensities representing the propagation speed of the outline at each point. This speed image usually depends on the edges in the image to be segmented, with high speed around 1.0 within the anatomical object and decaying quickly to low speed around 0.0 to stop the propagation near the edges of the object to be segmented. A possible procedure to create the speed image is to smooth the image to be segmented with an edge-preserving smoothing function such as anisotropic diffusion [65], followed by finding the gradient magnitude at each pixel and applying a sigmoid function:

$$\frac{1}{(1 + e^{-\frac{(x-\beta)}{\alpha}})}, \tag{4.9}$$

where x is the gradient magnitude at the pixel, and α and β are user-defined constants to control the sharpness of the decay. With a negative α, the sigmoid function maps high gradients to low speeds and smoothes the edges a bit so the speed decays near the edges instead of abruptly changing. The algorithm then marches along calculating $T(x, y)$, and we find the segmentation outline by thresholding the output $T(x, y)$ at a user-defined stopping time t_s.

4.2 Methods

4.2.1 Prospective registration

As mentioned in section 3, registration is first used during the image acquisition phase to locate the same sagittal slice for both images if the person must exit the scanner between scans.

We use a global rigid transformation, optimizing the mutual information metric using gradient ascent. After finding the relative transformation between the axial images, we find the desired sagittal slice position as illustrated in Figure 4.2. In the figure, O_{room} represents the origin of the coordinate system of the room, which remains constant from scan to scan. L_1 and L_2 are the operator-defined landmarks for the first and second scans, respectively. C_1 and C_2 are the centers of the first and second images, respectively. Note that they are recorded with respect to the corresponding landmark point L_1 or L_2. The sagittal slice in the first image passes through point s_1 (measured as SL_1 with respect to the first landmark). We want the sagittal slice in the second image to pass through s_2, which is the same anatomical position as s_1. s_2 is measured as SL_2 with respect to the second landmark. The inset at the lower left of Figure 4.2 shows the result d of the prospective registration, which gives the transformation needed to move from a pixel p_2 in the second image to the pixel p_1 representing the same anatomical point in the first image. The green dot p_1 in the lower image represents the same pixel p_1 in the upper image to show how the anatomical position differs. The scanner tells us the values of quantities SL_1, C_1, and C_2,

and the prospective registration gives d. We wish to know SL_2. Figure 4.2 illustrates that we may calculate SL_2 as:

$$SL_2 = SL_1 - d + C_2 - C_1.$$

We then input the right/left component of SL_2 into the scanner to aim the sagittal acquisition.

4.2.2 Segmentation

Segmentation serves us in two ways. Disc segmentation identifies the regions that we will be analyzing, and vertebral segmentation identifies the regions that move rigidly in the images, which guides the registration of the discs, as described in section 4.2.3.

Our segmentation process requires very little human interaction, and the required interaction is easily learned and not sensitive to error.

Seed placement Segmentation begins with the user indicating the anterior and posterior ends of each disc of interest, two points per disc, as shown in red in Figure 4.3.

The software then automatically selects two seed points within each disc and a centrally-located seed point within each vertebra. See Figure 4.4 for an example of disc seed positioning. The disc seeds (large red points in Figure 4.4) are selected by splitting into thirds the (green) line connecting the anterior/posterior (A/P) points (small red points).

Seeds for vertebrae with discs both above and below are calculated by taking the mid-

In this graphic,

$$s_1 = s_2 - C_2 - L_2 + L_1 + C_1 + \mathbf{d}$$

$$SL_2 = s_2 - L_2$$

Therefore

$$SL_2 = SL_1 - \mathbf{d} + C_2 - C_1$$
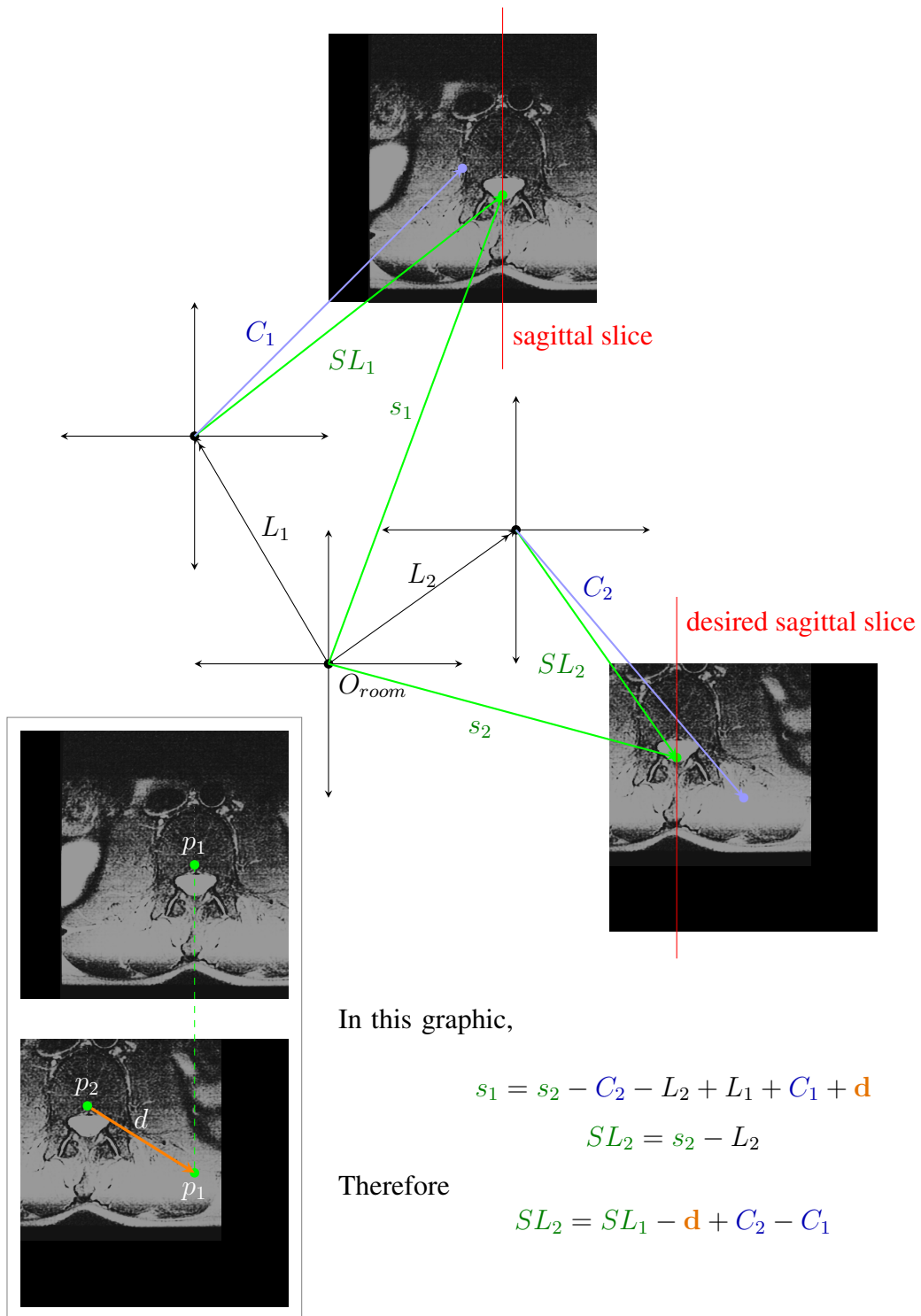
Figure 4.2: Prospective registration alignment. See page 56 of the text for details and variable definitions.
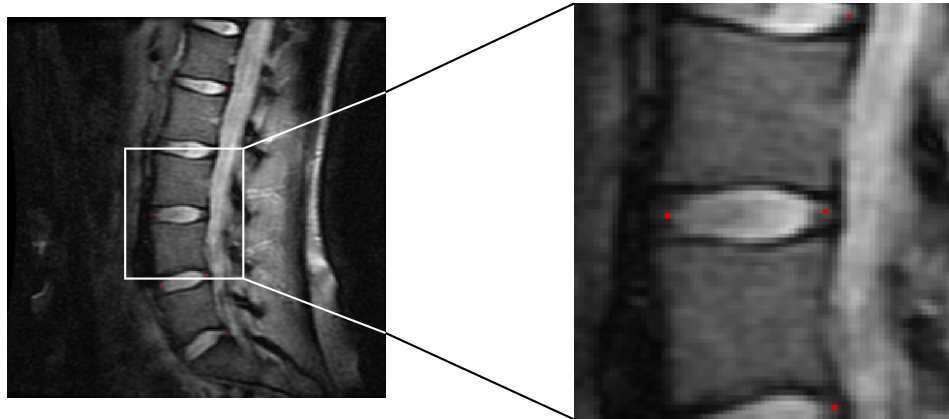
Figure 4.3: The user indicates the anterior and posterior points of the intervertebral discs. These points are shown in red.
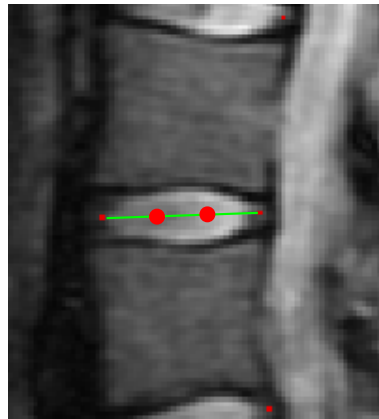
Figure 4.4: The software positions seeds within the intervertebral disc to use in segmentation. The locations of the disc seeds (large red dots) depend on the user-defined anterior/posterior points (small red dots).

point of the line connecting the anterior seed in the superior disc with the posterior seed of the inferior disc, as demonstrated in the left picture in Figure 4.5. The central seed located in a vertebra with a disc of interest either above or below (not both) is located along the normal (yellow) to the line (green) connecting the disc seeds. The normal is placed in the center of the green line, and the seed is located a distance of d away from the green line,

where d is the length of the green line, as shown in the right illustration in Figure 4.5. Note that the seeds in the vertebrae at the top and bottom of the image are located closer to the disc rather than central to the vertebra. This is to ensure that the inferior seed does not overshoot the vertebra due to the sharp curve of the sacrum and that the superior seed does not rise outside the image.
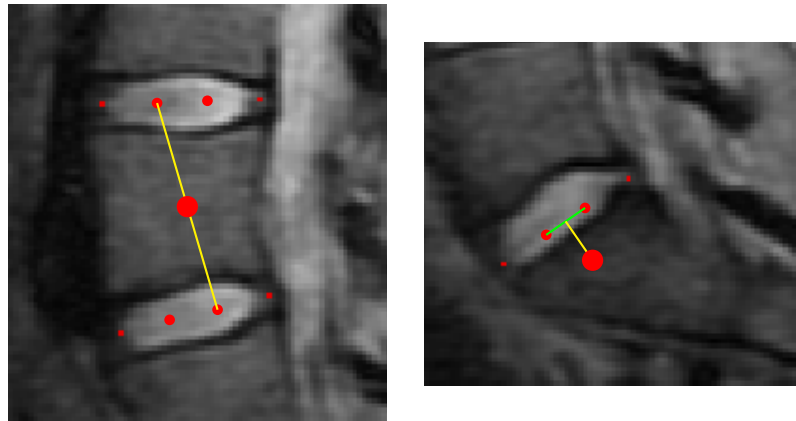
Figure 4.5: The software locates seeds (largest red points) within the vertebrae to use in segmentation. Vertebral seed locations depend on disc seed locations (medium-sized red points). The smallest red dots are the user-defined anterior/posterior points.

Disc segmentation Once the seeds are found, the software segments each disc. See Figure 4.6 for an overview of the procedure. Since the disc/vertebra contrast is high in the $T_{1\rho}$-weighted images, a threshold region growing algorithm is a good candidate for segmentation. Thus the major segmentation step is a region growing algorithm on each disc that grows from the disc seeds found above to include adjacent points that are above a threshold. Adjacency is defined as the 3x3 matrix around a pixel. The threshold is auto-

```
┌─────────────────┐
│  Pre-processing  │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│  Region growing  │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│  Fast marching   │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Voting operations│
│    (dilation,    │
│   hole-filling)  │
└─────────────────┘
```
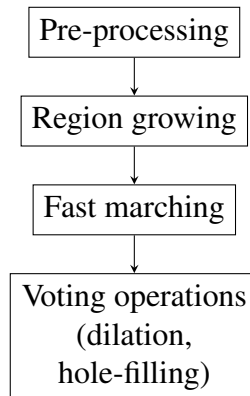
Figure 4.6: Overview of disc segmentation procedure.

matically determined by finding the minimum threshold for which the vertebra seeds are not included in the resulting segmented region. This threshold is then scaled by 1.1. The scaling factor is not strictly required, as many images do not need it, but it reduces bleeding in some images and does not hurt most images since the disc/vertebra contrast is high enough. Before the region growing step is a pre-processing step, the first part in which we eliminate the source of most of the region growing bleeding error, i.e. the spinal canal. Looking at the example images shown so far, we see that the discs have a noticeably higher intensity than their surroundings in the anterior, superior, and inferior directions. There is also a dark bar on the posterior edge separating the disc from the high-intensity spinal canal, but this bar is thin and is sometimes breached by the region growing algorithm. To eliminate this problem, we use the previously-defined points indicating the anterior and posterior boundaries of the discs. We blacken values anterior and posterior to the disc using lines orthogonal to the line connecting the A/P points, as shown in Figure 4.7.

The second and final part of pre-processing is to smooth the result, followed by the
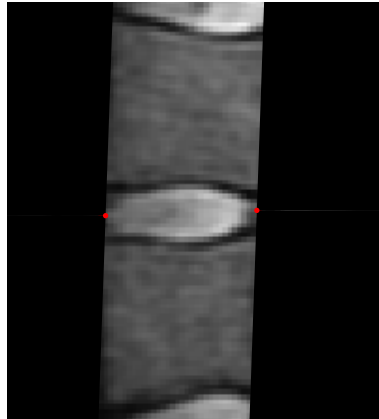
Figure 4.7: Blackening the regions anterior and posterior to the disc assists segmentation.

threshold region growing algorithm explained above. In some cases the region growing still leaks outside the discs, so we next use a fast marching level set method. Instead of basing the speed image on the original image as is typically done, we base it on the binary output of the region growing algorithm. The fast marching method will then march outward from the disc seeds. The resulting segmentation outline is determined by a pre-set stopping time. The stopping time allows the propagating front to march to the nearer edges, where it will stop. The farther edges are assumed to be errors due to region-growing leaking and are not included in the fast marching result. We make the binary output of the region growing algorithm into a speed image by finding the gradient magnitude and applying a sigmoid function (see equation (4.9)) to generate high speed within the disc and zero outside, with a short period of decay as we approach the edges. (Sigmoid parameters: maximum speed = 1.0, minimum speed = 0.0, $\alpha = -0.5$, $\beta = 12$.)

After the fast marching method, we dilate the results using a 3x3 matrix of ones as a

structuring element to recover the width of the outline in the fast-marching speed image. Finally, we apply a hole filling algorithm (voting threshold of majority+2) to deal with noisy regions of low intensity within the disc. See an example result in Figure 4.8.
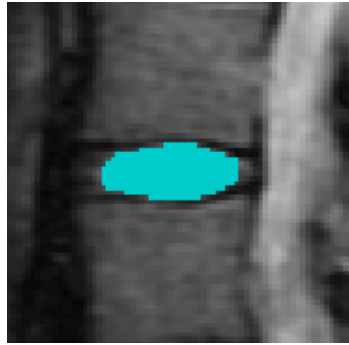
Figure 4.8: Example result of automatic disc segmentation.

Since there can be unexpected anatomical deviations from the anticipated shapes and intensities in medical images, we also include a manual touch-up feature where the user paints or erases pixels with the computer mouse as a paintbrush.

Vertebral (rigid body) segmentation The purpose of segmenting the vertebrae is to identify which portions of the image are expected to move according to a rigid transformation, as opposed to the elastic transformation of soft tissue. Using rigid transforms to guide the registration instead of using a fully deformable registration provides more accurate and faster results. Since we are concerned with analyzing the $T_{1\rho}$ values within the discs, the vertebral segmentation must be accurate near the discs. However, as will be explained in section 8, the anterior/posterior edges of the vertebrae can be (and are even desired to be) overestimated.

A fast marching level set process accomplishes the vertebral segmentation. As with the disc segmentation, we use a non-standard approach to creating the speed image. Since our goal is to find an interface that approaches but does not the discs, we base our speed image on the disc segmentation. We assign value one to pixels that were previously classified as discs and value zero to other pixels. Edge detection and a sigmoid filter transform this binary image into a speed image ranging from 0.0 to 1.0, as explained for the disc segmentation. The resulting zero-valued edges of the segmented discs prevent the vertebral segmentation from marching into the disc regions. For each vertebra, the fast marching procedure is initialized with the previously-determined central vertebra seed. The stopping value depends on the distance between the vertebra seed and the seeds within the bordering discs so that the fast marching interface extends all the way to the discs (though the zero-speed edges prevent it from entering the discs). The superior and inferior vertebrae use twice this distance since these vertebra seeds are situated closer to the discs, as described earlier. See Figure 4.9 for an example of the resulting segmentation.

4.2.3 Disc registration

After acquisition, we register the intervertebral discs so that we can statistically analyze $T_{1\rho}$ values from the same disc location in both images. For each pixel p_1 within a disc in the "fixed image" (position 1), we locate the pixel in the "moving image" (position 2) that corresponds to the same anatomical position as p_1. Registration uses $T_{1\rho}$-weighted images with a short spin-lock duration since the intensity contrast between the discs and vertebrae
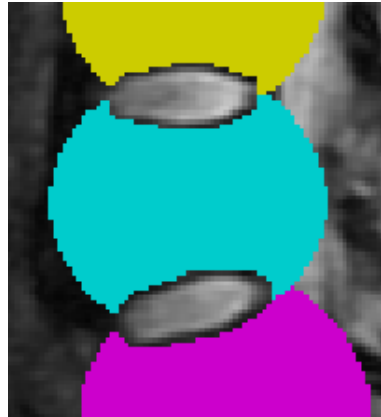
Figure 4.9: Example result of automatic vertebra segmentation. Anterior/posterior ballooning is justified in section 8.

is greater and the signal to noise ratio is higher for shorter spin-lock durations. See Figure 4.10 for an outline of the registration procedure.
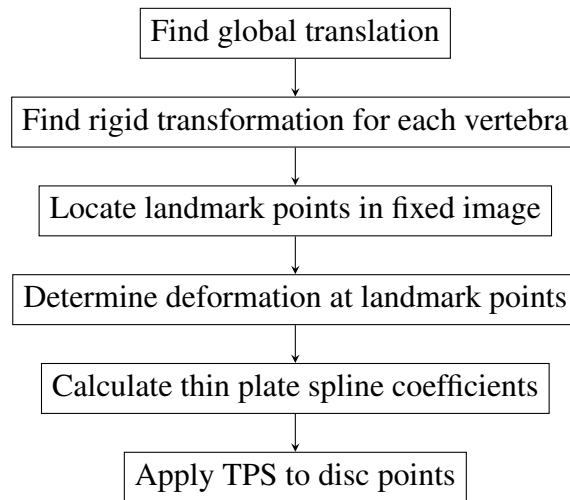
Figure 4.10: Outline of the disc registration procedure.

The first step is to apply a global translation to the moving image to get the spines in approximately the same location. We find the translation by optimizing the mutual information metric using gradient ascent. Mutual information is often used for multi-modality

registration, but it is also appropriate here because intensity values in $T_{1\rho}$-weighted images vary according to position within the magnet and with respect to the radio frequency coils. To improve robustness, we use a four level multi-resolution approach, where the images are smoothed and downsampled to create four levels of detail. Registration begins with the coarsest levels, which is less sensitive to noise but less precise. The subsequent finer levels improve precision.

After global registration, we register the discs with local precision. Discs are soft tissue and deform as the spine bends, so a deformable (non-rigid) transformation is most accurate. We use an elastic thin plate spline (TPS) transformation. Before the TPS may be computed, we must find landmark points $\mathbf{p_i}$, which are points at which we know the deformation. To find these points, we note that the discs are bounded above and below by vertebrae, which are rigid bodies, i.e. the vertebrae change position but maintain their shape as the spine bends. By choosing landmarks on the vertebrae, we can determine the deformation at these landmarks by finding the rigid transformation of the vertebrae, which is beneficial since rigid transforms have fewer degrees of freedom than most curved transforms.

We automatically determine landmark points at the edges of the vertebrae. Since closer landmarks have a greater influence on radial basis function deformation, we can have good control of the disc deformation without needing too many landmarks by selecting points only on the portions of the vertebrae that border the disc. To automatically find such points, we dilate a segmented disc until it overlaps with the two surrounding vertebrae. We then select evenly spaced points on this overlap to serve as landmark points. Sometimes the

ballooning vertebral segmentation includes the regions anterior and posterior to the disc, which are not truly part of the vertebra. We eliminate any landmarks in this erroneous section by rejecting any landmarks anterior and posterior to the disc, as done in the disc segmentation in section 4.2.2.

The user may determine the number of landmarks; more landmarks control the deformation more precisely but are more costly in computation time. To find a reasonable default number of landmarks to suggest to the user, we compare the registration results using different numbers of landmarks on a sample image. Figure 4.11 shows the landmark positioning for four examples: where the landmarks are evenly spaced with 2, 5, 10, and 20 pixels between them along the interface between the disc and surrounding vertebrae. Between the closest spacing (separation of 2 pixels) and the 5, 10, and 20 pixel spacings, the registration as measured by the mutual information between the transformed discs degrades by 0.08%, 0.33%, and 2.8%, respectively. The difference in registration was less than a quarter pixel (mean 1/50 pixel) between the 2 and 5 pixel separations, less than a half pixel (mean 1/30 pixel) between the 2 and 10 separations, and less than a half pixel (mean 1/10 pixel) between the 2 and 20 separations. The computation time to find the landmarks, apply the deformation, and calculate the deformation field drops from almost 14 seconds at 2 pixel spacing to about 9 seconds at 5 pixel spacing to almost 8 seconds at 10 pixel spacing to about 6 seconds at 20 pixel spacing. We select a 5 pixel spacing as the default since the degradation of registration is minimal compared to the reduction in time. However, even the time required for 2 pixel spacing is acceptable considering the calculations do not need

to be done in real time.

(a) Separated by 2 pixels

(b) Separated by 5 pixels

(c) Separated by 10 pixels

(d) Separated by 20 pixels
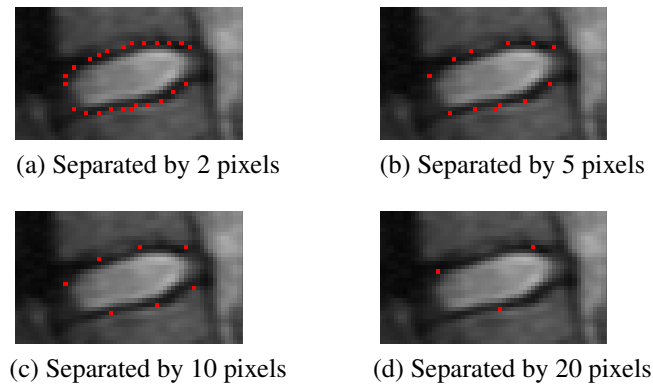
Figure 4.11: Illustration of TPS landmark positioning with different numbers of landmarks. Landmarks are separated by 2, 5, 10, or 20 pixels along the interface between the vertebral segmentation and the disc segmentation. There are places where the points are less evenly distributed because the interface is not of uniform thickness, so walking through n pixels does not necessarily yield consistent spacing.

To find the rigid transformation (translation and rotation) for each vertebra, we begin at one end of the spine and establish a bounding box that contains the entire segmented vertebra. Using gradient ascent to optimize mutual information, we find the appropriate rotation and translation. The rotation is performed around the center of the bounding box. Moving to the next vertebra, we initialize the transformation by adjusting the previous transformation to use the coordinate system of the new bounding box. This sequential method of initialization lets the algorithm succeed even if there is a lot of relative flexion/extension in the spine.

After calculating the thin plate spline weights, we apply the deformable registration to the pixels within the intervertebral discs in the $T_{1\rho}$ maps. The moving image is transformed to the fixed reference frame via inverse mapping. Specifically, each pixel in the transformed

image is calculated as follows: (1) The pixel $\mathbf{p}$ is viewed as belonging to the fixed reference frame. If it is not within one of the segmented discs, the value is set to zero. Otherwise, we continue: (2) The deformation $\mathbf{d}(\mathbf{p})$ is calculated using the TPS to yield the pixel $\mathbf{p_m} = \mathbf{p} + \mathbf{d}(\mathbf{p})$ in the moving reference frame. (3) The intensity value in the moving image at $\mathbf{p_m}$ is estimated with linear interpolation and entered into the final transformed image at pixel $\mathbf{p}$. The resulting transformed images are passed to the statistical analysis function for quantitative $T_{1\rho}$ value calculation and comparison.

Even though this method has so far performed reliably, there's the chance that anatomic variations or poor image quality could foil the algorithm. To cover this case, we include a manual touch-up feature where the user can adjust the rigid parameters of each vertebral transformation. To give the user visual feedback for these adjustments, we apply a hybrid transformation to the entire image, where points within a rigid body are transformed according to the corresponding rigid transformation, and points outside the rigid bodies are transformed according to the thin plate spline.

Chapter 5

Statistical analysis

To study the effect of disc compression on quantitative $T_{1\rho}$ values, one needs to analyze data across many people. But first we need to understand the changes in each of those individual people, which is the goal of the statistical analysis developed in this dissertation. The purpose of aligning the intervertebral discs between images is to be able to locally compare the $T_{1\rho}$ values. Here we describe a method of comparison to determine if there are statistically significant changes between images of the same subject.

5.1 Background

Hypothesis testing A standard procedure when deciding whether differences viewed in groups of measurements are real differences or just due to chance is to set up a hypothesis test. One first decides what is going to be tested. For example, is a cluster of $T_{1\rho}$ values from a volunteer's disc different in a neutrally positioned spine than in a flexed spine?

The null hypothesis is typically defined as there being no real difference. Under the null hypothesis, the values in the two clusters were drawn from the same distribution, and any measured difference is due to random chance. The alternate hypothesis is defined as there being a real change. The values in the two clusters were drawn from distributions that differ in some way, perhaps in the mean or the shape of the distribution, depending on how the test is arranged. The test proceeds by calculating the probability p that we would observe a difference at least as large as the measured difference due to random chance if the null hypothesis were true. The lower the probability p, the less likely the null hypothesis is true. If p is less than some pre-set significance level, often 0.05, the null hypothesis is rejected, though it does not prove that the alternate hypothesis is true. If p is greater than the significance level, the null hypothesis is not rejected. The evidence supports the null hypothesis over the alternate hypothesis, though the null hypothesis is not proven.

In a one-tailed hypothesis test, the alternate hypothesis is that the test statistic of group one (s_1) is not only different from that of group two (s_2) but that it is larger (or smaller, but the decision must be made before the test begins). The p value is then an estimate of the probability that we would observe the difference $\Delta = s_1 - s_2$ to be greater (or less) than the observed difference. A significance level of $\alpha = 0.05$ means that the area under the probability distribution $Pr(d)$ along $d = [\Delta, \infty)$ (i.e. one tail) must be less than 0.05. A two-tailed hypothesis test looks for absolute differences between the test statistic and is used if it is unknown or unimportant which group has the larger test statistic. The p value is then an estimate of the probability that we would observe the difference Δ to have a greater

magnitude than the observed difference. Now a significance level of $\alpha = 0.05$ means the area under $Pr(d)$ along $d = (-\infty, -\Delta] \cup [\Delta, \infty)$ (i.e. two tails) must be be less than 0.05.

If p is less than the predetermined significance level, we reject the null hypothesis. However, for non-zero probabilities p, there is the possibility that an observed difference is due to chance. When this does occur, we have rejected the null hypothesis even though it was true. This is called a type I error. Failing to reject the null hypothesis when it is in fact false is called a type II error. When we reduce the likelihood of one type of error, we often increase the likelihood of the other. Depending on the type of experiment, one type of error may be more critical than the other. For example, when looking for an effect of treatment, one usually tries to reduce type I errors since they suggest the treatment has an effect when it does not.

Resampling statistics   In parametric frequentist hypothesis testing, one makes assumptions about the underlying probability distribution from which samples were randomly drawn. A common assumption is that samples from two tested groups are distributed normally with the same variance. The mean of the distributions is the statistic that is tested, with the null hypothesis being that the means are the same and any difference in the means of the measured realization is due to chance. These types of tests provide mathematical formulae for quickly calculating the significance of an outcome, but the results rely on the validity of the assumptions on the distributions. Another approach, called resampling, is to use the samples themselves to estimate the probabilities needed for the hypothesis test.

No assumptions about the shapes of underlying distributions are required, and since the test statistic is directly calculated, it is straightforward to compare statistics other than the mean. The resampling null hypothesis is more general, stating that the two samples were drawn from the same population of possible samples (often called the same "universe" of samples), i.e. that there is no difference between the two groups other than which samples in the universe were randomly drawn. Resampling was introduced in the 1930's but has recently gained more popularity since modern computers can handle the high computational cost of Monte Carlo type simulations. See [66] for an introductory text on resampling.

To estimate p values using resampling, we recall that p is an estimate of the likelihood of seeing a difference in the test statistic at least as large as the observed difference, assuming the difference is due to random chance, i.e. assuming the null hypothesis is true. If the null hypothesis is true, both groups of samples were drawn from the same distribution, so swapping samples between groups should make no difference. Two main types of resampling used for hypothesis tests are permutation and bootstrap. In a permutation test, samples are randomly assigned to the two groups and the test statistic (e.g. the mean) is calculated for each of the two new groups. This is repeated hundreds of times and the p value is the fraction of trials that the difference in the trial test statistic is at least as extreme as the original observed difference. In a bootstrap test, the underlying universe of samples is estimated with an infinite number of copies of each of the observed samples in both groups. If there were m samples in group one and n samples in group two, we create a trial by drawing m samples with replacement from the pool of all samples from both groups to

assign to trial group one, and drawing n samples with replacement to assign to trial group two. Like the permutation test, the trial is repeated hundreds of times, and p is the fraction of trials where the difference in the test statistic is at least as extreme as the actual observed difference.

Geographically localized hypothesis testing   A standard hypothesis test does not include information about the geographical location of the samples. Samples may be taken from different locations in order to draw conclusions about the effect of location on a parameter, but usually there is only one outcome – is there a statistically significant difference between the two groups or not. We would like to create a map of statistically significant locations within the intervertebral discs. We generate a p value for each pixel based on the local pixel intensities.

Investigating spatially varying data has been done under the name geographically weighted regression (GWR), where information about the geographical location of samples is incorporated in regression analysis. For GWR details, see [67, 68]. Since parameter-based hypothesis testing can be formulated as a linear regression problem, GWR is a way of determining local statistical significance of differences between images. Points contributing to the regression are weighted according to their proximity to the point being tested. The weights are higher for closer points and lower or zero for farther points. The weighting function often has a smooth transition between high and low weights to reduce the impact of single points. If there is a semi-nearby high-intensity point, a weighting function with

a sharp cutoff changes drastically if the bandwidth of the weighting function is increased from just excluding to just including the influential point, whereas a weighting function with a gradual decline would incorporate this point more smoothly.

5.2   Methods

We explain our method of comparing the registered $T_{1\rho}$ values, testing whether or not the values at anatomically corresponding points in the two images were likely drawn from the same distribution, where differences between them are likely due to random chance. A simple approach would be to compare the means of all pixels in an entire disc, but this approach encompasses the values within both the nucleus pulposus and annulus fibrosus, which are expected to be different [18, 7] and ignores any local changes, such as a differ- ent response within the nucleus and annulus or differences between anterior and posterior locations caused by position changes or local pathology. Another approach is to select re- gions of interest (ROI), separating the nucleus and annulus [18], manually selecting an ROI at approximately the same location within the nucleus [7, 5], or manually selecting many regions of interest within the disc, trying to aim for the same regions in both images [69]. Since we have already registered the discs in the two images, matching anatomical loca- tions is done. We take into account local variations by running hypothesis tests centered at each disc pixel, including in the calculation only disc values that fall within a neighborhood of user-specified size. Allowing the user to specify the size gives the flexibility of exploring

the scope of the differences, including covering the entire disc. Only pixels within the segmented disc are considered for the statistical analysis, so even neighborhoods at the edge of the discs include only disc pixels, though the number of pixels is less than neighborhoods in the center of the disc. Increasing the size of the neighborhood gives us more samples for calculating the statistical significance but reduces the local quality of the computation.

Figure 5.1 demonstrates how neighborhood size affects the outcome of hypothesis testing. We show the statistical significance for pixels using neighborhoods of 3, 7, and 11 pixels at two-tailed significance level $p < 0.05$. Small neighborhoods yield noisier results (Figure 5.1a) and large neighborhoods spread out the areas of significant change, de-localizing the measure (Figure 5.1c). The effect of neighborhood size on the result of hypothesis testing is explored more in depth in section 6.1.

(a) Neighborhood 3 pixels   (b) Neighborhood 7 pixels   (c) Neighborhood 11 pixels
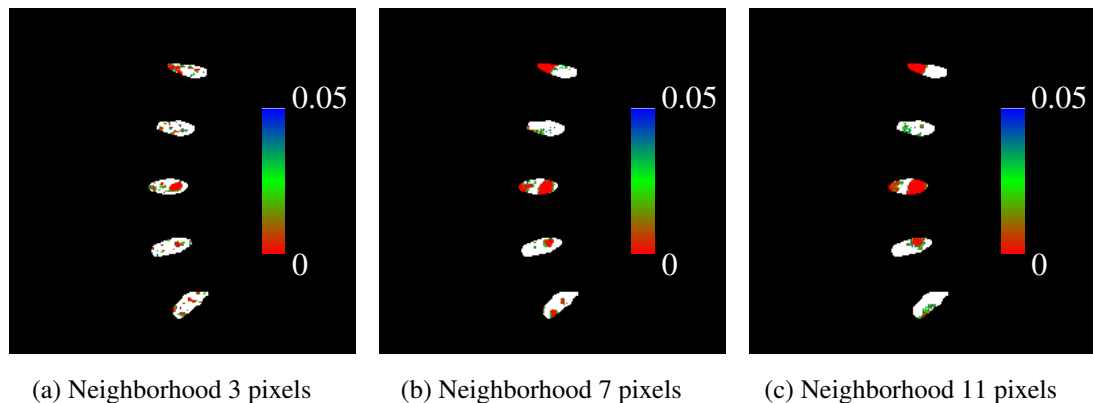
Figure 5.1: Example of how neighborhood size affects hypothesis testing. This example uses the $T_{1\rho}$ maps from volunteer I (see Chapter 7), registered from a flexed to neutral position. Pixels that are considered significant are colored according to their two-tailed significance level.

For the hypothesis testing, we use a resampling technique rather than a parametric test such as Student's t-test. Parametric tests require assumptions about the underlying prob-

ability distribution shape and variance, whereas resampling techniques are not based on such assumptions. Looking at histograms of the $T_{1\rho}$ disc values in the images we acquired, we see that the distributions vary from Poisson-like to uniform-like, but these distributions include pixels from many different locations in the discs, meaning they show the sum of many different distributions. Thus the underlying distribution is not clear. If we were using a parametric test, we could assume the Gaussian distribution since it appears often in nature and is a good approximation for many other distributions under certain conditions, but we chose to use resampling instead of making this assumption. We use the bootstrap method rather than the permutation method because it is more robust to differences in population variances. For each pixel, we follow the bootstrap procedure outlined in section 5.1. One group of samples is defined as the neighborhood around the pixel in question in the fixed image. The second group is the neighborhood around the corresponding pixel in the moving image after registration into the fixed reference frame. The null hypothesis is that the neighborhood pixels in the two images were drawn from the same distribution. Our test statistic is the mean. We use the standard significance level of $p < 0.05$, although we also report the actual value of p to give the investigator a more complete picture of the strength of the significance. Since it is not known which state of compression will yield higher $T_{1\rho}$ values, and different portions of the same disc may experience compression or decompression depending on spinal curvature, we chose a two-tailed test.

If the local neighborhood is small, the number of samples used by the bootstrap to approximate the underlying universe is relatively small, so we may question the quality of

the p-value estimate. To give an estimate how confident we are in the p-value, we modified a bootstrap method that is typically used for estimating the sample size required for a desired accuracy in a study. We estimate the probability that we committed a type I or type II error at a given pixel by estimating the probability that a random draw from the bootstrapped universe represented by the observed samples would lead us to a different conclusion about statistical significance. This error estimate is available to the investigators to help interpret the results of a data set.

To estimate the probability of error at a given pixel, we first collect the values within the local neighborhood of size n pixels from the first image in an array a_1, the n values from the second image into an array a_2, and the combined $2n$ values from both images in an array a_{both}. We find the minimum separation of means Δ_m considered statistically significant at the 0.05 level using the bootstrap technique described above. To do so, we randomly draw with replacement n values from a_{both} to represent a trial resample of the first image and calculate the mean of these n values. We draw another n values to represent a trial of the second image, calculate the mean, and subtract the two means. We repeat this hundreds of times, entering the absolute value of the difference of the means into a histogram. The minimum separation Δ_m is at the 95% quantile of the histogram. Next we calculate the percentage of time we expect to find means that are at least Δ_m apart if the two universes are represented by the samples we have. We use the bootstrap idea again to create the universes, but this time drawing with replacement only from the samples in one group rather than the pooled samples from both groups. Specifically, we draw with replacement

n values from a_1 and another n values from a_2 and compute the difference in the means of the two groups. After repeating the trial hundreds of times, the percentage P_{lower} of trials with difference in means lower than Δ_m indicates how often we would conclude that the difference is insignificant, assuming the universe is similar to the measured values. Thus if the original pixel pair was *not* deemed significantly different, we expect to make a type I error $(100 - P_{lower})\%$ of the time, and if the original difference *was* deemed significant, we expect to make a type II error $P_{lower}\%$ of the time. As a way of visualizing this information, the user can choose to have pixels highlighted only if their error estimate e is lower than a specified threshold (or equivalently the confidence level $(100 - e)\%$ is greater than a specified value).

Chapter 6

Validation

Validation of registration methods on *in vivo* medical images is difficult because the true transformation is not available for comparison. Nevertheless, researchers must use some measure to indicate the performance of algorithms. As summarized in [45], there are several items to consider in registration validation, which we will address in turn, including:

- Precision
- Accuracy
- Robustness/stability
- Reliability
- Resource requirements
- Algorithm complexity
- Assumption verification
- Clinical use

6.1 Precision

Agreement between repeated measurements is the typical definition of precision in engineering and science. For image registration, precision is defined in [45] as "the typical systematic error that can be obtained when the registration algorithm is supplied with idealized input." This definition emphasizes the properties of the system rather than the specific cases since the input is considered (close to) ideal. The input is often a phantom or a sample affected by known motion. Our evaluation of precision is two-fold, testing both the registration and statistical analysis components of the algorithm.

We evaluate the registration in two ways. First, we use the known global transformation obtained from prescribing an offset on the scanner while scanning a volunteer. We use axial images to test the prospective registration and sagittal images to test the main in-plane registration. Second, we transform vertebrae from a volunteer image, creating a realistic phantom with known vertebral transformations.

Figure 6.1 shows the result of the prospective registration on a known offset of 5mm to the left. This offset was artificially created by programming the scanner to aim the two acquisitions accordingly. The prospective registration program output was 4.97mm. The 0.03mm error is lower than the 0.1mm precision available for programming the scanner position and lower than the 0.31mm in-plane pixel size.

Figure 6.2 shows the result of the in-plane registration on a known global offset of 5.5mm to the right and 3.1mm up, i.e. rotation by 0.0 radians and transformation of (5.5mm, -3.1mm). The offset was artificially created by programming the scanner to aim

(a) First image

(b) Second image, offset by 5mm

(c) Pre - registration

(d) Post - registration
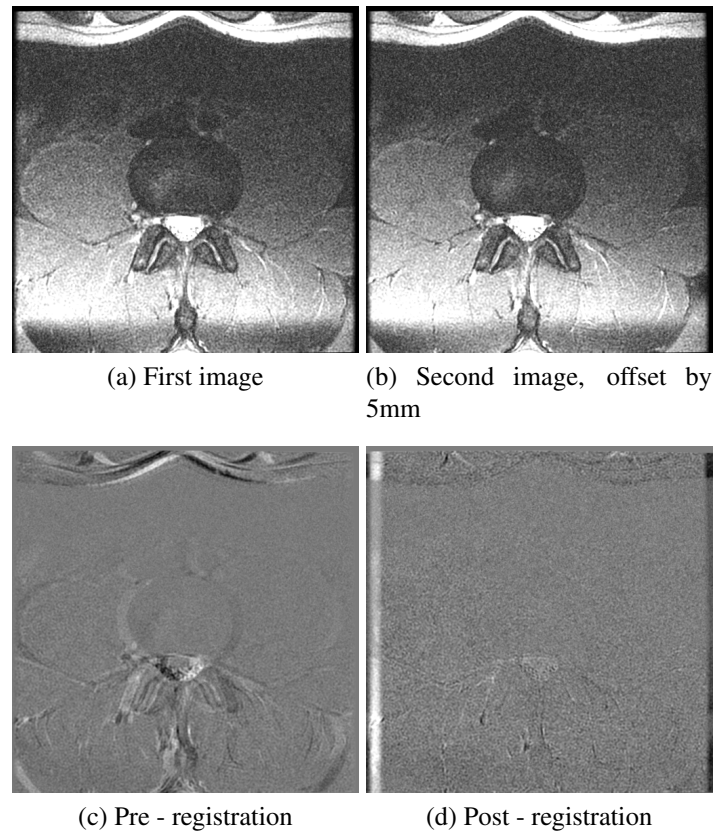
Figure 6.1: We test the precision of the prospective registration used for selecting the same anatomical location for the sagittal slice in two separate exams. The offset between images (a) and (b) was set to be 5mm to the left, and the extent of the misalignment is visible in the difference image (c). The registration algorithm found an offset of 4.97mm. Image (d) shows the difference in position after applying the registration.

the two acquisitions accordingly. The registration program output is shown in Table 6.1 as rotation and translation parameters organized by vertebra. Differences between the program output and the known transformation ranges from 0 to 0.01 radians in rotation and 0.03mm to 0.28mm in translation. These differences are lower than the 0.78x0.78mm pixel size. In fact, judging from the difference images in Figures 6.2d and 6.2e, the difference between the program output and the known global motion may be an improvement rather than an error, for perhaps the volunteer moved very slightly between acquisitions.

Vertebra	Angle (rad)	Translation X (mm)	Translation Y (mm)
Lumbar 1	0.00	5.53	-2.91
Lumbar 2	0.00	5.54	-2.93
Lumbar 3	0.00	5.57	-2.83
Lumbar 4	-0.01	5.78	-2.83
Lumbar 5	0.00	5.60	-2.84
Sacral 1	0.00	5.57	-2.90

Table 6.1: Results of automatic registration of each vertebra for a known global translation of (5.5mm,-3.1mm).

Next we look at a realistic phantom with known vertebral transformations. To create the phantom, we began with an image A of a volunteer. After segmentation, we chose transformations for each vertebra to simulate spinal flexion. The two positions are shown in Figure 6.3, (a) and (b). We applied these transformations to the original image A (creating phantom P_A) and also to a second image B acquired immediately following and in the same position as image A (creating phantom P_B). Phantom P_A is an ideal situation where we precisely know the true transformation. Phantom P_B allows us to test the registration while taking into account the intensity variations that occur between image acquisitions.

(a) First image   (b) Second image

(c) Pre - registration   (d) Post - registration   (e) Registration with the known global transformation
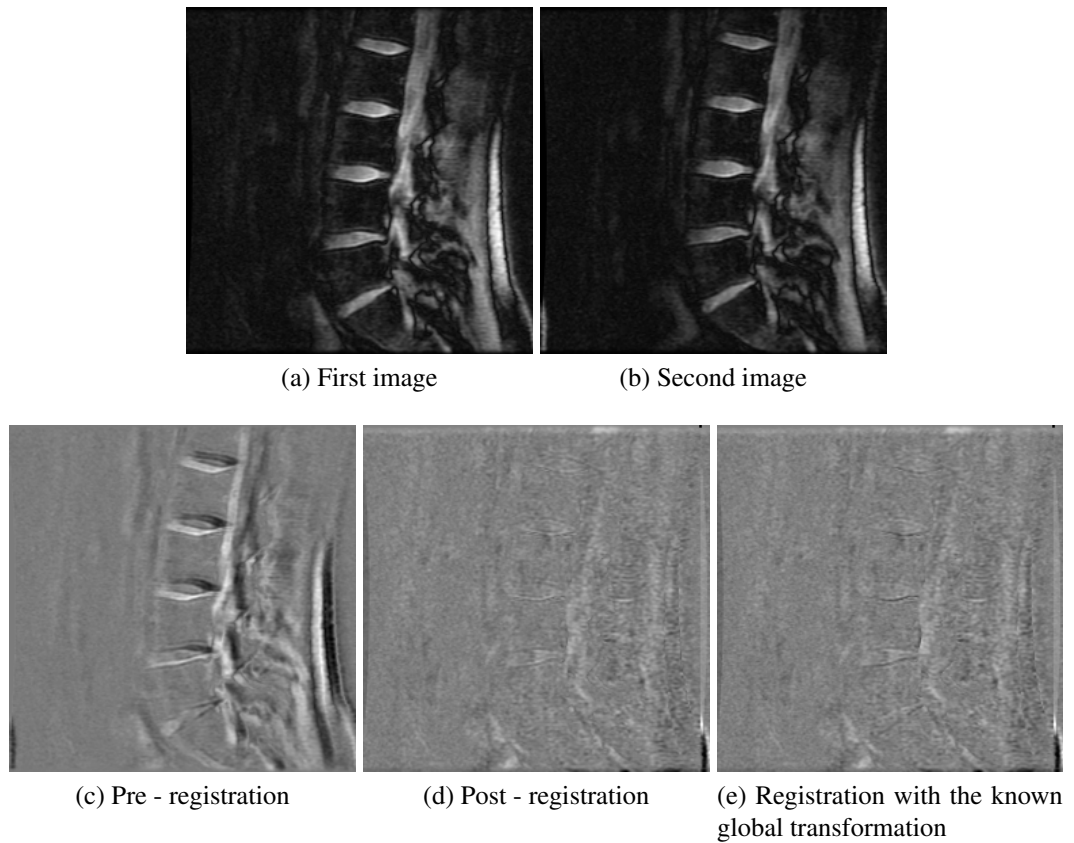
Figure 6.2: We test the precision of the in-plane registration using a known global offset. The offset between the acquisition of images (a) and (b) was set to be 5.5mm to the right and 3.1mm up. The extent of the original misalignment is visible in the difference image (c) and the improved alignment is seen in (d). Applying the known offset yields difference image (e).

Since there was minimal movement between images A and B, we consider the true transformation to be the same for both phantoms. After running the segmentation/registration algorithm on phantom P_A, we evaluate the results qualitatively and quantitatively. The result looks accurate by visual inspection, as shown in Figure 6.3d. We ran the registration three times, and the registration error is summarized in Table 6.2. Both the table and the figure show that the registration error is low. The highest error is in the disc between the first and second lumbar vertebrae, which is likely due to inconsistencies at the edge of the image, such as rotating out of the image and not segmenting the partially-included T12-L1 disc.

Phantom P_A			
	Error vector magnitude	Angle	Translation
Avg	0.08 pix (0.06mm)	0.003 rad	0.06 pix (0.04mm)
Max	0.2 pix (0.2mm)	0.008 rad	0.1 pix (0.1mm)

Phantom P_B			
	Error vector magnitude	Angle	Translation
Avg	0.2 pix (0.2mm)	0.005 rad	0.2 pix (0.1mm)
Max	1 pix (0.8mm)	0.01 rad	0.7 pix (0.6mm)

Table 6.2: Registration error summary for the phantoms described on page 82. Error vector magnitude (column 1) refers to lengths of vectors created by subtracting the true transformation vector field with the one determined by the automatic registration. Only vectors within the intervertebral discs are calculated since that is the region of interest. Angle (column 2) and translation (column 3) refer to the absolute value of the difference between the known and the calculated rigid transformation parameters for each vertebra.

The goal of the algorithm is not just to have a valid registration but to indicate where

(a) Fixed image   (b) Moving image

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) Phantom P_A (d) Phantom P_B
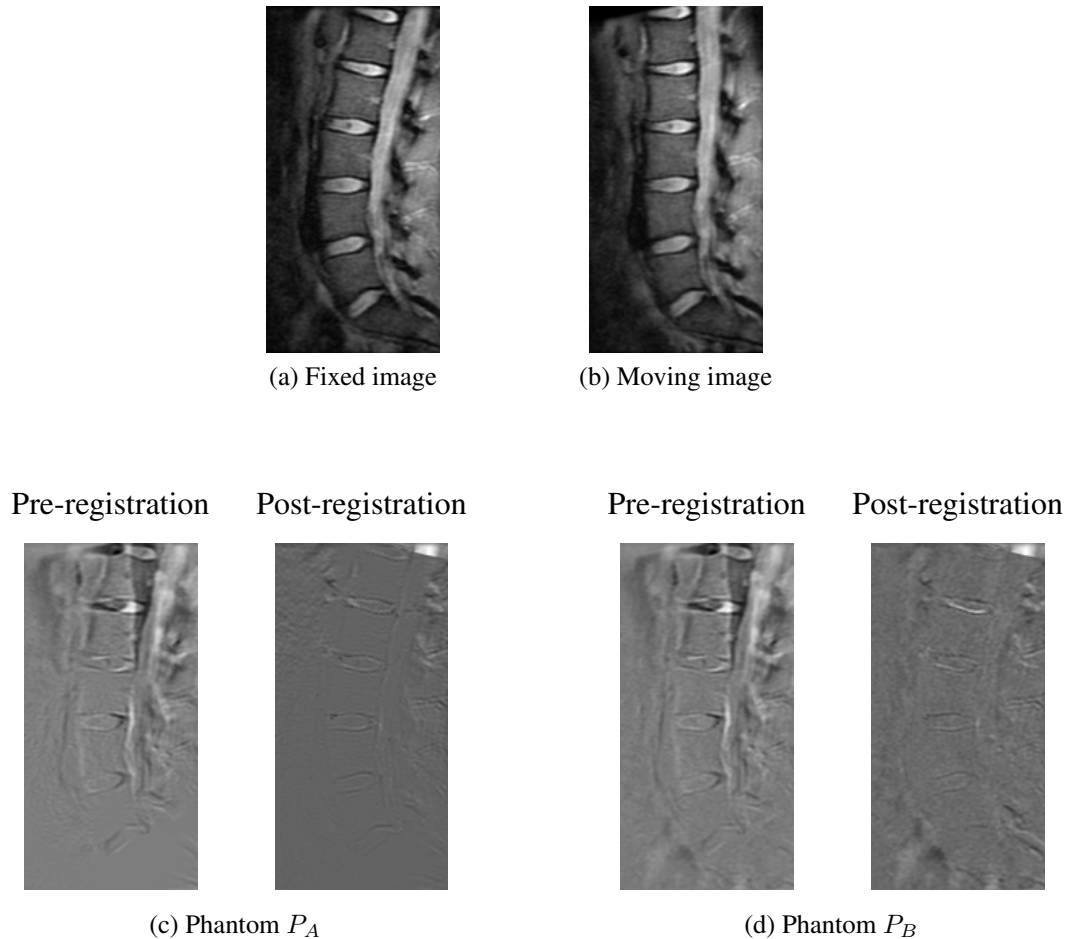
Figure 6.3: We test registration precision using a computer-generated phantom created by transforming the vertebrae of a volunteer image (a) to simulate spinal flexion, shown in (b). The second row demonstrates the registration on the phantom created by transforming image (a). The registration is quite successful as seen in image (d). The small differences seen within the disc are mostly due to interpolation error since we used linear interpolation to create a smoother, more realistic-looking phantom. The third row demonstrates the registration on the phantom created by using the same transformation but applying it to a different image of the same volunteer in the same position. As shown in image (f) and discussed on page 82, the registration is accurate with most error occurring near the superior boundary of the image.

changes in $T_{1\rho}$ values are statistically significant. Thus we also estimate the precision of the calculation of the significance level p-value. Using software phantoms, we calculate the number of pixels where we commit type I and type II errors. Since we created the phantom, we know which pixels were drawn from the same distribution and can count those for which the null hypothesis was incorrectly accepted or rejected.

We also compare the p-values with those calculated using the well-known Student's t-test. As discussed in section 5, the t-test makes more assumptions about the distribution than we are comfortable making for real data, but it is appropriate for comparisons using phantom data created with Gaussian-distributed noise. By counting the number of type I and II errors committed by the t-test, we compare the precision of our technique with the widely accepted t-test.

The phantom consists of six ovals representing intervertebral discs. Three discs (left column) are uniform in true $T_{1\rho}$ intensity, and three (right column) contain a central circle of different true intensity to model local $T_{1\rho}$ changes. Data is generated with two sources of noise. Zero-mean Gaussian distributed noise is first added to the true $T_{1\rho}$ values described above to generate noisy $T_{1\rho}$ map values m. The values m determine the $T_{1\rho}$-weighted values v at spin lock duration T_{SL} via the equation $v = \exp\left(-T_{SL}/m\right)$. Zero-mean Gaussian distributed noise is then added to the $T_{1\rho}$-weighted values v. The resulting phantom is determined by finding the best-fitting exponential parameter from these dual-noise adjusted values v as described in chapter 3 for real data. The true pixel intensities and the noise variances were chosen to be similar to those found in real data. Additive Gaussian noise is

chosen because of its preponderance in nature and its relationship to the t-test, which we use for quantitative comparison with the resampling method.

We show results from three realizations/cases of the above phantom. In case 1, the background values (i.e. the values within the discs but not in the central circle) are at mean 70ms for both images, and a central circle of mean 120ms appears in the second image. Standard deviations of the two noise sources are selected such that the local standard deviation in the resulting phantom is about 35% of the mean. (Measurements on actual data show local standard deviations of 20-30% of the mean and global standard devations of 20-45%.) Case 2 is the same as case 1 except the standard deviation is set to about 50% of the mean to investigate a worse-case situation. In Case 3, disc background intensity decreases from mean 90ms to 60ms and central circle intensity decreases from mean 150ms to 70ms. $T_{1\rho}$ values in the nucleus pulposus of healthy discs are typically 80-150ms, whereas degenerative discs have lower values, around 50-80ms. See Table 6.3 for a summary of the calculated type I and II errors. The data in the table are given in the form $A(B)$, where for type I, A is the percentage of pixels mislabeled as significant out of all truly non-significant pixels in the image. B is the percentage of pixels mislabeled as significant out of pixels that are both truly non-significant and whose local neighborhood was truly drawn from the same distribution. Similarly, for type II, A is the percentage of pixels mislabeled as non-significant out of all pixels that are truly significant, and B is the percentage of pixels mislabeled as non-significant that are both truly significant and are truly drawn from the same distribution (homogeneous neighborhood).

Table 6.3 shows that across the board, resampling and t-test results are similar, with slightly more type I errors for resampling and slightly more type II errors with the t-test. Looking at the significance maps in Figures 6.4, 6.5, and 6.6, the difference between the resampling and t-test is imperceptible. Comparing results with different sized neighborhoods in Table 6.3, we see that the type I errors increase with size and the type II errors decrease. We expect a trade-off in accuracy with neighborhood size, for increasing the size improves the hypothesis test by providing more samples per pixel, yet larger sizes reduce the local property of the measure by including more regions with potentially different means. We get an idea of the latter effect by comparing the percentages A (all pixels) and B (only neighborhoods that span homogeneous regions). In contrast with the increasing A type I errors, the B type I errors decrease or stay about the same as neighborhood size increases, suggesting that as long as the pixels in the neighborhood are drawn from similar distributions, larger neighborhoods increase or maintain the accuracy of the statistical significance calculations. The B calculation is something we can only do if we already know which locations in the image are drawn from the same distribution, which we do not for *in vivo* data. Thus the investigator should be aware of the precision trade-off with neighborhood size and view results at several sizes before drawing conclusions on a given data set. However, the numbers in Table 6.3 do not tell the whole story. For example in case 1, even though the type I errors are greater at neighborhood size 7 and the type II errors are greater at size 3, looking at the significance maps in Figure 6.4, the shape and patterns of the significant pixels at sizes 3 and 7 are very similar. The different number of errors

is mostly manifested in speckled pixels, maintaining the overall visual impression of the

locations of significant differences between the two $T_{1\rho}$ maps. Thus an increase in number

of erroneously classified pixels does not necessarily lead to a corresponding increase in

human interpretation of locations of significant changes. Finally, we note that as the level

of noise increases compared with the contrast in the image pairs (the order of decreasing

contrast to noise ratio is case 1, case 3, case 2), the number of errors increases, as expected.

Neighborhood side length (pixels) $\rightarrow$		3		5		7	
Error type $\rightarrow$		I	II	I	II	I	II
Case 1	resampling	7 (6)	20 (3)	10 (5)	4 (0)	12 (3)	0 (0)
	t-test	4 (3)	27 (5)	7 (3)	4 (0)	10 (2)	0 (0)
Case 2	resampling	6 (6)	37 (14)	9 (7)	14 (0)	12 (7)	5 (0)
	t-test	3 (3)	46 (18)	6 (4)	16 (0)	10 (5)	6 (0)
Case 3	resampling	N/A	34 (36)	N/A	6 (8)	N/A	1 (2)
	t-test	N/A	44 (47)	N/A	8 (10)	N/A	2 (2)

Table 6.3: Summary of errors in classification of statistical significance in three realizations of a phantom. Numerical values in the table represent percentage of pixels exhibiting the error. Data in each cell is presented as $A(B)$, where A is calculated using all pixels in the simulated discs, and B is calculated only at pixels where the local neighborhood was drawn from the same distribution. Data is organized to show differences between the resampling and t-test hypothesis tests and between neighborhood size. Level of significance is $p = 0.05$, two-tailed. Rejecting a true null hypothesis is a type I error, and not rejecting a false null hypothesis is a type II error. See page 87 for a discussion of this table.

(a) First image

(b) Second image

(c) Resampling, neighborhood with side length 3 pixels

(d) Resampling, neighborhood with side length 7 pixels

(e) t-test, neighborhood side length 3 pixels

(f) t-test, neighborhood side length 7 pixels
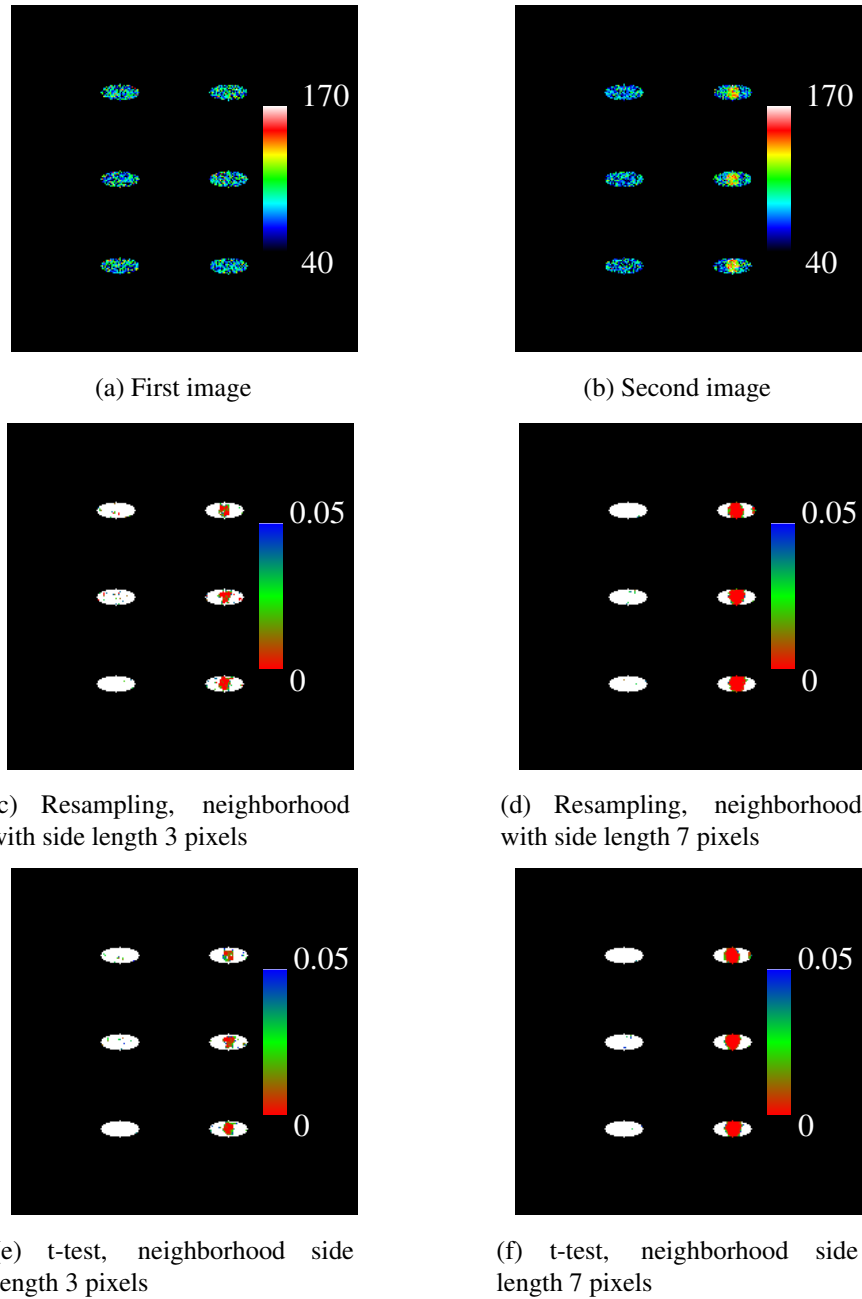
Figure 6.4: Statistical significance calculations on the phantom, case 1, as described on page 87. The first row shows the generated $T_{1\rho}$ values, where a high-intensity bright center appears in the right column of the second image. The second row shows the statistical significance values calculated using resampling. The third row shows the significance values calculated using Student's t-test for comparison. Pixels that are not significant are colored white, and the rest are colored according to their p-value. The correct statistical significance map has insignificant ovals (white) and significant central circles in the right column.

(a) First image

(b) Second image

(c) Resampling, neighborhood side length 3 pixels

(d) Resampling, neighborhood side length 7 pixels

(e) t-test, neighborhood side length 3 pixels

(f) t-test, neighborhood side length 7 pixels
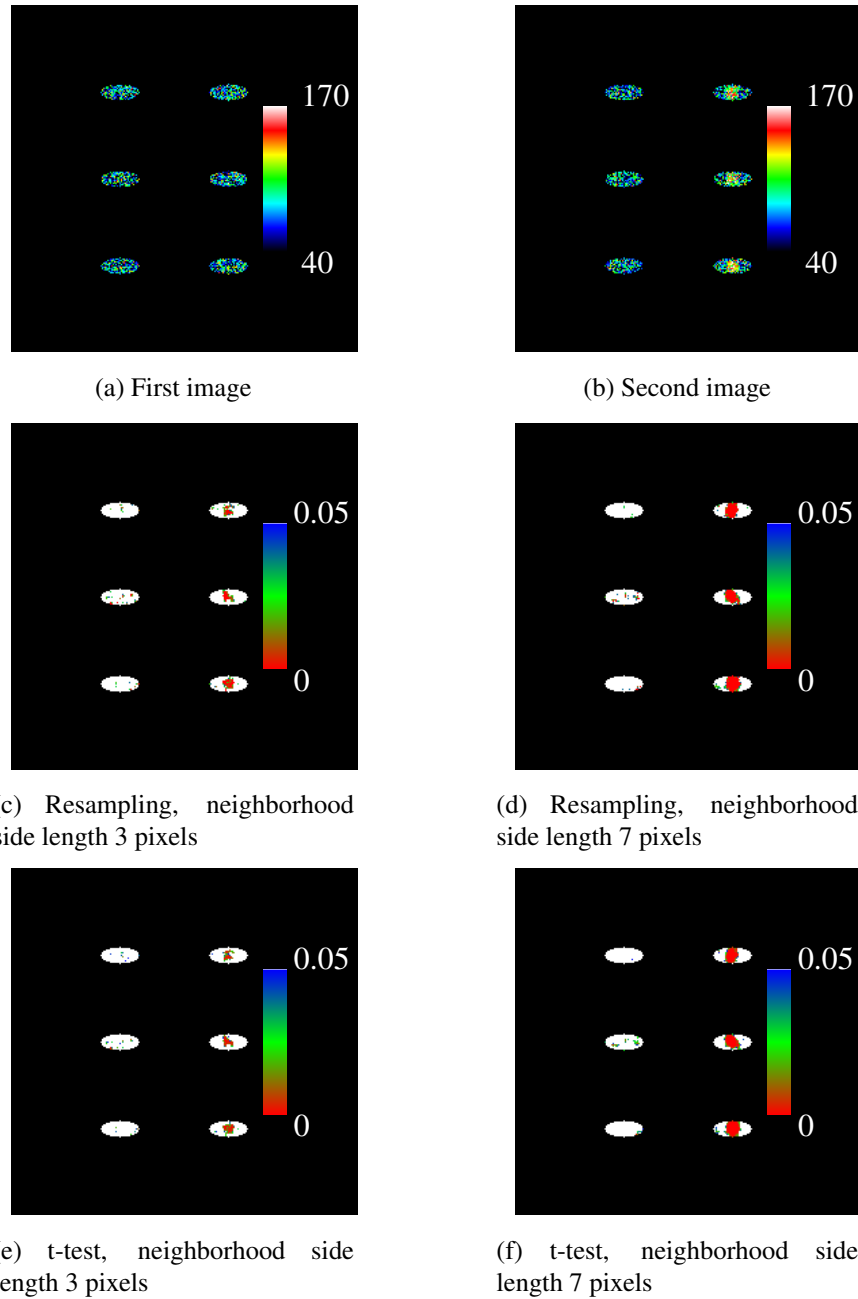
Figure 6.5: Statistical significance calculations on the phantom, case 2 as described on page 87. The first row shows the two generated images of $T_{1\rho}$ values, where a high-intensity bright center appears in the right column of the second image. The second row shows the result of resampling. The third row shows the statistical significance values calculated using Student's t-test. Pixels that are not significant are colored white, and the rest are colored according to their p-value. The correct statistical significance map has insignificant ovals (white) and significant central circles in the right column.

(a) First image

(b) Second image

(c) Resampling, neighborhood side length 3 pixels

(d) Resampling, neighborhood side length 7 pixels

(e) t-test, neighborhood side length 3 pixels

(f) t-test, neighborhood side length 7 pixels
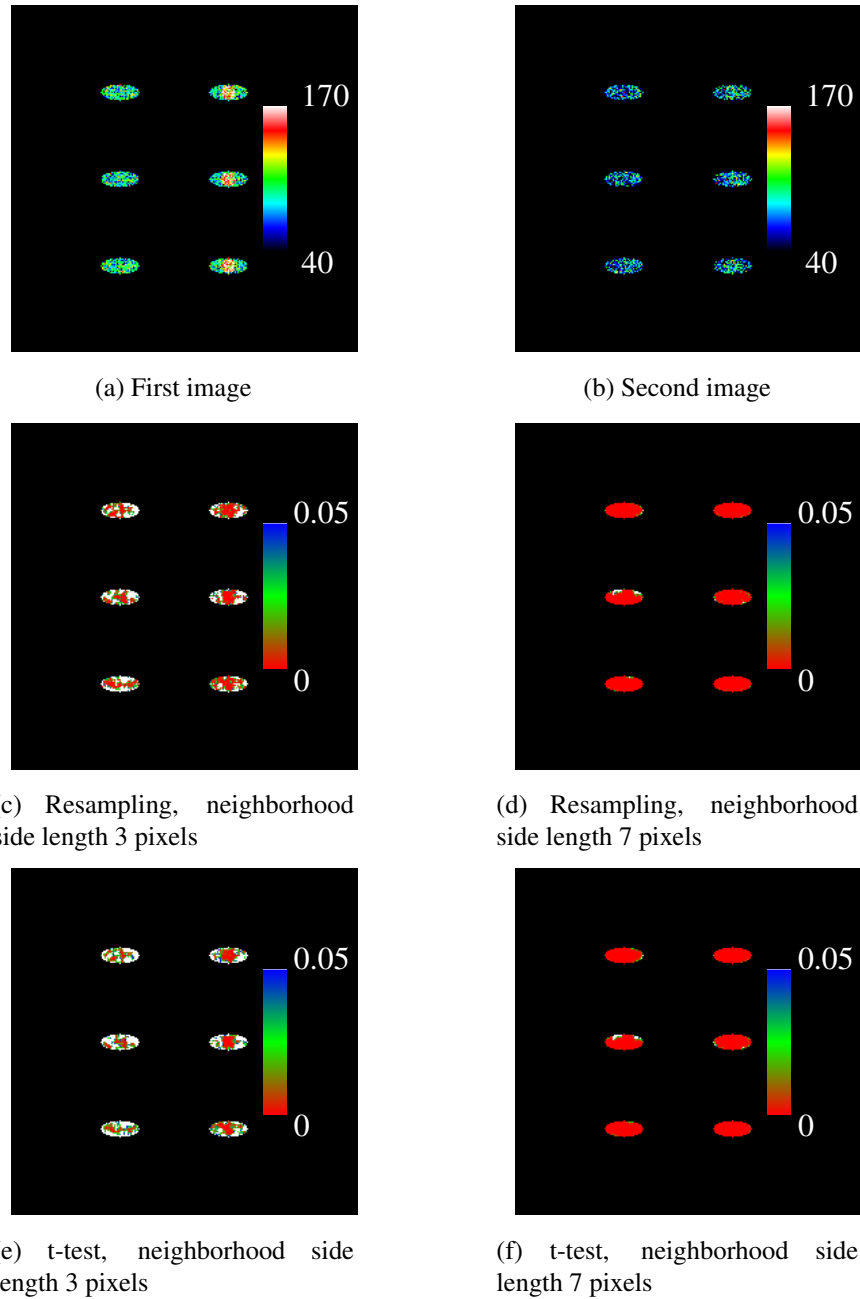
Figure 6.6: Statistical significance calculations on the phantom, case 3 as described on page 87. The first row shows the two generated images of $T_{1\rho}$ values, where the intensity decreases from (a) to (b) in both the high-intensity center and lower intensity oval. The second row shows the result of resampling. The third row shows the statistical significance values calculated using Student's t-test. Pixels that are not significant are colored white, and the rest are colored according to their p-value. The correct statistical significance map shows all pixels within the ovals as significant.

6.2 Accuracy

Accuracy refers to the error measured in a specific instance on real input. This measure is useful in evaluating the algorithm (as we are doing now) and during a clinical/research application (to give an idea of how much to trust the results for each particular case).

Evaluating registration accuracy with real images is difficult because the true transformation is unknown. For a test example, we compare the automatic result with a human observer's manual registration. However, as the manual process is tedious and defeats some of the purpose of having a computerized algorithm, our main measure of accuracy is a qualitative evaluation in the form of difference images.

The manual registration for comparison was done with the touch-up feature of the graphical user interface that we developed. The operator uses arrows to control the rotation and translation of each vertebra. S/he has visual feedback by switching between views of the fixed image, transformed image, and difference image. We had three operators each register one spine twice. The process took about an hour per operator for the two less experienced operators and about 30 minutes for the more experienced operator. Table 6.4 shows the mean and maximum difference in transformation vectors between each user's two attempts and the automatic result. For each operator, the difference between his/her two attempts is on the same order as the difference between one attempt and the automatic result. We conclude that the automatic algorithm agrees well with manual registration, at least for this example.

When used in the course of a study, the software includes a graphical user interface that

Operator		Run 1 vs Auto	Run 2 vs Auto	Run 1 vs Run 2
1	mean (pix):	0.0055	0.0058	0.0029
	max (pix):	0.51	0.52	0.57
2	mean (pix):	0.011	0.011	0.017
	max (pix):	0.84	1.3	1.6
3	mean (pix):	0.011	0.0074	0.0062
	max (pix):	1.3	0.85	1.4

Table 6.4: Comparison of manual and automatic registration. Three operators each registered the same spine twice. This table shows the mean and maximum differences between the resulting transformation vectors.

displays the difference between the fixed and moving images before and after registration, so the user has visual feedback on the accuracy of the registration for the particular case they are analyzing. Figure 6.7 gives a few examples of difference images before and after registration. Other examples are shown in chapter 7 and appendix A. Difference images are a qualitative measure of registration accuracy. The difference image of perfectly aligned identical images would be uniformly gray. Misalignments or inconsistencies between the fixed and moving images show up whiter or darker. These examples show that the algorithm yields visually very accurate registration.

To evaluate the prospective registration, we look at the difference image before and after the global registration on a pair of axial images, shown in Figure 6.8. The two images are acquired during different exams, so the subject has exited and re-entered the scanner. It is clear from parts (a) and (b) of Figure 6.8 that the images are not aligned in the superior/inferior (S/I) direction, i.e. the images shown are clearly not of the same anatomical axial plane. Part (c) is the original difference image, showing that we also start off with a

(a) Case 1: pre-registration   (b) Case 1: post-registration

(c) Case 2: pre-registration   (d) Case 2: post-registration

(e) Case 3: pre-registration   (f) Case 3: post-registration
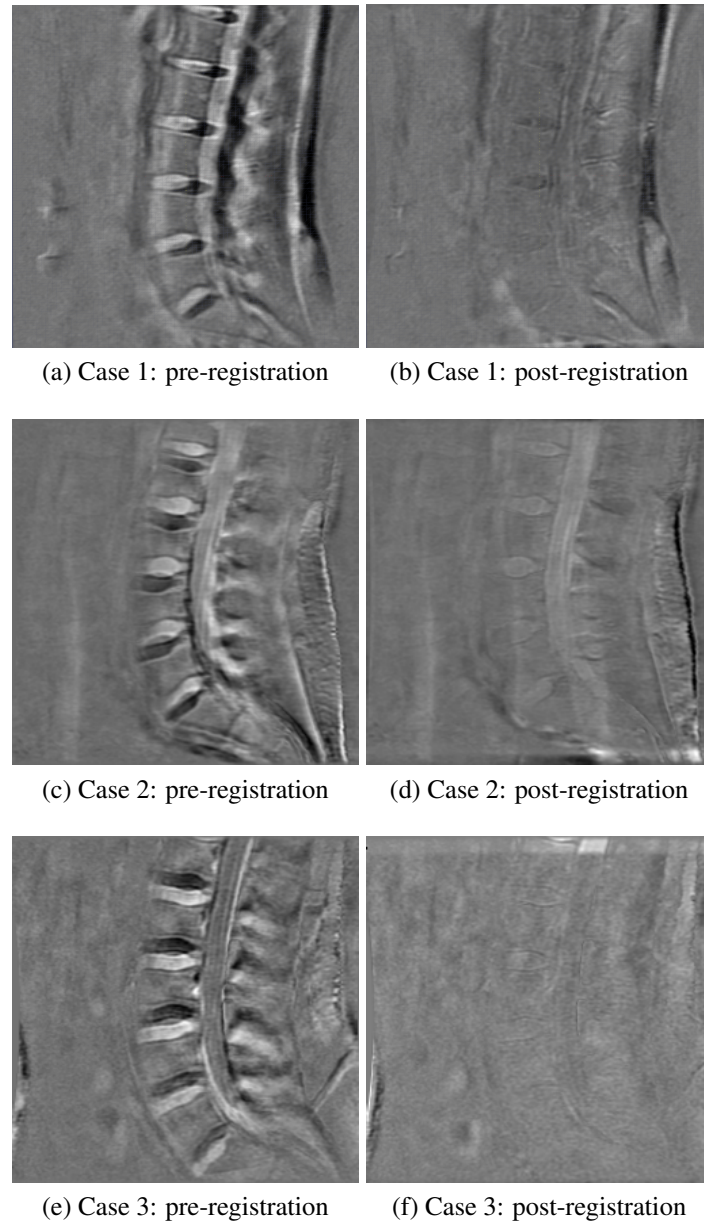
Figure 6.7: Here we show the difference between fixed and moving images before and after registration for three different volunteers. The moving image is subtracted from the fixed image, so the bright intervertebral discs from the fixed image show up whiter in the difference image, and the discs from the moving image show up darker. In the pre-registration images in the left column, the mismatch of discs and vertebrae is clear from the bright and dark regions. The alignment is visually very good in the post-registration images in the right column, with most of the darker and lighter areas due to differences between the images rather than misalignment.

right/left (R/L) and anterior/posterior (A/P) misalignment. The R/L and A/P misalignment is most obvious in the vertebral canal since the canal is bright and the vertebra is dark, making the difference image dark in the upper left of the canal and bright in the lower right. (Neutral gray represents the desired difference of zero.) Part (d) is the difference image after registration. The bright and dark intensities around the vertebral canal are symmetrical, indicating good alignment in the R/L and A/P directions. There are still many dark and light areas outside the vertebrae, indicating that the S/I alignment is still poor, which is expected since the resolution in the slice direction is about ten times larger than the in-plane resolution. However, since the purpose of the prospective registration is to aim a sagittal slice, only the R/L alignment is used, which is one of the directions that is visually accurately registered.

We next look at the automatic disc segmentation. As we will see in the robustness section (Section 6.3), registration is relatively insensitive to segmentation accuracy. However, accurate coverage of the disc is convenient for visual analysis. Thus we give a brief qualitative evaluation of the segmentation accuracy by showing several examples of varying image quality in Figure 6.9. We see that the segmented regions adequately cover most discs, with the exception of the L4-L5 disc in part (c) of the figure. There are several places where the segmented region is too large, especially in parts (b), (e), and (f) of the figure. Since the significant pixels are shown overlaid on an anatomical image, these misclassified pixels are obvious to the user. If this were an actual study, the investigator could erase those incorrect portions of the disc segmentation with a few swipes of the mouse, which is quicker than

(a) First image   (b) Second image

(c) Pre - registration   (d) Post - registration
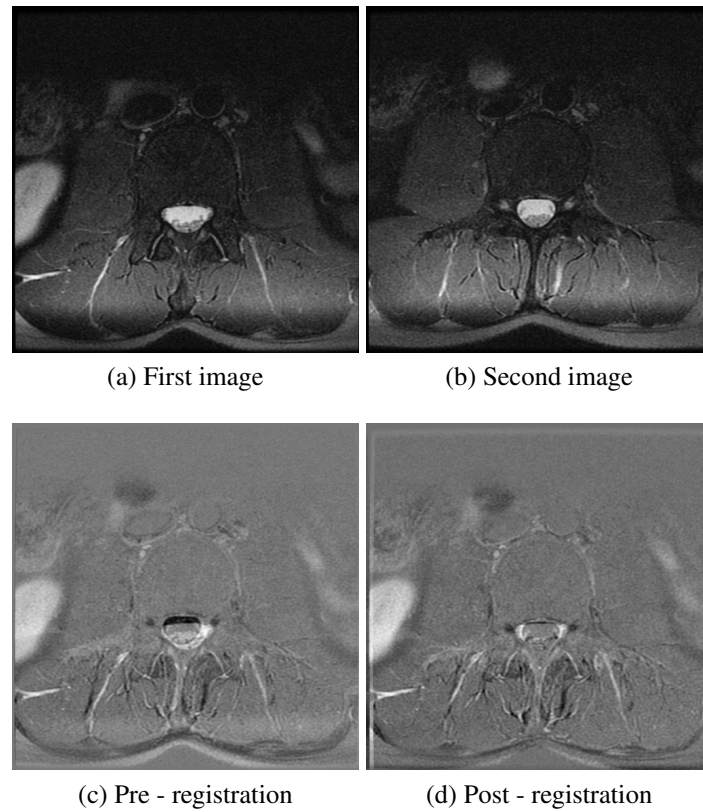
Figure 6.8: We demonstrate the accuracy of the prospective registration. Images (a) and (b) were acquired during two different exams. The volunteer exited the scanner between exams and the scanner landmark was reset. The pre-registration difference (c) shows in-plane misalignment, which is particularly visible in the dark and light regions around the vertebral canal, and out-of-plane misalignment, which is clear from the dark and light structures outside the vertebra. The post-registration difference image (d) indicates good in-plane alignment, shown by the symmetrical patterns around the vertebral canal. The out-of-plane alignment is still about the same, but that is unimportant since only the right/left (in-plane) component of the transformation is considered when selecting a sagittal slice, which is the purpose of the prospective registration.

coloring every disc.

We would also like to have an estimate of the accuracy in the decision whether to reject the null hypothesis. Since these are *in vivo* experiments, we do not know the true distribution from which the samples are drawn, so we cannot know for certain the error in the p-value estimate. Instead we use a resampling technique described in section 5.2 to estimate the probability that we committed a type I or type II error at a given pixel, providing a measure of the confidence we have in our estimate of p and therefore in our decision about rejecting the null hypothesis.

This accuracy estimate helps the investigator interpret the results of a given data set. However, before applying it to real data, we get a feel for the measure by applying it to the phantom from page 86. We then reject the null hypothesis only for pixels that are both under the $p = 0.05$ significance level *and* are estimated to be in error less than X% of the time, i.e. at confidence level 100-X%. In Table 6.5, we show the number of type I and type II errors on the case 1 phantom at confidence levels 30%, 50%, 80%, and 95% using a neighborhood of 5 pixels. The estimated error was less than 70% for all pixels, so the 30% confidence level is identical to the situation without accuracy estimates, i.e. the case 1 resampling size 5 cell in Table 6.3. As we increase the confidence level, we trade off type I errors for type II errors since we are rejecting the null hypothesis less often. Recall the error percentage in parentheses in the table is calculated using only pixels where the neighborhood is truly drawn from the same distribution. These errors are less affected by the confidence level since they are less prone to error. The error-prone pixels cluster at the
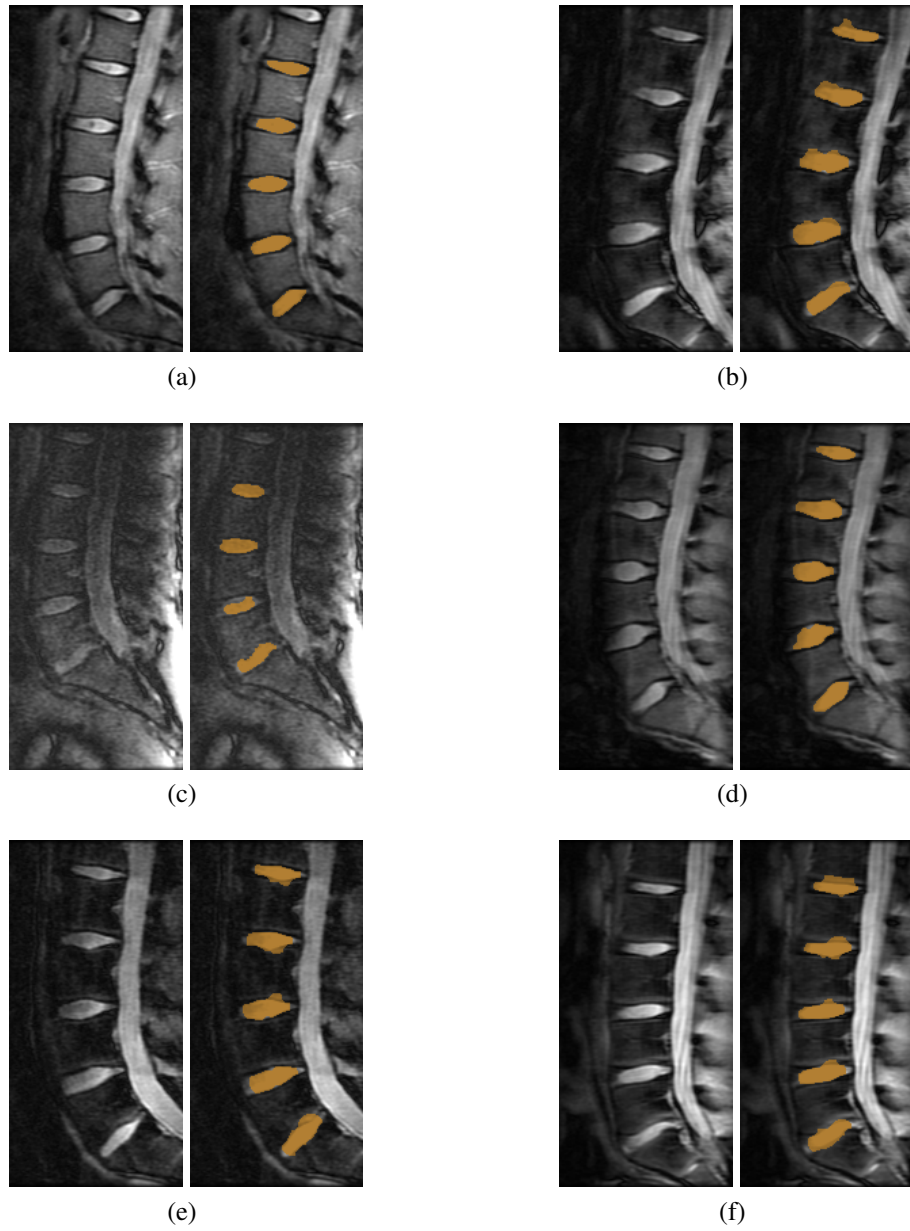
Figure 6.9: Examples of automatic disc segmentation. For each image pair in this figure, the image to be segmented is on the left, with the resulting segmentation overlaid in brown on the right.

boundaries between universes, i.e. at the border of the high-intensity central circle, as seen

in Figure 6.11.

Confidence level	30%		50%		80%		95%	
Error type	I	II	I	II	I	II	I	II
Error percentage	10 (5)	4 (0)	9 (5)	4 (0)	5 (3)	9 (0)	3 (3)	27 (0)

Table 6.5: Error count for hypothesis testing on a known phantom. Data is presented in the same format as Table 6.3. Numerical values in the table represent percentage of pixels exhibiting the error. Data in each cell is presented as $A(B)$, where A is calculated using all pixels in the simulated discs, and B is calculated only at pixels whose local neighborhood was drawn from the same distribution. The A values trade type I for type II errors as the confidence level increases. The B values are less affected by the confidence level. See page 98 for details.

While analyzing a particular data set, the researcher can opt to threshold the significant

pixels at a selected confidence level. Figure 6.11 compares the effect of no thresholding

with the effect of thresholding at the 80% and 95% confidence levels. Type II errors in-

crease at higher confidence levels, but type I errors decrease, and often type I errors are

considered more harmful since they suggest change where there is none.

6.3   Robustness/stability

In registration, robustness/stability is the property where small variations in the input

yield small variations in the result. The three important areas of stability in our algorithm

are robustness of segmentation given variation in manual initialization, robustness of reg-

istration given errors in the segmentation, and robustness of statistical analysis given errors

in the registration.
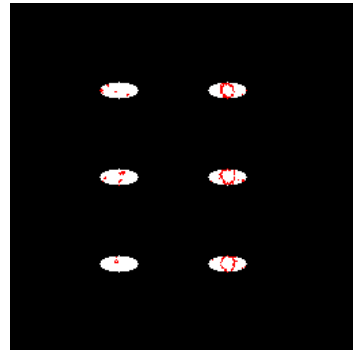
(a) More than 5% error is estimated for the red pixels
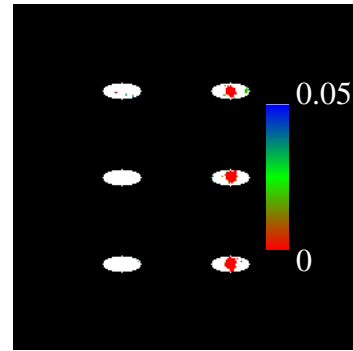
(b) Colored pixels are significant at the 95% confidence level
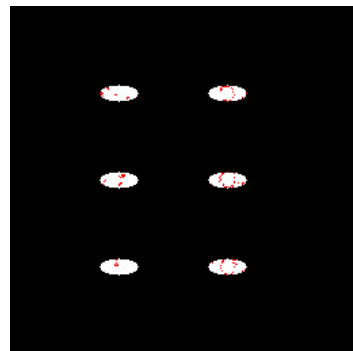
(c) More than 20% error is estimated for the red pixels
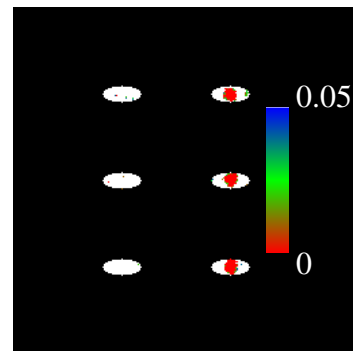
(d) Colored pixels are significant at the 80% confidence level

Figure 6.10: We demonstrate the effect of rejecting the null hypothesis only if we expect to be correct more than 5% (top row) or 20% (bottom row) of the time. Pixels in red in the left images, (a) and (c), have p-values below the significance level of 0.05, but since the confidence level is lower than the specified threshold, we do not consider them significant. Colored pixels in the right images, (b) and (d), are statistically significant with estimated error lower than the specified threshold.

(a) No error thresholding   (b) 80% confidence level   (c) 95% confidence level
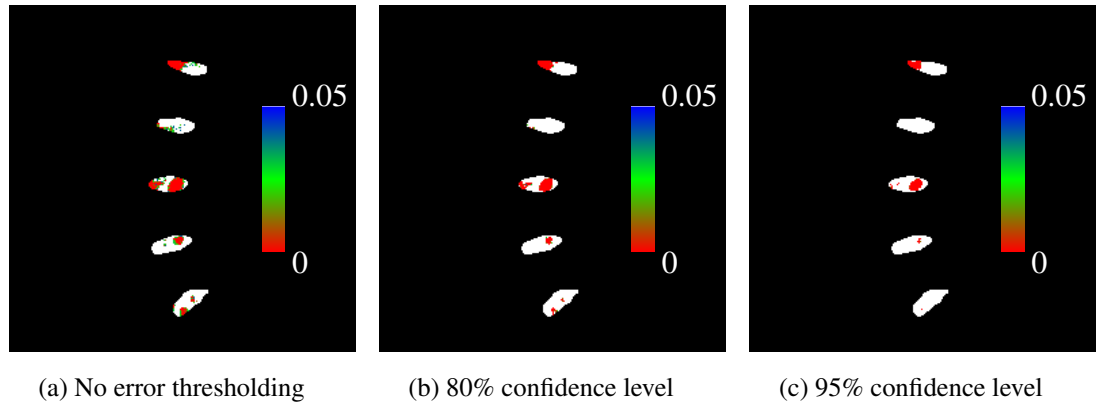
Figure 6.11: We show the effect of thresholding the statistical significance p values according to the amount of error expected in the p value calculation. This example takes the registered $T_{1\rho}$ maps between a neutral and flexed position (see Figure 7.27) and runs hypothesis testing with a neighborhood of side length 7 and significance cutoff $p < 0.05$, two-tailed. Results are in (a). If we do not reject the null hypothesis for pixels with confidence lower than 80% and 95%, the statistically significant pixels change as shown in (b) and (c), respectively. Thresholding at higher confidence levels decreases the type I error but increases the type II error.

We test robustness of segmentation with respect to initialization by initializing with intentionally misplaced points. Test points were selected directly at the anterior/posterior (A/P) points of the disc, slightly within the disc, largely within the disc, slightly outside the disc, or largely outside the disc. We visually evaluated the resulting segmentations on two volunteers, shown in Figure 6.12. As the A/P points get farther apart (rows 2 and 3 in the figure), the segmentation does not change much. Minimal change is expected since the points are still close enough to the disc that the derived seed points are still within the disc (as described on page 56). The exact seed placement within the disc is slightly different, meaning the automatic threshold may end up different, accounting for the slight segmentation differences in volunteer 2 (right half of the figure). As the initialization points get closer together (rows 4 and 5), the resulting segmentation gets narrower, as is expected

since the A/P points designate the extreme points of the segmentation. All of these changes in segmentation are reasonable and expected given the different input points, indicating a high level of robustness. As long as the points are fairly close to the disc and located exterior to the portion of the disc that is being evaluated, the resulting segmentation is reasonably consistent.

To test the registration robustness, we created intentional mis-segmentations where the bottom of each disc was either eroded or dilated by 3 pixels, as shown in Figure 6.13. We compare the registration deformation vector field automatically obtained using the correct, eroded, and dilated disc segmentations. This example registers the neutral and flexed positions of volunteer I (see Figure 7.10). All three vector fields match within a quarter pixel (mean difference less than a tenth of a pixel). The close match suggests slightly poor disc segmentation does not affect the registration much. Therefore the main penalty for a slightly poor disc segmentation is in the statistical analysis step, where some non-disc pixels would be included or some disc pixels would be excluded. The investigator can easily locate these areas on the disc segmentation overlay in the graphical user interface and can either manually touch up the segmentation or mentally discard extraneous pixels as unreliable.

To get an idea of the robustness of the statistical analysis, we took two images that were originally in different positions, found the automatically determined registration parameters, perturbed the correct registration slightly, and compared the resulting significance levels before and after the mis-alignment. Global translation by a tenth of a pixel resulted
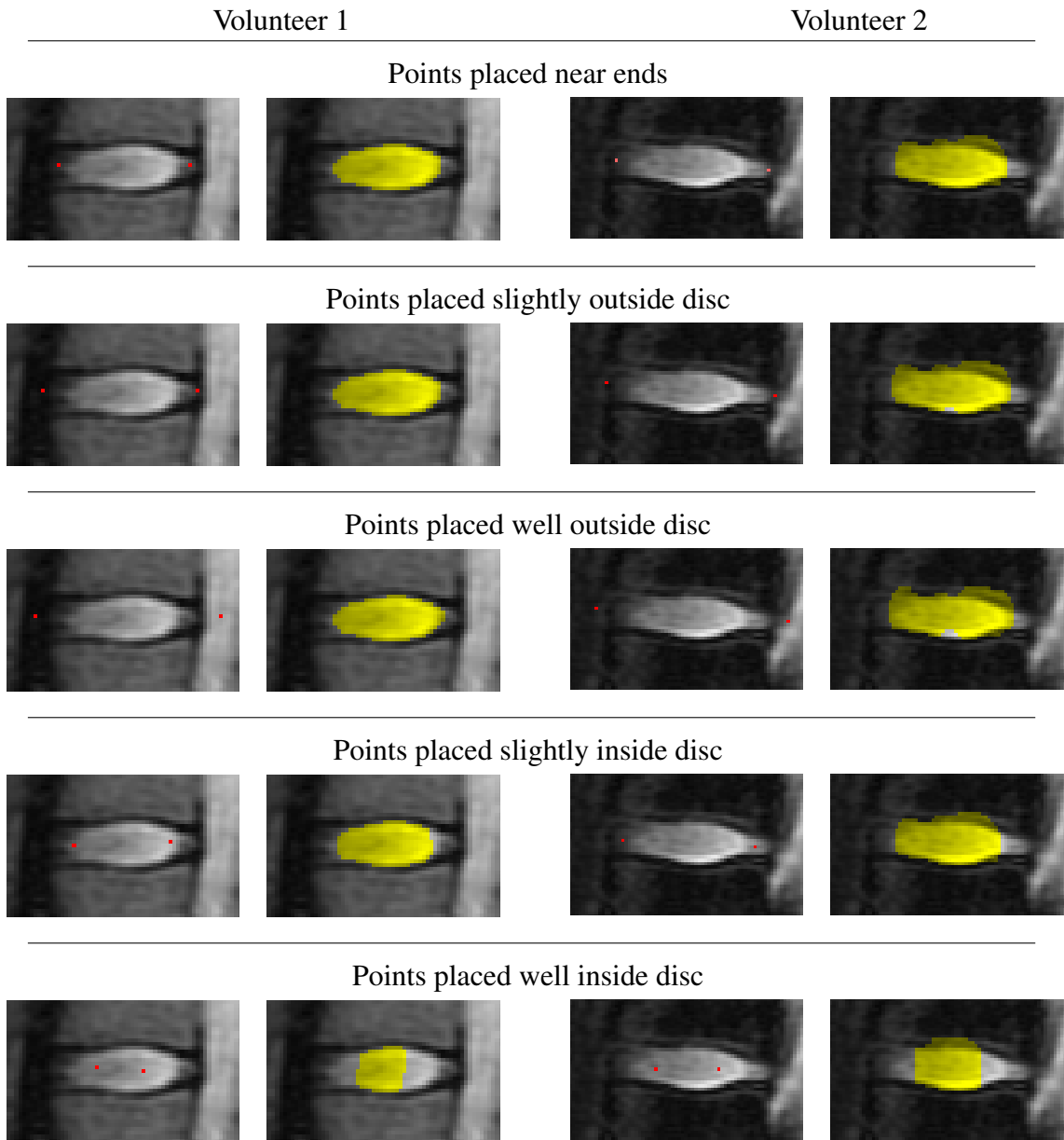
Figure 6.12: We explore the robustness of the disc segmentation with respect to the user-placed initialization points. The user is instructed to place points at the anterior and posterior ends of the disc. The figure shows the resulting segmentation with intentionally misplaced points. We show representative discs from two volunteers. The two columns on the left show a volunteer whose discs were bright and easy to segment. The two columns on the right show a volunteer whose vertebrae had bands of higher intensity, confusing the segmentation. The initialization points are shown in red and the resulting segmentation in yellow. The first row in the figure shows initializations that we may reasonably expect from a typical user. Each of the following rows show a different type of point placement error.
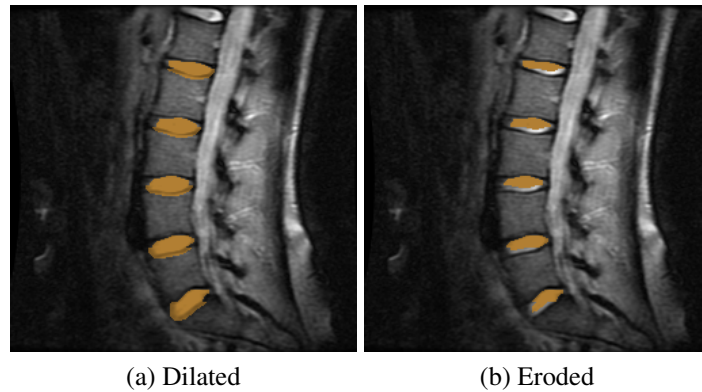
(a) Dilated   (b) Eroded

Figure 6.13: We intentionally corrupted the disc segmentation to test the robustness of the registration with respect to errors in the automatic segmentation. The bottom of the discs were dilated by three pixels in (a) and eroded by three pixels in (b).

in 4% more pixels marked as significant at the $p < 0.05$ level. Translation by a quarter pixel and one full pixel led to 9% and 55% more pixels marked as significant, respectively. The random nature of the resampling hypothesis test yielded about a 2% variation in number of significant pixels even without any registration difference. Figure 6.14 shows that the changes in significant pixels for the quarter pixel translation is barely noticeable, whereas many additional significant regions appear after the one pixel translation. The figure also shows the difference image after the quarter and one pixel translations. The mis-registration is clearly seen for the one-pixel translation (black and white lines above and below each disc), meaning the investigator can see this registration error and either correct it using the user interface touch-up feature or mentally take it into account when analyzing the meaning of the significance maps.

(a) Difference image using the correct registration

(b) Difference image with mis-registration of 0.25 pixel

(c) Difference image with mis-registration of 1 pixel

(d) Significant pixels using the correct registration

(e) Significant pixels with mis-registration of 0.25 pixel

(f) Significant pixels with mis-registration of 1 pixel
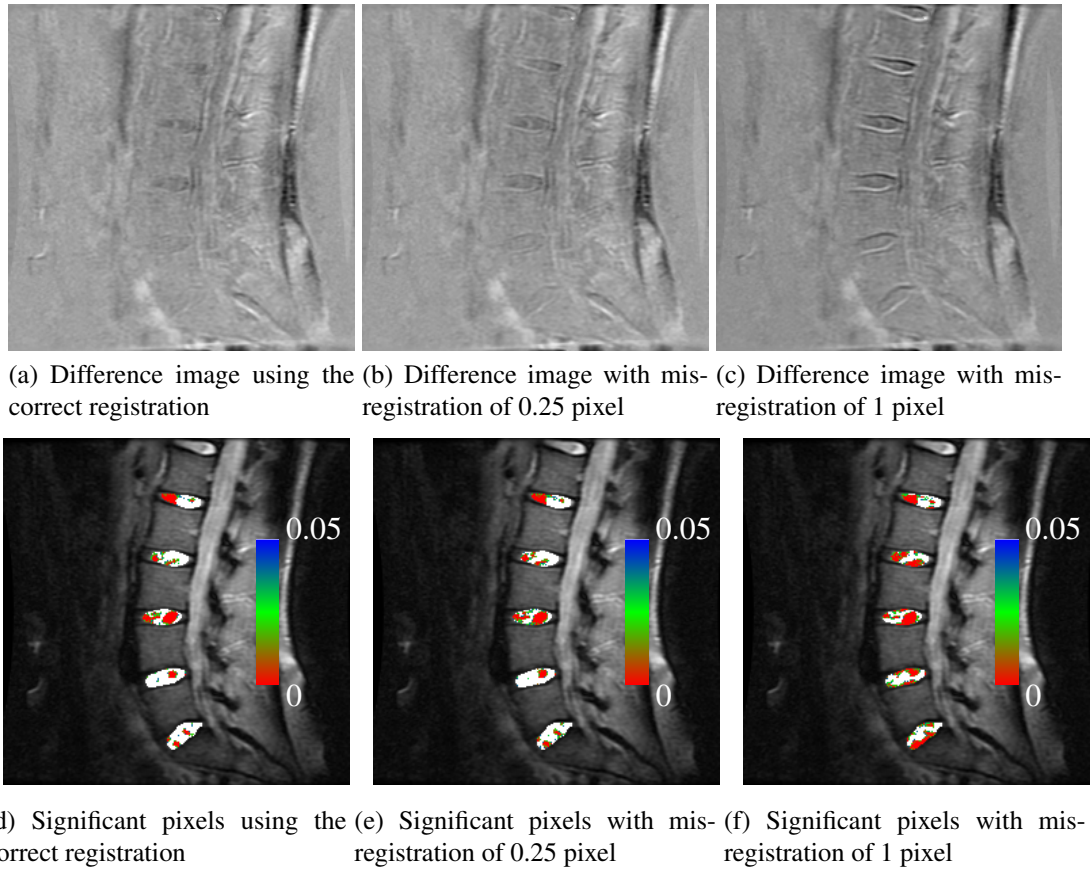
Figure 6.14: We demonstrate the effect of small registration errors on the results of hypothesis testing. The left column considers the "correct" registration, i.e. the output of the automatic registration algorithm. The other columns consider the intentionally incorrect registration obtained by increasing the "correct" translation parameters by 0.25 pixel (0.2mm) upward for the middle column and by 1 pixel (0.78mm) upward for the right column. In each column, the upper image is the difference between the fixed and transformed moving images to give an idea of the extent of mis-alignment. The lower image shows the result of the hypothesis test. We see that the 0.25 pixel translation is barely noticeable in both the difference image and the significance map, whereas the one pixel translation makes a visible difference. The analyses in this figure use the neutral/flexed positions of volunteer A with hypothesis testing on a neighborhood of 5 pixels and two-tailed significance level of $p < 0.05$.

6.4 Other validation measures

Reliability A reliable algorithm will perform acceptably given the expected range of input, i.e. a reasonable amount of anatomical variability and noise levels. Such variations are difficult to predict. In the four different patients in the five different compression situations presented here, the segmentation and registration are visually successful. We did not test the algorithm on any extremely degenerate cases. To deal with any extreme situations that foil the algorithm, there are manual touch-up features as explained in sections 4.2.2 and 4.2.3.

Resource requirements The material and effort involved in the analysis should be commensurate with the clinical or research benefit of the algorithm. To run the algorithm on a given case, a human being must make two semi-accurately aimed clicks per disc. The computation can take several minutes to complete, but the user can be doing something else during that time. The user must then interpret the results in the context of the overall study. The computer must be reasonably fast with enough memory to compile the C++ code, but most modern computers in research or clinical institutions meet this requirement so no extra hardware is required. The code must also be maintained, but no more so than any other development software. The use of the National Library of Medicine's Insight Registration and Segmentation Toolkit and the C++ Standard Template Library help make the code readable and maintainable by other researchers. In exchange for these required computing resources, we are released from some man-power requirements. Manual seg-

mentation and registration is time-consuming and prone to error especially for users with less medical experience. Reducing the human resource requirements improves the accuracy and speeds up overall analysis time, making a systematic study of *in vivo* quantitative parameters a reasonable proposition.

Algorithm complexity   The computation requires several minutes on a 2GHz Intel processor with 2MB RAM running Ubuntu Linux. Since the analysis does not need to be done in real time, this is a reasonable amount of time. Computation time does not differ much with amount of initial mis-alignment. Larger neighborhoods in the statistical analysis noticeably increase the running time but only on the order of a minute.

Assumption verification   During algorithm development, we assumed that that the disc deforms according to a thin plate spline function. We cannot know the exact true disc deformation of *in vivo* images. Instead we address this assumption by exploring different deformations consistent with the rigid transformations of the surrounding vertebrae. We compare results using different radial basis functions and stiffness values, using the data from volunteer I (see Chapter 7), registering between neutral and flexed.

First we apply the four different kernels presented in section 4.1.1: two-dimensional thin plate spline, three-dimensional thin plate spline, volume spline, and elastic body spline. The difference in registration using each of these basis functions is minimal. The transformation vectors are all within a quarter pixel of each other (mean less than a twentieth of a pixel). The thin plate spline kernel $r^2 \ln r$ has the best match according to the mutual

information metric, but the improvement is less than 0.5%. The small magnitude of the differences suggest that as long as the non-rigid deformation scheme we choose is in agreement with the surrounding rigid vertebral transformations, our use of the two-dimensional thin plate spline is satisfactory.

Next we look at the stiffness parameter in the TPS, a value usually around $0 - 0.1$ that permits the TPS to approximate rather than interpolate between landmarks, providing a deformation that is smoother and less sensitive to landmark errors [56]. Even with the relatively large difference in stiffness between 0 and 1, the mutual information improves only 0.007%, with less than a thirtieth of a pixel difference in deformation vector magnitude (mean less than 1/500 pixel). The minuscule difference suggests that introducing a stiffness parameter does not noticeably improve the match between the TPS deformation and the true anatomical deformation, so we set the stiffness parameter at zero, making the TPS an interpolating rather than an approximating spline.

Clinical use   This algorithm is designed for research use to improve the understanding of the behavior of $T_{1\rho}$ values in different situations. We must understand the changes in $T_{1\rho}$ if we are to study the relationship between $T_{1\rho}$ and early disc degeneration. Studying early disc degeneration may lead to early treatments and decreased morbidity in spine disease.

Chapter 7

Experiments

We demonstrate our registration and statistical analysis methods with three sets of examples. Each set constitutes a small preliminary experiment on $T_{1\rho}$ behavior in lumbar discs under differing compression states. The studies are so small that while they may give us insights into $T_{1\rho}$ behavior, they are not meant to draw general conclusions. We give suggestions for further study in Chapter 8 to expand on these investigations.

7.1 Methods

7.1.1 Data acquisition

For each of the three experiments, we used the fast spin echo $T_{1\rho}$ sequence described in Chapter 3. The sequence details are repeated here for reference in Table 7.1. Note that the echo train length (ETL) affects the calculated $T_{1\rho}$ relaxation values. We chose a long

ETL of 16 to shorten the sequence to a reasonable duration for examining $T_{1\rho}$ dynamics. Since the ETL does not change in these experiments, we can make relative comparisons within these studies, but the $T_{1\rho}$ values should not be directly compared with studies using a different ETL (e.g. an ETL of 2 [8]).

FOV (field of view)	20 cm
Imaging matrix	192x128
Slice thickness	8 mm (one central sagittal slice)
In-plane pixel size	0.78 x 0.78 mm
TR (repetition time)	2000 ms
ETL (echo train length)	16
BW (bandwidth)	244.1Hz
NEX (number of excitations)	2
Flip angle	90°
Spin lock frequency	300 Hz (0.007 mT)
Spin lock durations	0/20/50/90 ms
Scanning time	2:11 minutes

Table 7.1: $T_{1\rho}$ Imaging parameters

We recruited volunteers without back pain for these preliminary studies for several reasons. Healthy discs are expected to change more with compression, as discussed in Chapter 1, so we chose healthy volunteers for a better chance of seeing changes. In addition, the compression procedures are potentially painful for someone with back pain, so we initially chose healthy volunteers to try the procedures. Recruited volunteers were aged 21 to 38 years, with mean age 30.5 years.

In the first experiment, we looked at dynamic disc decompression while the volunteer relaxed supine for one hour. Initial compression occurred due to normal weight-bearing activities plus approximately 30 minutes of standing wearing a 20 lb backpack immediately

preceding the exam. The exam consisted of repeated $T_{1\rho}$ maps over the course of one hour. Four healthy volunteers participated. One volunteer complained of back soreness at the completion of the procedure. The exams occurred on the General Electric Signa 3 Tesla scanner at the University of California, San Francisco Radiology Imaging Center at China Basin Landing. The receiver coil was a 4-channel spine coil located on the scanner bed posterior to the volunteer.

For the second experiment, we created compression changes by using a leg-press-like device, where hanging weights compress the spine by pulling a shoulder harness toward a plastic footplate, as shown in Figure 7.1. We applied and released approximately 55 lbs as the volunteer lay supine. The volunteer remained in the scanner, instructed not to move during the approximately 45 minute exam. The procedure was well tolerated by the four healthy volunteers who participated. The exams occurred on the General Electric Signa 3 Tesla scanner at the University of California, San Francisco Mission Bay Campus. The receiver coil was an 8-channel spine coil located on the scanner bed posterior to the volunteer.

A single healthy volunteer alternately flexed and relaxed his/her spine in the third experiment. The volunteer lay supine with knees resting flat on the scanner bed for the neutral (relaxed) images. For the flexed image, he or she used leg and abdominal muscles to round the spine into the scanner bed. The volunteer held each position for the duration of the 2:11 minute scan, breathing shallowly to minimize motion artifacts. These images were taken during the same exam, and the volunteer did not exit the scanner between images. Holding

Figure 7.1: The compression device consists of a shoulder harness, straps, foot plate, and weights.

the positions without moving was difficult. The exams were taken on the General Electric Signa 3 Tesla scanner at the University of California, San Francisco Mission Bay Campus. The receiver coil was an 8-channel spine coil located on the scanner bed posterior to the volunteer.

The images using the compression device were acquired as part of the study titled "Advanced MR Imaging in Patients with Painful, Degenerative Disc Disease: A Pilot Study." This study is being conducted at the University of California at San Francisco under principal investigator Sharmila Majumdar, PhD. It is funded by the United States National Institutes of Health, and was approved by the UCSF Institutional Review Board (Committee on Human Research), with approval number H6513-26624-03. The images without the compression device were acquired as part of the study titled "Assessment of Spinal Stenosis and Characterization of the Intervertebral Disc," also conducted at UCSF with principal investigator Sharmila Majumdar, PhD, and funded by the National Institutes of Health. It

was approved by the UCSF Committee on Human Research with approval number H6513-19070-06.

7.1.2 Data analysis

We demonstrate the quality of the automatic registration by subtracting the fixed and moving images before and after registration, as described in section 6.2.

To analyze the changes in $T_{1\rho}$, we show the statistically significant local differences in $T_{1\rho}$ for each pair of consecutive images. The statistical tests were run with a neighborhood size of 5 pixels and significance level $p < 0.05$, without confidence thresholding. Statistically significant changes in the mean value of a neighborhood are indicated by coloring the central pixel in the neighborhood with the amount of the change, reported in milliseconds. Positive change is colored in red and indicates that the $T_{1\rho}$ parameter increased from the first to second image (i.e. $T_{1\rho}$ relaxation is slower in the second image since the time is longer). To relate the $T_{1\rho}$ with physical evidence of compression or relaxation, we calculated changes in disc height between each two images. We selected pairs of points at the superior and inferior border at the anterior, central, and posterior part of each disc. Then using the transformation parameters automatically determined during the registration process, we transformed these points from the earliest image of a volunteer into each of the later images and measured the length of the lines connecting each superior/inferior point pair. We then measured how much the line length changed between images. The difference in length is reported in percent change, with positive numbers indicating an increase

in length (expansion) and negative numbers indicating a decrease in length (compression). To test the repeatability of the disc height measurements, we repeated the registration four times and repeated the point selection four times for two volunteers. With five discs per volunteer and three disc height measurements per disc, this gives us 7 (=4+4-1) repetitions of 30 (=2*5*3) measurements. The average standard deviation of the percent change is 0.7 %.

We display the $T_{1\rho}$ map value for each individual image to show the actual values (as opposed to the difference in values) in each disc. $T_{1\rho}$ values are interesting since lower values are associated with degeneration [7, 8, 5]. Using $T_{1\rho}$ maps and T_2-weighted images, a radiologist identified pathological discs. We compared the amount of fluctuation of $T_{1\rho}$ in healthy versus degenerate discs, using the Student's t-test for this comparison. As a measure of the $T_{1\rho}$ fluctuation, we computed the standard deviation over the repeated images of the $T_{1\rho}$ mean throughout a given disc. We also ran a one-way analysis of variance test to see if the fluctuation depends on the disc type (i.e. L1-L2, L2-L3, etc.)

We analyzed the changes in $T_{1\rho}$ by examining the images showing the local statistically significant changes, looking for patterns in $T_{1\rho}$ that correlate with disc height and the expected compression or relaxation related to the compression method. This analysis method finds local changes and is appropriate given the unique responses expected between individuals and the complexity of the way the spine handles a load. However, due to the pixel-by-pixel nature of the method, it yields too many numbers to cleanly report a finding, and our summaries are subjective. Thus we also analyzed the changes quantitatively over

regions of interest to find significant changes over all discs and all patients. We used four automatically-determined regions of interest: the entire disc, the anterior half, the posterior half, and the central portion in the anterior/posterior direction. The anterior and posterior halves are defined as the midpoint between the anterior / posterior points selected by the user for segmentation. The central portion of the disc is defined as the central third between those same user-selected portions. All regions include the full superior / inferior height of the disc.

To study the $T_{1\rho}$ values in these regions of interest, we first plotted the progression of the mean $T_{1\rho}$ within each of these regions of interest. We then calculated the variation of $T_{1\rho}$ values using the coefficient of variation (CV) measure, which is defined as the ratio of the standard deviation over the mean. The CV of the backpack experiment gives an estimate of the amount of change in $T_{1\rho}$ that is expected during relaxation, which we compared with the $T_{1\rho}$ differences induced by the compression device to see if the device produced more change than what is expected from uninterrupted relaxation. We checked for significant changes in $T_{1\rho}$ within the regions of interest with a paired Student's t-test. The significance value p depends on a the value t, which is calculated as $t = \bar{d}\sqrt{N}/\sigma_{\bar{d}}$, where N is the number of samples, $\bar{d} = \sum_{i=1}^{N}(x_i - y_i)/N$ is the mean of the difference in $T_{1\rho}$ between the two images (x_i and y_i are the mean values of $T_{1\rho}$ in sample i in the first and second images, respectively), and $\sigma_{\bar{d}}$ is the standard deviation of the differences in means over all disc pairs. Since we compared differences between means of pixel values, it is reasonable to approximate their distribution as normal, making the t-test an appropriate statistical test.

We perform the paired t-test on each of the discs separately (i.e. L5-S1, then L4-L5, and so on) to look for changes specific to the disc location as well as on all of the discs combined to look for changes global to all discs. We first compared pairs of images within each of the three experiments to look for effects unique to each protocol. Next we checked to see if the difference between the first and last images of a patient was significant compared with the differences between consecutive images. Finally we tested the significance of the differences between compressed and relaxed images using the compression device compared with the differences between consecutive images in the backpack experiment to see the effect of the compression device beyond pure supine relaxation.

7.2 Results and discussion

7.2.1 Qualitative assessment of registration accuracy

Figures 7.2, 7.3, 7.4, and 7.5 show difference images for each volunteer in the backpack experiment, comparing consecutive images over time. In each case, the post-registration difference image shows good visual alignment. As for how much the registration was needed, we see that with some image pairs, the pre-registration images are already well aligned, such as the 6-10 minute pair in Figure 7.2. In other cases, the volunteer has shifted slightly, which can be seen by thin black and white lines on opposite edges of the intervertebral discs, such as the 10-16 minute pair in Figure 7.2. Other cases show larger initial misalignment, caused by moving the image field of view between acquisitions, as

with the 17-25 minute pair in Figure 7.3.

The difference images for the compression device experiment are shown in Figures 7.6, 7.7, 7.8, and 7.9.The images are labeled indicating the status of the consecutive image pairs. "Relax" means the image was taken without weight applied, and "compress" means the volunteer was supporting weight during image acquisition. "Relax-relax" indicates that both images were taken without changing the compression state, with a few minutes separating the two acquisitions. "Relax-compress" means the compression was applied between the images, with a few minutes separating the two acquisitions. As was the case in the previous experiment, the post-registration alignment is visually very good. Between images without a state change (i.e. "relax-relax" or "compress-compress"), the alignment before registration is also good (though sometimes still slightly misaligned). Application of the weight (i.e. "relax-compress") tends to shift the entire spine inferiorly in addition to changing the relative positions of the vertebrae. To a lesser extent, release of compression (i.e. "compress-relax") tends to shift the spine superiorly as well as alter the relative vertebral positions. In addition, misalignment was caused by shifting the field of view in order to compensate for the compression-induced shifts of the spine for volunteer F in the third row of Figure 7.7 and in all the rows in the continuation of Figure 7.7.

The difference images before and after registration in the third (flexion/extension) experiment are shown in Figure 7.10. The pre-registration mis-alignment is clear when comparing flexed and neutral images and is minimal between the two final neutral images, as is expected. Again, we see a drastic visual improvement in alignment after registration.

Pre-registration   Post-registration   Pre-registration   Post-registration

(a) 6-10 min   (b) 10-16 min

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) 16-25 min   (d) 25-30 min

Pre-registration   Post-registration   Pre-registration   Post-registration

(e) 30-34 min   (f) 34-38 min

Figure 7.2: Volunteer A. This figure shows difference images between pairs of consecutive images before and after registration. The label "xx-yy min" indicates that the two images were taken xx and yy minutes after the 20 lb backpack was removed.

Pre-registration   Post-registration   Pre-registration   Post-registration

(a) 17-25 min

(b) 25-34 min

Pre-registration   Pre-registration

(c) 34-39 min

Figure 7.3: Volunteer B. This figure shows difference images between pairs of consecutive images before and after registration. The label "xx-yy min" indicates that the two images were taken xx and yy minutes after the 20 lb backpack was removed.

Pre-registration   Post-registration   Pre-registration   Post-registration

(a) 12-22 min   (b) 54-65 min

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) 65-68 min   (d) 68-78 min

Figure 7.4: Volunteer C. This figure shows difference images between pairs of consecutive images before and after registration. The label "xx-yy min" indicates that the two images were taken xx and yy minutes after the 20 lb backpack was removed.

Pre-registration   Post-registration   Pre-registration   Post-registration

(a) 5-11 min

(b) 11-17 min

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) 17-22 min

(d) 22-28 min

Pre-registration   Post-registration   Pre-registration   Post-registration
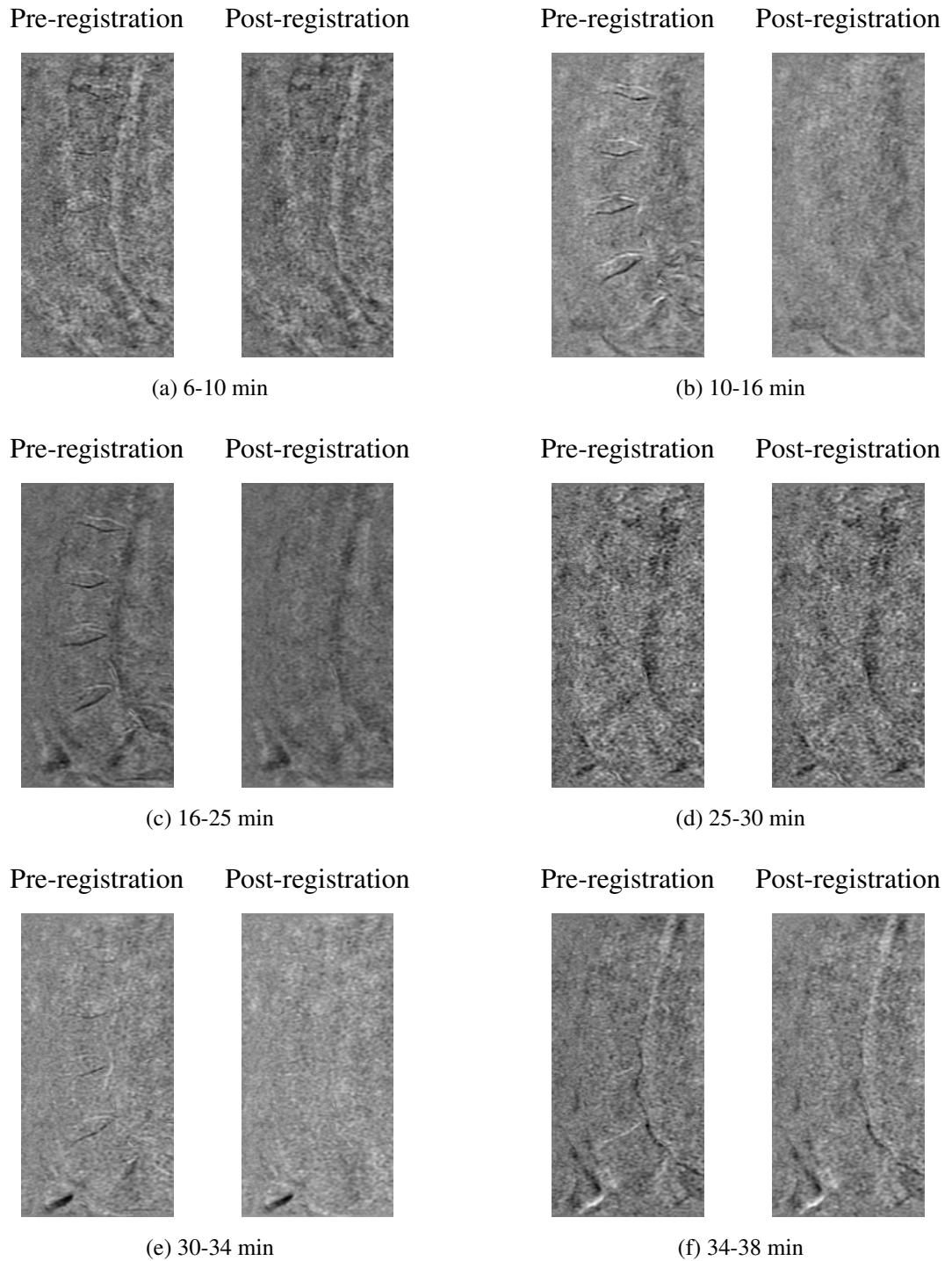
(e) 28-33 min

(f) 33-38 min

Figure 7.5: Volunteer D. This figure shows difference images between pairs of consecutive images before and after registration. The label "xx-yy min" indicates that the two images were taken xx and yy minutes after the 20 lb backpack was removed.

Pre-registration   Post-registration   Pre-registration   Post-registration

(a) relax-compress

(b) compress-compress

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) compress-relax

(d) relax-relax

Pre-registration   Post-registration   Pre-registration   Post-registration

(e) relax-compress

(f) compress-compress

Figure 7.6: Volunteer E. This figure shows differences between consecutive images before and after registration. Weight was applied between the two images in the pairs labeled "relax-compress" and removed between the images in the pairs labeled "compress-relax".

Pre-registration   Post-registration   Pre-registration   Post-registration

(a) relax-compress

(b) compress-compress

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) compress-relax

(d) relax-relax

Pre-registration   Post-registration   Pre-registration   Post-registration
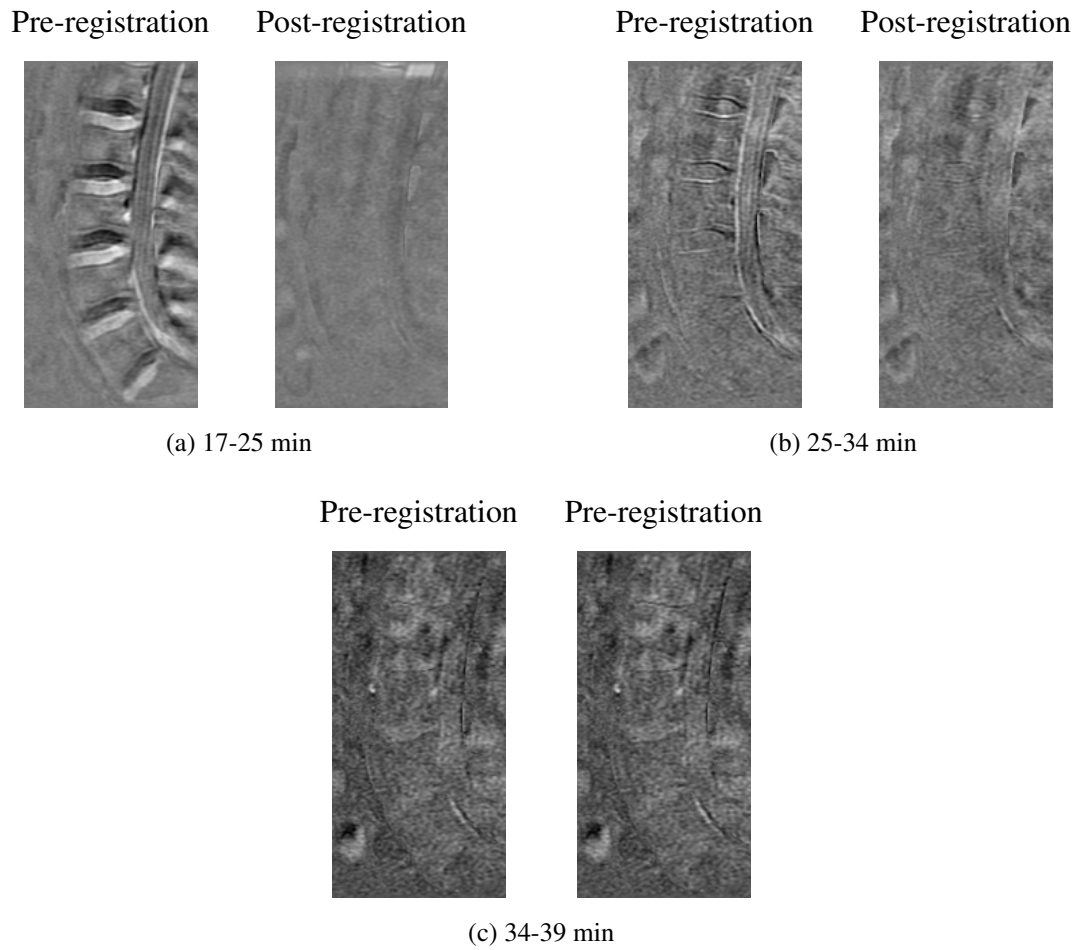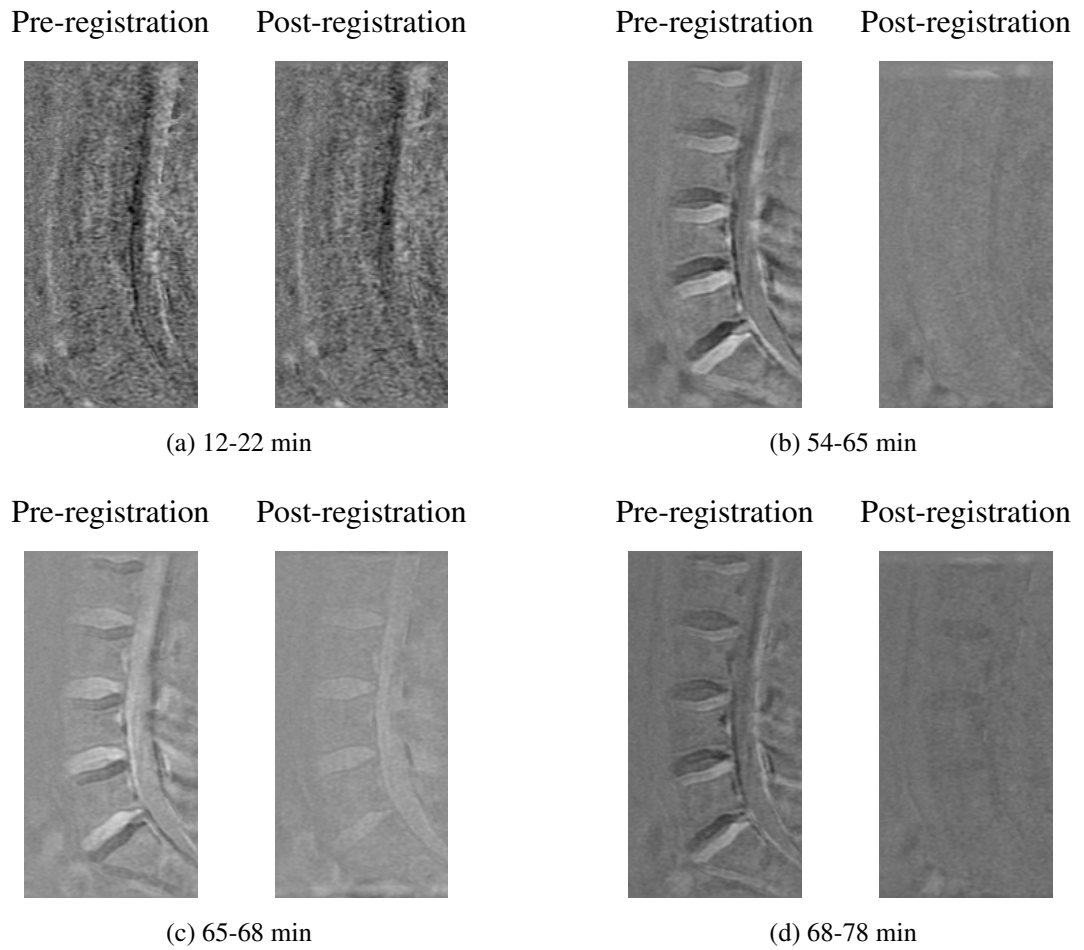

(e) relax-compress

(f) compress-compress

Figure 7.7: Volunteer F. This figure shows differences between consecutive images before and after registration. The volunteer was supporting weight during the images labeled "compress" and was lying without weight during the images labeled "relax".

Pre-registration Post-registration

Pre-registration Post-registration

(a) relax-relax

(b) relax-compress

Pre-registration Post-registration

Pre-registration Post-registration

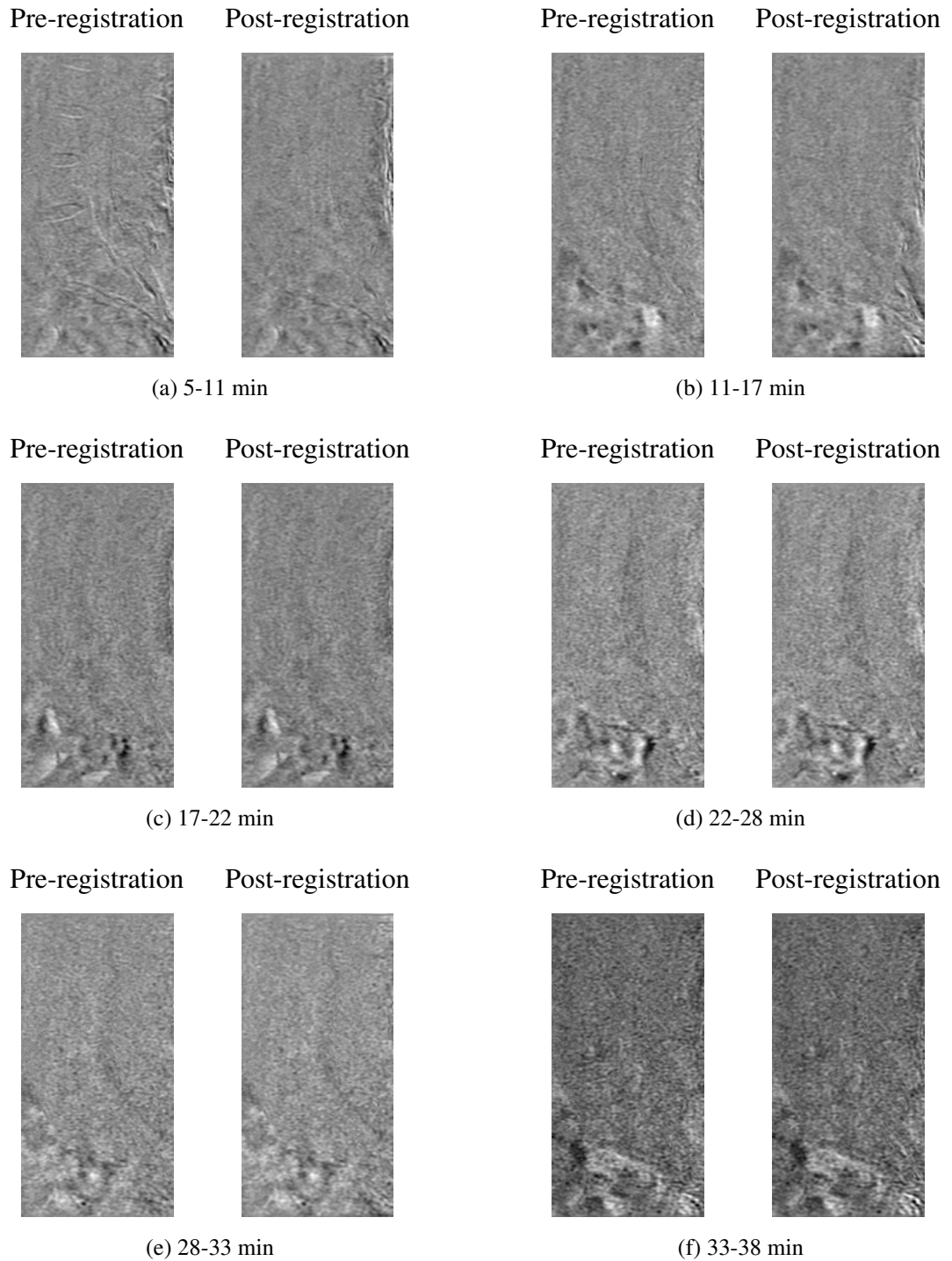(c) compress-relax

(d) relax-relax

Pre-registration Post-registration

(e) relax-compress

Figure 7.8: Volunteer G. This figure shows differences between consecutive images before and after registration. Weight was applied between the two images in the pairs labeled "relax-compress" and removed between the images in the pairs labeled "compress-relax".

Pre-registration Post-registration

(a) compress-relax

Pre-registration Post-registration

(b) relax-relax

Pre-registration Post-registration

(c) relax-compress

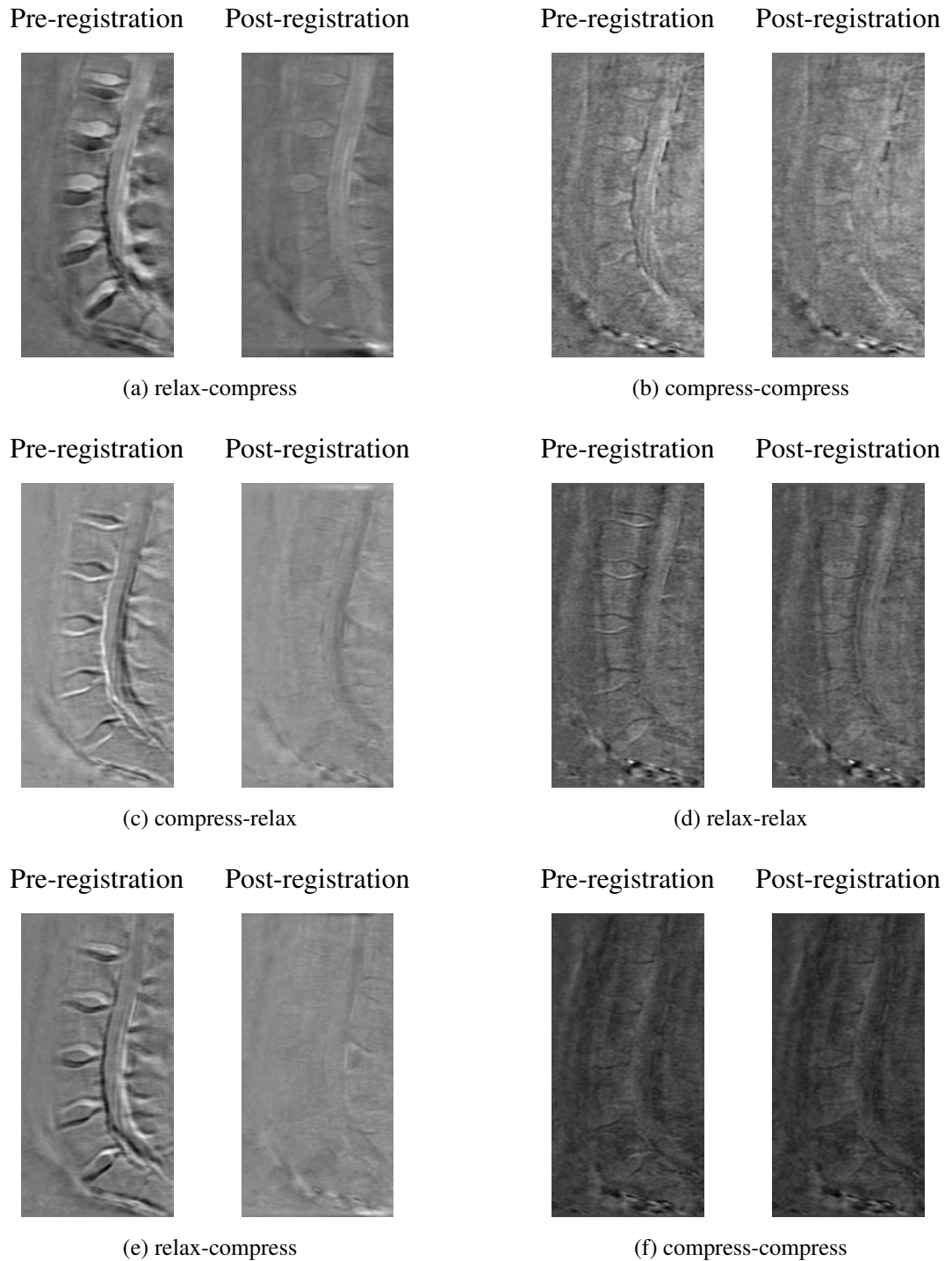Pre-registration Post-registration
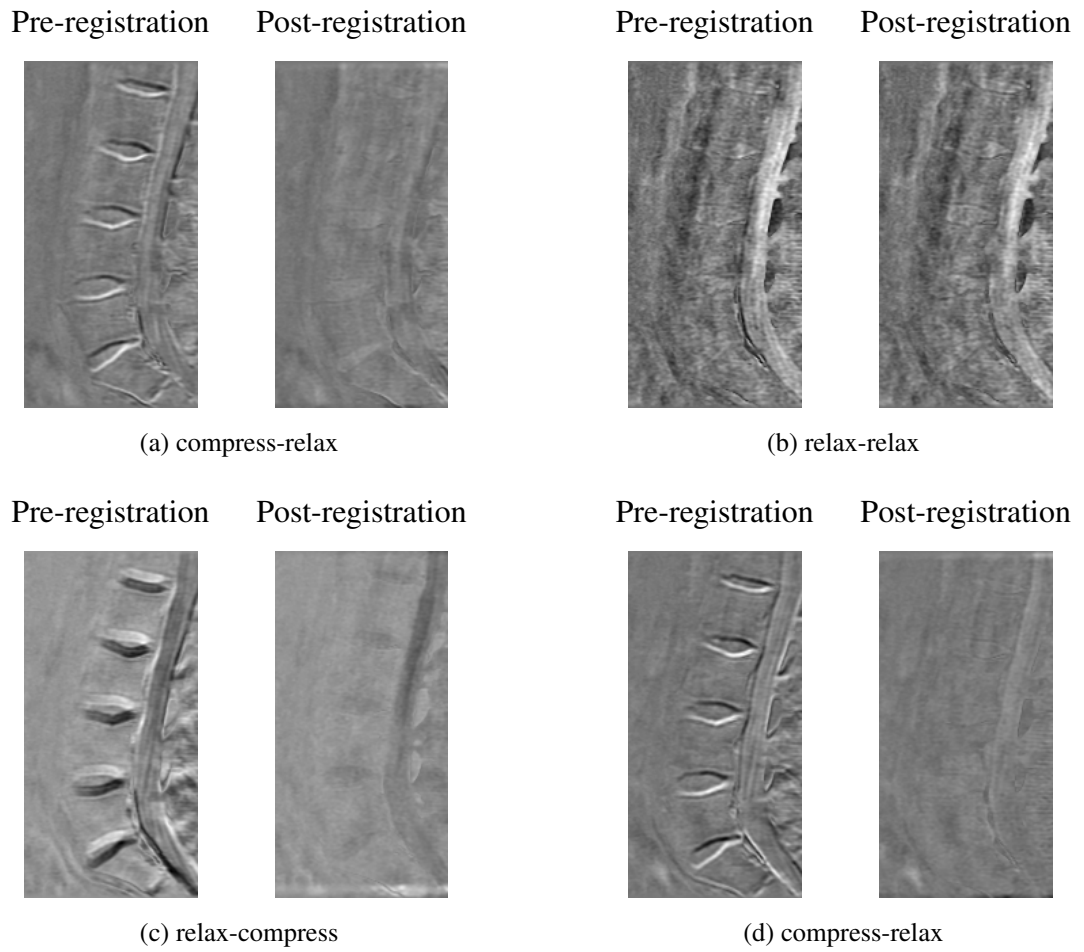
(d) compress-relax

Figure 7.9: Volunteer H. This figure shows differences between consecutive images before and after registration. Weight was applied between the two images in the pairs labeled "relax-compress" and removed between the images in the pairs labeled "compress-relax".
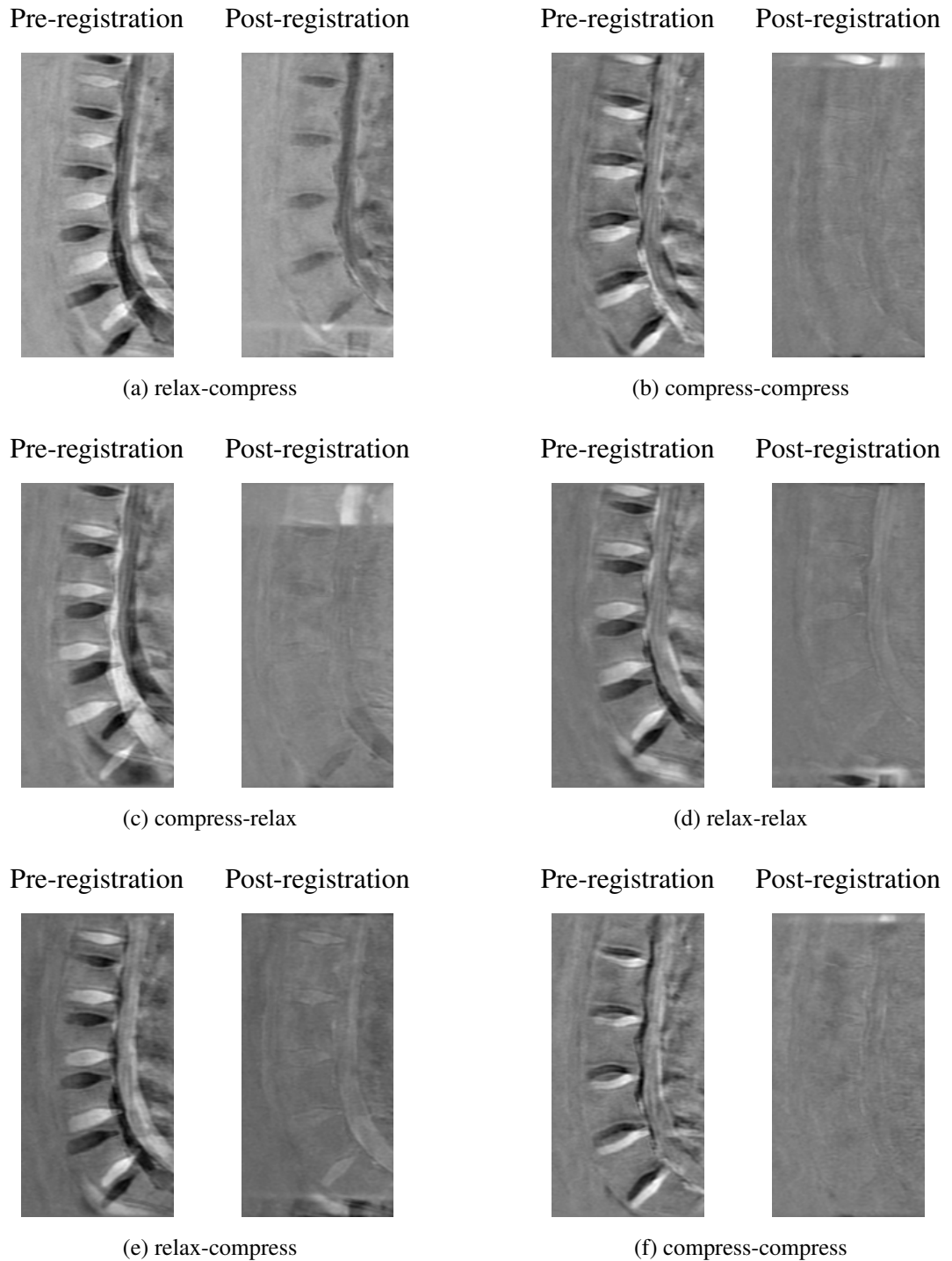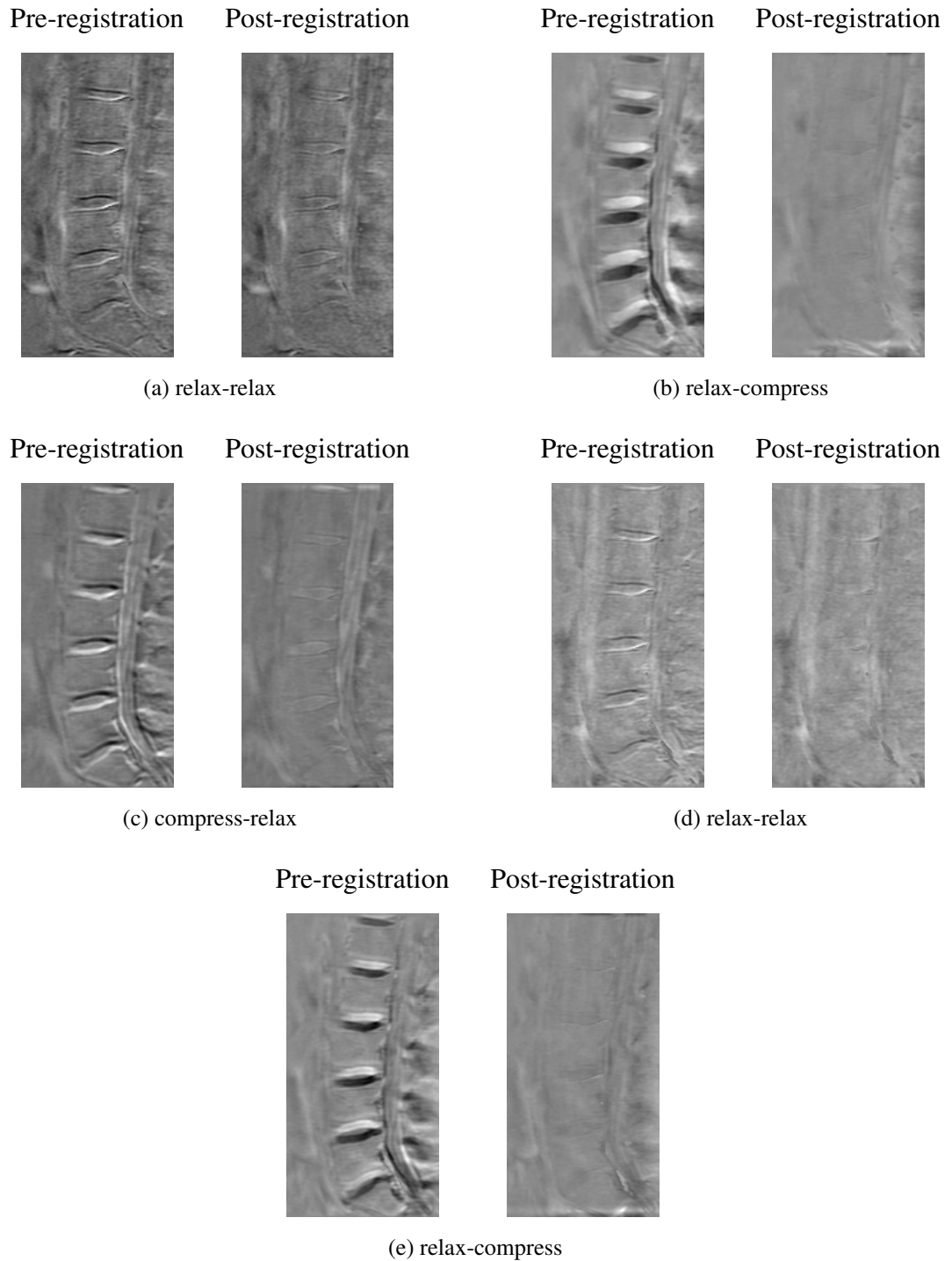
Pre-registration   Post-registration   Pre-registration   Post-registration

(a) neutral-flexed

(b) flexed-neutral

Pre-registration   Post-registration   Pre-registration   Post-registration

(c) neutral-flexed

(d) flexed-neutral

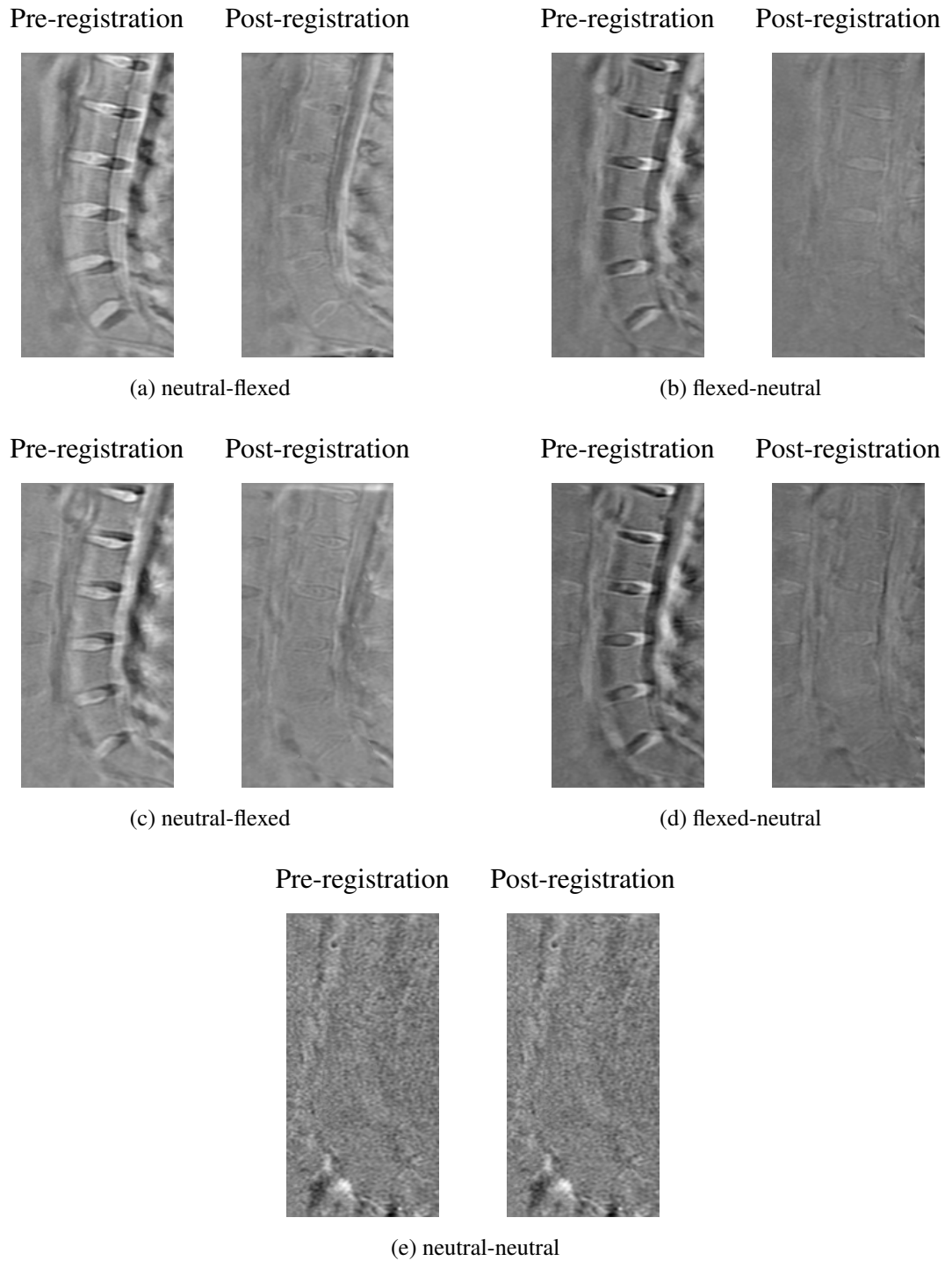Pre-registration   Post-registration

(e) neutral-neutral

Figure 7.10: Volunteer I. This figure shows differences between consecutive images. The volunteer was relaxing with knees straight for the neutral images and rounding his/her back for the flexed images.

7.2.2 Assessment of local changes in $T_{1\rho}$

The $T_{1\rho}$ values progressing through time during the backpack experiment are shown in Figures 7.11, 7.13, 7.15, and 7.17. Statistically significant local changes in mean $T_{1\rho}$ for this experiment are colored according to the amount of change, shown in Figures 7.12, 7.14, 7.16, and 7.18. The rows of numbers alongside each image gives the percent change in disc height. Each row represents one vertebra and the A (anterior), C (central), and P (posterior) columns indicate the location of the point pair. Image labels are in the form xx-yy min, meaning the first image was taken xx minutes and the second image taken yy minutes after the 20 lb backpack was removed.

For all volunteers (Figures 7.12, 7.14, 7.16, and 7.18), we note that there are statistically significant changes in $T_{1\rho}$ between each pair of consecutive images with magnitudes up to 75% of the map value. However, the changes do not go in a consistent direction as time passes. The disc height measurements do not go in a uniform direction as time passes either, nor is there a correlation between the directions of $T_{1\rho}$ and disc height changes. Vertebral motion due to decompression from resting for one hour may be too small to measure relative to the normal slight movements of a living volunteer from breathing, heartbeat, and muscular inability to hold completely still. Perhaps visible differences in $T_{1\rho}$ and disc height between standing and lying occur more quickly than we were able to capture since the MRI setup and landmarking takes several minutes before the first images are captured, and the sequence itself takes more than two minutes. On the other end of the time spectrum, it is known that after many hours of sleep, the spine measurably decompresses, so perhaps

(a) 6 min   (b) 10 min   (c) 16 min   (d) 25 min

(e) 30 min   (f) 34 min   (g) 38 min

180 ms

0 ms

Figure 7.11: $T_{1\rho}$ maps for volunteer A as time progresses. $T_{1\rho}$ parameter values are overlaid on an anatomical image. Minutes indicate the length of time since the 20 lb backpack was removed.

6-10 min	A	C	P
	3.8	-0.1	-5.9
	-3.1	-0.7	2.4
	0.6	-0.5	-3.9
	-1.9	-1.2	-0.5
	1.0	0.4	-0.8

10-16 min	A	C	P
	-5.4	1.8	16.4
	5.0	1.6	-1.9
	-2.0	-0.2	1.5
	-2.0	-0.8	1.9
	3.3	1.4	-2.0

16-25 min	A	C	P
	6.3	-1.6	-14.5
	0.2	0.3	1.3
	0.4	-0.8	-2.8
	1.2	0.6	-0.4
	-2.4	-0.8	3.4

25-30 min	A	C	P
	2.9	0.6	-2.0
	-3.6	-0.7	3.2
	5.5	2.4	1.3
	-1.0	-0.1	1.9
	-1.9	-1.3	-1.7

30-34 min	A	C	P
	-6.2	-0.9	7.4
	3.1	1.2	-0.6
	-1.8	0.2	3.3
	1.4	0.6	-1.1
	2.6	0.8	-1.0

34-38 min	A	C	P
	9.3	0.6	-11.1
	-3.3	-0.8	2.6
	-3.9	-1.4	-1.5
	1.2	-0.3	-3.7
	-1.2	0.6	4.1

60 ms / -60 ms

The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.

Figure 7.12: Statistically significant changes in $T_{1\rho}$ for Volunteer A are colored according to the amount of change.

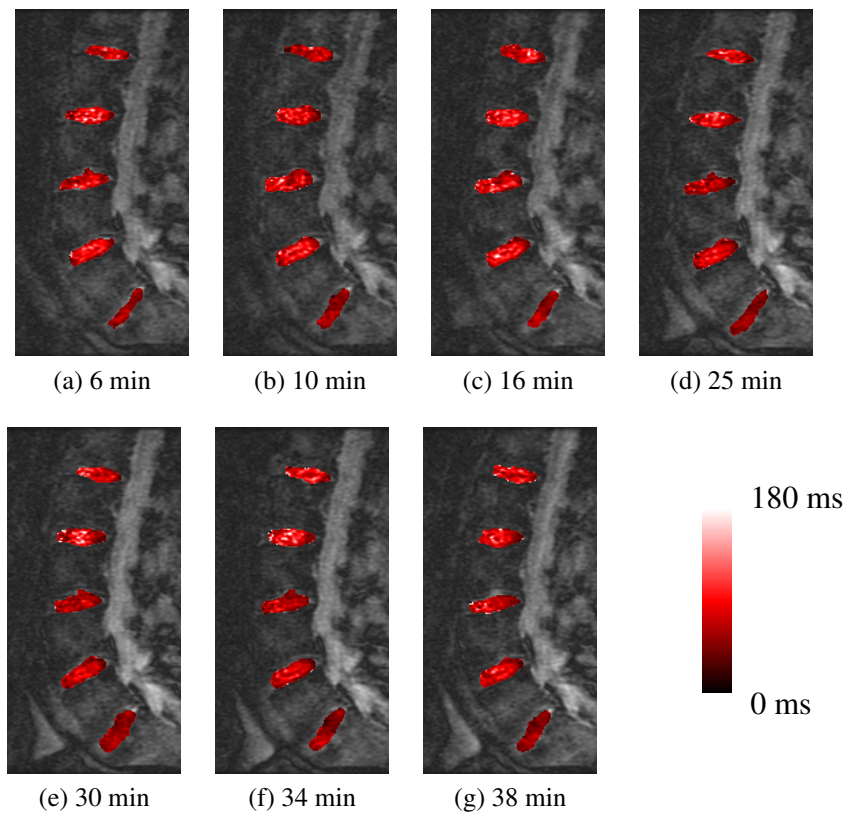(a) 17 min   (b) 25 min   (c) 34 min   (d) 49 min

Figure 7.13: $T_{1\rho}$ maps for volunteer B as time progresses. $T_{1\rho}$ parameter values are overlaid on an anatomical image. Minutes indicate the length of time since the 20 lb backpack was removed.
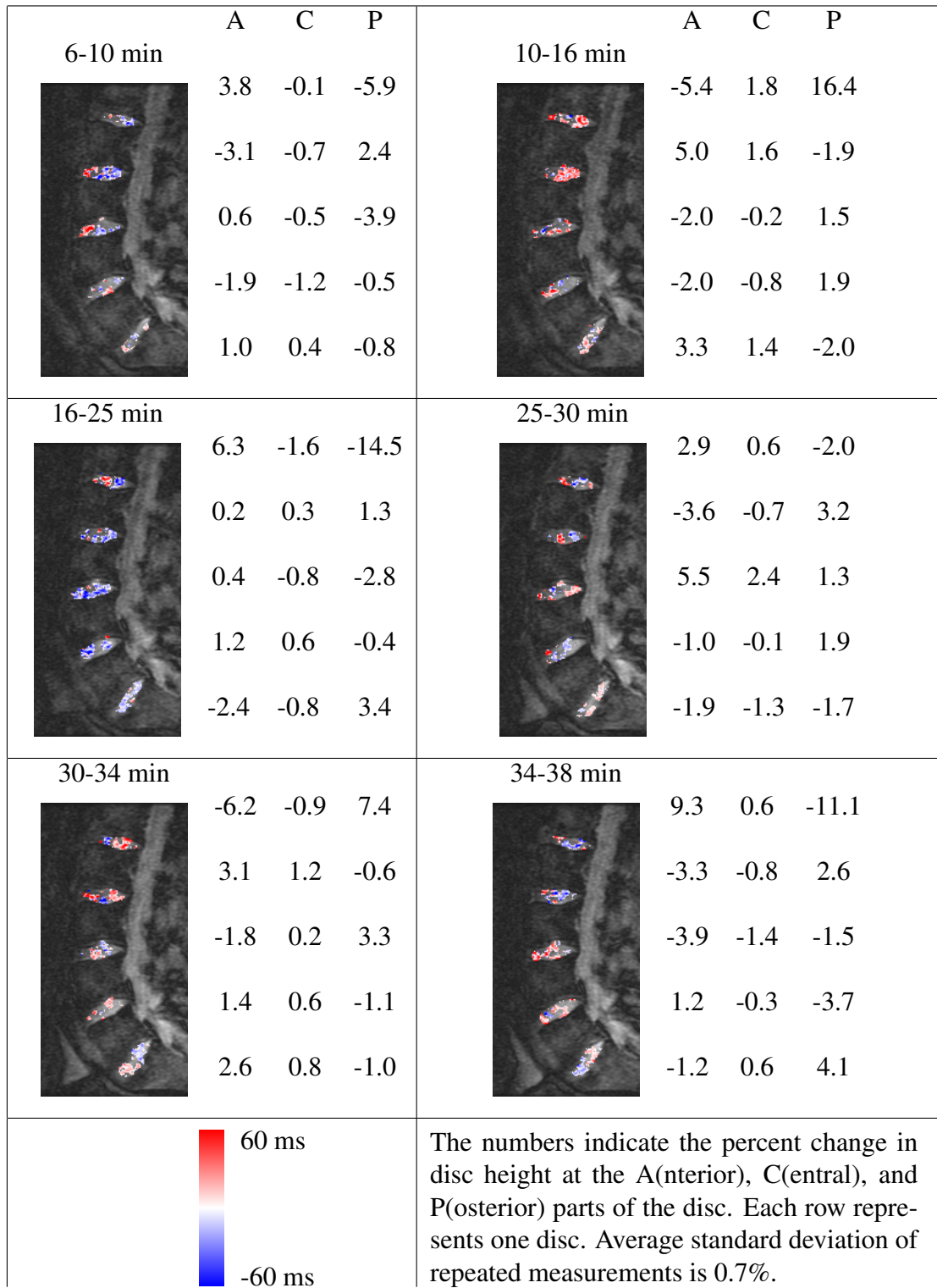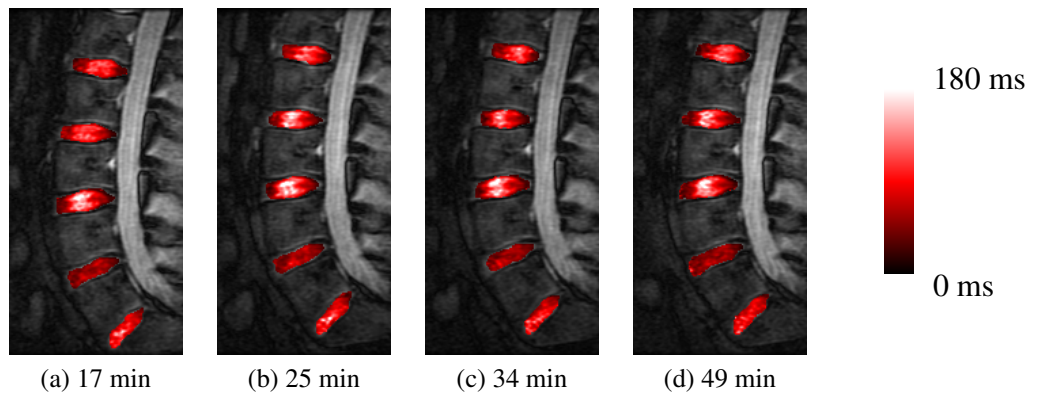
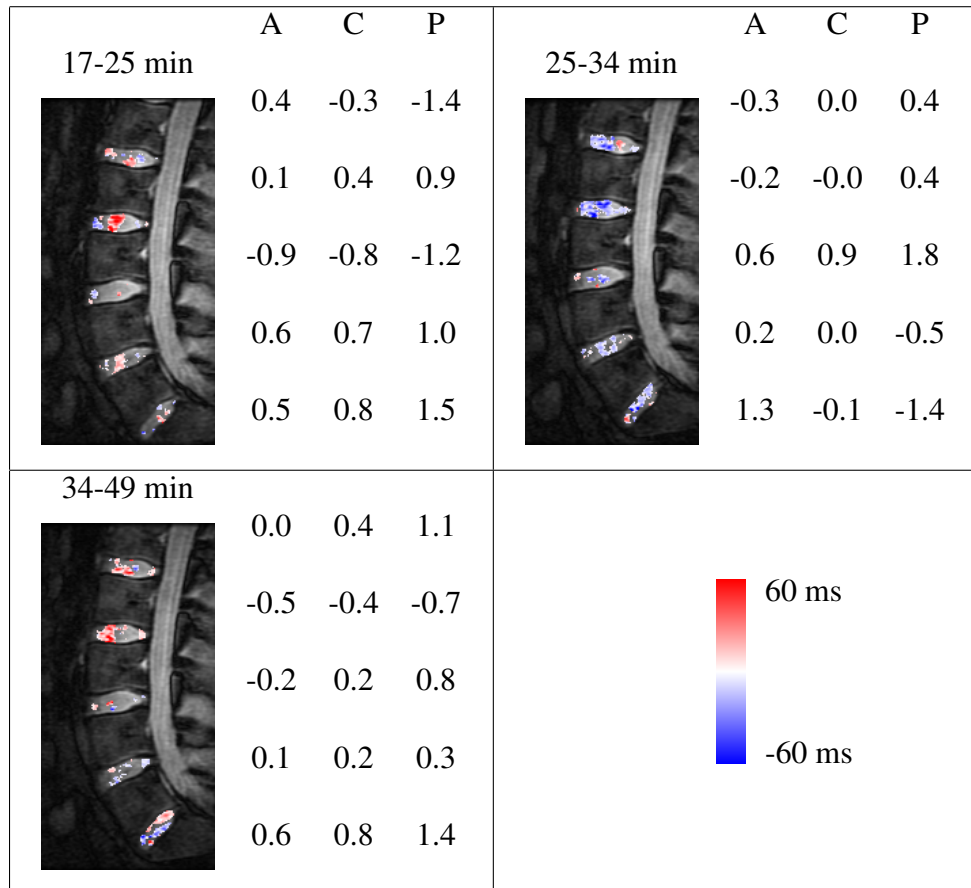|	A	C	P		A	C	P
17-25 min				**25-34 min**			
	0.4	-0.3	-1.4		-0.3	0.0	0.4
	0.1	0.4	0.9		-0.2	-0.0	0.4
	-0.9	-0.8	-1.2		0.6	0.9	1.8
	0.6	0.7	1.0		0.2	0.0	-0.5
	0.5	0.8	1.5		1.3	-0.1	-1.4
34-49 min							
	0.0	0.4	1.1				
	-0.5	-0.4	-0.7			60 ms	
	-0.2	0.2	0.8				
	0.1	0.2	0.3			-60 ms	
	0.6	0.8	1.4				

Figure 7.14: Statistically significant changes in $T_{1\rho}$ for Volunteer B are colored according to the amount of change. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.

(a) 12 min   (b) 22 min   (c) 54 min
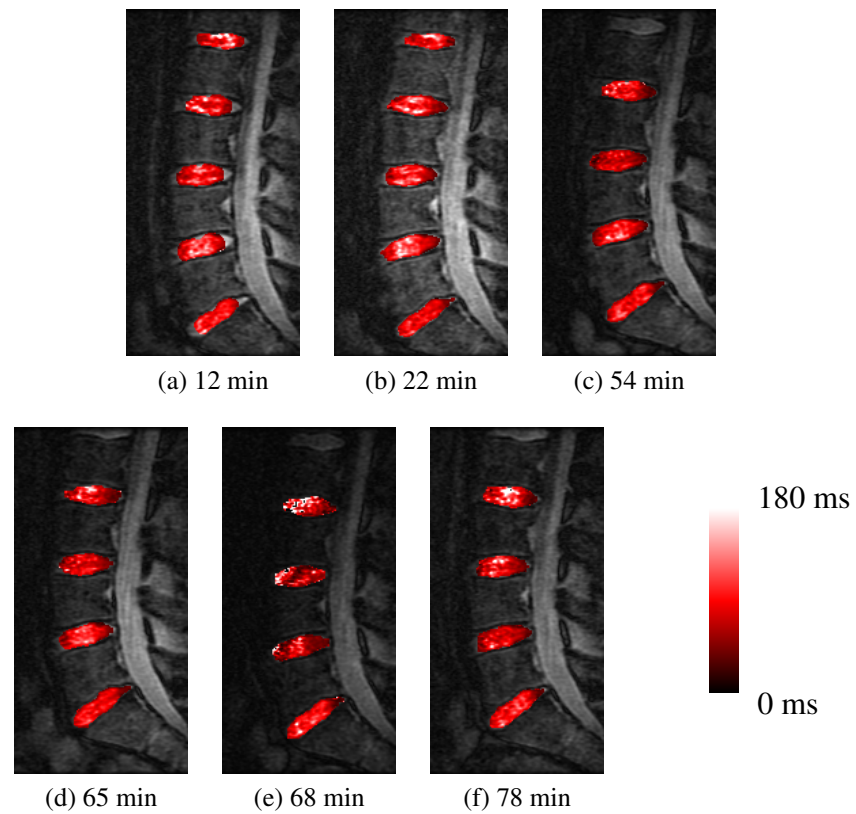
(d) 65 min   (e) 68 min   (f) 78 min

Figure 7.15: $T_{1\rho}$ maps for volunteer C as time progresses. $T_{1\rho}$ parameter values are overlaid on an anatomical image. Minutes indicate the length of time since the 20 lb backpack was removed.
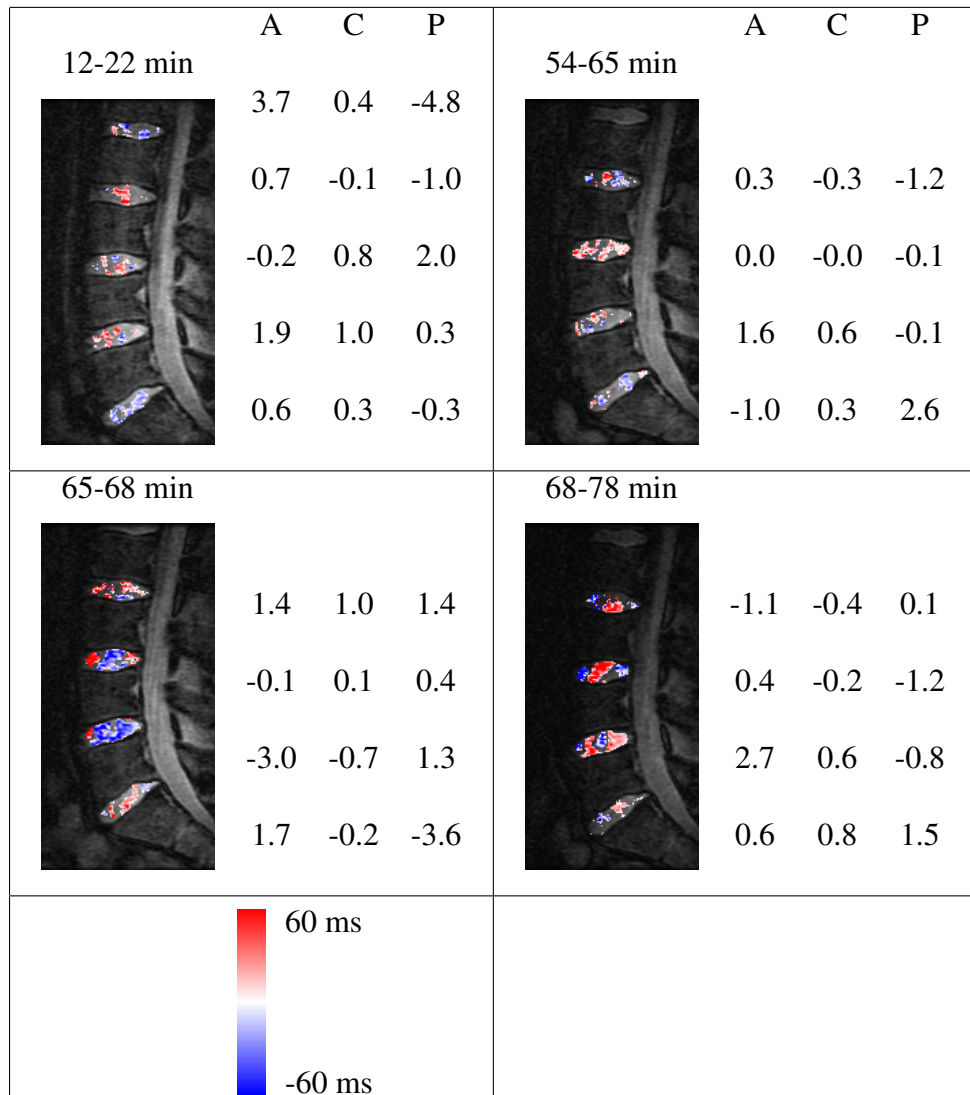
	A	C	P			A	C	P
12-22 min					54-65 min			
	3.7	0.4	-4.8					
	0.7	-0.1	-1.0			0.3	-0.3	-1.2
	-0.2	0.8	2.0			0.0	-0.0	-0.1
	1.9	1.0	0.3			1.6	0.6	-0.1
	0.6	0.3	-0.3			-1.0	0.3	2.6
65-68 min					68-78 min			
	1.4	1.0	1.4			-1.1	-0.4	0.1
	-0.1	0.1	0.4			0.4	-0.2	-1.2
	-3.0	-0.7	1.3			2.7	0.6	-0.8
	1.7	-0.2	-3.6			0.6	0.8	1.5

60 ms

-60 ms

Figure 7.16: Statistically significant changes in $T_{1\rho}$ for Volunteer C are colored according to the amount of change. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.
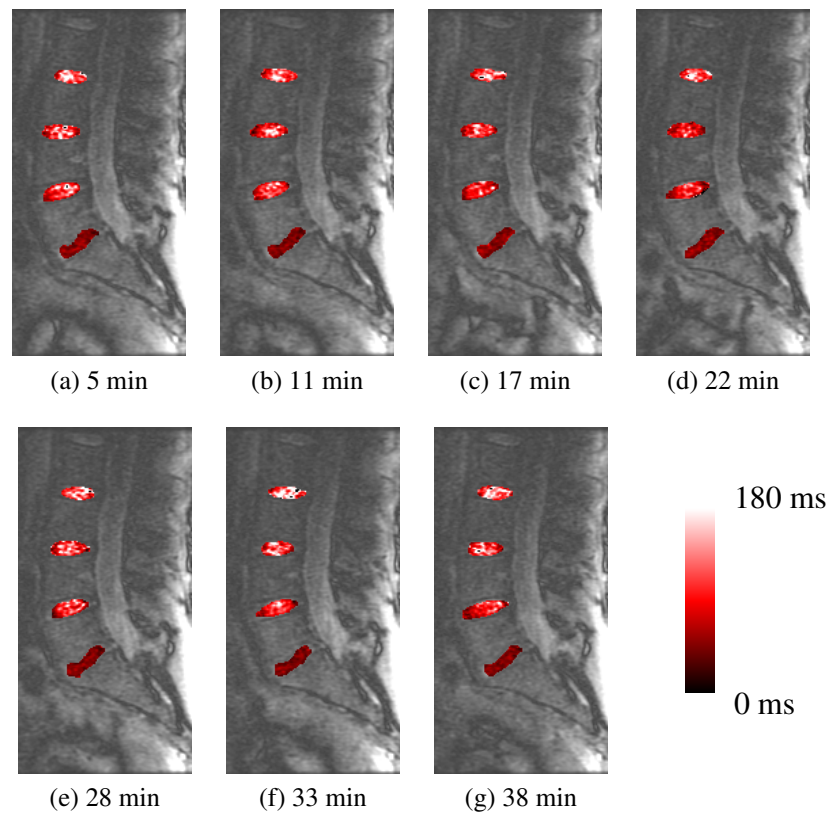
(a) 5 min   (b) 11 min   (c) 17 min   (d) 22 min

(e) 28 min   (f) 33 min   (g) 38 min

Figure 7.17: $T_{1\rho}$ maps for volunteer D as time progresses. $T_{1\rho}$ parameter values are overlaid on an anatomical image. Minutes indicate the length of time since the 20 lb backpack was removed.

	A	C	P			A	C	P
5-11 min					**11-17 min**			
	-1.1	1.1	4.2			2.8	-0.6	-4.4
	0.9	0.7	0.4			-0.6	0.6	2.5
	-0.9	-0.3	-0.2			2.3	1.0	0.6
	1.1	1.0	1.0			-1.4	-1.3	-1.2
17-22 min					**22-28 min**			
	-4.2	-0.5	2.7			3.0	1.5	2.2
	0.9	1.0	1.6			-0.4	-1.3	-2.7
	0.5	-0.1	-0.7			-0.4	0.1	0.6
	-1.1	-0.5	0.3			-3.7	-2.6	-1.2
28-33 min					**33-38 min**			
	-0.5	-0.9	-2.3			-1.3	0.2	1.7
	0.8	0.6	0.8			2.0	0.5	-1.2
	0.8	0.8	1.6			1.9	0.4	0.4
	3.6	3.5	4.1			-2.6	-2.1	-1.5

60 ms

-60 ms

The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.
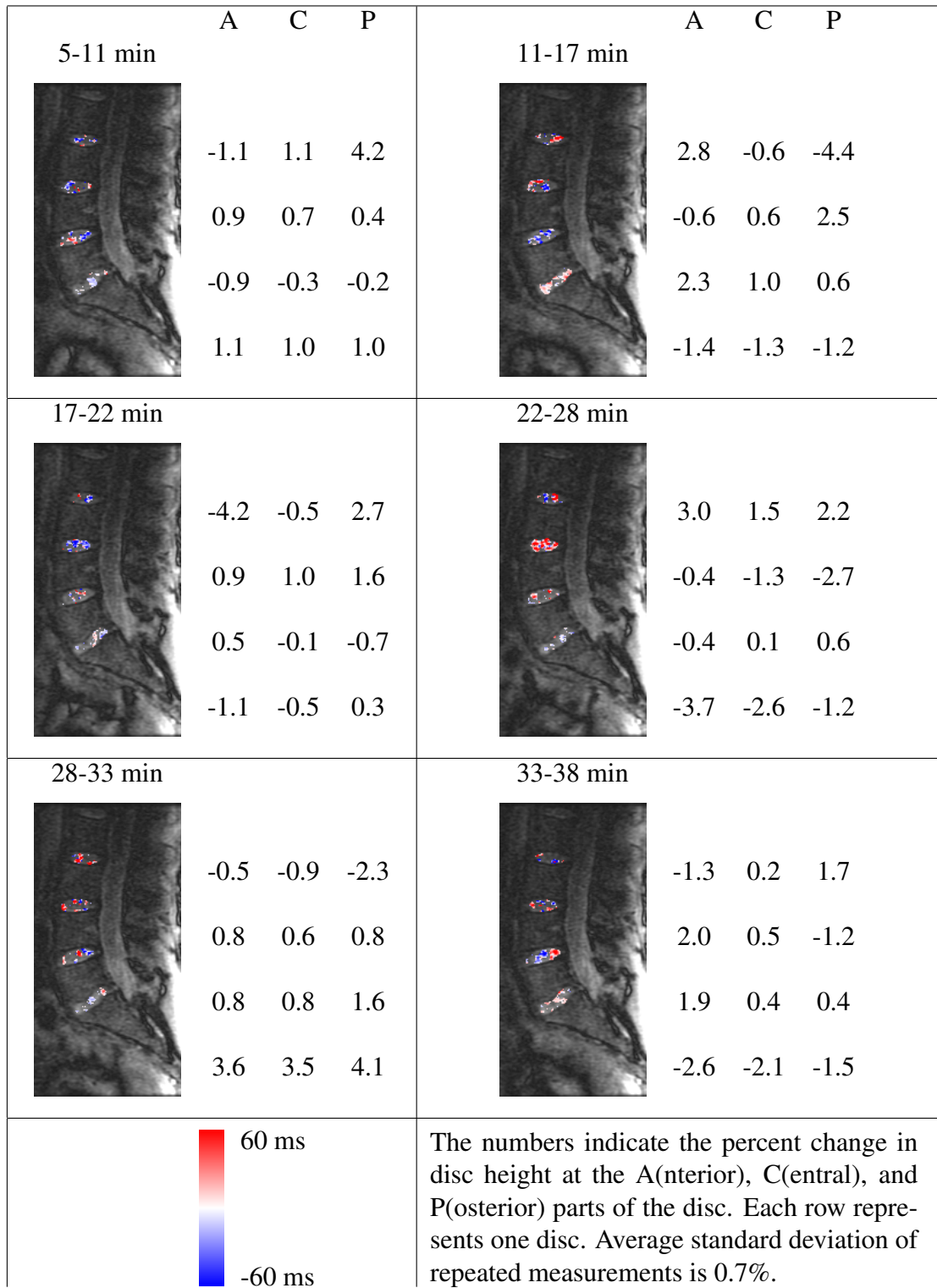
Figure 7.18: Statistically significant changes in $T_{1\rho}$ for Volunteer D are colored according to the amount of change.

an hour was not long enough to show a discernible pattern.

On the other hand, there is a consistent relationship between the magnitude of $T_{1\rho}$ change and the magnitude of the $T_{1\rho}$ map values. The more inferior discs, especially L5-S1, have less variation in $T_{1\rho}$ than other discs and also have markedly lower $T_{1\rho}$ values, which are associated with degeneration. The first question we ask is whether this decrease in variation is an artifact from noise or if it is a true pattern. A lower $T_{1\rho}$ value means that the exponential $T_{1\rho}$ relaxation curve decreases faster, so more of the measurements will be lower. We expect some level of noise in each measurement. Assuming the noise level is approximately the same for each measurement point, the relative amount of noise is larger for lower measurements, meaning we expect the noise to affect the lower $T_{1\rho}$ values more, which is the opposite of what we observed. In addition, discs with lower $T_{1\rho}$ values often have lower initial measurements at spin lock time of 0ms. Once again, we would expect these lower values to be *more* sensitive to noise, which is the opposite of what we are seeing. This analysis suggests the relationship between less variation and low values is a true observation. Since low values are associated with degeneration, the reduced amount of $T_{1\rho}$ change agrees with other studies showing less T_1 and T_2 variation in degenerate discs compared with healthy discs between morning and evening measurements [19, 27, 28]. Another factor contributing to less change in the inferior discs may be anatomical positioning. The arrangement of the sacrum, pelvis, and fifth lumbar vertebra (L5) restricts the movement of L5 compared with the other vertebrae. The curvature of the spine is different at this level as well, meaning that the pressures on the disc are different, possibly

decreasing the amount of decompression that occurs while the volunteer is resting.

When analyzing the $T_{1\rho}$ changes, we must recognize that we did not use prospective registration to select the sagittal slice from image to image. Since the volunteer was lying still, we expected minimal movement, which was indeed the case as the pre-registration difference images in Figures 7.2, 7.3, 7.4, and 7.5 all show minimal misalignment, with the exception of 7.3a and 7.4b-(d), where the larger misalignments were due to repositioning the field of view. The pre-registration misalignment is less than about one in-plane pixel (0.78mm). We have shown in Section 6.3 that this amount of in-plane misalignment affects the outcome of the statistical analysis and thus must be corrected with registration. On the other hand, since the sagittal slice thickness is 8mm, we expect the out-of-plane misalignment to be relatively small with minimal effect. Since prospective registration takes up to 10 minutes and we were looking for high temporal resolution for dynamic analysis, we opted not to take the time for prospective registration. However, the actual effect of potential out-of-plane misalignment was not tested and should therefore be considered when interpreting these results. None of the other examples in this chapter used prospective registration either, for these same reasons.

We now show the $T_{1\rho}$ maps for the compression device experiment in Figures 7.19, 7.21, 7.23, and 7.25. Statistically significant differences in $T_{1\rho}$ and changes in disc height are displayed in Figures 7.20, 7.22, 7.24, and 7.26.

One thing to keep in mind when studying the differences between compression and relaxation is that these processes are dynamic and take time. For example, cadaver discs have

(a) Relax (b) Maintain relaxation (c) Compress (d) Maintain compression
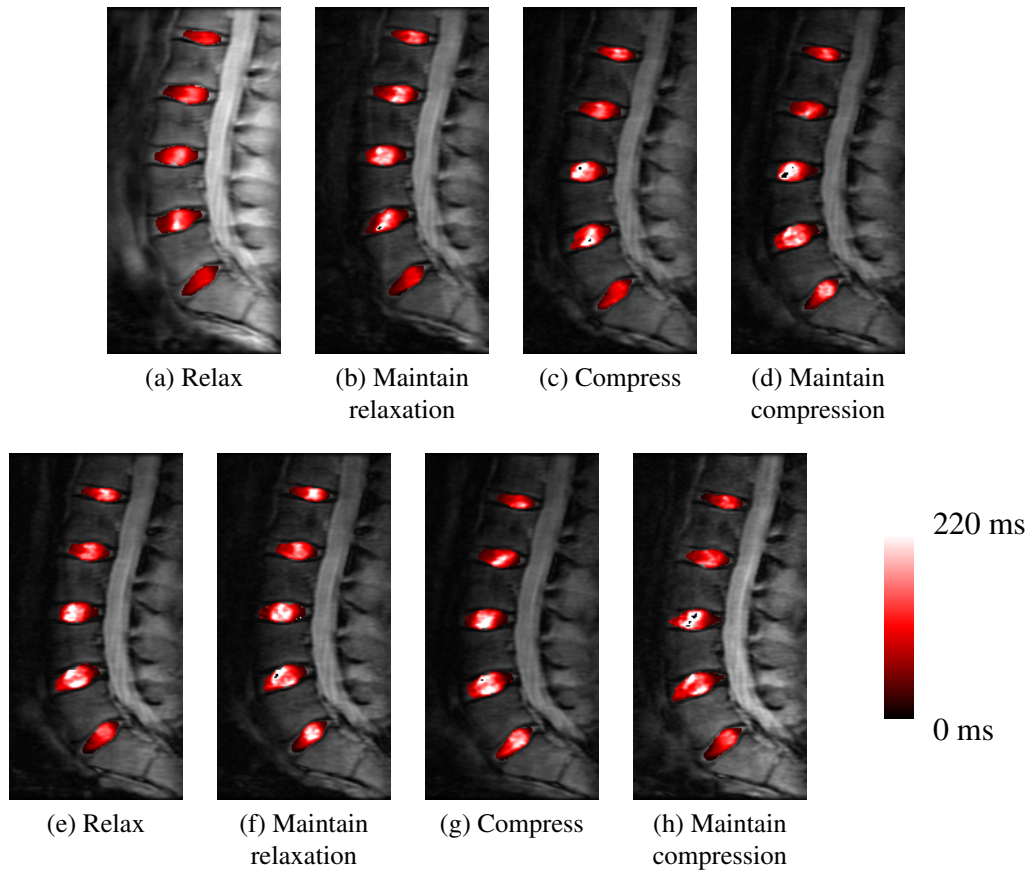
220 ms

0 ms

(e) Relax (f) Maintain relaxation (g) Compress (h) Maintain compression
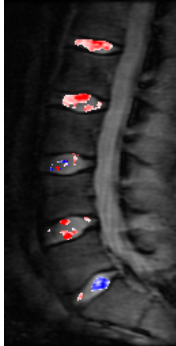
Figure 7.19: $T_{1\rho}$ maps for volunteer E
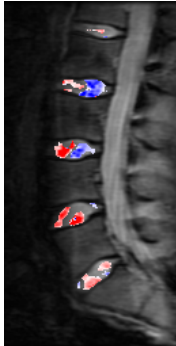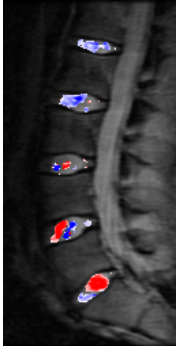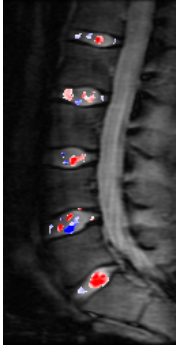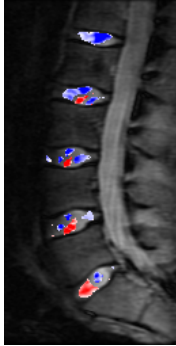
	A	C	P		A	C	P
relax-relax				relax-compress			
	2.0	0.8	0.1		-2.0	-0.1	0.8
	-0.4	-0.0	0.5		2.0	-0.1	-2.1
	0.8	0.4	0.0		4.6	-0.4	-5.6
	0.3	0.7	1.4		2.5	0.3	-2.5
	1.4	0.2	-2.0		-4.6	-1.1	5.2
compress-compress				compress-relax			
	-0.1	0.4	1.4		2.1	0.7	-0.3
	1.1	-0.2	-1.7		-0.5	0.6	1.9
	0.1	0.3	0.7		-5.6	-1.1	3.2
	0.7	0.2	-0.2		-0.8	-0.4	-0.1
	0.6	0.8	1.0		2.5	0.9	-1.8
relax-relax				relax-compress			
	-1.0	-0.6	-0.7		-3.4	-0.7	1.6
	-2.1	-0.8	0.1		4.2	1.2	-1.0
	2.4	1.2	0.5		3.7	-0.0	-3.7
	0.8	1.0	2.1		-1.2	-0.7	-0.5
	-0.8	-0.7	-0.5		-0.7	-0.8	-1.2

Figure 7.20: Statistically significant changes in $T_{1\rho}$ for Volunteer E are colored according to the amount of change. See continuation on the next page.
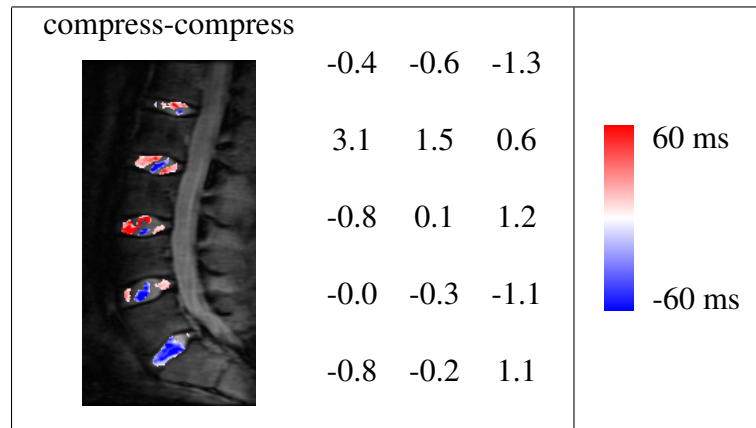
Figure 7.20: Statistically significant changes in $T_{1\rho}$ for Volunteer E, continued. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.

been shown to continue to decrease in height after four hours under a static load [9]. Thus we would expect the disc compression process to be complicated in a living volunteer with alternating application and release of compression every 10 minutes. Relaxation / compression will not be complete when the load is toggled. Also complicating the analysis, the distribution of the weight and the way the body supports it is different between upright standing (i.e. the pre-scan weight-bearing position) and the compression device. In addition, the way the spine takes the load (and consequently the distribution of compression over the discs) will vary from person to person, from weight application to weight application of the same person, and perhaps even over the course of holding the weight if the volunteers consciously or subconsciously adjust how they are supporting the weight.

Keeping these complexities in mind, we survey the $T_{1\rho}$ comparisons for Volunteers E-H. Volunteers E, F, and G (Figures 7.20, 7.22, 7.24) all began the study relaxing, followed by resisting weight, relaxing, and resisting again. Volunteer H (Figure 7.26) had already

(a) Relax (b) Maintain relaxation (c) Compress (d) Maintain compression
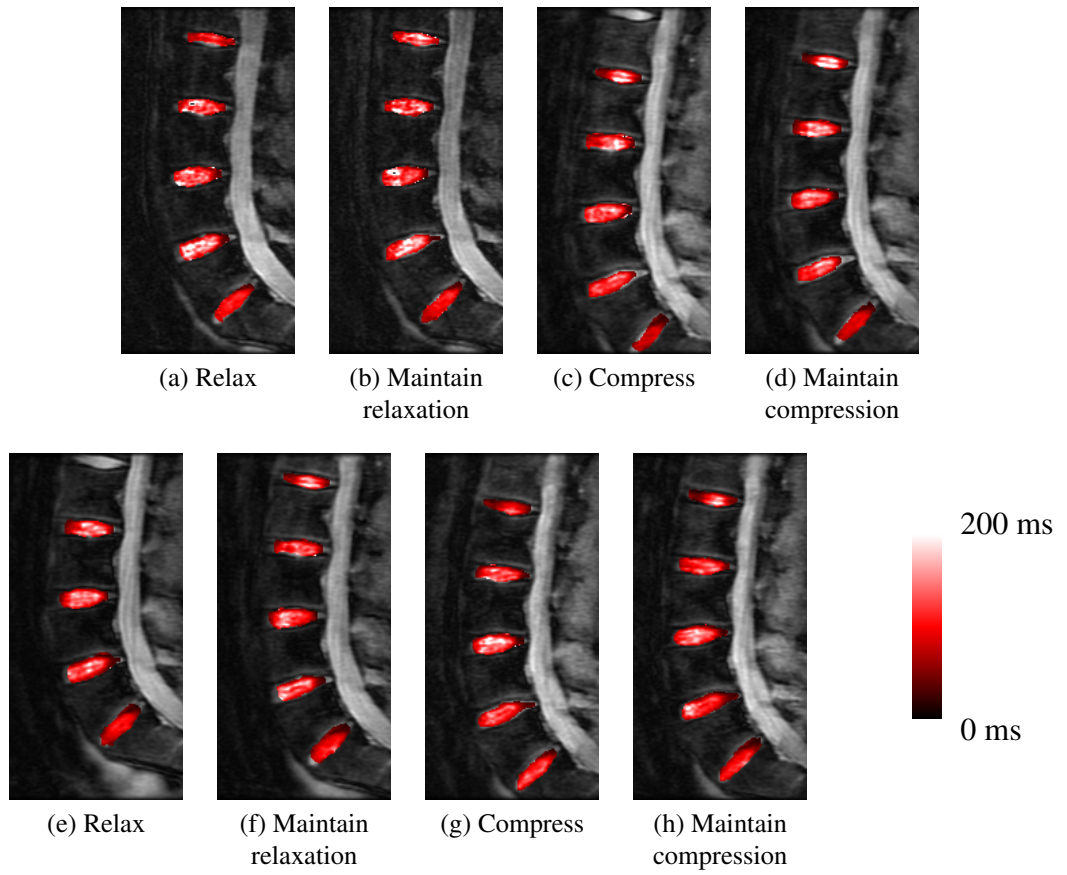
(e) Relax (f) Maintain relaxation (g) Compress (h) Maintain compression
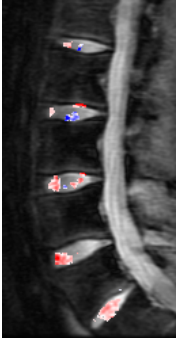
Figure 7.21: $T_{1\rho}$ maps for volunteer F

	A	C	P		A	C	P
relax-relax				**relax-compress**			
	-0.8	-0.3	-0.3		4.4	1.2	-0.9
	1.8	0.5	-1.0		1.1	0.4	-0.4
	-0.1	0.2	0.9		6.0	1.1	-7.5
	0.0	-0.1	-0.6		-1.3	-0.0	2.9
	-0.6	0.0	0.9		-6.8	-2.5	0.4
compress-compress				**compress-relax**			
	-2.8	-1.4	-1.8				
	1.6	-0.1	-2.8		-1.9	0.0	3.4
	-2.7	-1.1	0.4		-0.4	0.6	4.7
	0.0	-0.3	-1.2		-2.5	-1.0	1.2
	2.5	1.6	1.6		4.3	0.9	-1.8
relax-relax				**relax-compress**			
					-7.2	3.4	24.0
	-1.8	-0.2	1.8		2.3	0.7	-1.0
	-0.2	0.7	4.3		2.1	-1.1	-10.3
	1.6	0.8	-0.6		-0.2	0.2	1.5
	2.0	0.4	-0.8		-6.6	-2.1	1.0

Figure 7.22: Statistically significant changes in $T_{1\rho}$ for Volunteer F are colored according to the amount of change. See continuation on the next page.

Figure 7.22: Statistically significant changes in $T_{1\rho}$ for Volunteer F, continued. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.

been lying in the scanner for about 10 minutes for other scans, then began the study resisting weight, followed by relaxing, resisting, and relaxing again.

The dominant color in the $T_{1\rho}$ comparison images appears to be red, representing a $T_{1\rho}$ increase, from start to finish for each of these volunteers, especially E, F, and G. There is likely an underlying decompression over time due to the brevity of the compression sessions, so we suggest this observation points to a correlation between decompression and an increase in $T_{1\rho}$.

Looking in more detail, we begin with the first pair of images for Volunteer E. Both images were taken during relaxation (relax-relax). Most of the measurements of disc height change are positive, indicating height increase, which is reasonable since the discs are decompressing. There are many bright red areas, meaning increasing $T_{1\rho}$, which agrees with our earlier observation. The magnitude of the change is much lower in the L5-S1 disc than the other discs, which agrees with the findings in the backpack experiment. For the second

(a) Relax   (b) Relax   (c) Compress

180 ms

0 ms

(d) Relax   (e) Maintain relaxation   (f) Compress
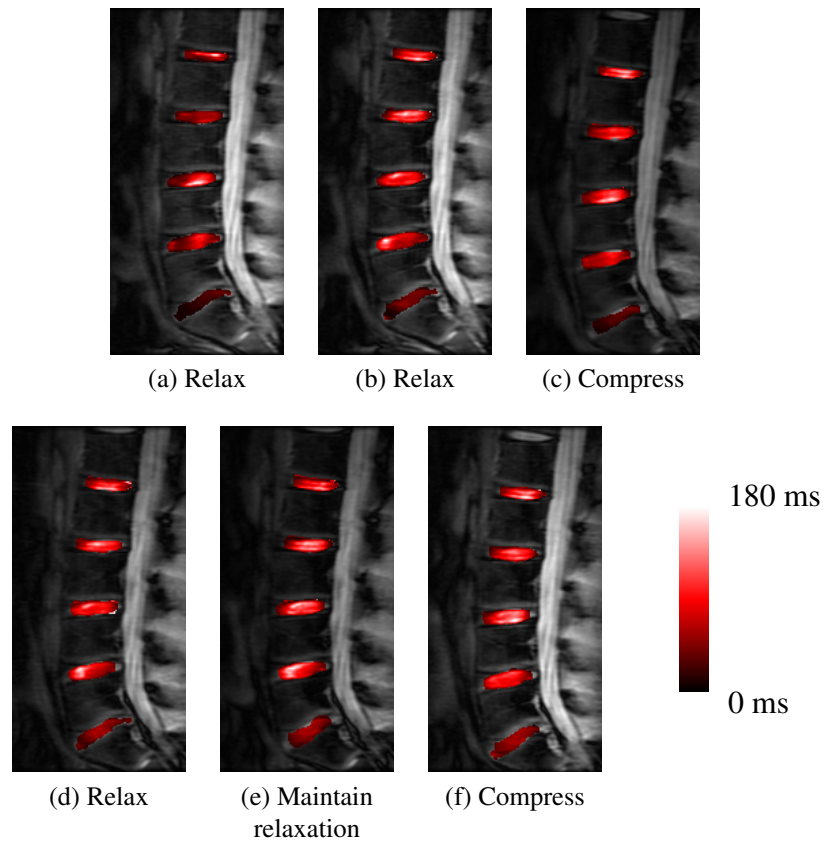
Figure 7.23: $T_{1\rho}$ maps for volunteer G

	A	C	P
relax-relax			
	0.4	0.1	-0.2
	1.7	1.0	0.4
	0.2	0.2	0.1
	0.4	0.1	-0.1
	0.7	0.7	0.5

	A	C	P
relax-compress			
	-0.6	0.4	2.0
	2.7	-0.2	-3.3
	3.4	1.2	-0.2
	0.2	-0.2	-0.7
	0.8	-0.7	-2.7

	A	C	P
compress-relax			
	-1.2	-0.6	-0.1
	-1.5	-0.2	1.4
	-2.8	-0.9	0.8
	0.8	0.3	-0.2
	1.0	0.4	-0.5

	A	C	P
relax-relax			
	0.4	-0.5	-1.5
	0.2	0.7	0.9
	1.2	-0.0	-1.5
	-1.0	-0.3	0.2
	1.3	0.5	-0.9

	A	C	P
relax-compress			
	2.2	0.5	-1.9
	-0.9	-0.7	-0.2
	-0.5	0.4	1.0
	1.9	0.8	-0.0
	-3.0	-0.6	4.6

60 ms

-60 ms

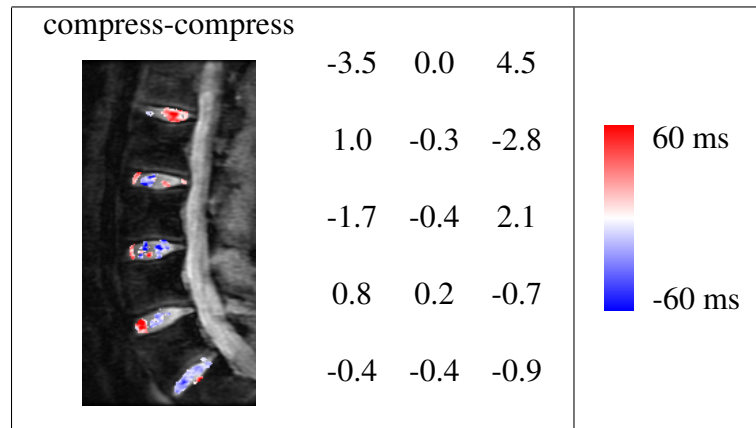Figure 7.24: Statistically significant changes in $T_{1\rho}$ for Volunteer G are colored according to the amount of change. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.

(a) Compress   (b) Relax   (c) Maintain relaxation

(d) Compress   (e) Relax

200 ms

0 ms

Figure 7.25: $T_{1\rho}$ maps for volunteer H

	A	C	P			A	C	P
compress-relax					relax-relax			
	0.8	0.3	-0.1			0.9	0.0	-0.7
	1.4	1.6	5.6			-1.7	0.2	2.8
	1.8	2.6	6.9			1.1	0.1	-1.1
	0.5	0.8	2.3			1.4	1.2	1.8
	5.9	2.0	-3.4			0.4	0.2	0.0
relax-compress					compress-relax			
	-0.6	0.0	0.4			0.1	-0.3	-0.7
	-0.1	-0.3	-1.2			1.2	0.7	1.8
	-1.5	-0.7	-0.4			1.4	0.5	0.6
	1.3	1.0	-0.4			-2.5	-0.9	1.2
	-1.0	-1.0	-1.7			1.9	1.1	0.4

60 ms

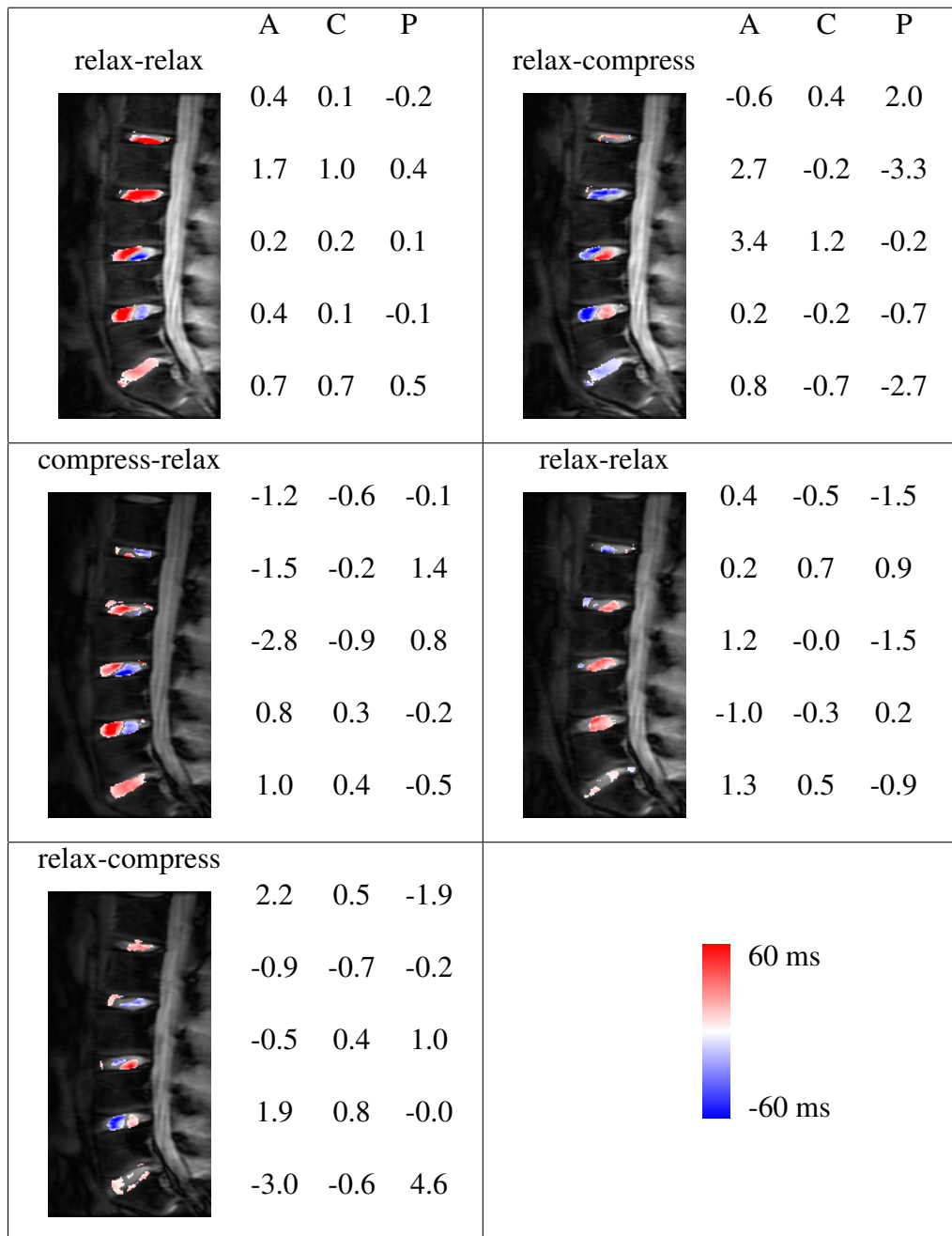-60 ms

Figure 7.26: Statistically significant changes in $T_{1\rho}$ for Volunteer H are colored according to the amount of change. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.
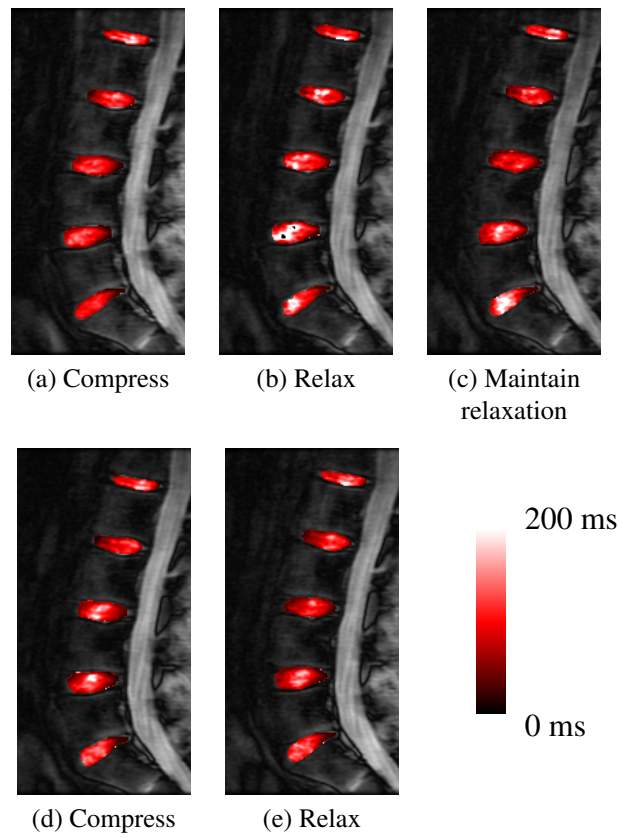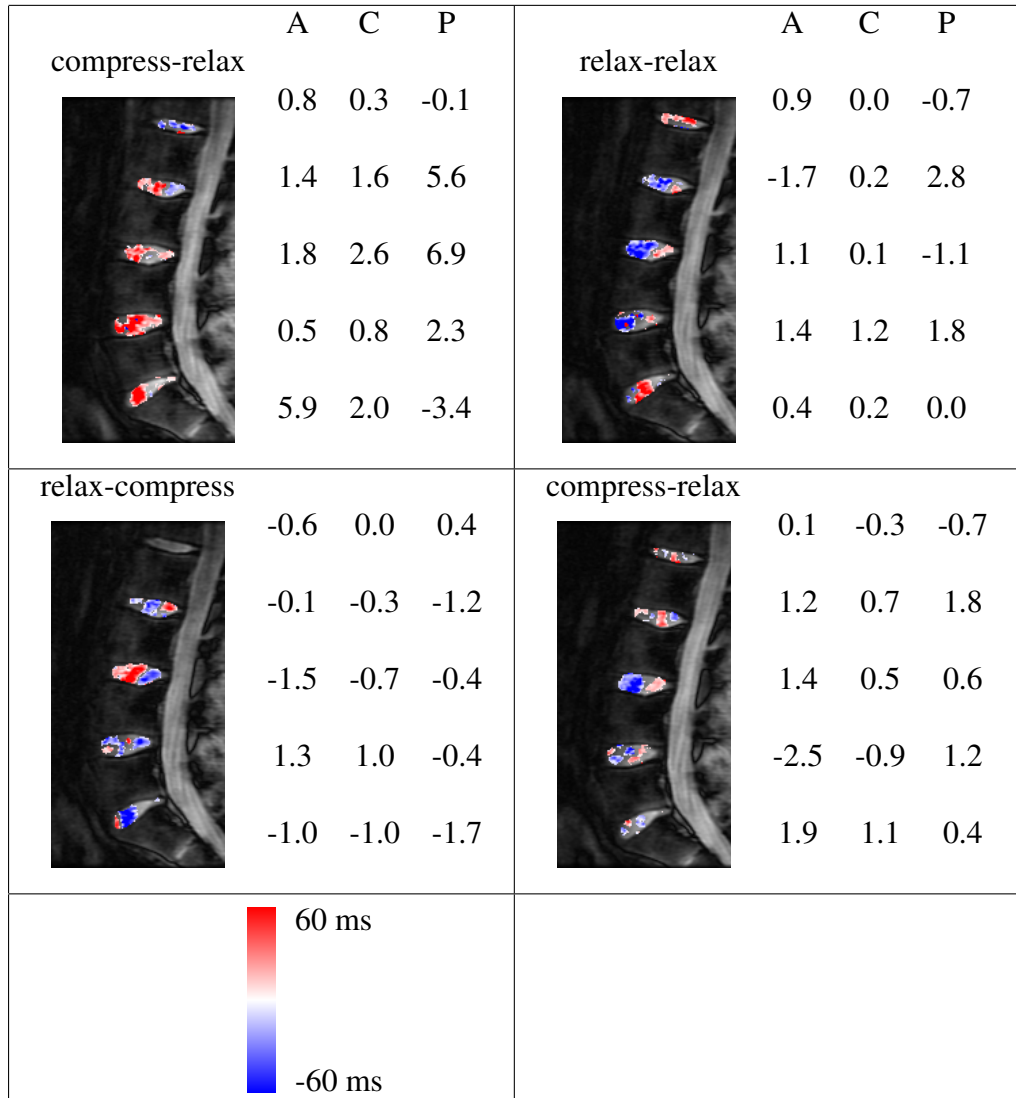
image pair, weight was applied between the two images (relax-compress). The L5-S1 and L1-L2 discs decreased in height anteriorly and increased in height posteriorly, while the others increased in height anteriorly and decreased in height posteriorly, showing how the vertebrae rotated differently to support the applied weight. In discs L2-L3, L3-L4, and L4-L5, the change in $T_{1\rho}$ follows the increase-decompression pattern suggested earlier, with increases in the anterior sections (where an increase in disc height indicates decompression) and decreases in the posterior sections (where compression is marked by a decrease in disc height). The L5-S1 disc does not follow this pattern, as the $T_{1\rho}$ increases throughout most of the disc even though the anterior portion decreased in height. On the other hand, this disc does follow the pattern where $T_{1\rho}$ changes are smaller in magnitude than for the more superior discs. The third image pair compares the two images taken after compression was applied. The disc height differences are much smaller, as the compression state remained the same. $T_{1\rho}$ decreases in the upper discs and increases in the lower discs, corresponding to a tendency toward disc height decrease in the upper discs and increase in the lower discs. Disc L5-S1 now shows a large $T_{1\rho}$ change, contrary to our earlier observations. Perhaps while resisting compression over several minutes, the posture shifted, releasing some of the pressure on the pelvic region. The fourth image pair compares before and after removing the weight (compress-relax). Disc height measurements resemble the opposite of the relax-compress pair, which is expected since the spine is returning to a relaxed state. $T_{1\rho}$ predominately increases, following the pattern of relaxation, though it decreases in parts of the L5-S1 disc. In the fifth pair (relax-relax), the disc heights change to a lesser degree

than the previous pair. $T_{1\rho}$ changes are also minimal, with some increase in the L5-S1 disc. The disc height changes in the relax-compress (sixth) pair follow a similar pattern as the earlier relax-compress pair. There are more regions of $T_{1\rho}$ decrease than the earlier relax-compress pair – perhaps the compressive effect of the load is larger since the spine has had more time to relax in the supine position before applying the load this time. There are regions of increasing $T_{1\rho}$ in most of the discs, suggesting how different parts of the disc may be affected by the load. Finally in the seventh pair, maintaining compression (compress-compress) leads to some disc height increase and $T_{1\rho}$ increase in the upper discs and disc height decrease and $T_{1\rho}$ decrease in the lower discs.

We move on to Volunteer F. The largest $T_{1\rho}$ changes for this volunteer are decreases in the first relax-compress pair, which is again consistent with the previous volunteer. The magnitude of the $T_{1\rho}$ changes in the other image pairs are much lower and do not have an obvious pattern. The disc heights follow a general pattern of squishing the posterior and stretching the anterior parts of discs L2-L3, L3-L4, and L5-S1 when the weights were applied and doing the reverse for weight removal, though the magnitude of the changes are different.

Volunteer G shows the same patterns fairly distinctly. The first pair (relax-relax) is dominated by strong $T_{1\rho}$ increase and disc height increase, which is the pattern seen in earlier for initial relaxation. The L5-S1 disc has both lower $T_{1\rho}$ values and reduced $T_{1\rho}$ changes compared with the other discs, again following the pattern for relaxation before applying the load. The second pair (relax-compress) follows the patterns for compression by show-

ing mainly $T_{1\rho}$ decreases and disc heights that decrease posteriorly and increase anteriorly. $T_{1\rho}$ changes with subsequent relaxation (pair three) mimic the initial relaxation with mostly increases, though at lower magnitude, suggesting the magnitude of the relaxation between standing and lying is greater than the relaxation between applying and removing weight in the compression device, which might be expected. Disc heights for L2-L3 and L3-L4 reverse their directions from the initial compression, while the other discs behave differently. The following continued relaxation (pair four) again follows the $T_{1\rho}$ change pattern of increasing for relaxation. The disc height changes are low in magnitude and do not follow any particular pattern. Disc heights for the next compression (pair five) do not follow the pattern either and are also mostly low in magnitude. $T_{1\rho}$ changes for pair five are lower in magnitude than the other pairs and show both increases and decreases.

Volunteer H followed a slightly different procedure. Since he or she had been in the scanner for about 10 minutes before compression, some relaxation had already occurred. The compression/relaxation pattern began with compression, unlike the other three volunteers who began with relaxation. Finally, he or she spent only about 4 minutes in the second set of compressed and uncompressed states as opposed to the approximately 15 minutes that the other volunteers remained in each state, leading to less time for the discs to adjust to each new compression state. The initial compress-relax pair follows our relaxation $T_{1\rho}$ change pattern with mostly increases accompanied by mostly increases in disc height. The following relax-relax pair continues to show disc height increases, which is expected for continued relaxation. However, there is a lot of $T_{1\rho}$ decrease, which is the

opposite of what we expect given the previous patterns. The following compression yields disc height decreases as expected, many places of $T_{1\rho}$ decrease, also as expected, but also some regions of strong $T_{1\rho}$ increase. The next pair (compress-relax) shows a similar mixture of $T_{1\rho}$ directional changes. These mixtures of increases and decreases may be due to the short period of time in each compression state; some parts of the discs may not have gotten a chance to compress or relax.

Finally we show the $T_{1\rho}$ maps for the neutral vs. flexion experiment in Figure 7.27, with the statistically significant $T_{1\rho}$ changes and percentage changes in disc height reported in Figure 7.28.

The first comparison is between neutral and flexed positions. The upper four discs compress anteriorly and expand posteriorly according to the disc height measurements, which is expected given the spinal flexion. $T_{1\rho}$ changes continue to follow the above-mentioned pattern in these discs, increasing posteriorly (corresponding to expansion) and decreasing centrally and anteriorly (corresponding to compression). The L5-S1 disc behaves differently, with its height decreasing everywhere but mostly increasing in $T_{1\rho}$, which is the opposite of our hypothesized pattern. The second comparison is between flexed and neutral positions. The upper three disc heights reverse the effect of the previous neutral-flexed motion as expected, expanding anteriorly and compressing posteriorly, with corresponding $T_{1\rho}$ increases in central and anterior locations, except for a slight anterior decrease in L3-L4. The disc height decreases on L4-L5 indicate compression, with a corresponding decrease in $T_{1\rho}$. The L5-S1 disc height increases anteriorly and decreases posteriorly like
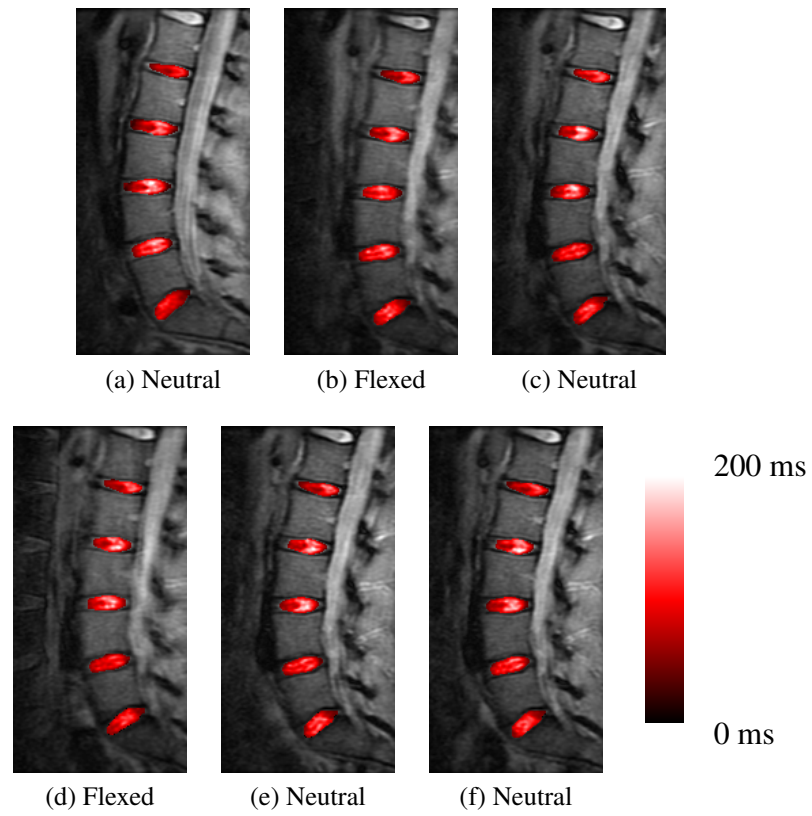
(a) Neutral   (b) Flexed   (c) Neutral

(d) Flexed   (e) Neutral   (f) Neutral

Figure 7.27: $T_{1\rho}$ maps for volunteer I

	A	C	P
neutral-flexed			
	-4.5	0.2	3.0
	-7.8	-1.1	4.1
	-8.5	-1.2	5.6
	-3.4	-1.0	1.6
	-4.7	-3.7	-3.4

	A	C	P
flexed-neutral			
	5.2	0.1	-2.3
	6.0	1.1	-2.3
	9.4	1.7	-4.3
	-0.5	-0.4	-0.2
	4.4	1.8	-2.6

	A	C	P
neutral-flexed			
	-0.0	-2.7	-7.5
	1.0	0.4	0.4
	-3.2	0.1	3.0
	-3.2	-1.0	0.9
	-12.0	-4.0	6.6

	A	C	P
flexed-neutral			
	0.2	2.6	6.8
	0.2	-0.8	-2.1
	-0.8	-1.9	-3.1
	6.0	2.1	-1.1
	8.5	2.2	-5.9

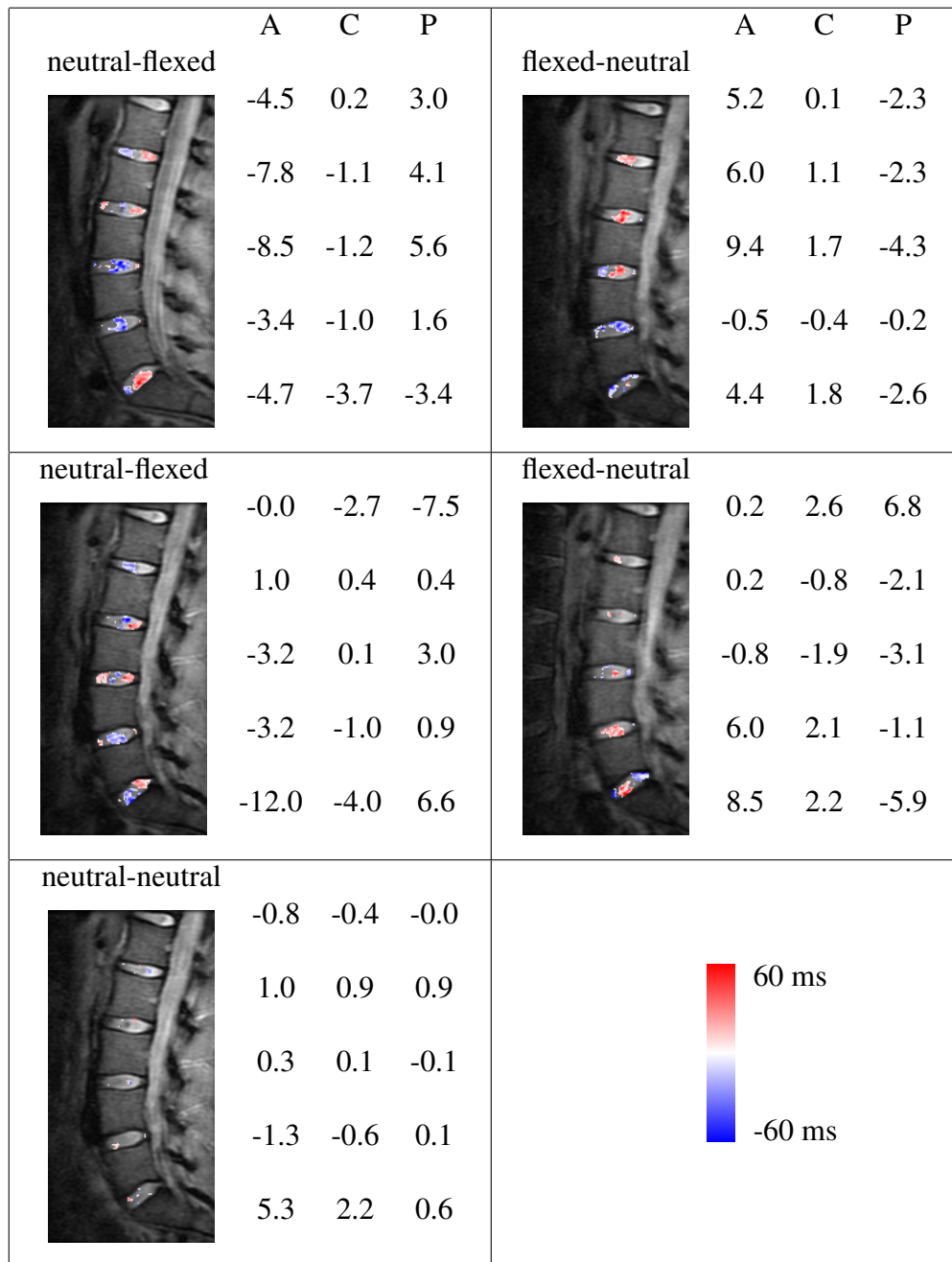|	A	C	P
neutral-neutral			
	-0.8	-0.4	-0.0
	1.0	0.9	0.9
	0.3	0.1	-0.1
	-1.3	-0.6	0.1
	5.3	2.2	0.6

60 ms

-60 ms

Figure 7.28: Statistically significant changes in $T_{1\rho}$ for Volunteer I are colored according to the amount of change. The numbers indicate the percent change in disc height at the A(nterior), C(entral), and P(osterior) parts of the disc. Each row represents one disc. Average standard deviation of repeated measurements is 0.7%.

the upper discs, but the $T_{1\rho}$ changes in this disc are minimal. The third comparison is again between neutral and flexed positions. The disc height changes resemble the first neutral-flexion comparison with some deviations, especially in the superior discs. The inferior two discs follow the pattern of $T_{1\rho}$ moving in the same direction as the disc height. The other three discs loosely follow this pattern. The fourth image set repeats the flexed and neutral position comparison. Disc heights are again mostly as expected, with anterior expansion and posterior compression. $T_{1\rho}$ changes are minimal, with some increase centrally and anteriorly in L4-L5 and L5-S1; these discs also have corresponding anterior expansion in disc height, which is consistent with our hypothesized pattern. The final comparison is between two neutral images with no position change and minimal time (less than 3 minutes) between them. As expected, there is very little change in $T_{1\rho}$, and only slight disc height differences for many of the discs, though L5-S1 expands anteriorly a fair amount.

7.2.3 Assessment of $T_{1\rho}$ changes within regions of interest

These calculations compare large regions of interest (ROI) within the discs (all, anterior, posterior, or central portions), so there is one value per disc for each of the four ROI's. These calculations are less local than in the previous section, but since there are fewer numbers, it is easier to report results.

We begin this section by discussing the reproducibility of the $T_{1\rho}$ measurements. The average coefficient of variation (CV) of the median $T_{1\rho}$ within discs between two repeated measurements for the $T_{1\rho}$-FSE sequence we used has been reported as 5.64% [8]. In an-

other study, the CV was 4.59% in vivo and under 3% for agarose phantoms using a sequence with the same $T_{1\rho}$ preparation but a spiral acquisition [18]. Using the data from our experiments, the average CV per disc is 7.6% for the backpack experiment, 13.9% for the compression device experiment, 6.9% for the flexion/extension experiment, and 10.4% for all the data. CV's are often used to report reproducibility, but here they also reflect any changes in $T_{1\rho}$ related to the state of compression as all the data in each experiment were averaged. The larger values for the compression device are likely because the compression device creates more differences in the spine than supine relaxation. The CV for the backpack experiment is closer to the previously reported CV's, indicating that the change in $T_{1\rho}$ over the course of an hour is not much different than the expected variation between measurements. The rest of this section explores these relative differences.

Let us look at the progression of $T_{1\rho}$ in each disc for each volunteer, as plotted in Figures 7.29 - 7.37. There are many fluctuations, but most are not significant using the paired Student's t-test. We did find a significant increase of 20% in $T_{1\rho}$ when comparing images under compression with those after compression was released in the anterior region of the L4-L5 disc ($p \approx 0.03$) and in the anterior and posterior regions and the entirety of the L5-S1 disc ($p < 0.02$). This test identifies significant changes with decompression compared with changes during repeated tests in the relaxation sequence. The relaxation sequence consisted of 5-7 separate measurements at about 7 minute intervals after the backpack loading. We used the standard deviation of this sequence in the t-test to represent the control distribution. This change in $T_{1\rho}$ follows the same formula of increasing with the release of compression

as the local patterns we pointed out in the previous section.

We now look quantitatively at our earlier observation that discs with lower $T_{1\rho}$ values have less variation in $T_{1\rho}$. Discs with lower $T_{1\rho}$ values are associated with degeneration. We compared the variation in the four discs identified as degenerate to the variation in the 32 healthy discs. We found the variation in $T_{1\rho}$ between successive images to be 48% lower ($p < 0.03$) in degenerate than in healthy discs.

Using a one-way analysis of variance, we checked whether $T_{1\rho}$ variation depended on the location of the disc (i.e. L2-L3, L3-L4, L4-L5, or L5-S1). We found no significant effect of disc identity on the amount of $T_{1\rho}$ variation.

We also noticed that the discs identified as degenerate using clinical T_2-weighted images had noticeably lower $T_{1\rho}$ values, as has been shown by other researchers [7, 8, 5]. It is not the main emphasis of this work, but we mention it here as an additional observation.

7.3   Summary

The misalignment in the pre-registration difference images demonstrates that registration is needed for multiple acquisitions of the same person, even if the images are acquired during the same exam. The post-registration difference images show that the registration is visually accurate for all of the image pairs in these experiments.

The data shown suggest that $T_{1\rho}$ values in lumbar intervertebral discs decrease when compressed and increase when expanded or relaxed. This pattern occurs in the local com-

Figure 7.29: Progression of $T_{1\rho}$ values in regions of interest for Volunteer A.

Figure 7.30: Progression of $T_{1\rho}$ values in regions of interest for Volunteer B.

Figure 7.31: Progression of $T_{1\rho}$ values in regions of interest for Volunteer C.
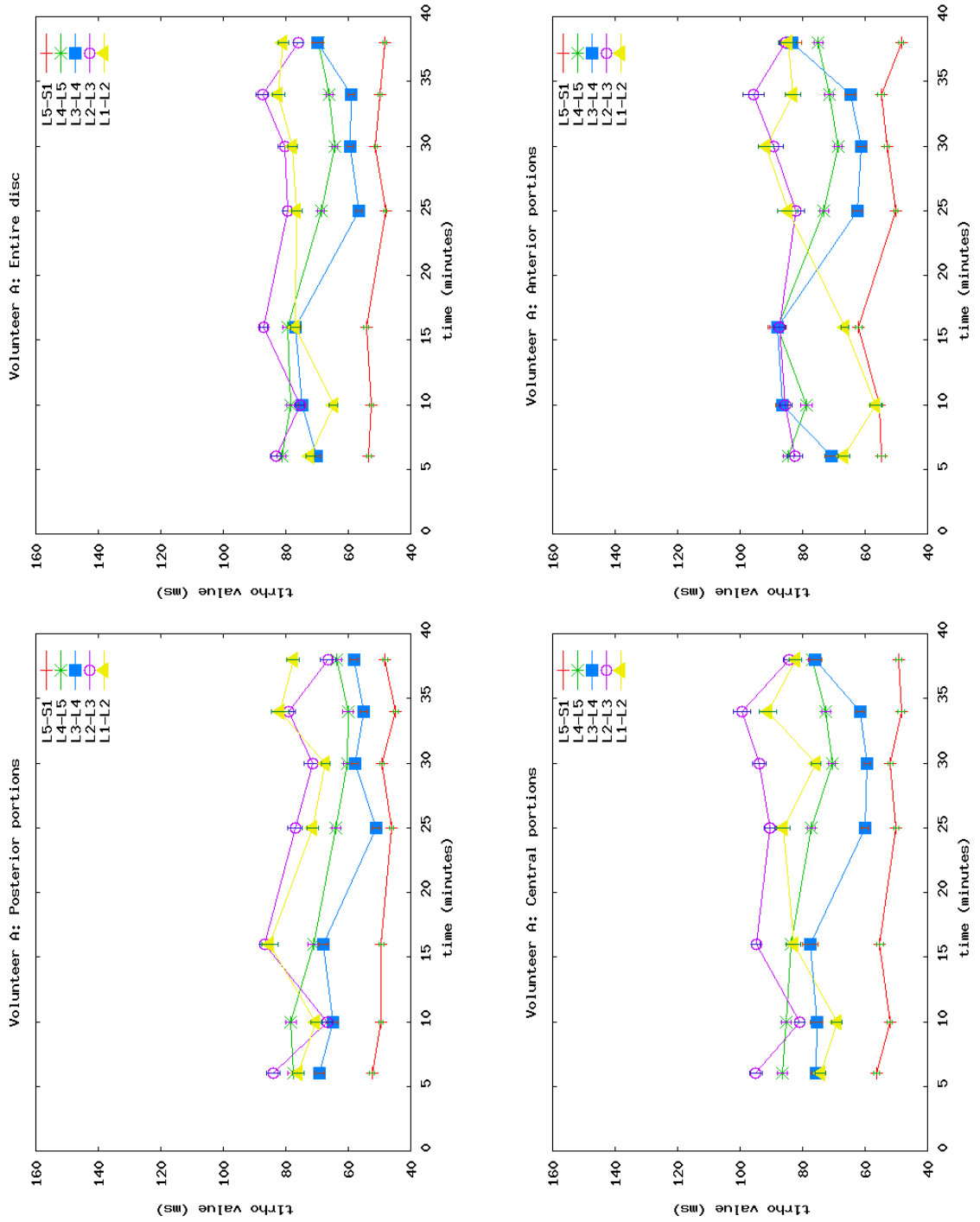
Figure 7.32: Progression of $T_{1\rho}$ values in regions of interest for Volunteer D.

Figure 7.33: Progression of $T_{1\rho}$ values in regions of interest for Volunteer E. Note that the $T_{1\rho}$ scale for the plots in this figure is slightly different than for the other similar figures.
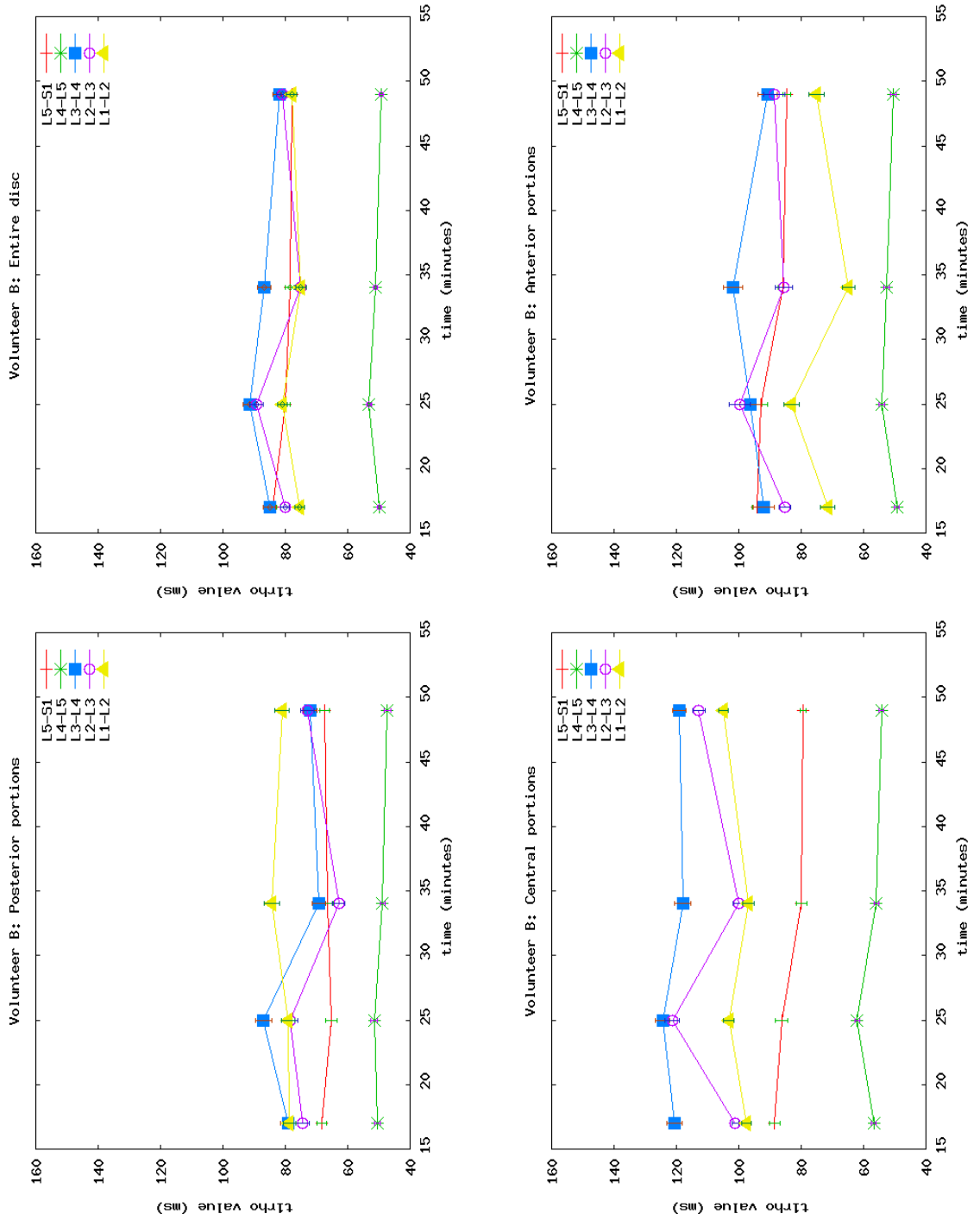
Figure 7.34: Progression of $T_{1\rho}$ values in regions of interest for Volunteer F.

Figure 7.35: Progression of $T_{1\rho}$ values in regions of interest for Volunteer G. Note that the $T_{1\rho}$ scale for the plots in this figure is slightly different than for the other similar figures.
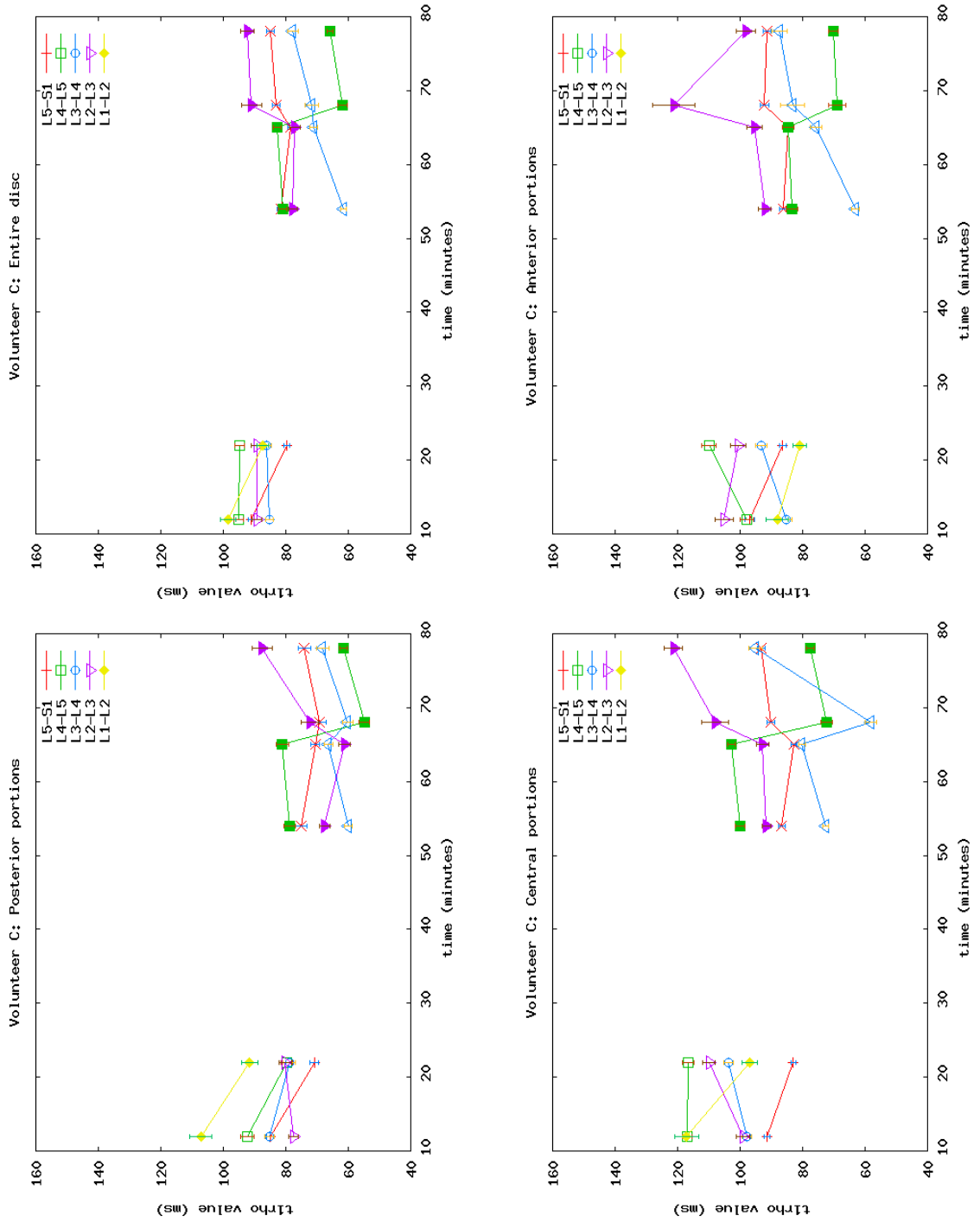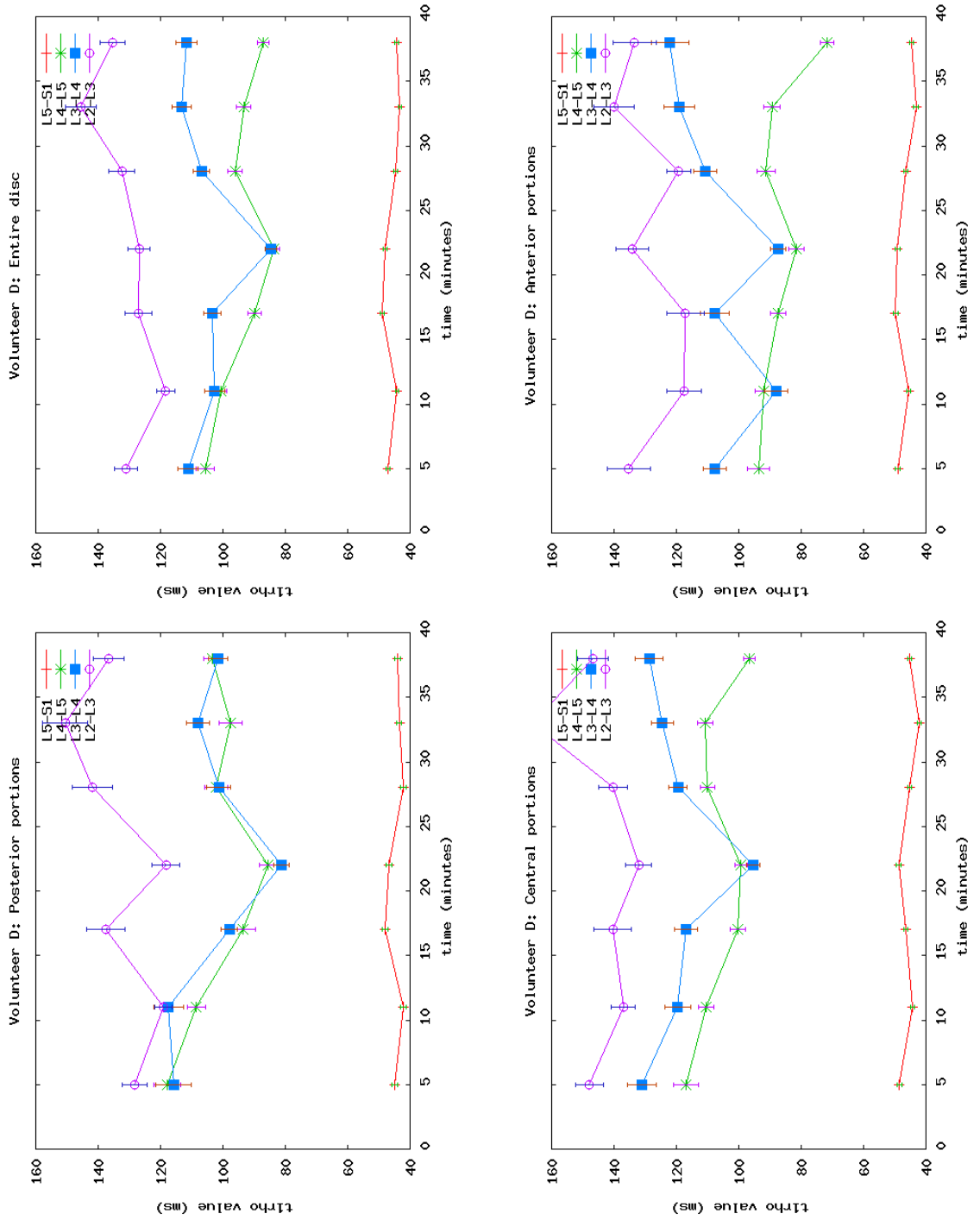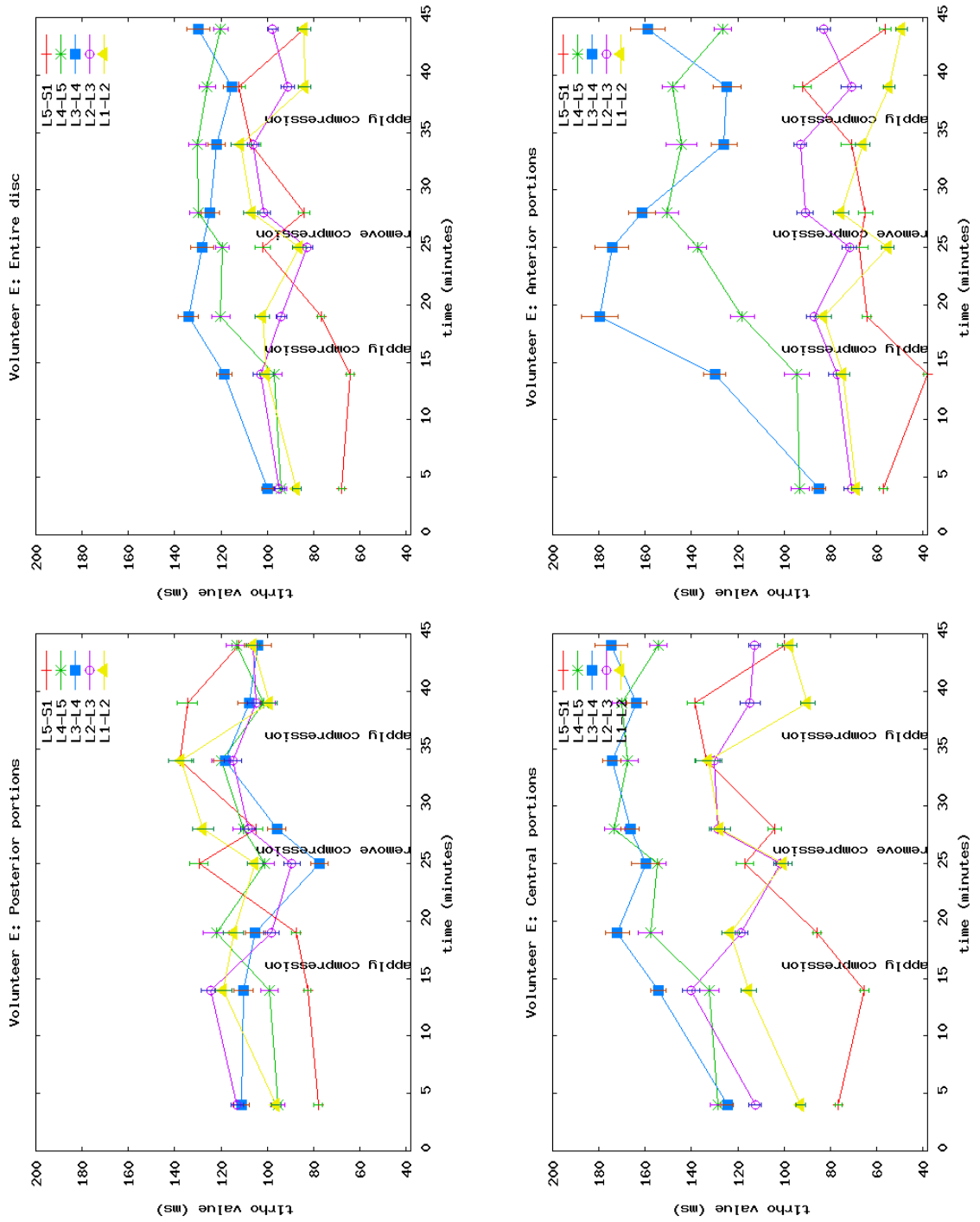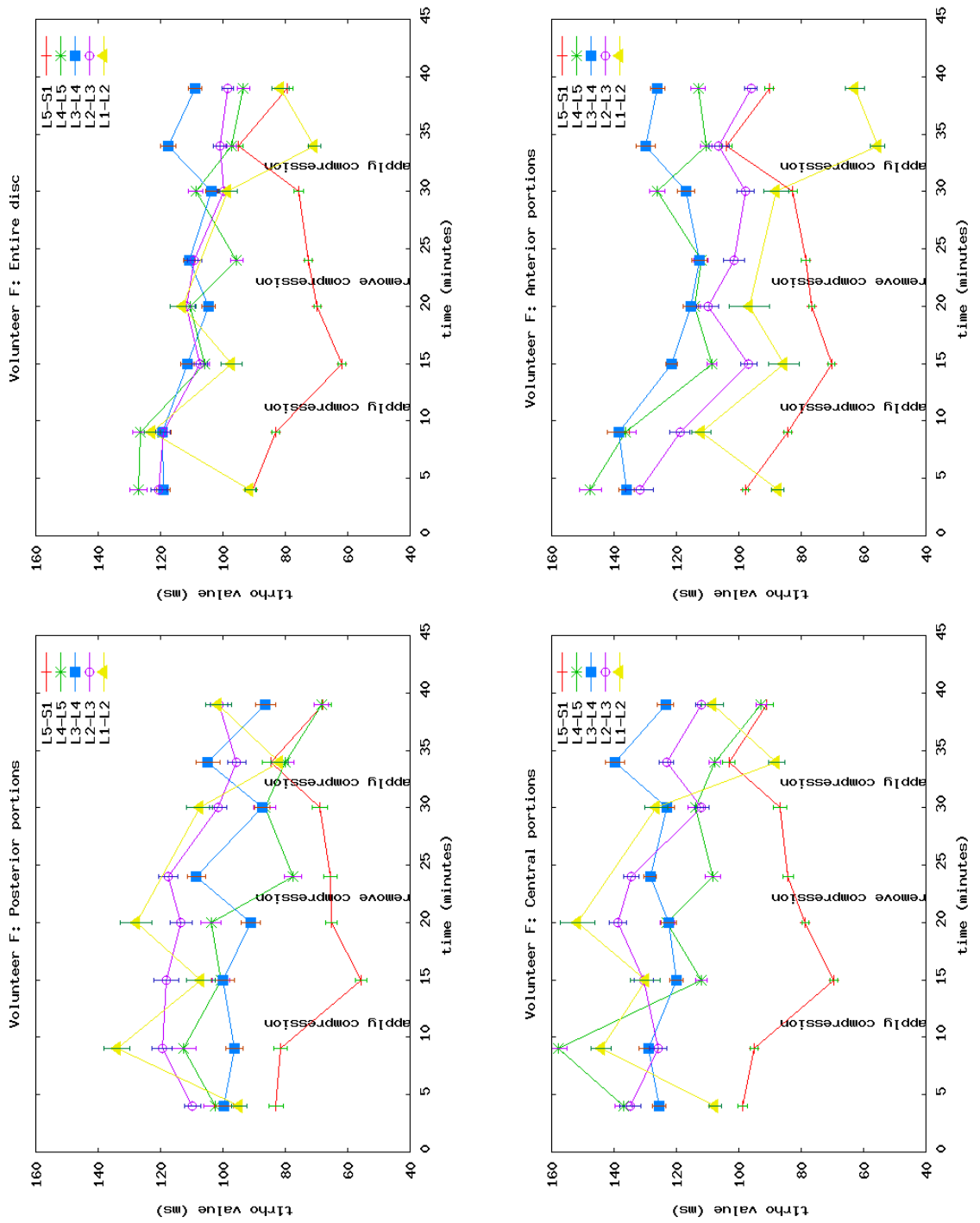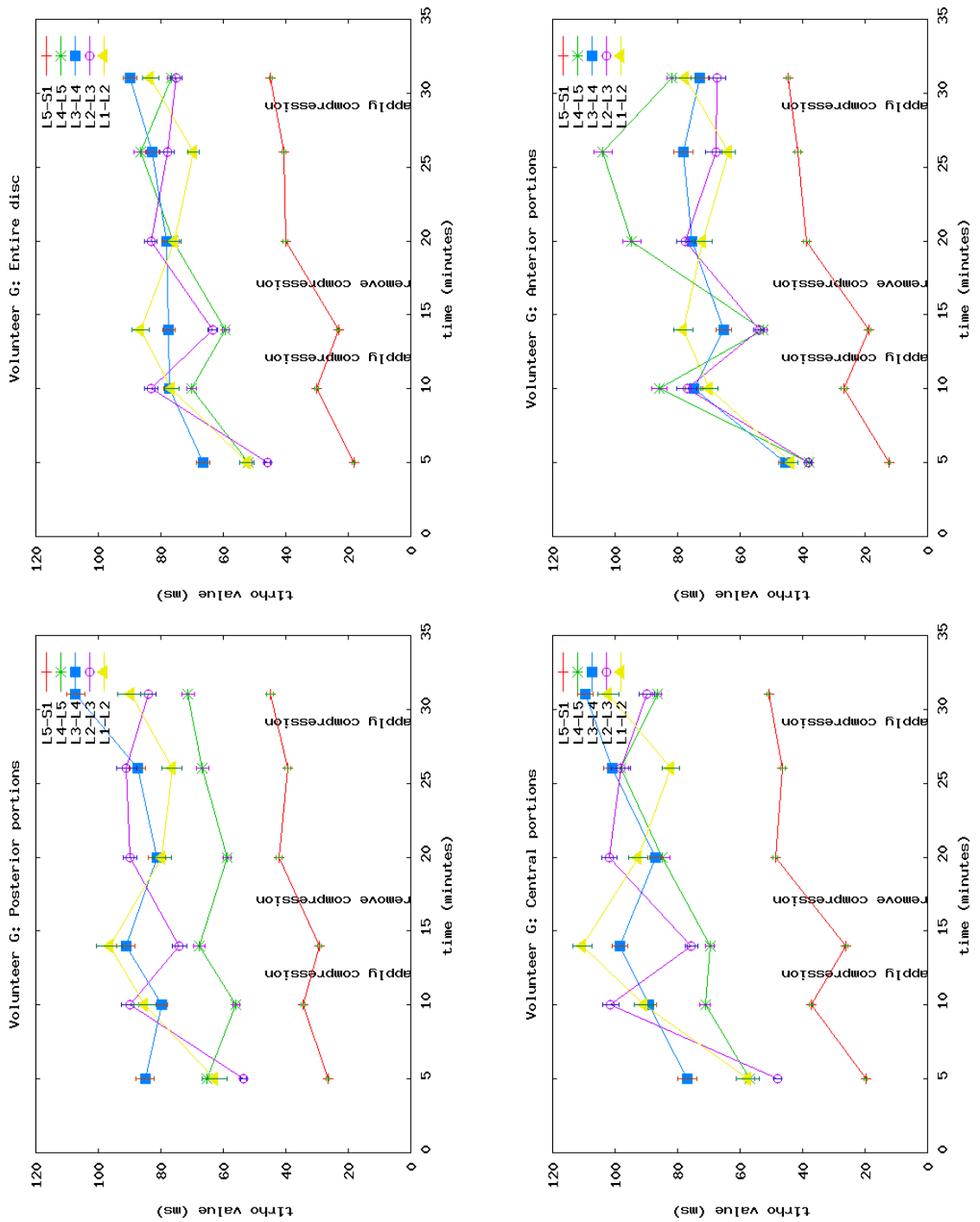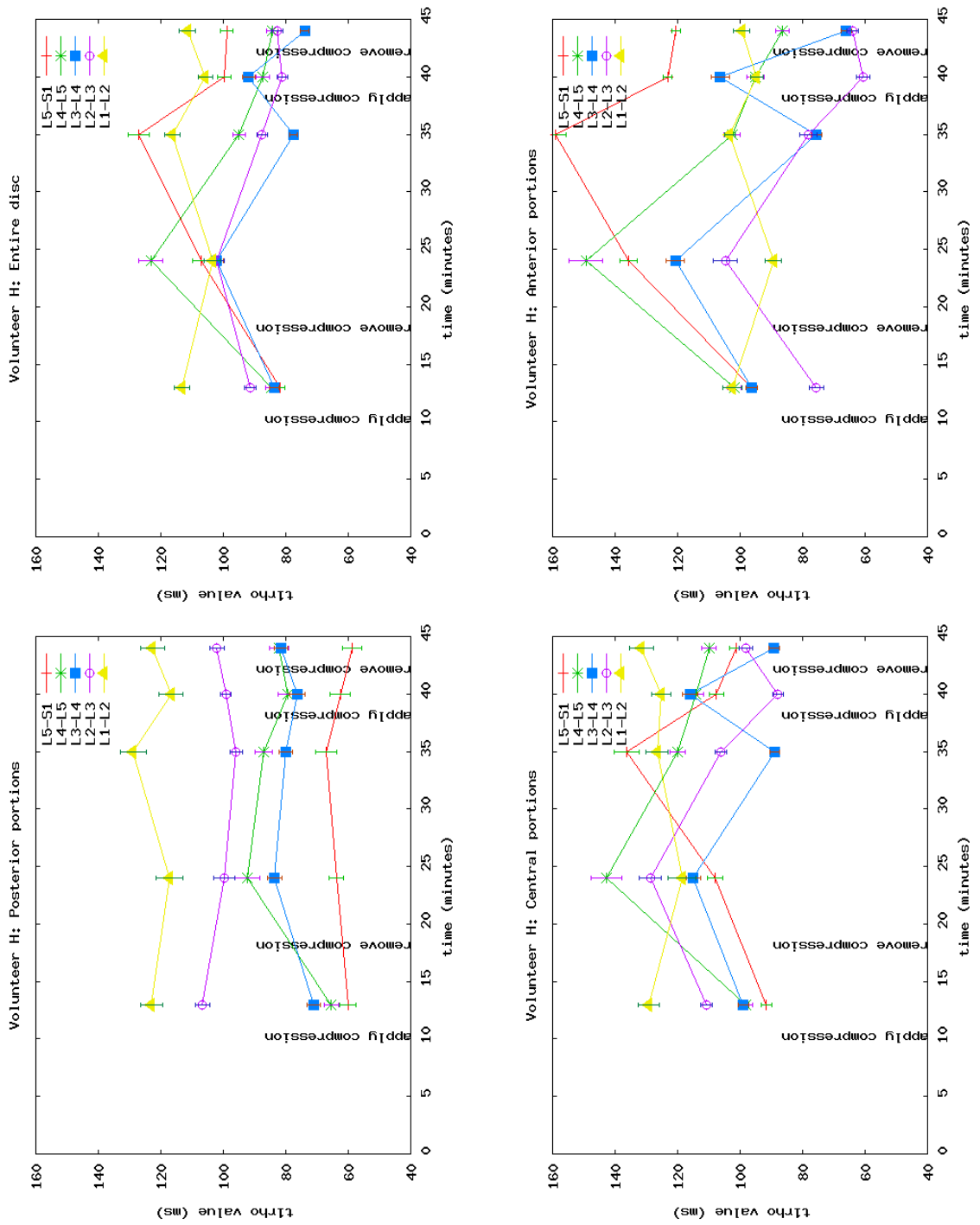
Figure 7.36: Progression of $T_{1\rho}$ values in regions of interest for Volunteer H.
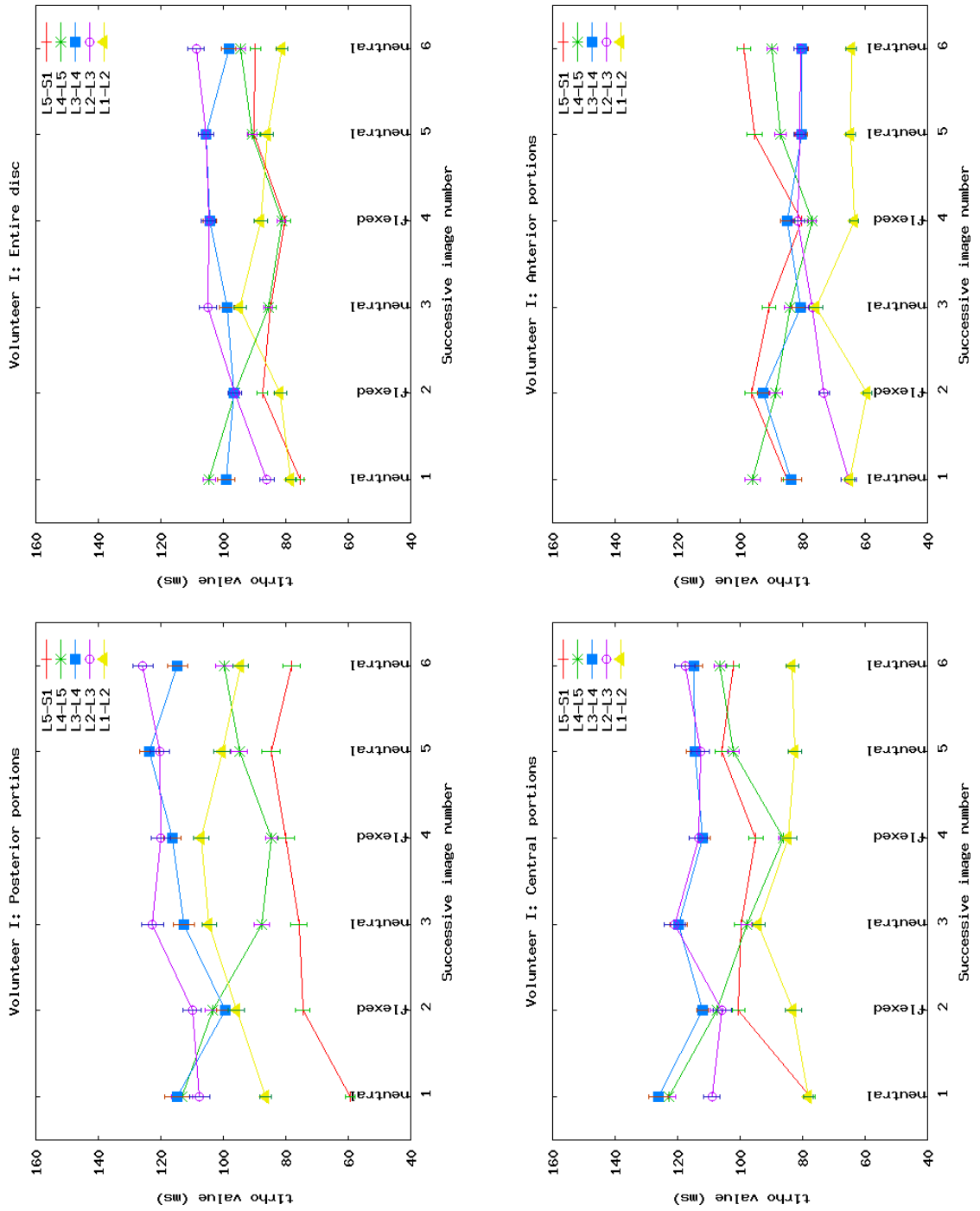
Figure 7.37: Progression of $T_{1\rho}$ values in regions of interest for Volunteer I.

parisons and is backed up by the statistically significant increase in $T_{1\rho}$ when compression was released during the compression device experiment. Note that lower $T_{1\rho}$ values are associated with degeneration. Compression seems to change the $T_{1\rho}$ values in the same direction as degeneration does, and relaxation seems to raise $T_{1\rho}$ higher, which is associated with healthy discs. Lower $T_{1\rho}$ (faster relaxation) under compression may be because the proteoglycans and water molecules are closer together, allowing more water-protein interactions and decreasing the molecular tumbling rate, which increases the $T_{1\rho}$ relaxation rate, as discussed in Section 2.3.

The above pattern occurs often in the data but is not entirely consistent. One does not expect a completely consistent pattern *in vivo* due to different ways of loading, which is related to individual anatomy, muscular support, etc. The way a load is distributed through a disc varies with the load situation, the person, and the anatomical location of the disc. Disc health also affects the way it responds to a load. Equilibrium after applying or removing a load takes many hours, so in our studies (under an hour) the response to a load is complicated by previous loads applied before or during the study. For example, in the compression device study the magnitude of $T_{1\rho}$ changes tended to decrease in the second round of compression / relaxation, perhaps from a residual condition from the first round of loading and unloading.

There was no significant difference in mean $T_{1\rho}$ over the course of the hour-long relaxation after the upright backpack loading. This protocol missed the first 5-10 minutes of relaxation between backpack removal and the first $T_{1\rho}$ acquisition due to the setup tasks

of helping the volunteer enter the machine and get comfortable, landmarking the anatomy of interest, acquiring a localization scan, and positioning the field of view. Any significant compression from the backpack and other pre-scan activities had either relaxed close to equilibrium by the time the setup was complete or was relaxing slowly over a much longer time period than is practical for MRI exams. The setup tasks are required for all exams and the exam duration is typical for scan sessions, so compression differences due to pre-scan activities is not likely to be a problem in most situations, reducing the need to control pre-scan activities before $T_{1\rho}$ studies.

Variation in $T_{1\rho}$ is lower in L5-S1 discs that have lower $T_{1\rho}$ values. This is shown in the local comparisons and by the significant difference in $T_{1\rho}$ variation between healthy and degenerate discs. This observation is consistent with studies of T_1 [19] and T_2 [27, 28] under compression. Lower $T_{1\rho}$ values are associated with degeneration, and degenerate discs are expected to lose water more quickly under compression and have fewer proteoglycans and water molecules than healthy discs. Thus we expect the difference in restriction and interactions of proteoglycan and water molecules under compression to be less pronounced with respect to healthy discs, leading to less change in $T_{1\rho}$ values.

Measurements of disc heights give us insight into how the spine handles an applied load. Disc heights between flexed and neutral positions responded as expected, with most discs compressing anteriorly and expanding posteriorly for flexion with the reverse for extension back to the neutral position. When a load was applied, discs tended to compress posteriorly and expand anteriorly. However, these observations are only general patterns, as the disc

height response to compression varied with the disc's anatomical location, with different volunteers, and with different trials of the same volunteer, as expected. Disc heights did not follow a pattern of decompression during the backpack study, likely because the amount of height change with supine relaxation is on the same order as changes due to small volunteer movement over time.

It is important to note that while we did see some tendencies and patterns in the $T_{1\rho}$ behavior under compression, this study is far from conclusive. The observed patterns were not seen in all of the examined discs, so there may be other observations that could be made in a larger study. The human body is quite complex and we do not yet completely understand the mechanisms that $T_{1\rho}$ measures, so there is likely to be more going on than these observations show. To help isolate the effect of compression, future studies could be done on animal tails so that compression can be applied on individual vertebrae for longer periods of time.

Chapter 8

Conclusion

Our aim is to quantitatively assess local changes in the $T_{1\rho}$ relaxation parameter within intervertebral discs under different amounts of compression. This section discusses the information from previous sections to address how well our method accomplishes this goal, how much human interaction is required, and how useful the method is likely to be.

We begin by looking at data acquisition. Since the imaging sequence is single-slice, we need to aim the second acquisition to contain the same anatomical points as the first. Even if a future sequence were multi-slice, the slice thickness is typically so large (4-8mm) that a slice in an incorrectly positioned second image could include points several millimeters away from the first image. As an alternative to prospective registration, a highly skilled technician may be able to locate a given anatomical landmark in both imaging sessions with acceptable accuracy, but there are limitations to this approach. First, the desired slice must be at a landmark distinctive enough for human visual identification. Second, to recall

where the slice was located in the initial exam, the technician must either have the images from the first exam on hand, with the sagittal slice marked on the axial image (which is unlikely to be easily available) or the slice must be a standard named landmark identifiable across patients. Third, the same technician must be available for both exams. Even if these conditions are met, the accuracy of manual positioning is difficult to quantify so the investigator will not know how much to trust the end statistical comparison of the images. The prospective registration tool increases flexibility, quantification, and possibly accuracy over technician localization.

The flexibility of the spine in bending and twisting means that global rigid transformation is not the only motion that can affect which anatomical points are imaged between exams. MRI acquisition methods cannot accommodate a twisting, curving path, so we must instead help the patient lie as straight as possible. The scanner bed keeps the hips and shoulders from twisting, and lateral bending can be minimized by asking the volunteer to lie with a straight spine. We have visually noted reasonably consistent positioning with respect to twisting and lateral bending as volunteers exit and reenter the scanner. However, we do not present a quantitative measure for assisting or evaluating this type of positioning.

The amount of initial misalignment is limited in a standard closed-bore MR scanner. However, we saw in section 6.3 that even a misalignment of a single pixel is enough to affect the local hypothesis testing, indicating that registration is worthwhile. If in the future we wanted to apply the algorithm to larger motions, such as those possible in open scanners, we would have a much greater possibility of twisting and out-of-plane bending, suggesting

that a sagittal slice is no longer sufficient. Rather, we would seek a three-dimensional acquisition with near isotropic voxels. Three-dimensional $T_{1\rho}$ sequences have been reported for the knee [70], but not for the spine and not with isotropic voxels at sub-millimeter resolution. Theoretically, our registration and analysis technique would still apply. I wrote the code using the C++ Standard Template Library with image dimension as one template parameter, so using the algorithm in three dimensions should be reasonably straightforward.

Next we examine the disc segmentation. As described in section 6.3, the required manual initialization is quick and insensitive to minor variations. Figure 6.9 gives examples showing that disc segmentation covered most of the disc with minimal leaking. This level of segmentation accuracy is sufficient to produce sub-pixel registration accuracy, which in turn is sufficient for accurate hypothesis testing within the disc (as demonstrated in section 6.3). A more accurate segmentation may be desired for the interpretation of the hypothesis tests since including portions of the vertebrae can include distracting, irrelevant information, and not filling in all of the disc could cause one to miss significant changes. Even where over-segmentation causes pixels in the vertebrae (rather than the disc) to appear statistically significant, since the hypothesis tests are done within local neighborhoods and the data is presented as an overlay on the anatomical image, it is obvious which of the regions marked significant are truly within the disc. If desired, the small number of incorrectly labeled pixels can be erased via the graphical user interface. As for under-segmenting, or not including all portions of the disc, in our examples, missed areas were small and at the edges of the disc, not interfering much with the statistical analysis.

While our examples include a few discs with low $T_{1\rho}$ (associated with degeneration), we did not image any severely diseased discs, where the disc is barely visible. We expect segmentation to be difficult in such cases; manual touch-ups could be done, but if the disc is nearly gone or the intensity is too low even with no $T_{1\rho}$ preparation, it is likely that $T_{1\rho}$ values cannot be reliably estimated, making the segmentation unnecessary.

We now address vertebral segmentation, which determines landmark locations and bounding boxes for the rigid transformations that guide the deformable registration of the discs. Since rigid registration of an individual vertebra is done using a bounding box around the vertebral segmentation rather than using the segmentation borders directly, the registration is very robust to errors in vertebral segmentation. In fact, slight over-segmenting (including more than the actual vertebra) in the anterior/posterior direction is beneficial in that it incorporates more features to guide the registration. The deformable registration of the disc is sensitive to landmark placement, so the vertebral segmentation must be more precise near the discs. This unequal need for precision led to the balloon-type shape of the vertebral segmentation demonstrated in section 4.2.2.

Next we look at the registration of the discs, which we use to ensure that the pixels from the first image are being compared with pixels in the same anatomical region in the second image for reliable comparison of local $T_{1\rho}$ values. Registration on rigid bodies can be simpler and more robust than deformable registration of soft tissue since it requires optimizing fewer parameters. By using a piecewise-rigid transformation to select control points for deformable registration, we take advantage of the constraints offered by the rigid

bodies (vertebrae) while still allowing non-rigid deformation within the discs, which is more anatomically accurate for flexible tissue. Using more tissue than just the discs to guide the registration is more reliable since we expect the pixels within the discs to change. Using surrounding tissue ensures that the transformation is based on physical motion rather than changing intensities. As discussed in chapter 6, we register with sub-pixel accuracy and high repeatability. Automatic results are similar to manual results but more consistent, and the process is less tedious and faster.

Now we consider the statistical analysis. One of the strengths of our statistical technique is the opportunity for the user to investigate statistically significant areas by changing the neighborhood size, significance level, and confidence level. Allowing exploration gives more information than an isolated report of statistical significance at the common but somewhat arbitrary cutoff of 0.05. First, hypothesis tests are probability-based so the significance level gives only a likelihood, not a certainty, of a change being due to chance. Several pixels will therefore be incorrectly labeled as statistically significant. To address this, the user may view results at different significance levels to explore the strength of the significance. Second, since neighborhood size affects the calculations, we made it easy for the user to adjust the size. Third, some error arises from using a limited number of samples (pixels) in the statistical estimate. We calculate a confidence estimate based on this type of error, and the user may opt to reject the null hypothesis only for pixels that are above a certain confidence level.

The statistical technique developed in this dissertation extends the geographically

weighted regression (GWR) idea to the image domain and alters it to use resampling techniques. Our geographical weighting function has a hard bandwidth (i.e. weights of one for pixels within a bandwidth and zero elsewhere), and since we are limited to a small number of finite-size square pixels, to increase the number of pixels within the bandwidth and for computational simplicity, our weighting function is not radially symmetric, i.e. we use a square neighborhood rather than a circle. We experimented using a radially symmetric function with a soft cutoff but found the results harder to interpret. The main advantage cited for using a weighting function with a smooth cutoff in GWR is to reduce the impact of single points at the edge of the bandwidth. This advantage is reduced when we are calculating the values for many adjacent pixels in a regular grid (as in an image), for even if one neighborhood includes an outlier, its neighbors will not, thereby limiting its influence. A gentle cutoff would result in a spatially smoother result, but we prefer ease of interpretation to smoothness.

The last comparison (neutral-neutral) in Figure 7.28 shows two consecutively acquired images between which we do not expect any significant changes, yet there are small speckles of pixels marked as significantly different. This gives us the opportunity to remind the reader that statistical significance is simply stating that such a large change occurring by chance is unlikely. At the $p < 0.05$ level, "unlikely" means that 5% of the time, we expect it to happen. Thus the investigator must keep in mind that there are specks of noise in the significance calculations and that not all statistically significant pixels are truly different. Exploring with varying significance levels, confidence thresholds, and neighborhood sizes

can help one understand the true character of $T_{1\rho}$ changes.

We reported results of small preliminary studies of $T_{1\rho}$ under varying states of compression. The main reason to present them in this work is to provide an example of our algorithm in action. The algorithm was successful in this capacity, providing visually accurate registration results and maps of statistically significant differences in $T_{1\rho}$ in a manner that allowed us to analyze these differences in all locations within the disc. Another reason to include these results is to inspire further investigation of the behavior of $T_{1\rho}$ within intervertebral discs subject to varying loading situations. Our preliminary finding is that $T_{1\rho}$ tends to decrease with compression and increase with relaxation / expansion and that $T_{1\rho}$ varies less in degenerated discs than in healthy discs. $T_{1\rho}$ changes vary depending on the location within the disc, likely due to the unequal distribution of weight throughout the disc. These suggested patterns are not conclusions because the study population is small and the complexities of the load bearing are great. Further studies are warranted to see if these ideas hold.

Using a compression device increased the coefficient of variation of $T_{1\rho}$ values between repeated measurements to 13.9% compared with 7.6% for uninterrupted supine relaxation. There was a significant increase ($p < 0.05$) in $T_{1\rho}$ when compression was released. Changes were not significant during an hour of supine resting, indicating that the pre-scan compression from normal daily activities does not considerably complicate $T_{1\rho}$ measurements.

These experiments also demonstrate that the registration algorithm can be used to mea-

sure changes in disc height. Points indicating the top and bottom of the disc in the anterior, central, and posterior regions of the disc are manually selected on one of the images and transformed to each of the other images via the transformation determined by the registration algorithm. When compared with the measurement method of manually placing points on each image and comparing the disc height from each pair, the automatic method is faster, more convenient, and potentially more accurate. It is faster because only one set of points needs to be manually placed. The entirely manual method requires a very experienced operator to be able to identify the same anatomical points in each image, so the automatic version is more convenient because it can be done by a less-trained operator still with reasonable accuracy. Also, it is possible that the vertebral endplates do not have sharp enough landmarks to place three sets of repeatably-identifiable points, even by a trained operator, in which case the automatic method may be more accurate.

We also note that while the examples in this work all deal with $T_{1\rho}$ in varying states of compression, the registration could be used for any other studies of $T_{1\rho}$ that require two or more images of the same person, not just in a compression context. Also, the algorithm could be used with other quantitative parameters. For example, we show a successful example of registration of T_2 maps in Appendix A, Figure A.6.

In summary, we have developed an automatic method of comparing $T_{1\rho}$ values in intervertebral discs between different images of the same person, intended for investigating the behavior of $T_{1\rho}$ when the disc is under various types of loads. The method includes registration to deal with spinal movement, including large intentional position changes,

small unintentional movement, field of view repositioning, and exams on different days. Understanding $T_{1\rho}$ behavior under load could provide insight into the reason for the correlation between low $T_{1\rho}$ and disc degeneration, could help people interpret or control for compression-related changes when using $T_{1\rho}$ to evaluate disc health, and could provide another measure in itself for assessing disc health.

Further work   We recommend that future studies look further at the tendency of $T_{1\rho}$ to increase under compression and decrease under extension / relaxation. We suggest using more volunteers, more controlled loading methods, and including volunteers with documented disc degeneration. We also propose comparing $T_{1\rho}$ values between morning and evening measurements since disc compression as well as other diurnal differences may produce discernible tendencies in $T_{1\rho}$.

Bibliography

[1] H.-R. Guo, S. Tanaka, W. E. Halperin, and L. L. Cameron, "Back pain prevalence in US industry and estimates of lost workdays," *American Journal of Public Health*, vol. 89, no. 7, pp. 1029–1035, 1999.

[2] P. M. Kent and J. L. Keating, "The epidemiology of low back pain in primary care," *Chiropractic and Osteopathy*, vol. 13, 2005.

[3] C. Leboeuf-Yde and K. Kyvik, "At what age does low back pain become a common problem? a study of 29,424 individuals aged 12-41 years," *SPINE*, vol. 23, no. 2, pp. 228–234, 1998.

[4] X. Li, E. T. Han, C. B. Ma, T. M. Link, D. C. Newitt, and S. Majumdar, "In vivo 3T spiral imaging based multi-slice $t_{1\rho}$ mapping of knee cartilage in osteoarthritis," *Magnetic Resonance in Medicine*, vol. 54, pp. 929–936, 2005.

[5] J. D. Auerbach, W. Johannessen, A. Borthakur, A. J. Wheaton, C. A. Dolinskas, R. A. Balderston, R. Reddy, and D. M. Elliott, "In vivo quantification of human lumbar disc degeneration using $t_{1\rho}$-weighted magnetic resonance imaging," *Eur Spine J*, vol. 15, no. Suppl. 3, pp. S338–S344, 2006.

[6] A. Borthakur, E. Mellon, S. Niyogi, W. Witschey, J. B. Kneeland, and R. Reddy, "Sodium and $t_{1\rho}$ MRI for molecular and diagnostic imaging of articular cartilage," *NMR in Biomedicine*, vol. 19, pp. 781–821, 2006.

[7] W. Johannessen, J. D. Auerbach, A. J. Wheaton, A. Kurji, A. Borthakur, R. Reddy, and D. M. Elliott, "Assessment of human disc degeneration and proteoglycan content using $t_{1\rho}$-weighted magnetic resonance imaging," *SPINE*, vol. 31, no. 11, pp. 1253–1257, 2006.

[8] X. Li, E. T. Han, T. M. Link, and S. Majumdar, "In vivo MR t1rho mapping in lumbar vertebral discs at 3T: a preliminary study," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, (Seattle, USA), May 2006.

[9] E. J. Chiu, D. C. Newitt, M. R. Segal, S. S. Hu, J. C. Lotz, and S. Majumdar, "Magnetic resonance imaging measurement of relaxation and water diffusion in the hu-

man lumbar intervertebral disc under compression in vitro," *SPINE*, vol. 26, no. 19, pp. E437–E444, 2001.

[10] T. C. Dunn, Y. Lu, H. Jin, M. D. Ries, and S. Majumdar, "T2 relaxation time of cartilage at MR imaging: comparison with severity of knee osteoarthritis," *Radiology*, vol. 232, no. 2, pp. 592–598, 2004.

[11] V. Haughton, "Imaging intervertebral disc degeneration," *Journal of Bone and Joint Surgery*, vol. 88, pp. 15–20, 2006.

[12] J. Perry, V. Haughton, P. A. Anderson, Y. Wu, J. Fine, and C. Mistretta, "The value of t2 relaxation times to characterize lumbar intervertebral disks: preliminary results," *Am J Neuroradiol*, pp. 337–342, February 2006.

[13] D. Burstein, A. Bashir, and M. L. Gray, "MRI techniques in early stages of cartilage disease," *Investigative radiology*, vol. 35, no. 10, pp. 622–638, 2000.

[14] S. M. Kealey, T. Aho, D. Delong, D. P. Barboriak, J. M. Provenzale, and J. D. Eastwood, "Assessment of apparent diffusion coefficient in normal and degenerated intervertebral lumbar disks: initial experience," *Radiology*, vol. 235, no. 2, pp. 569–574, 2005.

[15] R. R. Regatte, S. V. S. Akella, A. Borthakur, J. B. Kneeland, and R. Reddy, "In vivo proton MR three-dimensional $T1\rho$ mapping of human articular cartilage: Initial experience," *Radiology*, vol. 229, pp. 269–274, 2003.

[16] R. R. Regatte, S. V. S. Akella, J. H. Lonner, J. B. Kneeland, and R. Reddy, "T1rho relaxation mapping in human osteoarthritis (OA) cartilage: Comparison of $T_{1\rho}$ with T_2," *Journal of Magnetic Resonance Imaging*, vol. 23, pp. 547–553, 2006.

[17] R. Regatte, S. Akella, A. Borthakur, and R. Reddy, "High resolution $t_{1\rho}$ relaxation and dispersion imaging of the intervertebral disc," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, (Kyoto, Japan), p. 1544, May 2004.

[18] G. Blumenkrantz, X. Li, E. T. Han, D. C. Newitt, J. C. Crane, T. M. Link, and S. Majumdar, "A feasibility study of in vivo $t_{1\rho}$ imaging of the intervertebral disc," *Magnetic Resonance Imaging*, vol. 24, pp. 1001–1007, 2006.

[19] N. Boos, A. Wallin, T. Gbedegbegnon, M. Aebi, and C. Boesch, "Quantitative MR imaging of lumbar intervertebral disks and vertebral bodies: Influence of diurnal water content variations," *Radiology*, vol. 188, pp. 351–354, 1993.

[20] R. R. Regatte, J. H. Kaufman, E. A. Noyszewski, and R. Reddy, "Sodium and proton MR properties of cartilage during compression," *J Magnetic Resonance Imaging*, vol. 10, pp. 961–967, 1999.

[21] H.-J. Wilke, P. Neef, M. Caimi, T. Hoogland, and L. E. Claes, "New in vivo measurements of pressures in the intervertebral disc in daily life," *SPINE*, vol. 24, no. 8, pp. 755–762, 1999.

[22] J. P. G. Urban, S. Roberts, and J. R. Ralphs, "The nucleus of the intervertebral disc from development to degeneration," *Amer Zool*, vol. 40, pp. 53–61, 2000.

[23] W. C. Hutton, Y. Toribatake, W. A. Elmer, T. M. Ganey, K. Tomita, and T. E. Whitesides, "The effect of compressive force applied to the intervertebral disc in vivo: A study of proteoglycans and collagen," *Spine*, vol. 23, no. 23, pp. 2524–2537, 1998.

[24] W. C. Hutton, T. M. Ganey, W. A. Elmer, E. Kozlowska, J. L. Ugbo, E.-S. Doh, and T. E. W. Jr., "Does long-term compressive loading on the intervertebral disc cause degeneration?," *SPINE*, vol. 25, no. 23, pp. 2993–3004, 2000.

[25] J. C. Lotz and J. R. Chin, "Intervertebral disc cell death is dependent on the magnitude and duration of spinal loading," *SPINE*, vol. 25, no. 12, pp. 1477–1483, 2000.

[26] J. A. Malko, W. C. Hutton, and W. A. Fajman, "An in vivo MRI study of the changes in volume (and fluid content) of the lumbar intervertebral disc after overnight bed rest and during and 8-hour walking protocol," *Journal of spinal disorders and techniques*, vol. 15, no. 2, pp. 157–163, 2002.

[27] N. Roberts, D. Hogg, G. H. Whitehouse, and P. Dangerfield, "Quantitative analysis of diurnal variation in volume and water content of lumbar intervertebral discs," *Clinical Anatomy*, vol. 11, pp. 1–8, 1998.

[28] O. Karakida, H. Ueda, M. Ueda, and T. Miyasaka, "Diurnal T2 value changes in the lumbar intervertebral discs," *Clinical Radiology*, vol. 58, pp. 389–392, 2003.

[29] Various.
http://en.wikipedia.org/wiki/Anatomical_terms_of_location.

[30] Various. http://en.wikipedia.org/wiki/Vertebral_column.

[31] H. Gray, "Henry gray's anatomy of the human body, illustration 301."
http://www.bartleby.com/107/illus301.html [accessed 2007], 1918.

[32] H. Gray, "Anatomy of the human body, illustration 94."
http://www.bartleby.com/107/illus94.html [accessed 2007], 1918.

[33] J. P. G. Urban and S. Roberts, "Review: Degeneration of the intervertebral disc," *Arthritis Research and Therapy*, vol. 5, no. 3, pp. 120–130, 2003.

[34] W. Johannessen and D. M. Elliott, "Effects of degeneration on the biphasic material properties of human nucleus pulposus in confined compression," *Spine*, vol. 30, no. 24, pp. E724–E729, 2005.

[35] H. Grohn, *A study of $T_{1\rho}$ relaxation: From relaxation mechanisms to the magnetic resonance imaging contrast.* PhD thesis, University of Kiupio, Finland, June 2003.

[36] R. Sepponen, "Rotating frame and magnetization transfer," in *Magnetic Resonance Imaging* (D. D. Stark and W. G. Bradley, eds.), vol. 1, ch. 8, Mosby-Year Book, Inc., second edition ed., 1992.

[37] A. Virta, M. Komu, and M. Kormano, "$T_{1\rho}$ of protein solutions at very low fields: Dependence on molecular weighte, concentration, and structure," *Magnetic Resonance in Medicine*, vol. 37, no. 1, pp. 53–57, 1997.

[38] U. Duvvuri, A. D. Goldberg, J. K. Kranz, L. Hoang, R. Reddy, F. W. Wehrli, A. J. Wand, S. W. Englander, and J. S. Leigh, "Water magnetic relaxation dispersion in biological systems: The contribution of proton exchange and implications for the nonivansive detection of cartilage degradation," *Proceedings of the National Acadamy of Sciences*, vol. 98, no. 22, pp. 12479–12484, 2001.

[39] P. J. Roughley, "Biology of intervertebral disc aging and degeneration: involvement of the extracellular matrix," *Spine*, vol. 29, no. 23, pp. 2691–2699, 2004.

[40] R. R. Regatte, S. V. S. Akella, A. Borthakur, and R. Reddy, "Proton spin-lock ratio imaging for quantitation of glycosaminoglycans in articular cartilage," *Journal of Magnetic Resonance Imaging*, vol. 17, pp. 114–121, 2003.

[41] A. J. Wheaton, F. L. Casey, A. J. Gougoutas, G. R. Dodge, A. Borthakur, J. H. Lonner, H. R. Schumacher, and R. Reddy, "Correlation of $T_{1\rho}$ with fixed charge density in cartilage," *Journal of Magnetic Resonance Imaging*, vol. 20, pp. 519–525, 2004.

[42] N. M. Menezes, M. L. Gray, J. R. Hartke, and D. Burstein, "T2 and T1rho MRI in articular cartilage systems," *Magnetic Resonance in Medicine*, vol. 51, no. 3, pp. 503–509, 2004.

[43] A. J. Wheaton, G. R. Dodge, D. M. Elliott, S. B. Nicoll, and R. Reddy, "Quantification of cartilage biomechanical and biochemical properties via $T_{1\rho}$ magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 54, no. 5, pp. 1087–1093, 2005.

[44] A. Saifuddin, S. Blease, and E. MacSweeney, "Axial loaded MRI of the lumbar spine," *Clinical Radiology*, vol. 58, pp. 661–671, 2003.

[45] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[46] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physica in medicine and biology*, vol. 46, pp. R1–R45, 2001.

[47] K. Rohr, "Elastic registration of multimodal medical images: A survey," *Künstliche Intelligenz*, vol. 00, no. 3, pp. 11–17, 2000.

[48] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[49] B. P. Rogers, V. M. Haughton, K. Arfanakis, and M. E. Meyerand, "Application of image registration to measurement of intervertebral rotation in the lumbar spine,"

[50] J. A. Little, D. L. G. Hill, and D. J. Hawkes, "Deformations incorporating rigid structures," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 223–232, 1997.

[51] R. H. Huesman, G. J. Klein, J. A. Kimdon, C. Kuo, and S. Majumdar, "Deformable registration of multi-modal data including rigid structures," *IEEE Transactions on Nuclear Science*, vol. 50, no. 3, pp. 389–392, 2003.

[52] K. Wang, Y. He, and H. Qin, "Incorporating rigid structures in non-rigid registration using triangular B-splines," *Lecture Notes in Computer Science*, vol. 3752, pp. 295–306, 2005.

[53] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[54] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A physics-based coordinate transformation for 3-D image matching," *IEEE Transactions on medical imaging*, vol. 16, no. 3, pp. 317–328, 1997.

[55] M. Fornefett, K. Rohr, R. Sprengel, and H. S. Stiehl, "Elastic medical image registration using orientation attributes at landmarks," in *Proc. Medical Image Understanding and Analysis*, pp. 49–52, 1998.

[56] K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," *IEEE Transactions on Medical Imaging*, vol. 20, no. 6, pp. 526–534, 2001.

[57] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. V24, no. 2, pp. 137–154, 1997.

[58] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *Medical Imaging, IEEE Transactions on*, vol. 16, no. 2, pp. 187–198, 1997.

[59] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Transactions on Medical Imaging*, vol. 22, no. 1, 2003.

[60] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*. Kitware, Inc. ISBN 1-930934-10-6, http://www.itk.org/ItkSoftwareGuide.pdf, first ed., 2003.

[61] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of mutliresolution optimization strategies for multimodality image registration by maximization of mutual information," *Medical image analysis*, vol. 3, no. 4, pp. 373–386, 1999.

[62] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual information matching in multiresolution contexts," *Image and vision computing*, vol. 19, pp. 45–52, 2001.

[63] D. L. Pham, C. Xu, and J. L. Prince, "A survey of current methods in medical image segmentation," *Annual review of biomedical engineering*, vol. 2, pp. 315–338, 2000.

[64] J. A. Sethian, "Theory, algorithms, and applications of level set methods for propagating interfaces," in *Acta Numerica* (A. Iserles, ed.), Cambridge University Press, 1996.

[65] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[66] J. L. Simon, *Resampling: The new statistics*. Resampling Stats, Inc., 1997.

[67] C. Brunsdon, S. Fotheringham, and M. Charlton, "Geographically weighted regression – modelling spatial non-stationarity," *The Statistician*, vol. 47, no. 3, pp. 431–443, 1998.

[68] A. S. Fotheringham, M. Charlton, and C. Brunsdon, *Geographically weighted regression: The analysis of spatially varying relationships*. John Wiley and Sons, 2002.

[69] S. Rajasekaran, J. N. Babu, R. Arun, B. R. W. Armstrong, A. P. Shetty, and S. Murugan, "A study of diffusion in human lumbar discs: A serial magnetic resonance imaging study documenting the influence of the endplate on diffusion in normal and degenerate discs," *SPINE*, vol. 29, no. 23, pp. 2654–2667, 2004.

[70] A. Borthakur, A. Wheaton, S. R. Charagundla, E. M. Shapiro, R. R. Regatte, S. V. S. Akella, J. B. Kneeland, and R. Reddy, "Three-dimensional T1rho-weighted MRI at 1.5 Tesla," *Journal of Magnetic Resonance Imaging*, vol. 17, pp. 730–736, 2003.

[71] J. P. Hornak, "The basics of MRI." http://www.cis.rit.edu/htbooks/mri/ [accessed 2007], 1996.

[72] E. M. Marieb and J. Mallatt, *Human anatomy*. Pearson Education, Inc. as Benjamin Cummings, 2001.

[73] Various.
http://en.wikipedia.org/wiki/Cartilage#Types_of_cartilage.

Appendix A

Additional examples

We present several image pairs used for testing the registration accuracy that were not in the context of the experiments in Chapter 7. We show the difference between the fixed and moving images before and after registration to give a qualitative measure of the registration accuracy for these cases.
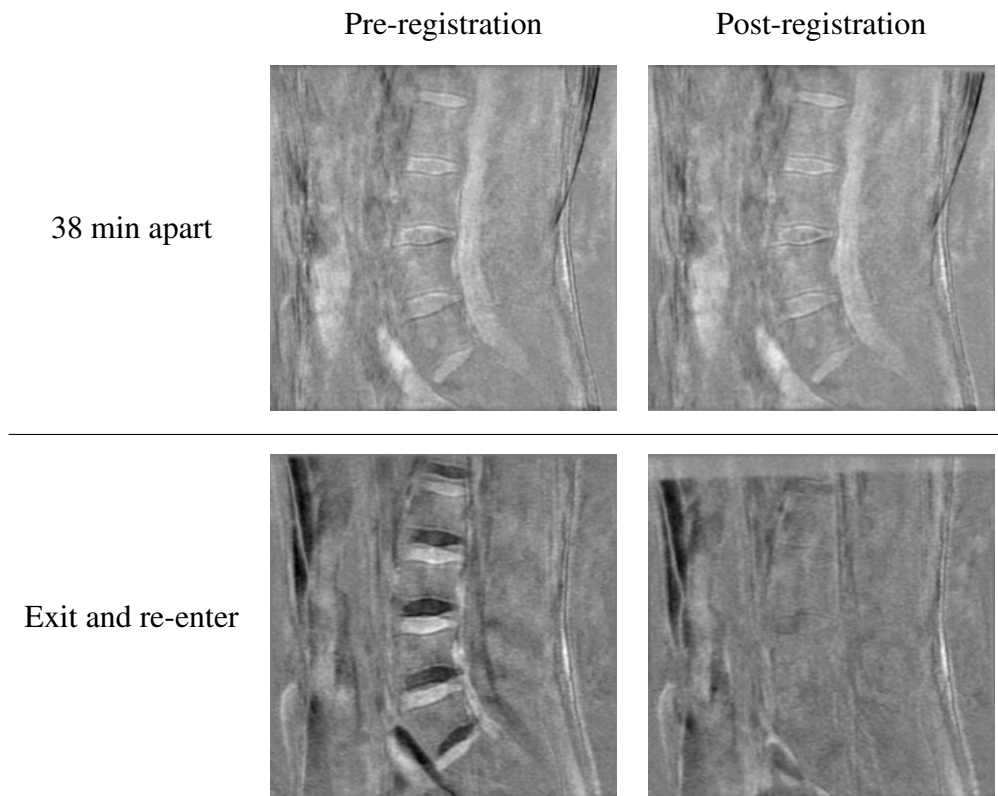
Pre-registration   Post-registration

38 min apart

Exit and re-enter

Figure A.1: Volunteer J. The fixed and moving images compared in the top row were separated in time, without any intentional volunteer movement. There is not much initial misalignment in this case. The bottom row compares images where the volunteer exited the scanner between scans. The prospective registration method found the correct slice to image when the volunteer re-entered the scanner. There is significant initial in-plane misalignment, which is corrected during registration.

Pre-registration   Post-registration
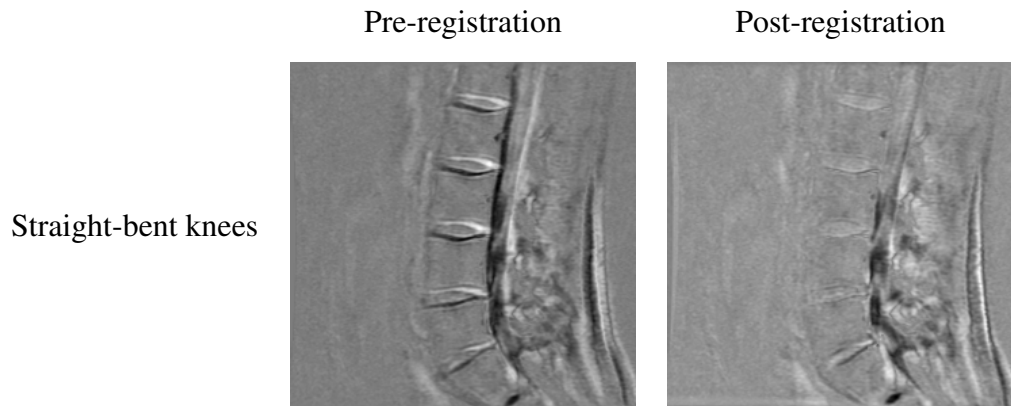
Straight-bent knees

Figure A.2: Volunteer K. For the fixed image, the volunteer has his/her knees straight, lying flat on the scanner bed. For the moving image, the knees are propped up by a pillow in a bent position. The initial misalignment is mostly corrected by the registration program, but some misalignment remains at the very bottom of the image. It is more difficult for the registration to work at the edges of the image, especially when a transformation causes the bounding box to exit the image. Nevertheless, the alignment is good, and the dark and light areas that remain after registration are mostly due to differences between the two images rather than misalignment. We did not use prospective registration here, so the slices are not exactly the same in the fixed and moving images.

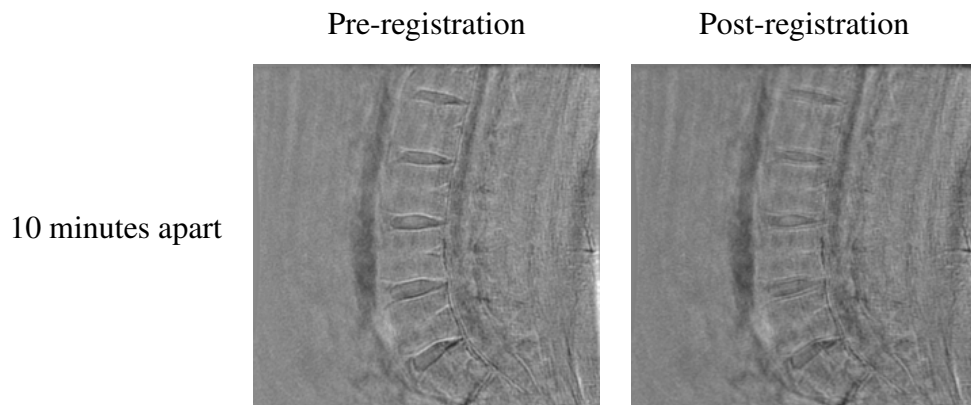Pre-registration   Post-registration

10 minutes apart

Figure A.3: Volunteer L. The moving image was acquired about 10 minutes after the fixed image, with no other change. There is only slight initial misalignment, mostly near the bottom of the image, which is improved after registration.

Pre-registration   Post-registration

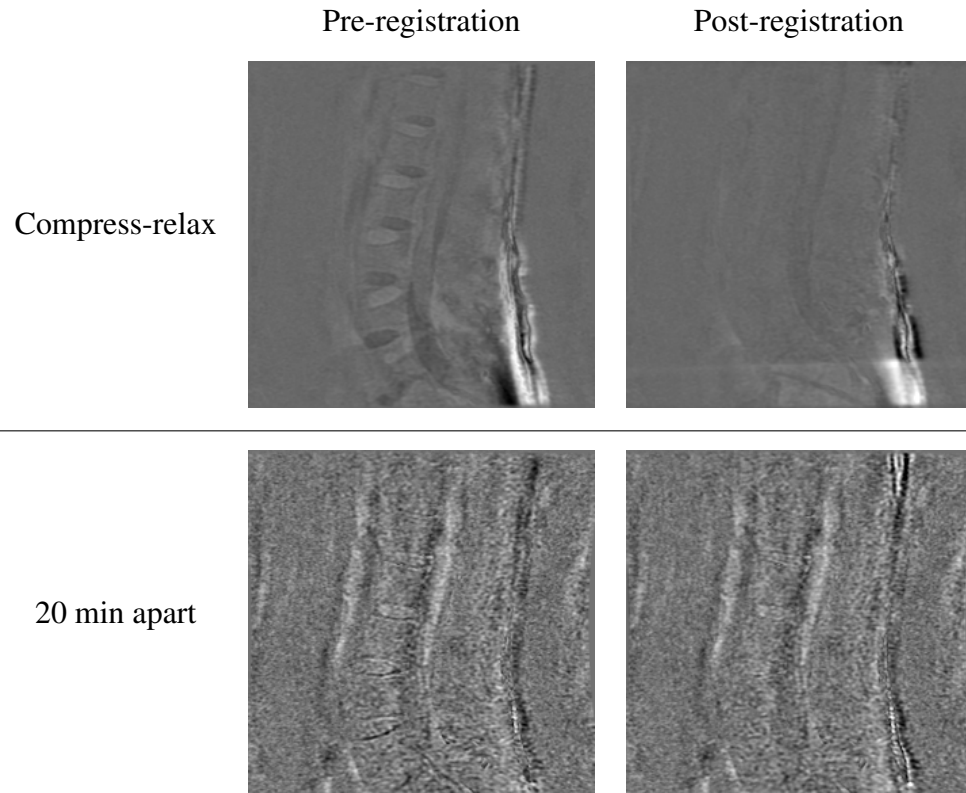
Compress-relax

20 min apart

Figure A.4: Volunteer M. This is another example using the compression device. The top row demonstrates compression (fixed image) compared with relaxation (moving image), and the bottom row compares two images taken 20 minutes apart, relaxing without weights during and between the acquisitions. This example is not included in the experiment in Chapter 7 because the imaging sequence was not working properly, so the $T_{1\rho}$ values are not reliable. However, it is still a useful case to examine the performance of the registration portion of the algorithm. Note that the misalignment was so large in the first row that the user needed to set a coarse initial global translation. Setting this initial translation is simple with our user interface.

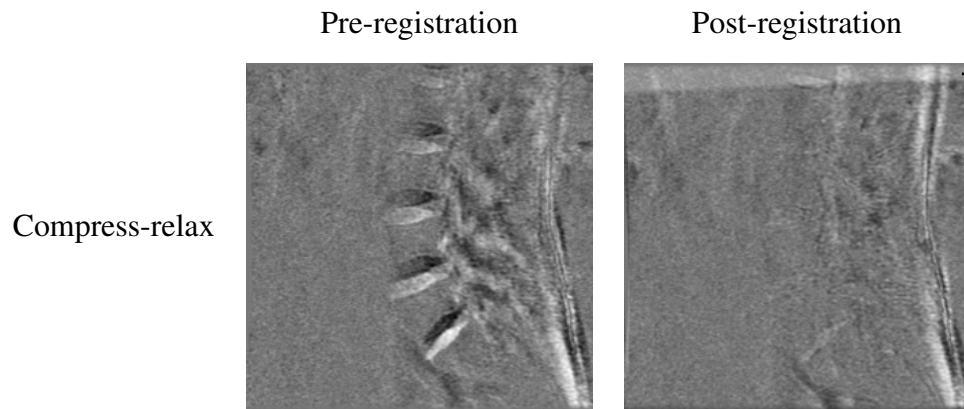Pre-registration   Post-registration

Compress-relax

Figure A.5: Volunteer N. Here is another example of compression/relaxation using the compression device. This case is not included in Chapter 7 because the $T_{1\rho}$ values are unreliable due to low signal from the receiver coil. We include it here as another example of the registration algorithm.
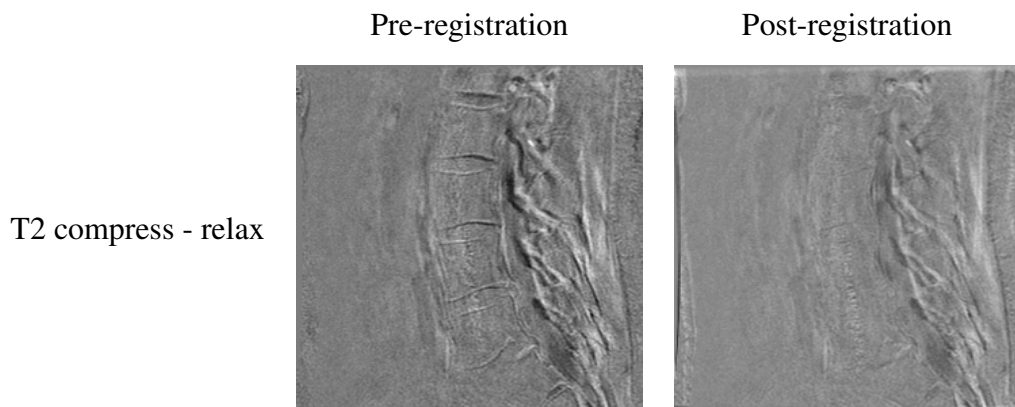
Pre-registration   Post-registration

T2 compress - relax

Figure A.6: Volunteer O. In this example, we register T_2 maps to demonstrate flexibility in analyzing quantitative MR parameters other than $T_{1\rho}$. For the fixed image, the volunteer sustained weight in the compression device (as in Chapter 7). The weight was removed for the moving image.