

**QUANTIFICATION AND SPATIOTEMPORAL
MODELING OF TISSUE APPEARANCE
IN MR (MAGNETIC RESONANCE)
STUDIES OF BRAIN
DEVELOPMENT**

by

Avantika Vardhan

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Bioengineering

The University of Utah

August 2016

Copyright © Avantika Vardhan 2016

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of Avantika Vardhan

has been approved by the following supervisory committee members:

<u>Guido Gerig</u>	, Chair	<u>10/27/2015</u> Date Approved
--------------------	---------	------------------------------------

<u>Marcel Prastawa</u>	, Co-Chair	<u> </u> Date Approved
------------------------	------------	--

<u>Preston Thomas Fletcher</u>	, Member	<u>10/28/2015</u> Date Approved
--------------------------------	----------	------------------------------------

<u>Edward Victor Rebok Dibella</u>	, Member	<u> </u> Date Approved
------------------------------------	----------	--

<u>Sarang Joshi</u>	, Member	<u>10/28/2015</u>
---------------------	----------	-------------------

<u>Robert S. MacLeod</u>	, Member	<u>10/28/2015</u> Date Approved
--------------------------	----------	------------------------------------

and by Patrick A. Tresco, Chair of
the Department of Bioengineering

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Magnetic Resonance (MR) is a relatively risk-free and flexible imaging modality that is widely used for studying the brain. Biophysical and chemical properties of brain tissue are captured by intensity measurements in T1W (T1-Weighted) and T2W (T2-Weighted) MR scans. Rapid maturational processes taking place in the infant brain manifest as changes in contrast between white matter and gray matter tissue classes in these scans. However, studies based on MR image appearance face severe limitations due to the uncalibrated nature of MR intensity and its variability with respect to changing conditions of scan. In this work, we develop a method for studying the intensity variations between brain white matter and gray matter that are observed during infant brain development. This method is referred to by the acronym WIVID (White-gray Intensity Variation in Infant Development). WIVID is computed by measuring the Hellinger Distance of separation between intensity distributions of WM (White Matter) and GM (Gray Matter) tissue classes. The WIVID measure is shown to be relatively stable to interscan variations compared with raw signal intensity and does not require intensity normalization.

In addition to quantification of tissue appearance changes using the WIVID measure, we test and implement a statistical framework for modeling temporal changes in this measure. WIVID contrast values are extracted from MR scans belonging to large-scale, longitudinal, infant brain imaging studies and modeled using the NLME (Nonlinear Mixed Effects) method. This framework generates a normative model of WIVID contrast changes with time, which captures brain appearance changes during neurodevelopment. Parameters from the estimated trajectories of WIVID contrast change are analyzed across brain lobes and image modalities. Parameters associated with the normative model of WIVID contrast change reflect established patterns of region-specific and modality-specific maturational sequences. We also detect differences in WIVID contrast change trajectories between distinct population groups. These groups are categorized based on sex and risk/diagnosis for ASD (Autism Spectrum Disorder). As a result of this work, the usage of the proposed WIVID contrast measure as a novel neuroimaging biomarker for characterizing tissue appearance is validated, and the clinical potential of the developed framework is demonstrated.

CONTENTS

ABSTRACT	iii
LIST OF TABLES	vii
LIST OF FIGURES	ix
ACKNOWLEDGMENTS	xii
CHAPTERS	
1. INTRODUCTION	1
1.1 Medical Imaging : An Introduction	1
1.2 A Brief Overview of Brain Imaging	4
1.3 Origins of MR Signal Intensity and Variations	8
1.4 WM-GM Intensity Variations in Neurodevelopment	11
1.4.1 Additional Sources of Intensity Variation	13
1.5 MR Studies of Early Brain Development	13
1.6 Appearance Studies in Neuroimaging	16
1.6.1 Relative Intensity-based MRI Studies	18
1.7 Goals of This Dissertation	19
1.7.1 Quantification of WM-GM Intensity Variations	20
1.7.2 Spatio-temporal Modeling of Intensity Variations	20
1.7.3 Modeling Multimodal WM-GM Intensity Variations	21
1.7.4 Spatial Heterogeneity in WM-GM Intensity Variations	21
1.7.5 Intensity Variation-based Infant Studies	22
1.8 Overview of Chapters	22
2. QUANTIFICATION OF WM-GM MR INTENSITY VARIATIONS ..	23
2.1 Methods to Study Appearance and Contrast	24
2.1.1 Contrast in Medical Imaging	28
2.2 Method	29
2.2.1 Framework for Infant Brain MRI Studies of Appearance	31
2.2.2 Joint 4D Registration-Segmentation Pipeline	31
2.2.2.1 Intrasubject Image Registration	33
2.2.2.2 Intersubject Image Registration	35
2.2.2.3 Longitudinal Image Segmentation	36
2.2.2.4 Brain Image Parcellation	37
2.2.3 Removal of Shape and Volume Information	39
2.2.3.1 Construction of Intensity Distributions	40
2.2.4 Computation of Distance Between Intensity Distributions	43
2.3 WIVID Measure: Directionality and Simulations	48

2.3.1	WIVID Measure Behavior for Shifted WM Distribution	48
2.3.2	WIVID Measure Directionality	49
2.3.3	Phantom Image Studies	51
2.4	Stability of WIVID Measure	52
2.4.1	Preprocessing Pipeline	54
2.4.2	Visual Inspection of Image Quality	57
2.4.3	Quantitative Analysis of WIVID	57
3.	STATISTICAL MODELING OF INTENSITY VARIABILITY IN DEVELOPING BRAIN MRI	61
3.1	Normative Models of Growth	61
3.2	Modeling Repeated Measures Data	63
3.2.1	Repeated Measures Data	63
3.2.2	Quantitative Analysis of Data Fit	65
3.2.3	Visual Analysis of Data Points	66
3.2.4	Nonparametric Models via Kernel Regression	67
3.2.5	Parametric Regression Models	69
3.2.6	Subject-Specific Modeling of Data	75
3.3	Mixed Effects Modeling	76
3.4	Nonlinear Mixed Effects Modeling	76
3.5	Extension of NLME to Multimodal Data	78
3.6	NLME Modeling on Synthetic Data	80
3.7	NLME: Analysis and Inferences	80
4.	TISSUE APPEARANCE ANALYSIS : EXPERIMENTS WITH IBIS (INFANT BRAIN IMAGING STUDY) DATA	83
4.1	Studies of Infant Brain Appearance Change	83
4.1.1	ACE-IBIS Study	85
4.1.2	MR Image Acquisition	86
4.1.3	Sample Used in Analysis	87
4.2	Computation of the WIVID Measure	87
4.3	Experiments for Finding the Best Model Fit	92
4.3.1	Choice of Parametric Fits for Modeling Data	92
4.3.2	Cross-Sectional Model for Fitting data	93
4.3.3	Cross-Sectional vs. Mixed Effects Models	93
4.3.4	NLME Models for ACE-IBIS Data	95
4.3.5	Multivariate Versus Univariate Modeling	96
4.3.6	Other Factors to Consider in NLME Modeling	96
4.3.7	Experiments for Modeling ACE-IBIS Data	101
4.4	WIVID Contrast Analysis for ACE-IBIS Data	102
4.4.1	Visual Analysis of WIVID Data Points	102
4.4.2	WIVID Contrast Change: Intermodal Variations	105
4.4.3	WIVID Contrast Modeling: Interlobe Variations	107
4.4.4	WIVID Contrast Variations Based on Sex	109
4.4.5	Effect of ASD Risk/Diagnosis Category on WIVID	114

5. STUDY OF BRAIN TISSUE APPEARANCE IN INFANTS	122
5.1 Brain Appearance in the First Year After Birth	122
5.1.1 Five Time Point Data Under Study	123
5.1.2 Intrasubject Image Registration	123
5.2 Computation of the WIVID Measure	125
5.3 Results and Discussion	126
6. DISCUSSION AND CONCLUSIONS	129
6.1 Major Contributions	129
6.1.1 Development of WIVID Measure	131
6.1.2 Normative Trajectories of Appearance Change in Infants	133
6.1.3 Population-wise Differences in Appearance Change	135
6.2 Limitations	136
6.2.1 Inaccuracies in Image Processing Procedures	136
6.2.2 Intensity Inhomogeneities	137
6.2.3 Limitations: Longitudinal Data	137
6.2.4 Limitations: Variable Scanning Protocols	138
6.2.5 Nonspecificity of MR Image Appearance	138
6.3 Future Work	139
6.3.1 Robustness to Segmentation Errors	139
6.3.2 Regional WIVID Measure	140
6.3.3 Brain-Behavior Studies using WIVID	140
6.3.4 Neurobiology of WM-GM Contrast	141
6.3.5 WIVID Across Multiple Scanning Protocols	141
6.3.6 WIVID for Other Neurobiological Processes	142
6.3.7 Application of WIVID to Nonbrain Data	142
APPENDIX: INTENSITY NORMALIZATION FOR INFANT MRI	143
REFERENCES	145

LIST OF TABLES

3.1	Assessment of Linear and Nonlinear Cross-sectional Model Fits.	72
3.2	Comparison of Gompertz and Logistic Cross-Sectional fits with and without grouping based on risk/diagnosis category.	74
3.3	AIC values for choosing optimal configuration of random Effects for NLME model.	79
3.4	AIC comparisons for separate and joint fits of multivariate synthetic data. . .	81
3.5	Multivariate NLME: Estimation of growth parameters and hypothesis testing.	81
4.1	AIC Values for Comparison of Cross-Sectional Model Fits to T1W WIVID Values from ACE-IBIS Data.	94
4.2	AIC Values for Comparison of Cross Sectional Model Fits to T2W WIVID Values from ACE-IBIS Data.	94
4.3	AIC Values for Comparison of NLME Model Fits to T1W WIVID Values From ACE-IBIS data.	97
4.4	AIC Values for Comparison of NLME Model Fits to T2W WIVID Values From ACE-IBIS Data.	97
4.5	AIC Values for Comparison of Univariate and Multivariate NLME Model Fits to T1W and T2W WIVID values From ACE-IBIS Data.	98
4.6	Parameters Corresponding to Multilevel NLME Logistic Modeling for T1W and T2W WIVID Values based on Healthy Male Controls.	108
4.7	Parameters from Multilevel NLME modeling of WIVID differences based on male/female categorization for T1W scans.	113
4.8	Parameters from Multilevel NLME modeling of WIVID differences based on male/female categorization for T2W scans.	113
4.9	Results from hypothesis testing for WIVID trajectory differences based on male/female categorization for T1W scans.	115
4.10	Results from hypothesis testing for WIVID trajectory differences based on male/female categorization for T2W scans.	115
4.11	Parameters from Multilevel NLME modeling of WIVID differences based on risk/diagnosis group for ASD for T1W scans.	118
4.12	Parameters from Multilevel NLME modeling of WIVID differences based on risk/diagnosis group for ASD for T2W scans.	119
4.13	Results from hypothesis testing of WIVID curve parameters based on risk/diagnosis group for ASD for T1W scans.	121

4.14 Results from hypothesis testing of WIVID curve parameters based on risk/diagnosis group for ASD for T2W scans.	121
---	-----

LIST OF FIGURES

1.1 Longitudinal, multimodal, series of infant brain images.	3
1.2 Medical imaging modalities.	5
1.3 Basic principles of MRI.	6
1.4 Origin of the MR signal and intertissue intensity variations.	9
1.5 Changing T1 and T2 relaxation values in the developing infant brain.	12
2.1 The Concept Of Contrast : Phantom Image Series	30
2.2 Framework outlining the registration-segmentation pipeline as well as the procedure to perform statistical analysis on the contrast.	32
2.3 Framework outlining the intrasubject and intersubject registration pipeline. . .	34
2.4 Framework outlining the longitudinal segmentation and parcellation pipeline. .	38
2.5 Illustration of Kernel Density Estimation Procedure for Generating Intensity Distributions.	42
2.6 WM and GM Intensity Distributions for a Longitudinal Series of T1W and T2W MR Scans.	43
2.7 Assigning Directionality to WIVID: Configurations of WM and GM Intensity Distributions.	49
2.8 Decreasing (Negative) and Increasing (Positive) Components of WIVID Change Curve During Infant Brain Development.	49
2.9 Longitudinal Series of phantom Images.	52
2.10 Magnitude and Direction of WIVID values for a Series of phantom Images. . .	52
2.11 Traveling phantom dataset scans: Phantom 1.	56
2.12 Trio scans of Phantom 1 from multiple scan locations.	58
2.13 T1W Allegra scans of Phantom 1 from multiple coils.	58
2.14 Stability Analysis of WIVID and signal intensity: COV (Coefficient of Varia- tion) Plots by Region.	60
3.1 Motivation for establishing normative models of changes in neurodevelopmen- tal biomarkers.	62
3.2 Model-selection process for longitudinal neuroimaging WIVID data.	64
3.3 Cross-sectional modeling of changes in WIVID contrast measure.	68
3.4 Kernel regression of WIVID data.	69

3.5	Cross-sectional fits to WIVID data.	71
3.6	Logistic cross-sectional fit with grouping of parameters based on ASD risk/diagnosis.	74
3.7	Subject-specific fits to WIVID data.	76
3.8	Multivariate modeling experiments on synthetic data.	81
4.1	Example of longitudinal, multimodal ACE-IBIS data of a single subject.	88
4.2	Intrasubject registration of ACE-IBIS data.	89
4.3	Longitudinal segmentation of subject-specific scans from ACE-IBIS dataset.	89
4.4	Cortical Parcellation of scans from ACE-IBIS dataset.	90
4.5	Intensity distributions of WM and GM for a longitudinal series of MR scans of a single subject from birth to 2 years of age.	91
4.6	Exploration of signed/unsigned representation of WIVID contrast measure.	91
4.7	Contrast computation with and without direction, and associated WIVID values.	100
4.8	WIVID contrast values across controls, for major lobes in the left hemisphere.	103
4.9	WIVID contrast values across controls, for major lobes in the right hemisphere.	104
4.10	Multilevel NLME modeling of WIVID contrast change in healthy male controls.	106
4.11	Plots of parameters corresponding to Multilevel NLME Logistic Modeling for T1W and T2W WIVID values based on healthy male controls.	108
4.12	Multilevel NLME modeling of WIVID differences based on male/female categorization for T1W scans.	110
4.13	Multilevel NLME modeling of WIVID differences based on male/female categorization for T2W scans.	111
4.14	Parameters from Multilevel NLME modeling of WIVID differences based on male/female categorization.	112
4.15	Multilevel NLME modeling of WIVID differences based on risk/diagnosis group for ASD for T1W scans.	116
4.16	Multilevel NLME modeling of WIVID differences based on risk/diagnosis group for ASD for T2W scans.	117
4.17	Parameters from Multilevel NLME modeling of WIVID differences based on risk/diagnosis group for ASD.	118
5.1	Sample scans from serial, multimodal infant brain images along with their tissue segmentations. Note: These scans are yet to be co-registered.	124
5.2	Multimodal, repeated measures dataset: Co-registered series of longitudinal images from a single subject taken at five time points.	125
5.3	Atlas-based parcellation of multimodal data.	125
5.4	Temporal modeling of WIVID values from birth to 2 years of age in the right frontal lobe: Individual subject-specific trajectories.	127
5.5	Temporal modeling of WIVID values from birth to 2 years of age in the right frontal lobe: Average population trajectory.	127

5.6 Temporal modeling of T1W (top), and T2W (bottom) WIVID values from birth to 2 years of age in the right frontal lobe : Average population trajectories across major brain lobes.	128
---	-----

ACKNOWLEDGMENTS

To begin with, this work would not have been possible without the support, insight, and encouragement provided by my advisor, Dr. Guido Gerig, and my co-advisor, Dr. Marcel Prastawa. Guido has given me great freedom during my PhD studies, and always encouraged me to chalk out my own path of interest. His expertise in exploring the intersections between different ideas such as early development and image processing has played a central role in this work. Marcel was instrumental in shaping the core ideas behind this dissertation at the earliest stages. The mathematical framework behind the measures developed is in large part due to his guidance. I also thank my committee members - Dr. Ed DiBella, Dr. Tom Fletcher, Dr. Sarang Joshi, and Dr. Rob MacLeod. Each played a unique role in shaping my PhD. Dr. Ed DiBella contributed a great deal to the central idea of studying MR appearance, with his expertise in MR acquisition and scanning technology. Our meetings provided me with valuable insight beyond what I could find after studying several MR textbooks. Dr. Tom Fletcher was a source of encouragement and enthusiasm, and gave me important suggestions related to the statistical aspect of this work. His discussions on longitudinal modeling have been critical to forming the modeling aspect of this project. I thank Dr. Sarang Joshi for his encouragement, for helping me ask the right scientific questions, and for exemplifying the spirit of scientific curiosity and enthusiasm. Dr. Rob MacLeod, who was a source of unfailing support and encouragement, shed considerable light on the foundations of research and the art of effectively communicating scientific principles. Early technical support provided by Sylvain Gouttard was invaluable, particularly in grasping proper usage of neuroimaging tools. Clement Vachet has helped optimize the image processing pipeline we employed and perfect this to the greatest possible extent. His dedication, patience, and attention to detail has shaped this work in innumerable ways. Finally, I thank my labmate Neda Sadeghi, who was generous with sharing both her time and knowledge of early brain imaging studies. I thank Neda for extensive informal discussions and her insights into early brain development, not to mention her friendship and several fun chats on innumerable topics. Other labmates on Guido's team, particularly Anuja Sharma and Bo Wang, have been constant supporters

on this journey. They were always ready to give suggestions regarding presentations and provide guidance related to research, and their encouraging presence in the hallways of the 4th floor will always be remembered.

I would also like to thank friends in Utah for providing the warm support of a family even though I was miles away from home. It is hard to capture easily the innumerable things they taught me during this journey of the PhD, but each of them really made this work possible. Some friends I would like to thank are Bindu, Prasanna, Gopal, Poonam, Shrikant, Joanita, Ksheeraja, Vishakha, Nachiket, Preethi, Nikhil, Nisha, Payal, Amrish, Uma, Radhika, Shravanthi, Raj, Shashi, Vimal, Amruta, Giri, Vellachi, Spurthy, Ganesh, Padma, Phani, and Sandeep. I would also like to thank two lifelong friends from outside Utah - Kurinji and Archana - for their constant support through these years.

Finally, I would like to thank my family who have been true pillars of support in completing this work. My family in the US has provided a home atmosphere whenever I needed a break, and direction in this journey. Aunts and uncles in India have supported and believed in me throughout, gently prodded me to complete my dissertation without delay :), and have themselves been outstanding examples of dedication. My cousins and cousins-in-law have been very encouraging, continually sent me good wishes, and shown me the meaning of excellence in education and in work. I thank my grandparents living in Chennai - it is insufficient to say that their blessings, well-wishes, and friendship have been invaluable on this journey. I also deeply thank my grandparents who are no more, who were a vital part of my childhood, and whose blessings are present everyday. I thank my mother-in-law, father-in-law, and brother-in-law for their strong interest in my education and their encouragement in this direction. I thank my mother, father, and husband for being an integral part of this endeavour - it would have truly been impossible to complete this without them, and any description of this would fall short. Finally, I thank my Guru - the source of everything positive in this work, without any doubt. His grace is simply immeasurable and cannot be captured in words.

I thank the reader of this work. I strongly believe that research in the field of brain disorders such as autism is the need of the hour and hope to have made some contribution in this direction.

CHAPTER 1

INTRODUCTION

1.1 Medical Imaging : An Introduction

Advancements in the field of medical imaging have provided both researchers and physicians with an unprecedented “inside view” of the human body. As a result, several questions of both a scientific and diagnostic nature can now be clearly answered. Some of these critical questions are related to shapes of anatomical regions and their physiological functions, biophysical and chemical properties of tissues, the detection of abnormalities such as lesions, and tracking of disease progression. In general, medical imaging helps answer these questions without being impeded by the procedural risks and inconveniences of invasive techniques. Tissues, mechanisms, and processes, which are not visible to the naked eye even during surgery, can be clearly observed using medical imaging techniques. In the light of advances over the last century, it would surely not be untrue to declare that medical imaging has transformed the very manner in which we understand ourselves.

The major medical imaging modalities include X-ray, CT (Computed Tomography), PET (Positron Emission Tomography), SPECT (Single Photon Emission Computed Tomography), and MRI (Magnetic Resonance Imaging) [1]. Each imaging modality provides unique information about specific bodily tissues, structures, and processes. The scans generated by some of these imaging technologies are capable of displaying high levels of anatomical detail, clearly defined biological structures, and sometimes even dynamic physiological changes. In addition, they can capture intensity inhomogeneities indicative of underlying tissue properties.

Information that is present in medical images, beyond visual observation, can be extracted using advanced image processing and machine learning algorithms. Volume, shape, microstructure, and image features are examples of some quantitative medical imaging measures. When used in conjunction with biostatistical and information processing methods, these quantitative measures are valuable tools for Computer Aided Diagnosis and imaging-based clinical research studies. Examples illustrating the effectiveness of this

quantitative approach include the automatic detection of lung nodules on chest radiographs and CT [2], MRI studies of hippocampal volumetry for diagnosis of AD (Alzheimer’s Disease) [3], and FDG (fluorodeoxyglucose-positron)-PET imaging for the detection of aortitis and large-vessel vasculitis [4, 5]. In this work, we too use a quantitative, interdisciplinary approach to explore the question of effectively studying intensity variations that constitute appearance of a medical image.

Appearance of an image can be understood as a synthesis of two characteristics: 1) shapes of structures and 2) patterns of intensity variations *within* the image [6]. While current medical imaging techniques include sophisticated tools to study the shapes of biological structures, the aspect of intensity variations is relatively less explored. Intensity variations in medical images can provide vital biological information, which can be seen particularly in MRI. MR images display intensity variations within and across scans of a subject, reflective of changes in properties of underlying tissues. However, the lack of calibration in nonquantitative MRI and variability due to scanning conditions create significant variability of signal intensity *between* scans, independent of changes in inherent properties of the scanned region. Additionally, the presence of an MR bias field causes intensity inhomogeneities *within* a single image. These issues make MRI an especially interesting platform upon which new methods for quantification of intensity variations may be tested.

MRI, being a relatively risk-free and flexible modality, is widely used for imaging the brain. The two major brain tissues, White Matter (WM) and Gray Matter (GM), have dissimilar intensities when imaged using MRI. This intensity gradient between WM and GM constitutes a major portion of the intensity variations seen in MR images. The primary goal of this work is to measure and analyze the intensity differences between WM and GM tissues in brain MRI, with applications to neurodevelopmental studies.

The intensity difference between WM and GM does not remain static over time. Instead, brain changes resulting in alterations in the structure and composition of cortical WM and GM manifest as changes in intertissue contrast, as observed in MRI. These brain changes can be attributed to natural processes related to neurodevelopment or aging, or other reasons such as progression of diseases, trauma, injury, infection, or the presence of psychiatric abnormalities. Hence, studying WM-GM intensity variation, within and across brain MR images, can provide vital clues for characterization of normal processes related to growth and aging, and for detection of abnormalities in these processes. Analysis of intensity differences between WM and GM in the developing brain might even have the potential

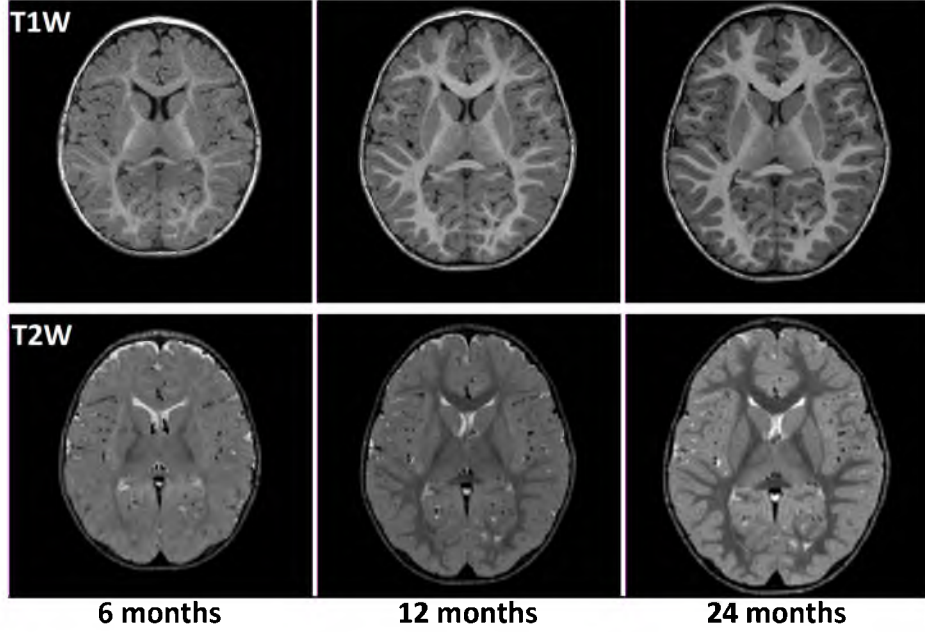


Figure 1.1: T1-Weighted (top) and T2-Weighted (bottom) MR images of the developing brain, scanned at 6 months (left column), 12 months (central column), and 24 months (right column) of age.

to provide a deeper understanding of the origin and basis of psychiatric and neurological disorders, subsequently enabling early diagnosis of these conditions. As an illustrative example, Figure 1.1 displays changes in WM-GM intensity differences during the course of normal neurodevelopment.

In this dissertation, we address questions related to intensity variations, specifically oriented towards analyzing intensity differences between WM and GM tissues in MR images. Some of these key questions are, “How can MR image appearance be effectively quantified in terms of relative WM-GM intensity variations *within* an MR scan?” and “Can we generate a normative model of changes in this quantified appearance measure *across* longitudinal series of scans, in order to characterize neurodevelopmental processes?”. To summarize, the broad goal of this work is effectively quantifying intensity variations between tissues *within* brain MR images, followed by creating a coherent spatiotemporal framework for modeling and analysis of this intensity variation *across* longitudinal neuroimaging datasets. Further, by application of this methodology to large-scale, longitudinal, imaging studies of the developing infant brain, we illustrate the substantial potential of spatiotemporal mapping of WM-GM intensity variations for making scientific and clinical inferences.

1.2 A Brief Overview of Brain Imaging

Within the extensive field of medical imaging, brain imaging is a relatively new discipline. Structural brain imaging investigates the structure of brain regions and tissues, and plays a key role in diagnosis of tumors, stroke, large-scale diseases, injury, and psychiatric disorders [7]. On the other hand, functional brain imaging relates to neural activity, connectivity between anatomical regions, and the role that different brain regions play in specific mental functions. Since its conception, brain imaging has continued to grow into a dynamic field, challenging everything from the way neurologists and neurosurgeons plan brain surgeries, to modern philosophies of the mind, consciousness, and perception [8].

From a clinical perspective, 10 highly important conditions in which brain imaging provides critical diagnostic aid are studied together as the “big 10.” These are 1) infarction, 2) hemorrhage, 3) infection, 4) tumor, 5) trauma, 6) dementia, 7) MS (Multiple Sclerosis), 8) epilepsy, 9) cranial neuropathy, and 10) neuro-ophthalmology [9]. In addition, measures derived from brain scans are used to assess changes accompanying neurological processes such as brain maturation, aging, and the progression of neurodegenerative disorders. Following recent advances, structural and functional brain imaging has even begun to play a vital role in the analysis of psychiatric conditions such as drug addiction, schizophrenia, bipolar disorder, and depression [10].

Histological studies were the earliest means used to actually observe the anatomical structure of the brain, and they remain highly informative, particularly in the detailed analysis of postmortem brains [11–13]. Recent studies of synaptogenesis have used advanced histological analysis based on semiquantitative blotting techniques for counting synaptic density [14]. Histological studies also serve as a reference for confirming neurobiological hypotheses generated from other types of imaging data.

Widely used imaging techniques applied to brain studies include CT, MRI, PET, SPECT, and ultrasound, as well as MR-related modalities such as DTI (Diffusion Tensor Imaging) and fMRI (functional Magnetic Resonance Imaging). Each modality conveys a specific type of information concerning the brain. Brain scans using some of these modalities can be seen in Figure 1.2.

A highly important modality for imaging the brain is CT. CT (interchangeably known as X-ray CT), which is based on reconstructing a 3D object from several 2D radiographs acquired at various angles, produces high-resolution brain images. CT can be applied for the diagnosis of headaches, abnormal development of the head or neck, hemorrhage or blood vessel abnormalities, cranio-facial trauma, stroke, fracture, and loss of sensory or motor

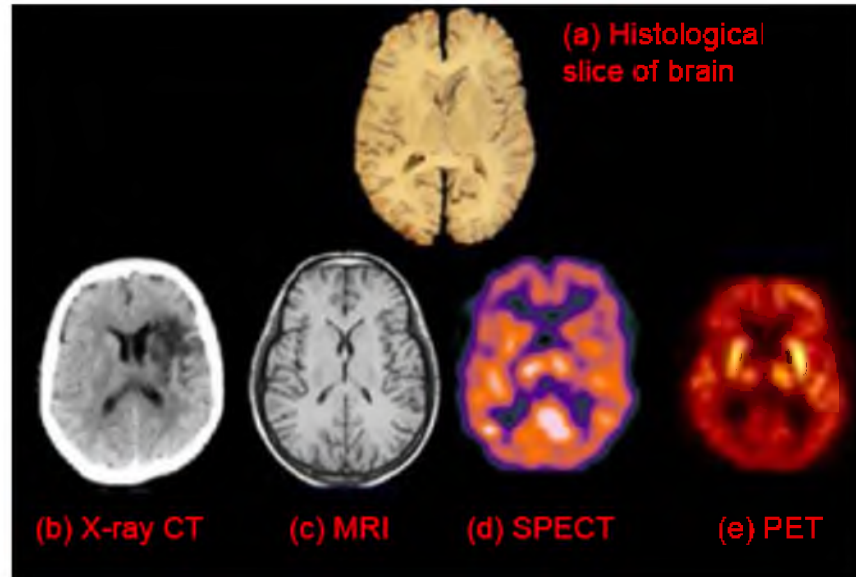


Figure 1.2: Multimodal imaging of the brain : (a) Histological study, (b) CT, (c) MRI, (d) SPECT, and (e) PET. Axial slices of the brain are displayed. Figure courtesy: Patrick J. Lynch and C. Carl Jaffe, Yale University, 2006.

function. While dramatic improvements in CT scanning speed and resolution have taken place, its greatest drawback is the medium-to-high radiation dosage it involves.

Among modalities that, unlike CT, are radiation-free, MRI is one of the most widely used for brain imaging. The principles of NMR are the basis for MR image generation, as illustrated in Figure 1.3. A high-field-strength magnet is used to create a large magnetic field, which we name B_0 . Protons (found in the hydrogen nuclei of water molecules in the brain) have associated spins, which align uniformly in the presence of this main magnetic field, B_0 [15]. A net equilibrium magnetization is reached, with all protons aligned along the direction of B_0 . This equilibrium can be disrupted by the application of a short duration RF field, B_1 , which results in selective excitation. This RF field causes individual protons to lose their equilibrium alignment, and they now precess and tilt at an angle known as the tip angle, pointing away from the direction of the main magnetic field. Once the RF field is turned off, the spins regain their original alignment, primarily through two processes known as T1 and T2 “relaxation.” The transverse component of the magnetization vector, precessing in the transverse plane, causes signal generation in the receiver coils by electromagnetic induction. The signals measured at the receiver coil form the resultant MR image. Properties related to protons in the brain - including their density, and the biophysical and chemical form in which they are present - manifest as variations in the

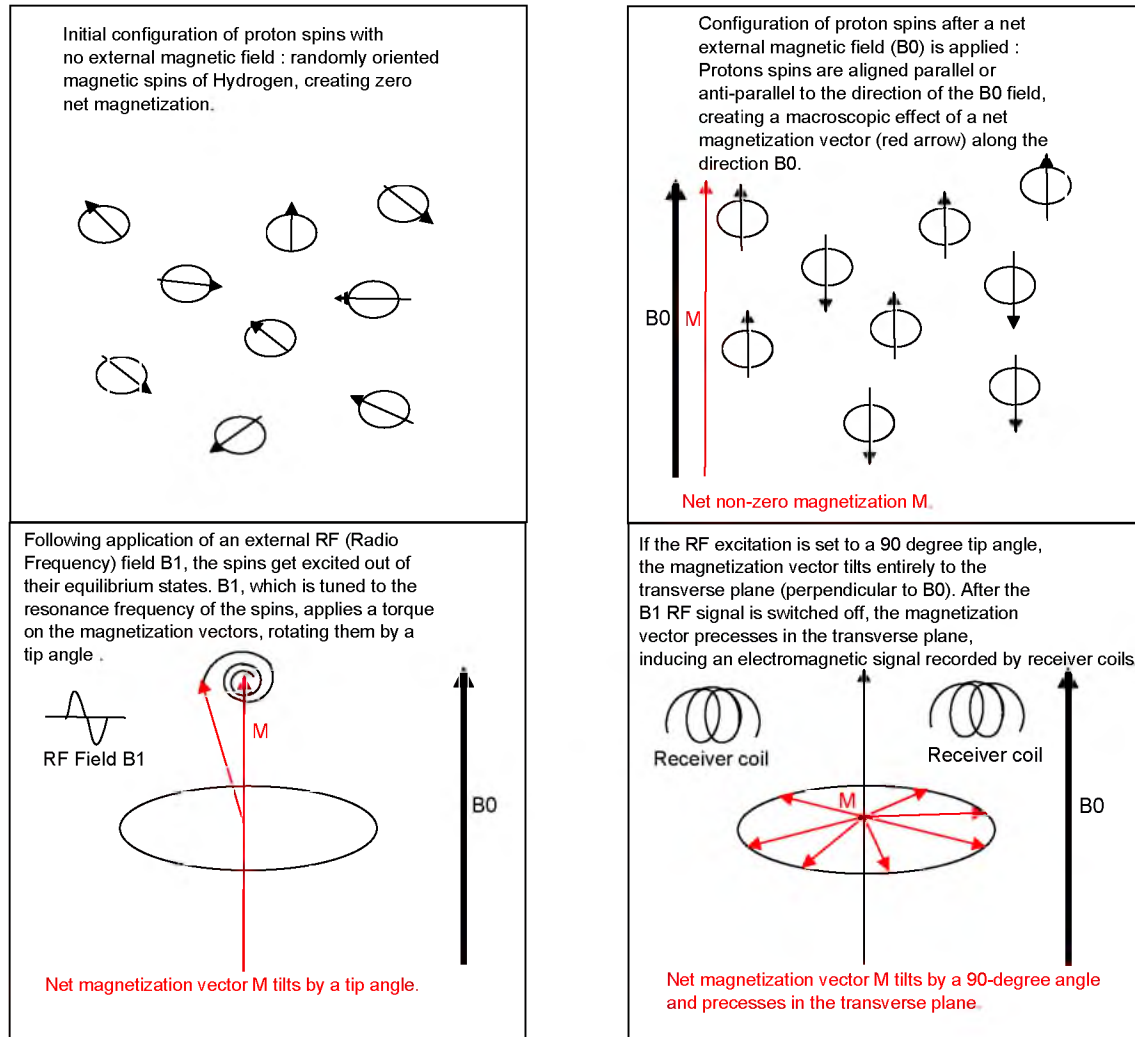


Figure 1.3: Basic principles of MRI.

received MR signal. Variations in tissue properties are captured in MR images in this manner.

The fact that MR imaging is not dependent on ionizing radiation, along with its high resolution, flexibility, and excellent soft tissue contrast, make it a popular modality, particularly for imaging the brain. The drawbacks of MRI are bulky instrumentation, high costs, and long scan times during which motion must be avoided. Several interesting MR scanning protocols and pulse sequences are currently being explored for imaging specific processes and tissues. The commonly accepted imaging protocols for brain MRI currently include a T1-weighted sequence in the sagittal plane (or a T1-weighted volumetric acquisition), and T2-weighted fluid-attenuated inversion recovery (FLAIR) and fast spin-echo or turbo-spin-echo (or equivalent) sequences in the axial plane [16].

In addition to conventional MRI, closely associated medical imaging techniques include DTI (Diffusion Tensor Imaging), fMRI (function Magnetic Resonance Imaging), magnetization transfer, and MWF (Myelin Water Fraction). DTI provides microstructural information related to the brain, by measuring the extent of diffusion of water molecules over a fixed time period [17]. By measuring diffusion, along with its associated directional information, it is possible to better understand the underlying structure of the brain, particularly details of axonal organization and white matter fiber pathways. On the other hand, fMRI, the functional counterpart of MRI, measures the change in magnetization between oxygenated and deoxygenated blood cells, which enables recording of BOLD (Blood Oxygen Level Dependent) contrast in the brain, which essentially maps neural activity according to changes in blood flow with time. Associated techniques such as MWF and magnetization transfer are novel methods for generating quantitative maps with tissue information, based on the usage of advanced pulse sequences and signal suppression, respectively.

In the last decade, PET and SPECT have emerged as powerful techniques in neurobiological studies of brain disorders [10]. Compared with CT and MR, PET and SPECT modalities both involve the introduction of a radioactive isotope into the bloodstream. In PET and SPECT imaging, the energy released by the radiotracer is detected and converted into a scan, enabling the detection of metabolic activity, rate of oxygen use, and blood flow. Since these scans can detect subtle metabolic changes even at the cellular level, as well as neurotransmitter activity, they can provide clues for the early detection of disorders much earlier than CT or MRI [18]. Although PET can be used to detect CNS disorders such as seizures, tumors, cancer, and even conditions such as AD (Alzheimer's Disease), the high cost of these scans makes their availability quite rare. SPECT is comparatively cheaper and has widespread applications in detecting injury, trauma, drug abuse, and psychiatric and neurological disorders.

Being a noninvasive technique with very low risks, ultrasound is used for specialized applications in brain imaging. In ultrasound, pulses of high-frequency sound waves that are generated using transducers are made to travel through the body. Complex tissues and organ boundaries reflect or scatter these sound waves, which then return back to the transducer. The transducer detects these echoes, converts the acoustic signals to electrical signals, and forms an ultrasound image. Although ultrasound results in relatively low-quality images, its popularity can be attributed to its risk-free nature and adaptability, along with the low cost and portability of its associated instrumentation. Ultrasound is used particularly for the examination of blood flow in brain imaging, and for fetal and

neonatal brain studies. Neurosonography is proven to be highly effective in detection of fetal brain abnormalities [19].

In addition to these popular modalities, MRS (Magnetic Resonance Spectroscopy) and NIRS (Near Infra Red Spectroscopy) have emerged as powerful optical imaging tools. MRS, which provides spatially encoded chemical information, has been used to study metabolic changes in diseases such as cancer [20,21] and AD [22] that affect the brain. NIRS has also demonstrated utility as a simple technique for measuring changes in blood flow associated with neural activation [23]. Other novel medical imaging techniques include elastography, tactile imaging, photoacoustic imaging, and advanced spectroscopic techniques.

Apart from these neuroimaging tools, EEG (Electroencephalography) and MEG (Magnetoencephalography) are two closely related techniques that capture underlying brain activity in terms of electrical and magnetic fields, respectively [24]. In EEG, an interface is created to the brain by placing electrodes at points on the scalp and recognizing voltage differences between different points, while MEG works on magnetic field measurements. These techniques have excellent temporal resolution but poor spatial resolution, and can be used in conjunction with imaging modalities such as fMRI for effective diagnosis of seizures, tumors, and functional abnormalities, in major cortical areas.

1.3 Origins of MR Signal Intensity and Variations

MRI is a radiation-free and relatively risk-free modality for imaging the brain. MR imaging is also highly flexible - by modifying the pulse sequences and scanning protocols used for image generation, specific properties of biological tissues that are of interest can be observed. In the last section, the fundamental principles of MR imaging were mentioned. Here, the origin of MR signal intensity and intertissue intensity variations will be elucidated.

Initially, given that the net external magnetic field is a static field \mathbf{B} , and γ is the gyromagnetic ratio of a nucleus, the magnetization vector \mathbf{M} associated with a proton varies according to the equation

$$\frac{d\mathbf{M}}{dt} = \mathbf{M} \times \gamma\mathbf{B}. \quad (1.1)$$

It is known that the magnetization vector \mathbf{M} , which contributes to the final MR image, varies as a function of relaxation times. This can be seen in Figure 1.4, which illustrates details of the MR signal formation process, as a continuation of Figure 1.3. After the RF field B_1 is switched off, the longitudinal component of the magnetization vector undergoes strengthening at a rate based on the longitudinal T1 relaxation time, while the transverse component of the magnetization vector undergoes decay at a rate based on the transverse

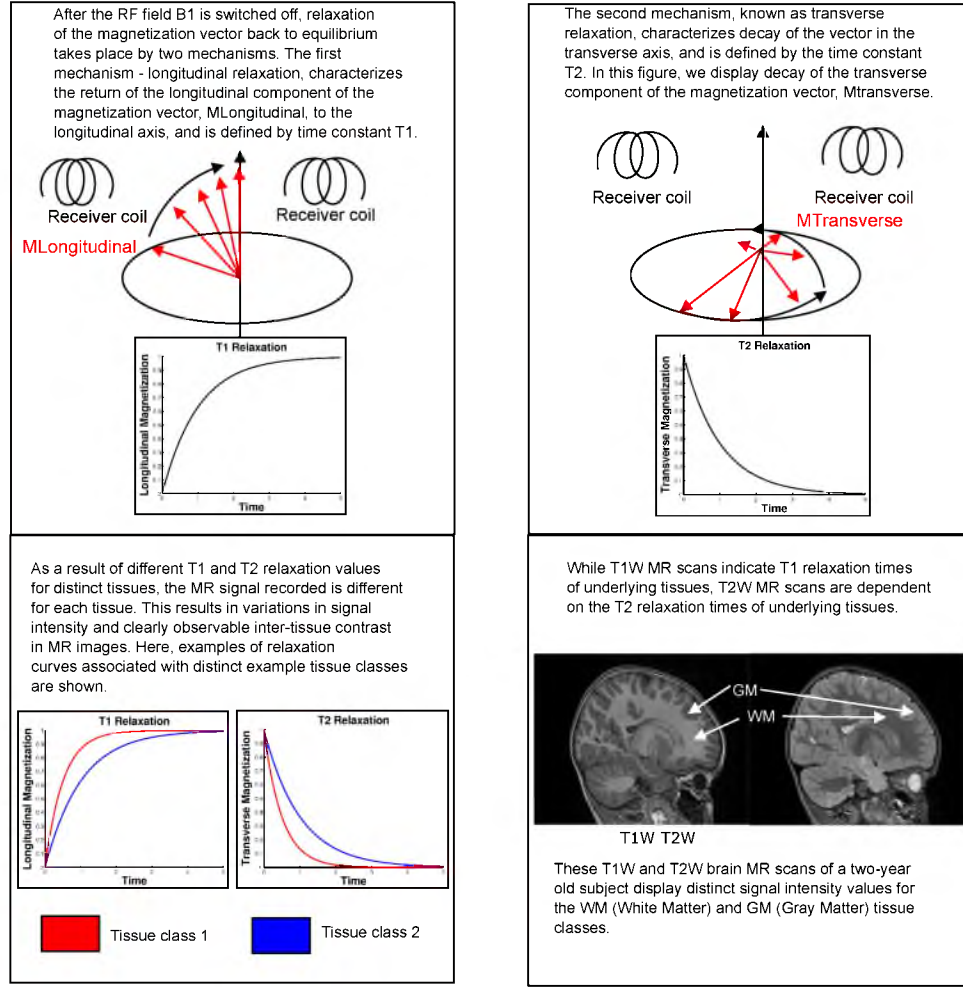


Figure 1.4: Origin of the MR signal in terms of T_1 and T_2 relaxation processes (top row), and mechanisms underlying intertissue intensity variations (bottom row).

T_2 relaxation times. Assuming that the RF field B_1 tilts the magnetization vector by a tip angle of 90 degrees, and given that M_0 is the magnetization vector at time $t = 0$, the longitudinal component of magnetization at a time t can be written mathematically as

$$M_{longitudinal}(t) = M_0(1 - e^{-\frac{t}{T_1}}). \quad (1.2)$$

Under the same assumptions as listed above, for a 90 degree tilt angle, the transverse component of magnetization at a time t can be written as

$$M_{transverse}(t) = M_0 e^{-\frac{t}{T_2}}. \quad (1.3)$$

As a result of the dependence of the magnetization vector on T_1 and T_2 relaxation times, the measured signal also varies as a result of changes in T_1 and T_2 . In this discussion,

we do not examine the complex details of the equations governing the received MR signal. However, for the purpose of our discussion, it is sufficient to know that the received MR signal depends on the signal oscillating in the transverse plane, and is hence dependent on both $T1$ and $T2$ constants, although the exact nature of this dependence might vary based on the pulse sequences and type of MR scanning used. In the simplified case of a spin echo saturation recovery sequence, repeated cycles of excitation and recording of the received signal take place. The received signal depends on two acquisition parameters : TE (echo time) or the time between the RF excitation pulse and measurement of MR signal, and TR (repetition time), i.e., the time between two excitation pulses in the constantly repeating sequence. In the final spin echo MR image, the intensity I of the received signal, can be written as

$$I = K\rho[1 - e^{-\frac{TR}{T1}}]e^{-\frac{TE}{T2}}. \quad (1.4)$$

Here, I is the MR Intensity, K is the gain constant, TR is the repetition time, TE is the echo time, ρ is the spin density, $T1$ is the longitudinal relaxation time, and $T2$ is the transverse relaxation time.

From this equation, it can be clearly seen that the received signal measurements, corresponding to the MR signal intensities, are dependent on $T1$ and $T2$ relaxation constants. TE and TR values are appropriately chosen while scanning to obtain $T1$ -Weighted ($T1W$) or $T2$ -Weighted ($T2W$) MR images. In standard $T1$ -Weighted images, the $T1$ relaxation constants contribute most dominantly to image intensity, whereas for standard $T2$ -Weighted images, the image intensity is mostly dependent on the $T2$ relaxation constants. Proton Density (PD), denoted by ρ in Equation 1.4, also contributes to image intensity in both $T1W$ and $T2W$ images.

The $T1$ or $T2$ relaxation constants for a small volume in the brain will vary depending on the physical and biochemical properties of the tissue, or tissues, constituting that volume. This intertissue difference in values of relaxation constants is hence a major source of the intensity variation observable in MR [25, 26]. Additionally, since different tissue classes have dissimilar densities of protons, a second factor contributing considerably to intertissue intensity variation is proton density [25]. Finally, the degree and nature of intensity variation can be modified by adopting advanced pulse sequences and scanning mechanisms such as fat suppression, radio frequency inversion, and magnetization transfer [26].

Based on the above discussion, it can be concluded that with all other parameters remaining constant and under ideal scanning conditions, variation in received signal intensity measurements for different points in a reconstructed MR image can be attributed to

differences in T1, T2, and PD values at these points [27]. In the human brain, WM and GM tissue have different values for these parameters. Therefore, for commonly used T1W and T2W scanning protocols, intensity variation is clearly observable between WM and GM regions in MR images. Additionally, by changing scanning parameters, images with varying degrees of WM-GM contrast and contrast direction can be produced [28].

Since WM-GM intensity variation changes as a result of maturation, aging, and neurological diseases, studying these changes can aid in neurological research. In the following section, biological underpinnings of intensity variations between WM and GM, as observed in MR-based neurodevelopmental studies, will be discussed.

1.4 WM-GM Intensity Variations in Neurodevelopment

Although early brain development involves a complex sequence of several rapid biophysical, chemical, structural, and functional changes, these changes occur in an extremely organized and predictable manner. A crucial component of these processes, known as myelination, consists of the formation of a myelin sheath around a nerve fiber [29]. Myelination of WM enables the effective transmission of neural impulses and occurs from birth until the end of the second year. Before myelination, lipid and water components are similar in gray and white matter, but as a result of myelination, bulk water content decreases from 88 % to 82 % at 6 months.

As shown in Figure 1.5, it is observed that as a result of maturation and associated myelinational processes, WM displays progressively higher signal intensities in T1W images, and progressively lower signal intensities in T2W images. These changes in T1W and T2W images result primarily from WM myelination, change in water content, and the subsequent shortening of T1 and T2 relaxation times with age, as well as changes in PD [27,30]. The T1 shortening that is observed during myelination is hypothesized to occur due to properties of the hydrophilic cholesterol and glycolipid components of the developing myelin sheath. T2 shortening is reported to occur at the time of tightening of myelin around the axon, and may correlate best with the development of myelination as determined by histological methods [29,31]. Plots of T1 and T2 shortening can be seen in Figure 1.5. These plots have been adapted from [32].

It should be noted that *both* GM and WM undergo shortening of T1 and T2 signals during brain maturation. However, the effect of this shortening is much more pronounced in WM tissue, resulting in highly noticeable changes in WM signal intensity. In comparison, the signal intensity of GM changes in a more subtle manner. As a result of these dual processes of WM and GM signal intensity changing with age, changes can also be seen in

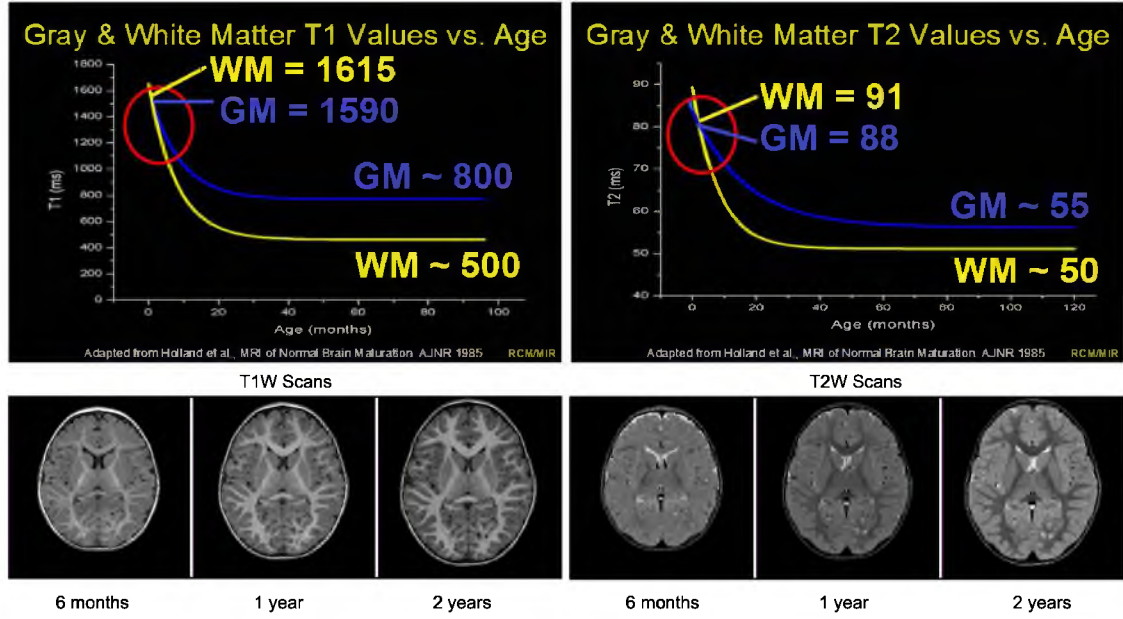


Figure 1.5: Changing T1 relaxation values (top left) and T2 relaxation values (top right) for WM and GM tissue in the developing infant brain. The T1 relaxation values for WM and GM are ≈ 1615 ms and ≈ 1590 ms, respectively, at the neonate stage; and ≈ 800 ms and ≈ 500 ms respectively, at 100 days of age. Similarly, the T2 values for WM and GM are ≈ 91 ms and ≈ 88 ms at the neonate stage; these values are ≈ 55 ms and ≈ 50 ms at 100 days of age. Consequently, observable changes in WM-GM intensity difference can be seen in longitudinal series of T1W scans (bottom left) and T2W scans (bottom right) of an infant scanned at approximately 6 months, 12 months, and 24 months of age.

the intensity variation between WM and GM. Intensity differences between brain WM and GM first decrease in the first few months after birth up to a point of iso-intensity. They then reverse in direction of intensity gradient, and finally keep increasing up to early childhood. Crucial neurodevelopmental processes involved in brain maturation, including myelination of WM fibers, can be tracked by observing changes in WM-GM intensity variation. Besides intensity analysis, actual T1 and T2 relaxation times have also been shown to have potential as in-vivo markers of tissue properties and brain tissue abnormalities [33].

The nature of WM-GM intensity variation is highly modality-specific. Although both T1 and T2 relaxation times change as a result of maturation, the sequence and timing of these changes are hugely different. T1 shortening takes place from birth until around 8 months of age, whereas T2 shortening is most prominent from around 6 months until 3 years of age. As a result of the timing of changes taking place, T1W images are useful in studying brain development in the first 6-8 months of life whereas T2W images are more useful after 6 months [30]. As a result of the underlying neurobiological causes for T1W and

T2W changes, each of these modalities supply useful and unique information about specific maturational stages and processes.

WM-GM intensity variation is also spatially dependent. Myelination is a heterosynchronous process, with different regions of the brain undergoing myelination at different times. It has been seen that myelination generally proceeds from central to peripheral, inferior to superior, and posterior to anterior brain regions [30,34]. This phenomenon leads to corresponding spatial differences in WM-GM intensity variations. Therefore, studies of WM-GM appearance must be undertaken in a regional or spatially localized manner, such that the spatially dependent nature of maturational processes can be better analyzed.

1.4.1 Additional Sources of Intensity Variation

Intensity variation, apart from inherent changes due to altered tissue characteristics, can be attributed to two causes: 1) variations within an image due to MR bias field and associated intensity inhomogeneities, and 2) variations between images due to differences in scanner type used for acquisition, scanning conditions, and minor changes in acquisition parameters [35]. A study of variations belonging to the latter category for a human phantom dataset will be explained in detail in Chapter 2.

As a result of these intensity variations due to undesirable sources, signal intensity or appearance-based MR analysis is relatively less common when compared to other analytical methods. Although undesirable intensity variations can be corrected in adult brain scans using advanced intensity normalization techniques as outlined in the Appendix, these corrections cannot be easily applied to infant brain scans due to inherent variations in tissue intensities in these scans, resulting from neurodevelopment. As a result, studies of brain development based on MRI have focused mainly on morphometric, volumetric, and microstructural changes rather than on appearance variations. In the following section, a survey of MR studies of early brain development will be presented.

1.5 MR Studies of Early Brain Development

Compared with other brain imaging techniques, neurodevelopmental changes can be observed using MR with greater sensitivity, high spatial resolution, and no risk of radiation dose [36,37]. Safety and radiation standards, which are especially important while imaging infants, are met by MR imaging. Moreover, MR images show developmental processes such as cortical folding, premyelination changes in white matter, myelination, iron deposition, and the growth of different brain regions, which cannot be imaged using CT or ultrasound [29].

Understanding the growth patterns of the brain in the first two years after birth is crucial to the study of neurodevelopment and neurological disorders. During this time period, the brain triples in size [38, 39] and also undergoes tremendous neuronal growth, increases in synaptic numbers, and myelination [40–42]. Since the creation of mature, complex cerebral circuitry depends to a large extent on neuronal growth right after birth, this is also a time in which adverse conditions can have a greater effect on brain growth and myelination [43, 44]. Therefore, apart from giving us important information about normative patterns of growth and cognitive development, studying the developing brain could play a key role in the early diagnosis of brain disorders and timely intervention. Several recent research studies have found links between neuropsychiatric disorders such as schizophrenia and quantitative measures obtained from early brain MRI [45, 46].

Most existing studies of the pediatric brain have focused mainly on volumetric and morphometric indicators [46–53]. In several of these studies, changes in volumes of gray matter and white matter tissues in major brain regions were analyzed, indicating rapid growth during the first year of life, followed by slower growth up to adolescence [47, 49, 52, 53]. Correlations between regional volumes and cognitive outcomes were also established, confirming the effectiveness of volumetric biomarkers for predicting cognitive disturbances [52]. Besides volumetric studies, nonlinear deformation-based mapping has been used to produce localized growth maps characterizing brain changes [48]. Fine-grained cortical thickness analysis has also been performed, indicating a decrease in cortical thickness with a back to front progression [54]. It should be noted that in several of these volumetric and morphometric studies, the usage of infant data posed unique challenges. Some of these challenges include extremely large changes in volumes of structures, low contrast to noise ratio, and intensity inhomogeneities due to myelination - all leading to difficulties in segmentation and estimation of cortical boundaries [55].

Several microstructural DTI (Diffusion Tensor Imaging) studies have recently been performed on the infant brain. These studies have focused on changes in diffusion parameters with age, thereby correlating changes in these parameters with brain maturation [56–58]. Since brain myelination results in restriction of water diffusion, the direction of myelinated fiber bundles can be detected by studying diffusion parameters. Studies of diffusion parameters across pediatric datasets consistently indicate directional restriction of water molecules with age, often showing changes even before myelination is macroscopically observable [59]. Correlations with behavioral parameters have also established the utility of diffusion parameters in brain maturation studies [60]. A drawback of diffusion measures,

however, is that the changes seen in these measures could be attributed to a variety of factors besides myelination. These factors, related to axon structure and fiber size, tract coherence, and membrane structure and permeability, affect diffusion values significantly, making it difficult to attribute changes in diffusion parameters to specific maturational processes such as myelination [61].

In addition to these methods, novel imaging techniques are now being used for studying early brain development. One such method, MRS (Magnetic Resonance Spectroscopy), derives information from protons present in nonwater molecules such as creatine, N-acetyl-aspartate (NAA), choline, and glutamate [60]. Studies of MRS have mapped changes in concentrations of cerebral metabolites with brain maturation [62], thereby measuring alterations in brain chemistry. Other methods for imaging brain maturation include Myelin Water Fraction (MWF) imaging, Magnetization Transfer (MT) imaging, and quantitative T2 imaging. Analysis of T2 relaxation in the central nervous system led to the discovery of two T2 components - a short component attributed to water trapped between myelin layers and a long component linked to intra- or extra-axonal water [63]. The Myelin-Water Fraction (MWF) measure was hence defined as the ratio of myelin water to total water, and has shown potential for identifying demyelination of brain tissue in MS (Multiple Sclerosis). MWF studies have also been extended to analysis of infant brains [61], resulting in successful mapping of myelination processes. MT imaging, a closely related technique, is based on suppression of signal and magnetization transfer between two components (semisolid and liquid) of a model. MT has shown efficacy in detection of lesions, WM integrity, and demyelination processes [64]. A third technique that is closely related is quantitative T2 mapping. Quantitative T2 imaging uses relaxometry techniques to map T2 relaxation values and has been shown to be a sensitive indicator of brain maturation [65]. The mapping of actual T2 values, rather than the acquisition of a weighted image, results in an actual quantitative indicator of the state of underlying tissues [33]. However, the uncertain reliability of T2 estimates, as well as longer scan times due to multiecho acquisition, are factors to be considered when employing this method for infant brain imaging.

In addition to these studies, histological analysis of the brain has been performed [40]. Postmortem studies involving staining for myelin revealed that white matter myelination takes place in an orderly sequence from inferior to superior, and from posterior to anterior of the brain [31]. A recent study also analyzed synaptic density in the prefrontal cortex using Western blotting methods [14]. While histological studies provide great insight into synaptic structure and organization of neurons at a scale that is far finer than most imaging

techniques, they can be performed only on postmortem brains. The results of histological studies might also be variable depending on the stains used for capturing histological features. These studies are, however, extremely valuable in validating the efficacy of newly developed imaging techniques, and in examining the biochemistry involved in neurological changes.

The biomarkers resulting from the volumetric, morphometric, and novel imaging techniques discussed above have proven highly effective in analyzing trajectories of early development. However, image appearance measures as biomarkers characterizing neurodevelopment are relatively less explored. Image appearance is known to be a valuable indicator of maturation in qualitative MR studies and presents scope for further analysis. Biophysical and chemical information, complementary to data from morphometric and volumetric analysis, can be obtained from appearance studies. Appearance analysis is also a viable alternative to other methods for obtaining biochemical information from the brain, such as diffusion, MWF, and MT, as these methods rely on modified techniques at the acquisition stage, and hence cannot be used for analysis of large retrospective studies. Therefore, appearance measures could be valuable neuroimaging biomarkers, particularly if computed in a manner that is stable with respect to changing conditions of scan and in-train image inhomogeneities. Additionally, appearance measures can be jointly analyzed along with other indicators for greater insight into neurodevelopmental mechanisms (for example, signal intensity and diffusion parameters can be jointly studied).

1.6 Appearance Studies in Neuroimaging

The earliest appearance studies in neuroimaging were purely qualitative in nature, and consisted of observations made by radiologists [29]. However, the subjective nature of qualitative studies may lead to interobserver variability and errors due to medical image quality, human errors, and perceptual factors [66].

Recently, quantitative appearance-based research studies restricted to MR signal intensity analysis of the pediatric brain have shown interesting results [67–71]. The goal of these studies was to analyze changes in tissue properties, as indicated by measured signal intensity values. One of the earliest intensity-based studies of early brain development jointly analyzed the changes in T1W and T2W signal intensity along with DTI parameters along major white matter fiber tracts [67]. A linear mixed model was fit to track changes in these parameters across a dataset of infants and to find correlations between diffusion and intensity measurements. In a similar pilot study, the utility of studying MR signal

intensities along with diffusion measures was shown on a dataset of infants [71]. Another study performed cross-sectional analysis of T1W and T2W images of 116 preterm subjects between 29-44 weeks gestational age (GA) [69]. Voxel-wise signal intensity change was modeled across age using kernel regression. The results of this study, which focused on changes in subcortical structures, showed that central and peripheral cortical regions display different rates of intensity change. In other work along similar lines, changes in WM signal intensities in infant brain MR images were modeled across time using the Legendre polynomial [70]. Following spatial and intensity normalization, regions displaying similar trajectories of intensity change were clustered in a data-driven manner using a Dirichlet Process Mixture Model (DPMM). For each cluster, growth curves displaying changes in intensity with age were generated. Along with T1 and T2 signal intensity, other properties such as diffusion parameters were also mapped across time, and clinical variables indicating rate of growth were developed. In a related research study, the signal intensity of T1W and T2W MR images of infants was analyzed and modeled with age in conjunction with diffusion parameters [68]. Analysis of these intensity and diffusion parameters was performed using three methods: 1) voxel-wise, 2) anatomical atlas-driven, and 3) purely data-driven. As a result of fitting these parameters across time with a Gompertz function, statistical indices characterizing changes in WM signal intensity, such as rate of growth, delay, and asymptote, were extracted - for voxels, anatomical regions, and data-driven regions. It was shown that nonlinear growth patterns observed in intensity data were characterized effectively using the Gompertz function.

It is important to consider that intensity normalization was performed on the infant brain images used in these studies in order to bring the intensities of all analyzed scans into a standardized range. A note on normalization techniques used in infant brain studies is presented in the Appendix. Intensity normalization on infant images is highly challenging due to several factors, including inherent intensity changes due to neurodevelopmental tissue changes. Intensity normalization algorithms should not remove these maturation-induced intensity variations by confusing them with image variations due to nonideal scanning conditions. Since the methods discussed above analyze absolute signal intensity values, they are limited by their crucial dependence on intensity normalization.

Spatial normalization (nonlinear registration) was also done in the studies discussed above, in order to enable voxel-wise comparison of signal intensity across images. Spatial normalization also has several limitations when applied to infant data, including low signal-to-noise ratios, low contrast, and large variabilities in brain shape of infants.

A recent study outlined a method for characterizing joint appearance and shape variability across neuroimaging datasets via manifold learning, with specific applications to infant brain scans [72]. This study utilized the Normalized Cross-Correlation (NCC) metric to measure similarity in the appearance of two distinct brain images. Although the NCC has the property of being invariant to affine transformations of underlying intensities and hence removing the need for intensity normalization, it requires point-wise spatial correspondence between images. Spatial normalization via deformations are embedded into the methodological framework, but the method is still dependent on accurate normalization for appearance characterization.

Intensity-based appearance analysis has recently gained importance for its role in ensuring accurate registration between longitudinal images of the rapidly changing developing brain. Recent studies have modeled intensity changes across longitudinal images of the early developing brain by fitting a linear or nonlinear model to voxel-wise intensity information [73,74]. Modeling intensity transformations between these images have subsequently improved interscan registration. In one of these studies, the nonlinear logistic model was used to create a spatiotemporal mapping of white matter intensities in brain MR scans of monkeys [74]. On similar lines, Intensity Growth Maps (IGM's) were created to model and correct for areas of low WM-GM contrast in regions of immature WM, with the object of improving the resultant tissue segmentation [75].

Although the intensity studies discussed above have begun exploration of MR signal intensity as an imaging biomarker, limitations associated with purely intensity-based analysis have proven a deterrent to extensive research in this direction. To summarize, some major limitations of intensity studies include variability of signal intensity due to external factors such as the type of scanner and scanning conditions, lack of calibration procedures directly linked to the MR signal, and crucial dependence on the use of accurate intensity normalization.

1.6.1 Relative Intensity-based MRI Studies

As an alternative to studies of absolute intensity, relative intensity differences or contrast between tissues has been explored as an indicator of brain maturation. In general, the contrast between WM and GM was measured as the difference in WM and GM mean intensities, divided by noise [25]. It was seen in an early research study that WM-GM contrast computed using this definition is better suited for the verification of increasing myelin density than the actual relaxation times. This conclusion was made since significant differences between regions undergoing early and delayed maturation were more clearly

indicated by contrast [27]. Reasons proposed for explaining the superior performance of contrast as a biomarker were amplification of both T1 relaxation and proton density differences in the contrast measure.

Besides neurodevelopmental research, contrast studies have been conducted for analysis of neurological processes such as aging and neural degeneration [76–79]. An early study based on the idea of WM-GM contrast as a biomarker used a subjective method to qualitatively “rate” the degree of WM-GM contrast, which was seen to decrease significantly with age [80]. This decrease in contrast was attributed to changes in the structure of myelin in white matter, and neuronal loss in gray matter. In a longitudinal, contrast-based study of the aging brain, a regional contrast ratio (rCR) given by $\frac{2[WM(x)-GM]}{[WM(x)+GM]}$ was computed, demonstrating that degenerative age changes in WM connectivity are captured by changes in contrast [77]. A recent study on subjects undergoing healthy aging measured the ratio of gray matter to white matter signal intensity (GWR) at each point along the cortical surface, showing strong, localized, significance of increasing GWR, and hence decreasing contrast, with age [76]. Another study showed that the GWR was significantly reduced in several regions in individuals with Alzheimer’s disease [78], also indicating that contrast is a unique measure which is complementary to other morphometric measurements. In a study of twins, it was seen that contrast as measured by the GWR is genetically influenced, and a high degree of heritability was seen for contrast in the major brain regions [81]. In addition, it was seen that contrast did not have significant genetic correlations with cortical thickness, thereby indicating the novelty and uniqueness of this measure.

These studies have shown that WM-GM contrast, also referred to as WM-GM intensity difference or WM-GM intensity variation, has potential as an indicator of healthy and abnormal trends in aging, and also as a biomarker in studies of imaging genetics. However, the methods for contrast computation that are listed above might not be easily applicable to infant brain studies, since they require estimation of the cortical surface for computation of WM-GM contrast. In the case of developing brains, the location of the cortical surface is highly uncertain, making these methods unreliable for infant studies.

1.7 Goals of This Dissertation

In this work, we propose a novel methodology with two objectives - first, to quantify WM-GM appearance variations by measuring the relative distance between their intensity distributions, and second, to provide parametric modeling of such tissue contrast distances as a function of time. The quantitative nature of this method ensures that the appearance

measure does not vary in a subjective manner and can be used to statistically test findings based on qualitative observations. In addition, since the measures we propose are not just measures of signal intensity, but *relative* indicators of signal intensity *variations*, instability due to external factors such as changing scanning conditions is minimized. Finally, we ensure that the method we develop is not dependent on intensity normalization, cortical thickness measurements, or other complex processing methodologies that are difficult to apply for the specific case of the infant brain. The major features of this framework, including details of spatiotemporal modeling of the developed appearance measure will be discussed, with the application of methods to several ongoing clinical infant neuroimaging studies.

1.7.1 Quantification of WM-GM Intensity Variations

In order to compute intensity variations between WM and GM, we implement a 4D longitudinal, registration-segmentation framework, followed by methods to generate intensity distributions in WM and GM tissues. Intensity variation is then evaluated in terms of the Hellinger distance-based “separation” between WM and GM intensity distributions, resulting in a measure named WIVID (White-gray Intensity Variability in Infant Development). The sensitivity of this measure of WM-GM intensity variation in capturing differential maturation patterns is analyzed. Since the WIVID measure is invariant to affine transformations of underlying signal intensities, it is less sensitive to external factors. We explore the possibility that this “relative intensity”-based measure can ensure greater stability with respect to changing scanning conditions, compared with absolute signal intensity measures. Higher stability will in turn allow usage of this method across large-scale neuroimaging studies that are undertaken across different scanner sites and scanning conditions without any changes in the analyzed biomarker.

1.7.2 Spatio-temporal Modeling of Intensity Variations

Following the development of WIVID to measure WM-GM intensity variation, we develop a methodology for longitudinal statistical modeling, inference, and analysis, based on this developed measure. The quantitative nature of the WIVID measure developed enables statistical analysis. We outline methods for modeling and analysis of changes in WIVID over time using advanced biostatistical techniques. The large-scale datasets under study are longitudinal in nature and consist of repeated scans of the same subject taken at different time points. Inherent challenges posed by the repeated nature of this data are tackled using NLME (Non Linear Mixed Effects) models. These NLME models account for correlation between repeated scans of a single subject, creating both population- and subject-specific

models to map temporal change. When used in conjunction with nonlinear growth models, NLME-based statistical analysis provides key information related to parameters underlying growth. The growth parameters that are estimated from NLME analysis are intuitive and easy to understand, in addition to holding clinically relevant information.

1.7.3 Modeling Multimodal WM-GM Intensity Variations

In the context of modeling of relative WM-GM intensity for an MR dataset, differences based on modality of scan and scanning protocols used should be considered. As can be seen in Figure 1.5, the very nature of the signal being recorded by the brain scan is determined by the scan modality: different MR imaging modalities display completely different image appearance and intertissue contrast [16]. For instance, T1W (T1-Weighted) and T2W (T2-Weighted) modalities display inverted WM-GM intensity gradients. That is, in T1W images of the mature brain, WM is of much higher intensity than GM, whereas the reverse is true for T2W images of the mature brain. In addition, based on the scanning protocols used, imaging datasets have different timing and sequential changes associated with brain maturation [29]. In the framework we develop, WM-GM intensity differences in terms of the WIVID measure are computed separately for each MR imaging modality. However, their statistical modeling is performed jointly across modalities, enabling consideration of temporal correlations between modalities in computation of the evolution of WM-GM appearance. Finding their temporal correlations also helps in understanding the physical tissue properties that each modality measures and the relationships between their growth trajectories.

1.7.4 Spatial Heterogeneity in WM-GM Intensity Variations

Even within a single scanning modality, the nature of brain appearance change is spatially heterosynchronous. WM-GM intensity differences vary considerably across cortical regions, particularly during neurodevelopment. Our framework involves the study of WM-GM intensity differences in a region-specific manner. Therefore, although trajectories of intensity variation can change considerably between different cortical regions, the region-specific framework we implement inherently models this variation. This region-specific framework, coupled with statistical analysis methods, can even test for significant differences in the trajectories associated with WM-GM intensity variations of distinct cortical regions.

1.7.5 Intensity Variation-based Infant Studies

We apply the complete framework developed for WM-GM intensity variation analysis to the study of longitudinal datasets of MR scans of the developing infant brain. As a result, we can study regional and spatially localized changes in WIVID across different stages of neurodevelopment, from birth to later stages of childhood. Further, the statistical framework we implement enables investigating the effectiveness of the appearance measures developed as biomarkers. The ability of these measures to distinguish between differential growth trajectories belonging to varied populations of interest grouped on the basis of 1) risk and diagnosis for brain disorders, and 2) sex, is analyzed.

1.8 Overview of Chapters

The remainder of this dissertation is organized as follows:

Chapter 2 discusses the methodology for quantification of WM-GM intensity variation in large cortical regions and evaluates the stability and effectiveness of the quantified measure.

Chapter 3 explores nonlinear mixed effects methods for longitudinal statistical modeling of WM-GM intensity variation across time, with extensions to multivariate data.

In Chapter 4, the complete framework for quantification and modeling of intensity variation is applied to a large-scale neurodevelopmental study of infants from 6 months to 2 years of age, and patterns of appearance change are analyzed.

Chapter 5 analyzes age-related appearance changes across a dataset consisting of 10 subjects scanned at five time points each, from birth to 12 months of age.

Chapter 6 summarizes the overall goals met by this dissertation and highlights significant contributions made in this work, followed by an analysis of related challenges and limitations and discussion of future work.

CHAPTER 2

QUANTIFICATION OF WM-GM MR INTENSITY VARIATIONS

Medical image appearance is a function of signal intensity at each point in the image. In medical images acquired using techniques such as X-ray CT and PET, the signal intensity value at a point corresponds to a physical quantity of interest, which characterizes the region being imaged. For example, the signal intensity values in an X-ray CT scan correspond to tissue attenuation as empirically defined by Hounsfield Units [25]. However, in the generalized case of MR images, the signal intensity value at a point does *not* correspond to an absolute physical quantity of interest. Instead, it is a complex function of the density of water molecules in the region under study, as well as T1 and T2 relaxation constants characterizing the region. The extent to which each of these parameters influences the signal intensities in an MR image depends on the type of weighted MR scan being acquired and on associated scanning parameters. In general, there is no single physical quantity of interest that can be determined from MR signal intensity, with exceptions being the case of specialized MR techniques such as MWF and MT that were discussed in Chapter 1.

In spite of not providing a direct quantitative measure, MR signal intensity has excellent potential as an indicator of underlying tissue characteristics. This has led to several MR signal intensity studies that are qualitative in nature, consisting of observations of MR image appearance made by radiologists [29]. However, qualitative studies have drawbacks such as the subjective nature of observations and errors due to image quality, image perception, and human factors [66]. Another drawback of purely qualitative analysis is the lack of resulting numerical quantities that can be modeled and analysed using statistical and machine learning techniques. Therefore, in order to effectively convert information based on the signal intensity of MR images into clinically useful findings, it is necessary to develop methods for *quantification* of MR image appearance.

It has been noted in the last chapter that the contrast between regions in an MR image is a useful imaging indicator of appearance [25], and could be a better measure than just signal

intensity. Specifically, it was mentioned that the contrast, or intensity variation, between brain WM and GM regions is of special interest in neuroimaging studies [27]. Underlying neurobiological reasons for WM-GM contrast appearance were also discussed in depth. From these discussions, it was concluded that methods for the quantification of intensity variation between WM and GM tissues can be of considerable potential in neuroimaging, particularly in studies of early brain development. One of the primary goals of this work is to develop a measure of the intensity variation between WM and GM regions.

For an effective measure of WM-GM intensity variation, it is crucial that this indicator must ideally be sensitive to changes in tissue properties, while remaining invariant to changing scan conditions. Interscan variability in conditions during MR signal acquisition introduces undesirable variations in signal intensity, in addition to variations of actual interest that are induced by changing tissue characteristics. This interscan variability, coupled with the lack of “standardization” of the MR signal (in terms of direct links to absolute physical quantities of interest), increases the importance of stability of MR appearance measures.

In this chapter, we will first discuss the concept of appearance from image processing and medical imaging perspectives, and then explore methods for its quantification. A stable measure of MR image appearance that captures intensity variations between WM and GM tissue in the brain will then be developed, and a detailed description of the pipelines for processing and computation of this measure will be given. Next, the question of whether the developed measure effectively captures maturation-related tissue changes will be explored. Finally, the relative stability of the developed measure will be tested.

2.1 Methods to Study Appearance and Contrast

In the context of image processing in general, the meaning of the terms “Appearance” and “Contrast,” which are used to describe medical images, will now be examined. The appearance of a medical image generally consists of a synthesis of two aspects - the shape or structure of the image and the texture or patterns of intensity variation in the image [6]. With complete knowledge of the appearance of an image, in terms of shapes and intensity patterns, the original medical image can be synthesized. Appearance variations between images can be defined in terms of variations in shape and intensity patterns. Methods for describing the shape component of appearance variations between images include deformable elastic models such as snakes [82], finite element methods [83], shape representation using geometric or rigorously parameterised models [84], and statistical

as well as diffeomorphic models [85, 86]. Other methods with extensive applications in vision that directly or indirectly utilize image intensity information include SIFT (via computation of image intensity gradients), the Harris corner detector, high-dimensional histograms representing texture based on combinations of spatial relationships between pixels, and moment invariants for joint shape and appearance representation of images [87].

Several methods jointly model both shape and intensity variation, with applications in computer vision¹. One such method used to match images is the representation of a 2D surface in a 3D format, with the third dimension corresponding to the intensity of the surface [90]. Another joint shape-intensity model uses the sum of squared differences in pixel-wise intensity between source and reference images to define texture [91]. Variants of this texture measure involve encoding pixel-wise intensity differences between image pixels and a reference atlas image [92, 93]. Advanced texture analysis methods use arrays of filter banks including Gabor filters, wavelets, and DCT (Discrete Cosine Transform), to name a few. These image analysis methods focus on characterizing the spatial relationships between pixels of certain intensities, or the frequency content of images. In most of these studies, the analysis generally requires perfect point-wise correspondence across images. In some cases, deformations to achieve accurate point-wise correspondence are estimated and applied in conjunction with these image analysis schemes. However, our interest in this work lies in methods for capturing only the intensity variation component of appearance, without considering shape variations.

In intensity-based studies of appearance, absolute intensity values are generally not commonly studied, unless they are quantitative indicators. Instead, intensity is usually measured in terms of its variation *across*, or *within*, images, or sometimes even both. Existing methods that capture intensity variations between images include image matching functions such as cross-correlation, mutual information, Euclidean distance, the L_1 -metric, and L_2 -metric [89]. Here, we briefly go over some of the major image matching techniques.

Consider two images, X and Y , both of size $N_x \times N_y$. The pixel values of each of these

¹Besides applications in medical imaging, intensity comparison methods between regions of images have been studied extensively in the field of object tracking and computer vision [88]. It is interesting to note that intensity comparison methods in the field of object tracking contend with some issues that are analogous to problems in medical imaging. Intensity variations studied in computer vision are attributed to illumination intensity differences, shadows, instrumentation gain and base level differences, and changes in light direction [89]. These issues are somewhat similar to interimage and intrainage intensity inhomogeneities present in medical imaging due to issues such as bias field, conditions during acquisition of MR scans, etc. While illumination invariance in vision applications is generally achieved by modeling light sources, in medical imaging, postprocessing techniques such as filtering of medical images and intensity normalization are adopted to remove these intensity changes.

two images are denoted by x_{ij} and y_{ij} respectively, with i and j being the 2D spatial indices ranging from 1 up to N_x , and from 1 to N_y , respectively. Some of the basic image matching functions can be defined over the paired pixel values x_{ij} and y_{ij} of these two images as follows [89] :

$$\sqrt{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (x_{ij} - y_{ij})^2} \dots \text{Euclidean metric.} \quad (2.1)$$

$$\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |x_{ij} - y_{ij}| \dots L_1 \text{ metric.} \quad (2.2)$$

$$\sqrt{\left(\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |x_{ij} - y_{ij}|^2 \right)} \dots L_2 \text{ metric.} \quad (2.3)$$

These metrics can be extended to p dimensions using the L_p metric, defined as follows :

$$\left(\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |x_{ij} - y_{ij}|^p \right)^{\frac{1}{p}} \dots L_p \text{ metric.} \quad (2.4)$$

Cross-correlation (CC) is another method by which intensities of different images can be matched, with notable applications in image registration [94, 95]. The cross-correlation between the two images X and Y with paired pixel values x_{ij} and y_{ij} , having mean intensities \bar{x} and \bar{y} respectively, and with standard deviations σ_x and σ_y , can be defined as :

$$CC(X, Y) = \frac{1}{N_x \times N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{(x_{ij} - \bar{x})(y_{ij} - \bar{y})}{\sigma_x \sigma_y}. \quad (2.5)$$

In the above normalized version of cross-correlation, subtraction of image intensities from their respective means (e.g., $x_{ij} - \bar{x}$), and division by their standard deviations ensures that affine transformations of the underlying intensities will not affect the similarity value. Normalized cross-correlation (NCC) between image pairs has been applied for characterization of appearance in infant brain images, primarily due to this feature of invariance to intensity scale changes [72].

Mutual information (MI), an information theoretic measure, has also been used to compare two images of different intensity [96, 97]. Consider two images, X and Y , with individual entropies $H(X)$ and $H(Y)$, respectively, and joint entropy $H(X, Y)$. The similarity between the two images in terms of their mutual information can be expressed as

$$MI(X, Y) = H(X) + H(Y) - H(X, Y). \quad (2.6)$$

Essentially, since the entropy of an image denotes the amount of information contained in it, this mutual information measure quantifies the amount of information in an image X that is

contained in image Y , or vice versa. Mutual information can be used to compare image pairs that might have undergone intensity changes (e.g., due to introduction of contrast agents), or image pairs belonging to different modalities. Unlike cross-correlation, which can be used to detect locally varying patterns, MI can be only used on large regions since it requires large sample sizes for statistical reliability. MI is also complex to define mathematically and model across time. In addition, the joint entropy estimation involves computation of the joint intensity distributions of corresponding voxels in both images, which requires voxel-wise correspondence for interimage appearance comparison.

As discussed in Chapter 1, the major drawback in current methods in appearance analysis is the need for spatial and intensity normalization. Although some of the methods, such as NCC, discussed above, capture intensity differences between images with little to no dependence on intensity normalization, they are still dependent on image shapes. That is, these methods assume that the two images X and Y being compared are already in perfect point-wise correspondence (i.e., the points x_{ij} and y_{ij} , as defined earlier in this section, are assumed to be in correspondence)². In this work, our goal is to develop a method in which this pointwise correspondence between image voxels or regions is not required for testing image similarity. The technique we use to accomplish this is intensity distribution based analysis - by extracting the intensity histogram of a region and extracting features of this histogram, intensity analysis can be performed without any dependence on spatial relationships.

Entropy as a measure, independent of MI-based computations, can also be used for assessment of appearance in a brain region. However, entropy calculations, besides being mathematically complex, are highly dependent on regional structures [98]. In a brain MR scan, the entropy of a region would depend on the ratio of WM and GM tissue volumes in that region. The structural dependence of entropy computations is a major drawback due to which it cannot be used to assess appearance in early brain MR studies.

Intensity distributions and histograms³ have been used in imaging and in some medical applications for characterization of appearance [99, 100].

²Alternately, these methods for computing the similarity between images are used in neuroimaging to obtain the best matching or registration between two images. This is done by computing the image registrations that maximise the similarity metric between them (e.g., to maximise Cross-Correlation (CC)) [94].

³The difference between the values of intensity histograms and intensity distributions lies in the quantities they represent. That is, intensity histograms generally contain frequency information related to occurrence of each intensity in the histogram's range. Intensity distributions, on the other hand, are normalized versions of these histograms and record the probability of occurrence of each intensity in the range of the distribution.

Features such as mean, median, and standard deviation can be computed from image histograms. These features can then be analyzed as quantitative markers that summarize information from high-dimensional histogram space. Multidimensional histograms can also be used for intensity analysis based on several images or image features. For example, joint histograms from images of an object obtained at different time instants have been used for characterizing the type of change taking place in the object [101]. An important point to consider while using histograms for appearance characterization is that studying histograms requires standardization of the intensity range of histograms under analysis (continuing dependence on effective intensity normalization). The accuracy of extracting region-specific and tissue-specific voxels, based on which the histograms are constructed, also plays a vital role in histogram analysis. While extraction of these voxels removes dependence on spatial locations of voxels, it could also introduce additional errors.

In this work, we characterize appearance in terms of pairwise relative differences between intensity distributions. Since features of the analyzed distributions are not independently extracted but studied in relation to another distribution extracted from the same image, the need for intensity normalization could be potentially reduced based on the methodology employed. Examples of pairwise analysis of distributions can be found in computer vision literature as well. An illustrative example consists of testing between similarity of image features by computing the normalized cross-correlation between intensity distributions of patches centered around the feature of interest [102].

Since we use this type of measurement of relative differences between two distributions for capturing intensity variations between two regions *within an image*, it could also be considered a form of contrast analysis. Contrast, although being closely related to appearance, focuses on patterns of intensity variation *within* a single image, rather than testing the similarity of intensities *between* images. Contrast in medical imaging, for example, is used to describe the apparent intensity differences between dissimilar tissue regions. In the following sections, contrast measurement within the context of medical imaging will be examined.

2.1.1 Contrast in Medical Imaging

In medical images, the definition of image contrast is highly dependent on the type of imaging system used. The metrics used to define appearance and contrast for a medical image can be chosen effectively only when the imaging modality used is taken into consideration. Signal intensity values in medical imaging modalities such as X-ray CT are directly proportional to an actual physical quantity (e.g., the attenuation, and increase

in atomic number, in X-ray CT). However, in MR the signal intensity values in images are quantities that are related in a complex manner to several factors, such as T1 and T2 relaxation times and proton density. Studies of contrast in NMR-based imaging have explored various mathematical expressions determining contrast between two regions, in terms of the magnetization between them [25]. The acquisition times of MR scans are also considered in this contrast computation. A phantom image constructed to demonstrate contrast between two tissues, A and B , is shown in Figure 2.1. For two interfacing tissues, A and B , the MR contrast can be expressed in terms of their local transverse magnetizations (defined in Chapter 1) as

$$C \propto M_{transverse}^A - M_{transverse}^B. \quad (2.7)$$

Given this definition of contrast, the contrast-to-noise ratio (CNR) can be written in terms of TR (Repetition time) and TS (Total scan time) as

$$CNR \propto \frac{TS}{TR} (M_{transverse}^A - M_{transverse}^B). \quad (2.8)$$

This equation highlights the dependence of contrast between tissues on the parameters of scans, in addition to the intrinsic tissue properties. Since the signal intensity of a point is a function of $M_{transverse}$, this definition of contrast in terms of MR physics can also be interpreted as being proportional to the difference between the signal intensity values of tissues A and B . Other representations of contrast include ratios of mean intensities of the regions A and B , denoted by $\frac{Mean(Intensity_A)}{Mean(Intensity_B)}$, and other functions of the mean intensities of two regions, e.g., $\frac{Mean(Intensity_A) - Mean(Intensity_B)}{Mean(Intensity_A) + Mean(Intensity_B)}$ [76, 77]. For our analysis, we adopt a method by which contrast is represented by divergence between the signal intensity distributions of the two regions being studied. This divergence is computed using the Hellinger distance, the reasoning for which will be discussed in the following sections.

2.2 Method

In the last section, a survey of methods for measurement of appearance and contrast from image processing and medical imaging perspectives was presented. In this section, we will adapt this knowledge for the development of an appearance measure relevant to MR-based neurodevelopmental studies.

One of the primary goals of this dissertation is the development of a stable method for quantifying appearance variations within a medical image that does not require intensity normalization. Specifically, the developed measure is designed to be an indicator of the intensity variation between WM and GM. In order to accomplish this, MR image appearance

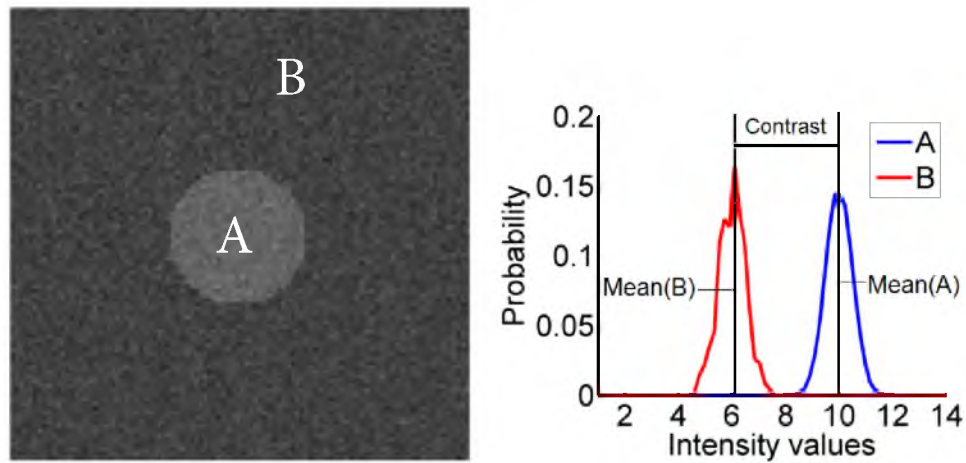


Figure 2.1: (Left) Phantom image illustrating the concept of contrast between two regions A and B, (right) Difference between mean intensities of regions A and B. The intensity distributions of regions A and B are separately normalized, thereby losing all volume information related to their spatial localizations, or the voxels they represent.

is measured in terms of the distance between intensity distributions of WM and GM tissue classes, computed using the Hellinger Distance. The developed measure is referred to by the acronym WIVID (White-gray Intensity Variation in Infant Development). The scientific reasoning behind the methodology used for computation of WIVID as well as detailed steps involved in this procedure are outlined below.

2.2.1 Framework for Infant Brain MRI Studies of Appearance

The framework for computation and modeling of WIVID consists of three major components. First, the entire image dataset is processed via a pipeline that performs spatiotemporal image registration and segmentation as well as parcellation of the brain into the major cortical regions of interest. This procedure segments the brain into major tissue classes and regions such that their intensity distributions and relative WM-GM intensity variation can be analyzed in a region-wise manner. The second step consists of quantification of this intertissue intensity variation for each major cortical region by finding the Hellinger Distance-based separation between gray matter and white matter intensity distributions, resulting in the measure referred to as WIVID. In addition to the magnitude of the WIVID measure, its direction is estimated based on its location on the contrast change curve in the early brain, which first decreases up to a minima point and then increases. Finally, the resulting WIVID values, which are now complete in their characterization due to both magnitude and direction being quantified, are modeled in a longitudinal manner using statistical modeling methods, constituting the third step. The parameters of the growth functions that result from the statistical modeling are then analyzed. The details of WIVID modeling are explored in later chapters. Figure 2.2 describes the entire framework outlined above.

2.2.2 Joint 4D Registration-Segmentation Pipeline

In order to study the changes in appearance that are seen in longitudinal datasets of the early brain, we implement a joint 4D registration-segmentation pipeline. This pipeline utilizes the longitudinal nature of the data to perform effective segmentation and parcellation of the infant brain - operations that divide the brain into major tissue classes and cortical regions, respectively. Obtaining accurate segmentation and parcellation of the brain is crucial to further analysis, particularly given the challenges facing these operations when applied to early infant brain images.

The results of implementing this pipeline include extraction of voxels belonging to each tissue class, as well as determination of the membership of these voxels to major cortical

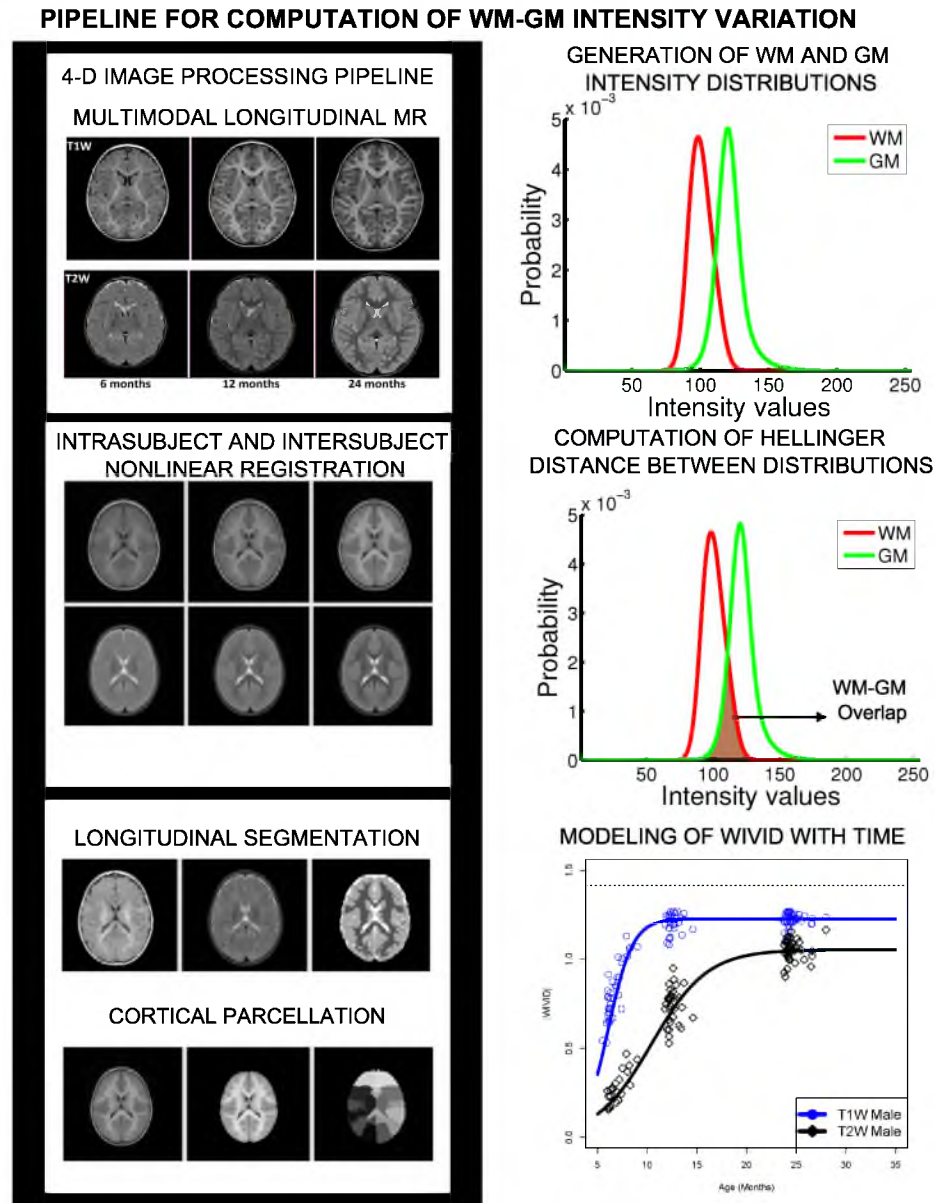


Figure 2.2: Framework outlining the registration-segmentation pipeline as well as the procedure to perform statistical analysis on the contrast.

regions. As a result, intensity distributions of tissue classes such as WM and GM can be generated for each major cortical region of the brain. Appearance variations between WM and GM for a specific cortical region can then be computed as the distance of separation between the WM and GM intensity distributions constructed from voxels in this region. Since spatial locations of the voxels used for generating distributions are not considered for computation of WIVID, and since normalization of the distribution belonging to each tissue class ensures removal of volumetric dependence, this appearance measure would not be affected by other factors such as shape, structure, and volume.

The first step in this procedure consists of bias correction and aligning the images rigidly with respect to a standard template image. Bias correction can be performed using a method such as N4-ITK [103] that estimates parameters of a smoothly varying bias field. Aligning with respect to a standard template ensures that the origins and coordinate of all images are standardized. Preliminary alignment is followed by intrasubject and intersubject registration between scans of the same subject acquired at different time points, the details of which are outlined in Figure 2.3, and described in the next sections.

2.2.2.1 Intrasubject Image Registration

Intrasubject image registration is the registration of all images belonging to the same subject, including those acquired at different time points, onto a common coordinate space. Intrasubject registration is accomplished using the ANTS algorithm based on symmetric normalization that uses Cross-Correlation (CC) as the metric of choice for matching of images within diffeomorphic maps [94]. The choice of this algorithm for intrasubject nonlinear deformation is motivated by several factors. Firstly, based on visual observation it was seen that this algorithm is highly effective, particularly in the case of infant image registration. Secondly, studies have indicated that this technique is consistently top-ranked in comparison to other popular registration methods [104]. Finally, ANTS is capable of jointly leveraging information from multiple modality scans for the purpose of achieving accurate brain image registration.

Intrasubject image registration can be expressed mathematically as follows. Consider a multimodal image dataset consisting of scanned modalities ranging from $m = 1, 2, \dots, M$. Scans of these modalities can be obtained for every subject $i = 1, 2, \dots, N$, at all K time points given by t_1, t_2, \dots, t_K ⁴. For a specific subject i , this series of multimodal

⁴For the sake of simplification, it is assumed here that each subject is scanned for the same number of time points (K), and at exactly the same time instants (t_1, t_2, \dots, t_K). This assumption is not applied in later chapters concerned with statistical modeling of this data.

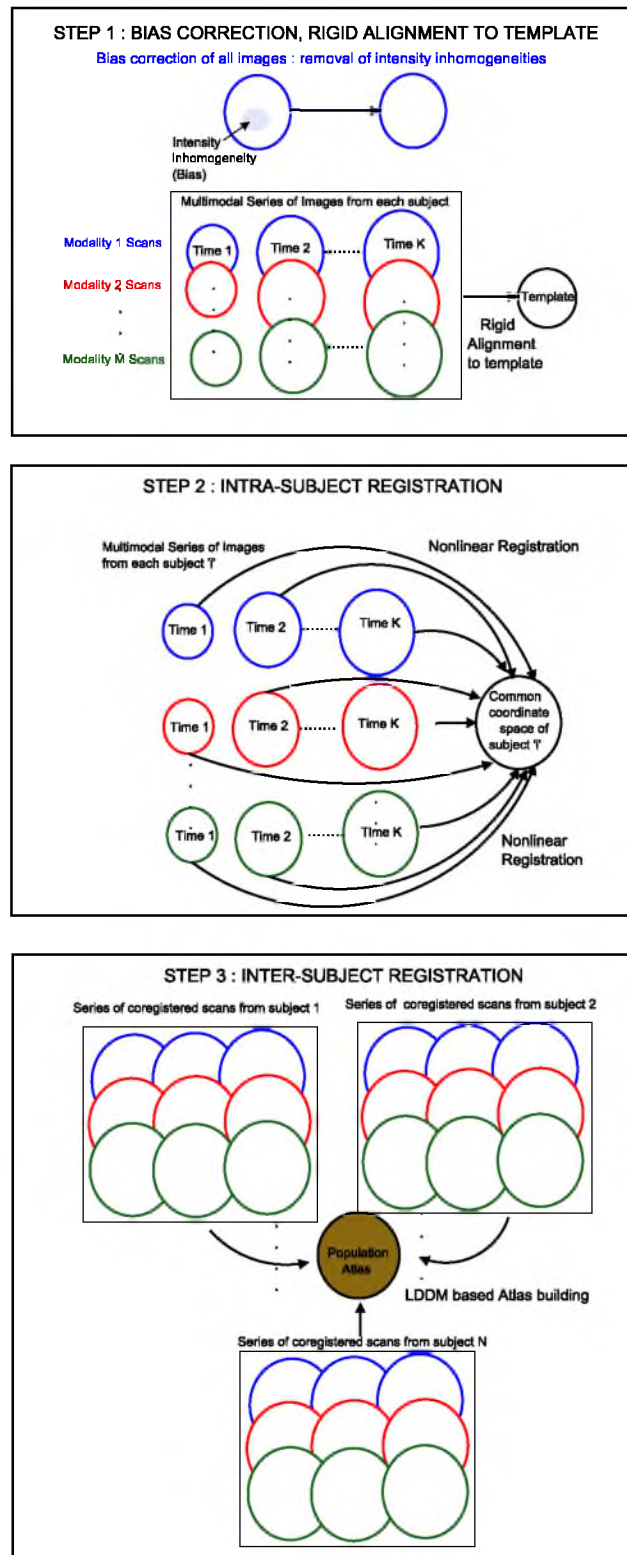


Figure 2.3: Framework outlining the intrasubject and intersubject registration pipeline.

images obtained at a time t_k is denoted by $I_{i,1}(t_k), I_{i,2}(t_k), \dots, I_{i,M}(t_k)$, and summarized as $\bar{I}_i(t_k)$. A series of such multimodal image sets from a subject i , obtained at *several* time points t_1, t_2, \dots, t_K can similarly be written as $(\bar{I}_i(t_1), \bar{I}_i(t_2), \dots, \bar{I}_i(t_K))$. During intrasubject registration, co-registration is performed across all scans belonging to a single subject i , obtained at different time points t and belonging to all m modalities. As a result, voxel x in a scan of modality m obtained at time t_k from subject i , denoted by $I_{i,m}(x, t_k)$, will correspond to the same voxel x in a scan of modality m' obtained from the same subject at time $t_{k'}$, denoted by $I_{i,m'}(x, t_{k'})$. This relation will be valid for all possible combinations of m, m', t_k , and $t_{k'}$, indicating that intrasubject registration is performed between scans of *all* modalities and between scans obtained at *all* time instants.

2.2.2.2 Intersubject Image Registration

Intrasubject registration is followed by optional intersubject image registration using an unbiased atlas building framework based on LDDMM (Large Deformation Diffeomorphic Metric Mapping) [105]. The choice of LDDMM is motivated by its unbiased method of spatial normalization and effectiveness in estimating large deformations. Datasets of infant brains always display large morphological variability, and their accurate standardization to an unbiased template space is possible with the large deformations estimated in LDDMM. Intersubject registration based on LDDMM results in the entire multimodal dataset being deformed to a common coordinate space, hence enabling population analysis of all the brain images in the dataset on a voxelwise basis and registration of the entire dataset to other reference atlases. Consider the voxel x from subject i 's scan belonging to the modality m and obtained at time t_k . Following intrasubject registration, this voxel will correspond to the voxel x in the m' th modality scan from another subject i' , obtained at time $t_{k'}$. That is, as a result of intersubject registration, voxels denoted by $I_{i,m}(x, t_k)$ and $I_{i',m'}(x, t_{k'})$ will have correspondence across all combinations of i, i', m, m', t_k , and $t_{k'}$.

The reason for applying this combined pipeline of intrasubject registration followed by intersubject registration is to improve co-registration quality by using the longitudinal nature of the dataset. Intrasubject registration usually consists of deformation of all images of a specific subject into the coordinate space of the latest time point images of that subject (it is assumed that the multimodal images belonging to the latest time point are co-registered before this). As a result, intersubject registration can be performed using the latest time point images of each subject alone (since all earlier time point images of each subject are already co-registered to this latest time point), as illustrated in Figure 2.3. Since the latest time point images are generally closest in appearance to the adult brain, they

have much better intertissue contrast and well-defined anatomical structures compared with images from earlier time points. As a result, intersubject registration based on the latest time point images will have improved quality compared with intersubject registration based on early time point images.

The above steps of intersubject and intrasubject image registration are not directly necessary for the appearance analysis framework, since computation of WIVID is done over large cortical regions and does not require voxel-wise correspondence across images in the dataset. However, these image registration steps enable effective segmentation and parcellation of the infant brain. Two facts support the hypothesis that accuracy of segmentation and parcellation operations on the infant brain, particularly on scans taken soon after birth, improve greatly by coregistration of the images and subsequent utilization of the longitudinal nature of the datasets. First, the brain at later stages of development (e.g., at 2 years of age) has much better intertissue contrast and less SNR compared with the brain at earlier stages, and second, the basic brain anatomy of an individual remains mostly the same across time in spite of volumetric changes due to very early gyrification [106, 107]. It can be concluded from these observations that segmentation and parcellation of early time point infant scans can be improved by using prior information from corresponding later time point scans of the same infant. Since accurate tissue segmentation and brain parcellation are essential for proper computation of WIVID, registration of all images obtained from a subject to the latest time point image of that subject is an important part of the appearance analysis pipeline.

2.2.2.3 Longitudinal Image Segmentation

Following coregistration, a segmentation procedure based on the expectation-maximization (EM) algorithm is implemented [108], which classifies each voxel in the image into one of the major tissue classes. In the case of brain images, segmentation of the images using EM algorithm results in every voxel in the image being classified into one of the major tissue classes $c_j = \{\text{white matter, gray matter, csf, non-brain}\}$. Consider the multimodal set of images denoted by $I_{i,1}(t_k), I_{i,2}(t_k), \dots, I_{i,M}(t_k)$ and summarized as $\bar{I}_i(t_k)$ that was introduced earlier. The segmentation procedure implemented is such that information from several modalities of the same subject is used to produce a single segmentation result. Since the result of segmentation is only a single set of tissue classification values for the entire tuple of scans belonging to subject i that were obtained at time t_k , this result can be denoted by the binary label map defining the presence or absence of a class c_j at a specific location x . This label map indicating membership to a tissue class c_j is given by

$L_i(c_j, x, t_k)$, for the image set $\bar{I}_i(t_k)$. In addition, priors based on the segmentation of later time point scans belonging to a specific subject could be used to get better segmentations of early time point scans from the same subject [106]. Consider the label map obtained from classification of the latest time point image into tissue class c_j - $L_i(c_j, x, t_{Latest})$. After intrasubject registration of all time point scans to the space of the latest time point image, the information from segmentation of the latest time point image, $L_i(c_j, x, t_{Latest})$, will be in correspondence with all the earlier time point images. When applied as prior in the Bayes theorem, $L_i(c_j, x, t_{Latest})$ can be used for computation of the label map at any earlier time point $L_i(c_j, x, t_k)$. The method of applying later-time point segmentations as probabilistic priors for improving the segmentation of early brain images is well established [106]. This procedure improves the accuracy of segmenting early time point images, and also ensures smooth transitions of segmented label maps of images of the same subject acquired across time. The details of this procedure are described in Figure 2.4.

2.2.2.4 Brain Image Parcellation

Since our analysis of contrast is region-specific, and based on anatomical regions of the brain, the final processing stage involves parcellation of the already registered brain images into major cortical regions. Parcellation atlases that have been obtained from past large-scale neuroimaging studies are deformed to the coordinate space of the set of scanned images, which can be done by deforming the template associated with it to the latest time point scan from each subject [50]. Alternatively, the template associated with a regional parcellation atlas could be deformed to the unbiased atlas template that was previously generated using the entire image dataset. A choice could be made between the two above methods based on the extent of intersubject shape variability in the dataset and the quality of the atlas template built.

Consider the same multimodal series of scans of a single subject acquired across time, which are coregistered such that voxel-wise correspondence has been established between each scan in the tuple $((\bar{I}_i(t_1), \bar{I}_i(t_2), \dots, \bar{I}_i(t_K)))$. Each voxel x of the parcellation atlas denoted by $Parcel(x)$ is now deformed such that it has correspondence with the same voxel x in all these co-registered scans (defined by $\bar{I}_i(x, t_k)$ in the general case). The membership probability of a voxel in the parcellation atlas template to a cortical region R_l can be given by the value $P_{Atlas}(R_l, x)$. After coregistration with the image set of a subject i , this atlas, which contains regional membership probabilities, can be redefined as $P_{Atlas,i}(R_l, x)$. Any voxel x in the image set $\bar{I}_i(x, t_k)$ will now be classified as belonging to a spatial region R_l , based on the value of $P_{Atlas,i}(R_l, x)$.

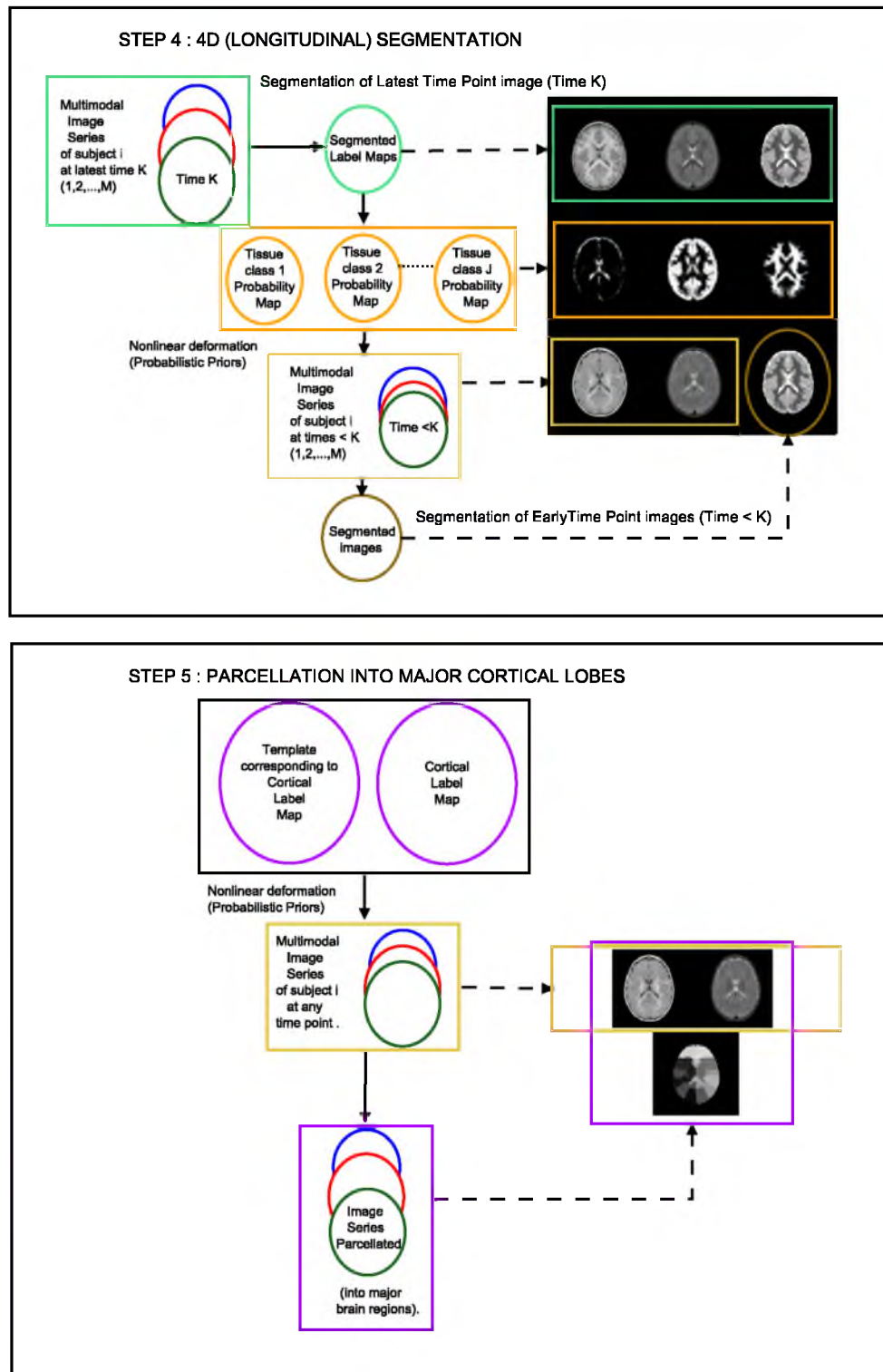


Figure 2.4: Framework outlining the longitudinal segmentation and parcellation pipeline.

As a result of the segmentation-parcellation pipeline outlined above, every voxel x in a brain image will have two associated values - the first defining membership to an anatomical region, and the second defining membership to a tissue class.⁴ Voxels belonging to a particular class and cortical region will then be used to generate intensity distributions for contrast analysis.

2.2.3 Removal of Shape and Volume Information

We require that any measure of appearance we develop should quantify *only* the variation in intensity between WM and GM tissues. Since this study is not concerned with volumetric and morphometric variations, the appearance measure developed must not be altered due to volumetric or morphometric changes. A possible method for achieving this invariance to morphometric and volumetric differences consists of removal of shape, structure, and volume differences via nonlinear image deformation. However, achieving accurate deformations that ensure voxel-wise correspondence is challenging, and is especially hard to accomplish given the large shape variations, low signal to noise ratios, and intensity variations present in infant brain image data. In order to maintain invariance with respect to volumetric and structural differences, we convert the spatially dependent intensity information contained in images into intensity distributions. The intensity distributions generated are simultaneously both region specific and tissue specific. Although the intensity distributions are region specific, they do not depend on the exact spatial position of contributing voxels. The image processing procedures for extracting voxels belonging to specific tissue classes and regions were discussed in detail above. As a result of these procedures, the WIVID appearance measure developed is computed solely as a function of WM and GM intensity distributions or histograms.

Intensity distributions are closely related to intensity histograms: while a histogram generally indicates the frequency of occurrence of each intensity value, the intensity distribution transforms this frequency into a probabilistic scale, indicating the probability of occurrence of each intensity value. Mathematically, for a range of intensities from Int_1 to Int_Q , with counts of occurrence of each intensity value Int_q denoted as $Count(Int_q)$, the

⁴The difference here between probabilistic maps associated with segmentation and parcellation must be noted. While the probabilistic maps associated with segmentation indicate the probability of each voxel in the image belonging to a tissue class such as WM or GM, the maps associated with parcellation indicate the probability of each voxel belonging to an anatomical brain region. As a result, each voxel has both membership to a tissue class and simultaneous membership to a brain region. This results in no dissonance, since the tissue membership is based on underlying biochemical properties that are captured via imaging, whereas regional membership is purely based on spatial localization to a brain region.

histogram value for an intensity Int_q is denoted by the number $Count(Int_q)$.

In the same scenario, the intensity distribution at each point can be expressed as the relative probability of occurrence of Int_q .

$$P(Int_q) = \frac{Count(Int_q)}{\sum_{q=1}^Q Count(Int_q)}. \quad (2.9)$$

We now extend this example to the case of several tissue classes. Each tissue class c_j will have an associated intensity distribution. The distribution value for intensity Int_q is denoted by $P(Int_q|c_j)$. This mathematically verifies that the intensity distribution of the tissue class c_j only retains information regarding the likelihood of a certain intensity value occurring for voxels belonging to that tissue class, while removing spatial information related to *where* the voxels are located. Similarly, normalization of the probabilistic distribution by the total number of voxels belonging to the tissue class ensures that the intensity distributions are not distorted by volumetric information.

Consider a series of individual scans of various modalities, obtained at a time point t_k , which can be given by $I_{i,1}(t_k), I_{i,2}(t_k), \dots, I_{i,M}(t_k)$. A function denoting the distribution of intensities is constructed for each modality, for all combinations of tissue class c_j and regions R_l . This function is computed based on the likelihoods of voxels belonging to this tissue class, defined previously as $L_i(c_j, x, t_k)$, and based on the membership of voxels to anatomical regions, given by $P_{Atlas,i}(R_l, x)$. As a result, this distribution is a function of the intensity values of image voxels x that belong to tissue class c_j and anatomical region R_l .

$$P_{i,m}(Int_q, t_k|c_j, R_l) = f(Int(I_{i,m}(x, t_k|x \in c_j, R_l))). \quad (2.10)$$

2.2.3.1 Construction of Intensity Distributions

As a part of the processing pipeline outlined above, intensity normalization can optionally be performed in the manner specified in Appendix A, although this is not strictly necessary. The only advantage with implementing intensity normalization at this stage concerns generation of intensity distributions. Heuristic values for parameters used in generation of intensity distributions could be set more easily across all subjects if the images have been normalized. Apart from this, intensity normalization should not affect the values resulting from the WIVID pipeline. The adoption of Kernel Density Estimation for construction of intensity distributions further minimizes the dependence of the WIVID pipeline on intensity normalization.

Construction of actual intensity distributions from voxels belonging to specific tissue classes and regions can be done using several techniques. The simplest method is to

parameterize the distribution such that it follows a pre-existing model such as a Gaussian, Beta, or Exponential distribution, for example. In this case, if the fitted model has N parameters, the intensity distribution $P(Int_q)$, which is originally $q - dimensional$, will be reduced to a much sparser, $N - dimensional$ representation. The intensity data from infant brain images were fit to major distribution models (including the Gaussian) and the resulting fit was verified using the Chi-squared test. It was concluded that based on empirical evidence from available image datasets, the intensity data failed to conform to any of the popularly used distribution models. As a result, the effectiveness of nonparametric representations of the distribution data was investigated instead.

Construction of $q - dimensional$ intensity distributions is an effective alternative if data fail to conform to parametric models. Earlier, a method for generation of an intensity distribution from a histogram, discretized to q -levels, was described. However, histogram-based distributions have drawbacks such as discretization errors associated with binning, and variability based on the choice of bins and binsize, as displayed in Figure 2.5. To overcome these drawbacks, we opt for a continuous, smoothly varying intensity distribution that is estimated using kernel-based methods.

Kernel Density Estimation (KDE) is based on the principle that rather than counting the number of voxels that belong to a certain intensity bin, the probability distribution at a point can be a cumulative sum of the weighted probabilities of voxels with similar intensities. That is, the intensity of a voxel determines a weight factor, thus influencing the probability distribution of intensities in its vicinity. If the weighting influence of each voxel's intensity on neighboring intensity values is smooth such as in the case of a Gaussian kernel, a smooth distribution results.

In this work, KDE using a Gaussian kernel G is used to obtain a smooth and continuous intensity distribution for each tissue class c_j and region R_l , belonging to each image of modality m - $I_{i,m}$. This distribution is generated from the voxels that belong to the tissue class and region under analysis. The intensity distribution corresponding to an intensity Int_q is denoted as $P_{i,m}(Int_q, t_k | c_j, R_l)$. This value, i.e., the probability of a signal intensity value Int_q being exhibited by voxels belonging to the tissue class c_j and region R_l , for a subject i 's scan of modality m is computed by the equation

$$P_{i,m}(Int_q, t_k | c_j, R_l) = \sum_{x \in c_j, R_l} G \left(\frac{Int_q - Int(I_{i,m}(x))}{h} \right). \quad (2.11)$$

Here, the intensity of a voxel x belonging to subject i and the modality m scan is given by $Int(I_{i,m}(x))$, and h is the bandwidth of the kernel. Using this equation, a continuous

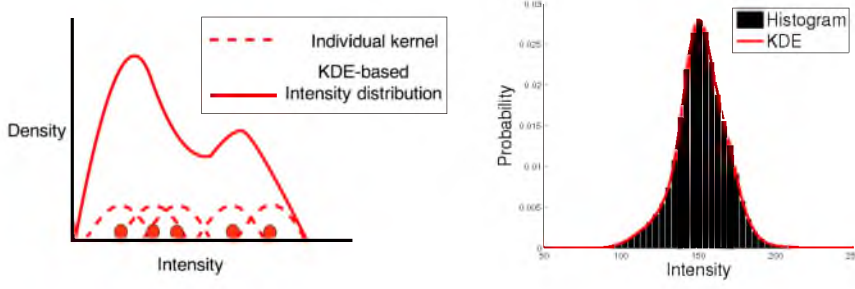


Figure 2.5: KDE procedure, and (b) Comparison between distributions from histogram-based and KDE analysis.

probability distribution $P_{i,m}(Int|c_j)$ is generated for each class c_j over all possible values of intensity Int , for a scan belonging to subject i and of modality m . An example of intensity histograms that are generated for the T1W and T2W modalities at 3 different time points can be seen in Figure 2.6. It should be noted that individual histograms are generated for each brain imaging modality.

Rather than using only voxels that have been classified as belonging to a tissue class in a binary manner, the label maps of the underlying voxels can be used for fuzzy membership-based generation of the KDE. Here, the contribution of the voxel x to the distribution of a tissue class c_j , will be weighted by its membership value $L_i(c_j, x, t_k)$ to this tissue class. In this case, the KDE equation can be redefined for computation of distribution at a single intensity value Int_q as

$$P_{i,m}(Int_q, t_k|c_j, R_l) = \sum_{x \in R_l} L_i(c_j, x, t_k) \times G\left(\frac{Int_q - Int(I_{i,m}(x))}{h}\right). \quad (2.12)$$

When this probability is computed over the entire range of intensity values, it results in a complete intensity distribution $P_{i,m}(Int_q, t_k|c_j, R_l)$. This representation could be advantageous since it reduces excessive dependence on binary label maps from segmentations, instead basing its results on probabilistic (or fuzzy) memberships to tissue classes. However, this approach is not adopted since the behavior of distributions computed in this manner could be highly complex and unstable, in addition to this representation being relatively less understood.

Based on empirical evidence, if the intensities under study ranged from 0 to 255, a Gaussian kernel of width 3 was used for generation of the KDE. Since these operations are quantized based on underlying intensities, the final KDE value is recorded for intensities separated by a stepsize of 0.1.

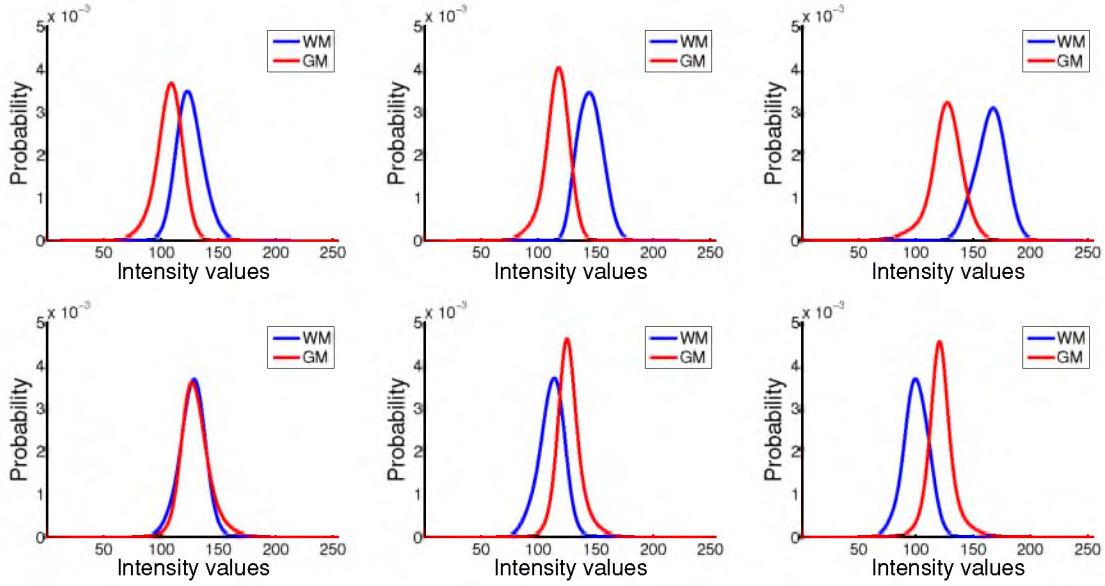


Figure 2.6: Intensity distributions of gray matter (red) and white matter (blue) changing with time for a single subject scanned at 6 months (leftmost column), 1 year (central column), and 2 years (rightmost column) of age, with the T1W images in the top row and the T2W images in the bottom row. Since each individual distribution has been normalized such that the area under it sums to one, the relative size of WM and GM distributions does not convey information related to their respective volumes.

2.2.4 Computation of Distance Between Intensity Distributions

Following extraction of intensity distributions, the computation of intensity variations is undertaken purely in the intensity domain - that is, by measuring the overlap between the probabilistic intensity distributions of WM and GM. The measure of distance adopted will be used to quantify the degree of separation between two intensity distributions, hence serving as an indicator of their intensity differences. An example of WM and GM intensity distributions computed for a single subject at different time points can be seen in Figure 2.6. A high degree of overlap between WM and GM intensity distributions will correspond to low contrast.

The primary role of a distance measure between two distributions is to summarize information from the high-dimensional space of the distribution to a lower-dimensional space. Only a subset of distance measures qualifies as distance *metrics*. The basic properties a distance measure must satisfy to qualify as a distance metric include symmetry, triangle inequality, non-negativity, and identity of discernibles. Considering two distributions, $P1$ and $P2$, these properties are briefly defined below for the distance measure D between them

$$D(P1, P2) = D(P2, P1) \dots \text{Symmetry.} \quad (2.13)$$

Considering a third distribution, $P3$,

$$D(P1, P2) \leq D(P1, P3) + D(P3, P2) \dots \text{Triangle Inequality.} \quad (2.14)$$

$$D(P1, P2) \geq 0 \dots \text{Non-negativity.} \quad (2.15)$$

$$D(P1, P2) = 0, \iff P1 = P2 \dots \text{Identity of Discernibles.} \quad (2.16)$$

The Hellinger Distance (HD), which satisfies all properties of a metric, is used to measure the overlap between the probability distributions of the distinct tissue classes [109]. Considering any two distributions $P1$ and $P2$ defined over a range of values y , the HD can be defined mathematically as [110]

$$HD(P1, P2) = \sqrt{\left[\int_y \left(\sqrt{P1(y)} - \sqrt{P2(y)} \right)^2 \right]}. \quad (2.17)$$

In the case of distributions that are defined for a discrete set of values indexed by y and ranging from 1 to Y , the HD can be redefined as

$$HD(P1, P2) = \sqrt{\left[\sum_{y=1}^Y \left(\sqrt{P1(y)} - \sqrt{P2(y)} \right)^2 \right]}. \quad (2.18)$$

Alternately, the Hellinger Distance can be generated in terms of the Bhattacharyya coefficient (BC). The Bhattacharyya coefficient can be defined for the continuous case as

$$BC(P1, P2) = \int_y \sqrt{P1(y)P2(y)} dy. \quad (2.19)$$

The Hellinger Distance can in turn be defined in terms of the Bhattacharyya coefficient as

$$HD(P1, P2) = \sqrt{2(1 - BC(P1, P2))}. \quad (2.20)$$

That is, the Hellinger Distance can alternatively be defined as

$$HD(P1, P2) = \sqrt{2(1 - \int_y \sqrt{P1(y)P2(y)} dy)}. \quad (2.21)$$

Finally, in the case that the distribution data we analyze is of Gaussian nature, a closed-form solution for the Hellinger Distance between two such Gaussian distributions exists. Consider two Gaussian distributions $P1$ and $P2$, with respective standard deviations $Std1$ and $Std2$, and Mean values $StdPooled = \text{Pooled Standard Deviation across the two distributions}$,

$Mean1$ = Mean of $P1$, $Mean2$ = Mean of $P2$. The Hellinger Distance between them is given by

$$HD(P1, P2) = \sqrt{2} \times \left(1 - \frac{(Std1^{\frac{1}{4}})(Std2^{\frac{1}{4}})}{StdPooled^{\frac{1}{2}}} \times e^{-\frac{(Mean1 - Mean2)^2}{StdPooled}} \right). \quad (2.22)$$

It follows from the above equations and the properties of probability distributions that the Hellinger Distance is bounded in the range from 0 to $\sqrt{2}$. The overlap measure based on the Hellinger Distance is an indicator of the divergence or separation in their intensity distributions, and is a bin-to-bin distance measure.

The major reasons for choice of Hellinger Distance for computing the divergence between intensity distributions are listed below. The metric properties of the Hellinger distance, along with other useful properties it possesses such as boundedness, make it suitable for the purpose of measuring WM-GM intensity divergence. Lack of symmetry in the KL distance computation, and triangle inequality in the Bhattacharyya distance computation, eliminate the possibility of using these distances in this work. The Hellinger Distance-based method outlined above, used for computation of intensity variation or intertissue contrast, is also not limited by Gaussian or parametric assumptions of nature of intensity distributions. For example, the Mahalanobis distance is generally applicable only to Gaussian distributions. It is also independent of the sample sizes used for estimating the distribution, a property that is not satisfied by other measures such as the Chi-squared distance [111]. Importantly, the Hellinger Distance computed remains invariant to affine transformations of the underlying distributions. This property is crucial to the measure we adopt since the measure we develop must be stable with respect to affine intensity transformations, particularly those caused by variations in external scanning conditions. Distances such as the L_1 -norm, L_2 -norm, and Earth mover's Distance (EMD) fail to remain the same with affine transformations of underlying intensities. Finally, it is also well known in areas of computer vision and image recognition that using a distance measure such as the Hellinger Distance to compare two histograms leads to better results than using other measures such as the regular Euclidean distance measure [112].

Following this reasoning, the measure of tissue intensity variation is then computed in terms of the Hellinger Distance between the intensity distributions of WM and GM tissue classes. In our application, the WIVID value is measured in a spatio-temporal framework - being both specific to each region of the brain being studied and to the time point of scan. The intensity distributions of the gray and white matter tissue classes are independently generated using KDE for each major cortical region of the brain. The WIVID value of

the k th scan of subject i , obtained at time point t_k , for the region R_l , is denoted as $WIVID_{i,m}^{R_l}(t_k)$. Since White-gray Intensity Variation in Infant Development, denoted by the acronym $WIVID$, measures the distance between WM and GM intensity distributions, it can be defined based on the equation given above for the Hellinger Distance as

$$WIVID_{i,m}^{R_l}(t_k) = HD(P_{i,m}(Int, t_k | c_j = WM, R_l), P_{i,m}(Int | c_j = GM, R_l)). \quad (2.23)$$

Some important properties of the WIVID distance measure are listed below, based on properties of the Hellinger Distance (for ease of representation, the distance measure is denoted just by the symbol $WIVID$, with references to region under study R_l , time point t_k , and the subject and modality details i and m , respectively, being excluded):

1) As the divergence between WM and GM distributions increases, the value of the WIVID measure also increases, and vice versa.

2) The WIVID measure between WM and GM intensity distributions remains the same even if WM and GM intensity distributions are interchanged (property of symmetry). This property is of importance particularly since intensity distributions of WM and GM are reversed during the course of early brain development.

$$WIVID = HD(P'(Int | c_j = WM), P'(Int | c_j = GM)). \quad (2.24)$$

In the above expression, $P'(Int | c_j = WM)$ corresponds to $P(Int | c_j = GM)$, and $P'(Int | c_j = GM)$ corresponds to $P(Int | c_j = WM)$.

3) Unless the intensity distributions of WM and GM are identical, the WIVID measure would not be 0 (property of identity), which ensures that only when the tissue appearance of WM and GM regions is alike would the WIVID measure between them be 0.

$$WIVID = 0, \iff P(Int | c_j = WM') = P(Int | c_j = GM'). \quad (2.25)$$

This property also emphasizes the principle that WIVID captures shape differences between intensity distributions - it will not be zero-valued unless there are no shape differences between intensity distributions, even if the means or medians of these distributions are equal. This property, that the measure not only reflects a shift between distributions but at the same time the difference between the shapes of the distributions, becomes important when interpreting results from longitudinal analysis.

4) The WIVID measure is always positive (follows from property of non-negativity). Modeling of the WIVID measure is hence restricted purely to the positive range, although

this restriction can be relaxed by considering the directionality of shifts between distributions by calculating as a sign (see following subsection).

$$WIVID \geq 0. \quad (2.26)$$

Besides these mathematical properties, the WIVID measure has two very important characteristics: it is computed in a region-specific and modality-specific manner. That is, the WIVID values are computed independently for each modality scanned and for each brain region within those scans. This characteristic results from our framework for computation of tissue intensity variation, rather than from the intrinsic nature of the Hellinger Distance measure. By defining a modality-specific WIVID value, the vast differences in tissue appearances that are seen in scans of different modalities are accounted for. The origin for this stems from the nature of the MR signal being recorded in each modality, which has been discussed in Chapter 1. The regional specificity of the computed WIVID measure ensures that variations in tissue appearance between different regions are represented and can be further analysed. Regional specificity is achieved by parcellation of the brain into the major cortical regions.

Finally, the WIVID measure as computed using the Hellinger Distance has the property of being invariant to affine transformations of the underlying intensity distributions [113, 114]. It is evident that since the Hellinger Distance measures values on the space of probabilities, it would be invariant to global translation of the underlying intensities. Global translation would result only in all the probabilities being shifted in terms of their range, while remaining the same in magnitude. Invariance to scaling follows from mathematical properties that are described by Gibbs *et al.* [113]. As a result of these properties, scaling and shifting the WM and GM intensity distributions would not change the computed WIVID value. This property is central to the concept of WIVID being stable and independent of intensity normalization techniques. Since WIVID is invariant to affine transformations of intensity distributions, normalization procedures involving simple scaling or shifting will not change the WIVID result. Similarly, any other intensity inhomogeneity that shifts or scales both WM and GM intensity distributions to an equal extent will not alter the WIVID result.

As an extension to this principle, the WIVID measure is also invariant to any invertible nonlinear deformation to the underlying intensity distributions [115]. However, in the practical case this principle would be true only dependent on the type of binning or quantization that is performed for generation of the intensity distribution of interest. In the case of an affine transformation, it would be easy to quantize or bin the data by appropriate scaling

or translation in such a manner that its essential variations and properties are preserved. However, in the case of a nonlinear transformation, a data point might be transformed nonlinearly, but due to binning and quantization methods it might not contribute to the same point on the intensity distribution. This type of transformation to the intensity distribution might result in a modified value of the Hellinger Distance. It is also important to note that only the application of *invertible* deformations to both WM and GM intensity distributions results in invariant Hellinger Distance values.

A further assumption we make in this case is that any nonlinear scaling or deformation that is applied to the tissue intensities is uniform across WM and GM tissue classes. If only a small subregion of the WM and GM tissue classes is affected by a transformation, a modified Hellinger Distance would result. In contrast, if a transformation is applied to the entire WM or GM region in a cortical lobe, the resulting Hellinger Distance remains the same. Since the cortical regions we consider are reasonably sized, this assumption would be applicable to our analysis.

2.3 WIVID Measure: Directionality and Simulations

Following the detailed description of the properties and method for computation of the WIVID measure, the performance of this measure will be evaluated in this section and directionality attributes will be assigned. Initially, behavior of the WIVID measure for artificially translated WM distributions will be analyzed. Based on this analysis, a method for characterizing the direction of WIVID will be described. Finally, a series of phantom images with differential variations in tissue intensity, similar to the tissue appearance differences observed during early brain growth, will be evaluated.

2.3.1 WIVID Measure Behavior for Shifted WM Distribution

In this section, we assess behavior of the WIVID measure in response to a simple experiment. WM and GM intensity distributions are extracted from a sample infant MR scan. While keeping the GM distribution as constant, the WM intensity distribution is shifted or translated to both left and right of its current position, as observable in Figure 2.7. The resulting WIVID curve is recorded in Figure 2.8. Typically, the pattern of WM-GM contrast variations seen in infants can be observed to be similar to this pattern, i.e, consisting of a decreasing portion followed by an increasing portion.

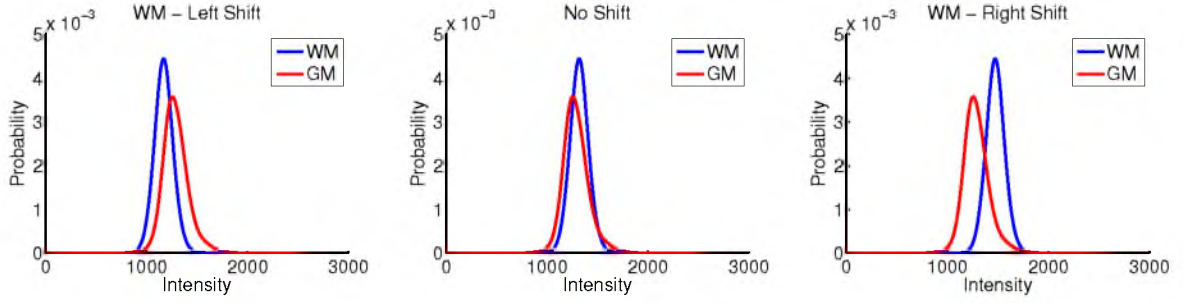


Figure 2.7: WM-GM configurations at different points shown in the previous Figure above: (Left) Plot with WM distribution shifted to the left from its original configuration, (Center) plot indicating original configuration of WM-GM distribution at point being studied, with no shifts in the distribution, and (Right) plot with WM distribution shifted to the right from its original configuration.

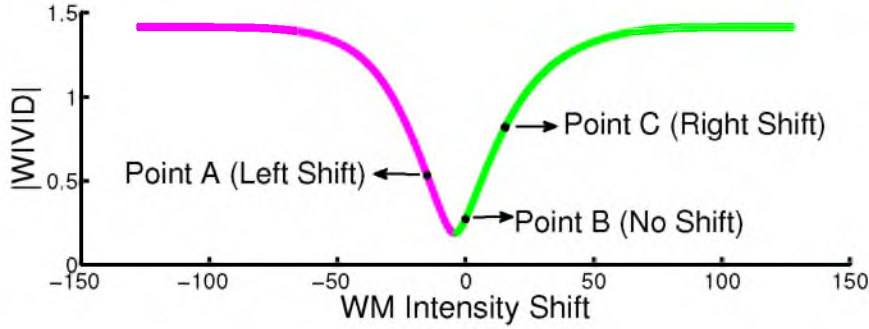


Figure 2.8: Plot indicating change in behavior of WIVID with shift of WM intensity distribution, keeping GM distribution as constant, with both positive (green) and negative (pink) gradient portions of the WIVID trajectory being displayed.

2.3.2 WIVID Measure Directionality

Consider the case that no directional attribute is assigned to the WIVID points analyzed in Figure 2.8. In this case, the WIVID values that result from left-shift of WM intensity and right-shift of WM intensity would be identical. However, it is known that the WM distribution underwent a shift in the process and converted from the left-shift configuration to the right-shift configuration. This difference between the left-shifted and right-shifted WM-GM intensity configurations is captured by the directional attribute.

Since we want our procedure for assigning a directional attribute to be generalizable for any modality studied, we study the common trajectory of WIVID contrast change over time in both T1 and T2 modalities, rather than the intensity gradient from WM to GM. The WIVID contrast in all modalities first decreases, reaching a minimum close to zero, and then increases over time, during early brain development. WIVID contrast values on either

side of the minimum have opposite WM-GM gradient directions. By the convention we assume, a WIVID contrast point with WM-GM configuration that maps to the decreasing portion of the curve is assigned a negative direction. If a WIVID contrast point has a WM-GM configuration that maps to the increasing portion of the curve, it is assigned a positive direction. This assumption is common to both T1W and T2W WIVID values.

We will now address the issue of how a particular configuration of WM-GM intensity distributions could be mapped onto the WIVID contrast change trajectory. Specifically, this can be done for a pair of WM and GM distributions by shifting WM distributions to both left and right while maintaining the GM distribution to be constant, as shown in Figure 2.7. The point of highest overlap between WM and GM distributions is the point at which the WIVID curve reaches a minimum. This point of lowest WIVID contrast can be determined by computing the WIVID measure for each WM-GM configuration generated by progressively shifting the WM.

To evaluate if the WM-GM configuration at a particular time has positive directionality, we can ask the question, “If the WM distribution is shifted to the left and right of its current position while keeping GM constant, does the point of lowest contrast lie on the left-shifted portion or right-shifted portion of the configuration?” If the point of lowest contrast lies to the left of the current configuration, it can be concluded that the WM-GM distributions currently lie on the curve with positive gradient and can hence be assigned a positive direction based on the convention defined above. Along similar lines, if the point of lowest contrast lies to the right of the current configuration, it can be concluded that the WM-GM distributions currently lie on the curve with negative gradient and can hence be assigned a positive direction based on the convention defined above.

The directionality attribute is assigned based on some assumptions. The major assumption, that the WM and GM distributions do not change vastly in shape with time such that the WIVID trajectory varies irregularly, is valid based on the behavior of WM and GM distributions as seen in the histogram images. From the behavior of WM and GM intensity distributions, the nature of the WIVID curve that we assume is reasonable, specifically since based on experiments it was shown that the infant brain has relatively constant GM intensity over time, while WM intensity shifts over time. Biological explanations for this observation have also been discussed in Chapter 1.

A point of scientific discussion at this juncture would be the choice of method for assigning directionality. Prior to implementing this solution, a signal intensity ratio based on ratio of mean or median values between WM and GM distributions was used to characterize

direction [116]. If the signal intensity ratio between WM and GM median values, for example, was greater than 1, the WIVID measure could be assigned a positive direction and vice versa. However, the mean or median values are incapable of representing the behavior of the entire distribution, particularly in cases of irregularly shaped distributions. In such cases the method described above could provide a more effective measure of directionality. Further, T1W and T2W modalities have opposite direction for WM-GM intensity gradient. Therefore, a method that assigns direction to WIVID contrast based on the ratio of signal intensities between WM and GM would be confusing to interpret for different modalities.

2.3.3 Phantom Image Studies

A series of four phantoms assumed to be longitudinal in time was created from Gaussian intensity models, as can be seen in Figure 2.9. Each phantom consists of the two regions *A* and *B* of the same shape, with the only difference being the intensities of these regions. The intensity variations between the two regions are high in Phantom 1 and lower in Phantom 2. Phantom 3 is of almost zero intensity variation as no contrast is seen between the two regions. Phantom 4 again has nonzero intensity variation, with the intensity gradient being of opposite direction when compared with Phantoms 1 and 2. That is, the intensities of the two regions underwent a reversal of relative intensities.

In the initial phase, Region *A* is of lower intensity than Region *B* in Phantom 1 and 2. In the middle phase, Region *A* and Region *B* have equal intensities. In the final phase, Region *A* is of higher intensity than Region *B*. These interregion patterns of intensity variation can be clearly observed in the intensity distributions displayed. Therefore, although the magnitude of the WIVID value for Phantom 1 might be the same as for Phantoms 2 or 3, the direction of their intensity gradient should be considered. The direction of the intensity gradient can be obtained by several methods, including finding the ratio between mean intensities of Regions *A* and *B*. However, in this study, we find the direction of intensity gradient by shifting one histogram while keeping the other constant. The exact procedure for assigning a direction sign to the WIVID value is discussed in a later section. To understand the phantom image series, it is sufficient to note that the direction of relative intensity gradient between the two regions *A* and *B* undergoes a reversal with time, changing from -1 , to zero at point of no contrast, to $+1$ for the last phantom image. This change in the WIVID value, and in its gradient, can be seen in Figure 2.10.

In conjunction with the absolute WIVID values, which capture the change in magnitude of relative intensity variation, the directional information related to the intensity gradient provides a complete picture of the appearance changes taking place in the phantom data.

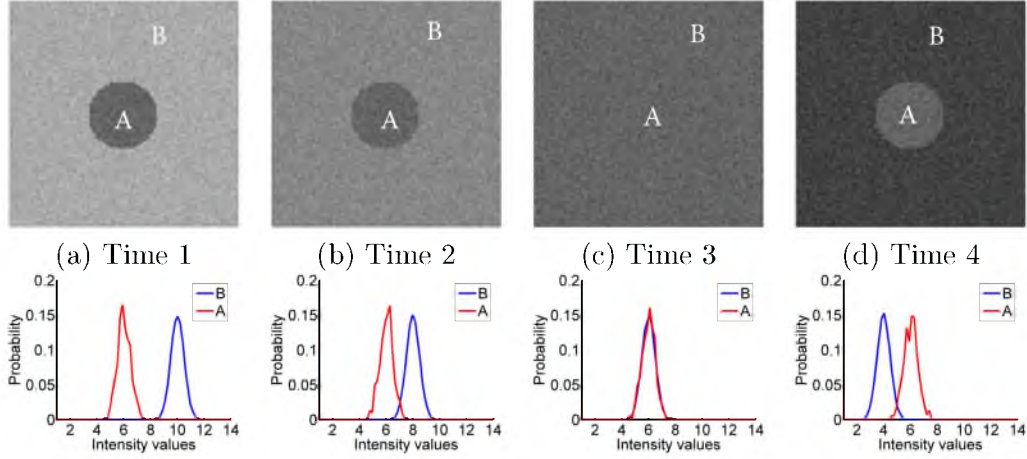


Figure 2.9: (Top row) Series of phantom images corresponding to longitudinal time series from a single subject. The phantom images have two distinct regions - A (circular foreground) and B (background). (Bottom row) Series of histograms of Region A (blue) and Region B (red) corresponding to each phantom image in the same column in the top row.

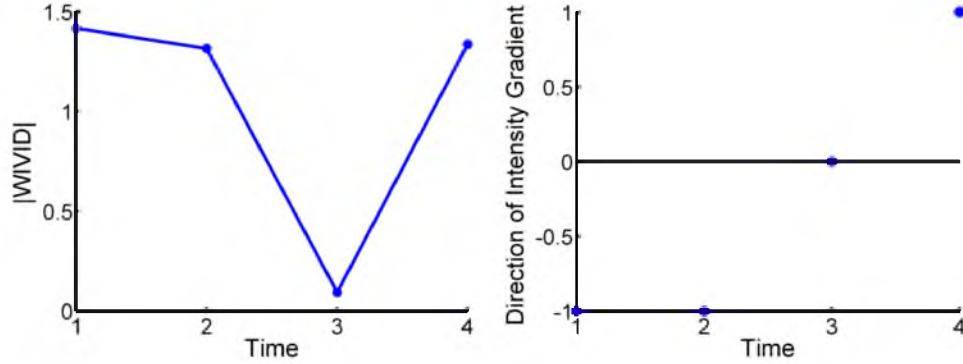


Figure 2.10: (Left) Absolute WIVID values for phantom Images over time, (Right) Direction of relative intensity gradient between two Regions A and B.

The addition of the direction information reflects the reversal in intensity gradient seen between Regions A and B. From this discussion it can be understood that the WIVID measure, which has both magnitude and direction, is capable of characterizing intensity variation between regions.

2.4 Stability of WIVID Measure

In this section, we will study the stability of appearance measures in 3D MR imaging. After conducting a literature survey of relevant papers in both statistics and neuroimaging, we will outline a methodology for stability analysis. We will then present results from systematically applying this methodology on a series of repeated scans of two human

traveling phantoms, acquired across different locations and scanner platforms. The primary goal of this analysis is to comparatively evaluate the stability of signal intensity and Hellinger-Distance based WIVID measures as regional appearance measures. With respect to the traveling phantom studies, however, it should be noted that it is not possible to compare signal intensity and the WIVID measure directly since they are entirely different quantities. Rather we independently compute the stability of intensity and WIVID in terms of their coefficients of variation and compare the results based on this coefficient.

Papers specifically dealing with the issues of stability, reliability and repeatability, especially in the context of neuroimaging, have become increasingly important in the context of large-scale, multisite, neuroimaging studies. A major multisite neuroimaging study of reliability [117] used a dataset similar to that of the human traveling phantom we employed - a single subject was scanned twice within a 24-hour time window. Further, these repeated scans were obtained at 5 different MR sites over a period of 6 weeks. The age of the subject (25 years), along with absence of physical and mental illness, suggests that the brain remained the same during the 6-week period. Three types of analysis were undertaken - an evaluation of reproducibility of different image analysis methods within a single site, between different sites with the same type of scanner, and between sites with different scanner types. Importantly, this study used COV (Coefficient of Variation) as a quantitative tool to study stability under varying conditions. Another research study compared different methods of tracking fibers using DTI [118], and performed repeatability analysis to confirm which method shows the greatest stability. This study also performed sensitivity analysis, exploring statistical methods to establish the method of obtaining DTI fibers that are most effective with respect to identifying microstructural alterations in MS patients.

Since we wish to apply our findings from the traveling phantom dataset to the study of infant brain development, it is necessary to fully understand how imaging of the infant brain is different when compared with that of the adult. A paper on pediatric neuroimaging [119] describes confounding factors that are specific to infant neuroimaging. Although the paper specifically addresses problems faced in functional neuroimaging, many of the same principles also apply to structural MR. Since magnets are shimmed to offer the highest signal at the center of the coil, which is generally out of reach for children (due to shorter necks), the resulting images have a lower signal to noise ratio. In addition, subjects who move more also show greater signal to noise ratio, thereby adding risk to studies dealing with children with attention and behavioral disorders. Although motion

correction algorithms are employed to correct motion-induced distortion, this is often not fully removed. Differences in skull thickness can also affect the signal to noise ratio of MR images. Finally, to study raw signal intensity, factors such as field inhomogeneity, head position, head motion, and other variables need to be carefully studied. Another publication summarizes these variations as occurring at two levels [117]. At the level of data acquisition, possible sources of variation could be patient positioning, scanner geometry, scanner intensity variation, and discrete image artifacts. At the level of image analysis, variation could occur due to different procedures for image registration, bias field correction, interpolation, and manual interaction. Finally, a study of stability related to cortical thickness measurements [120] also discusses the effect of the subject being scanned, such as hydration status of the subject, in addition to instrumentation. The importance of stability studies in the context of longitudinal data is also emphasized. Since longitudinal studies face challenges associated with both subject-related as well as instrument-related factors such as major scanner updates, it is critical to understand these thoroughly.

Depending on the type of study conducted, confounding factors that depend on the history of the subject and their treatment also need to be studied. For instance, in a critical review of neuroimaging-based ADHD studies, the effect of factors that might indirectly contribute to the imaging results is examined [121]. For example, the usage of psychotropic drugs for ADHD treatment might significantly reduce dopamine uptake in specific areas of the brain, thereby changing the appearance of scanned images.

Critical to our analysis is a paper by BIRN [122] which analyzes the calibration of multi-site structural MRI by accurately correcting for gradient-induced distortions, in order to allow cross-site comparisons of morphometry by minimizing dependence on on-site factors. Most relevant to our study of appearance measures is the quantification of per-voxel intensity variability defined in this paper, which is computed as $\text{std dev.}/\text{voxel mean}$. In this paper, the results of this variability are shown for a single subject test-retest both within and across sites, both before and after geometric distortion correction. This paper concludes, based on histogram analysis, that this per-voxel variability is highly reduced after geometric correction.

2.4.1 Preprocessing Pipeline

The traveling phantom study was designed to calibrate image data in a large multisite pediatric neuroimaging study, and includes two subjects (Phantom 1 and Phantom 2) who have undergone repeated scans at various imaging sites. Two scanners, a Siemens 3T Allegra head-only scanner and a Siemens 3T Tim Trio, were used in the study to estimate

the reliability of MR measures under changing conditions of scan [123]. The pulse sequences used were MPRAGE and high-resolution T2 (TSE). The two healthy male human phantoms of ages 26 and 27 were scanned at 4 sites within a week. Two repeated scans of the same phantom were obtained at each site, using the same scanner, within 24 hours. The age of the phantoms and their health status, as well as the short time period between repeated scans, indicate that subject-related changes were minimized. Since the entire image dataset for each phantom was acquired within a week, it is safe to assume that no major brain changes took place during this time period. The tuple of images belonging to each phantom p consists of a set of MR scans attributed to different modalities, locations, time points, and scanners. Seven co-registered, multimodal scans of Phantom 1, obtained at 4 scanning locations using 2 different scanner types, can be seen in Figure 2.11.

Initial preprocessing of the phantom images consisted of rigid registration to a template using the IRTK algorithm [124]. This was followed by bias correction and tissue segmentation, which were both computed in an iterative manner as part of the EM algorithm [108]. Prior to analysis of the traveling phantom images, we had to ensure that they were all co-registered in order to remove volumetric and morphometric differences. After co-registration by rigid transformation and bias correction, we created an unbiased atlas A_p from the set of T1W images from the Trio scans of Phantom p . Unbiased atlas building was done using an algorithm based on LDDMM (Large Deformation Diffeomorphic Metric Mapping) [105]. The T2W scans and images belonging to the Allegra scanner were then deformed onto the atlas created using a fluid-based deformation method [123]. As a result of the steps described above, the images from all scanners, obtained at all time points and locations, belonging to a phantom denoted by p , were co-registered to the corresponding atlas A_p . After being co-registered, the entire tuple of images obtained for a single phantom p could be denoted by \mathbf{I}_p , and the vector of intensities for a single voxel i could be given by $\mathbf{I}_{p,i}$. After atlas building, a parcellation map was registered to the generated atlas, and the major cortical regions were extracted. The final step in the processing pipeline consisted of intensity normalization of all images by linear scaling using the following normalization factors: (i) 90 percentile value of the fatty tissue region for T1W images, and (ii) 90 percentile value of the ventricular CSF region for the T2W images.

In order to measure the reliability of the WIVID measure, we calculate the COV (Coefficient Of Variation) across all scans to obtain a normalized measure of variation. The COV of a quantity Q with mean value $\mu(Q)$, and standard deviation $\sigma(Q)$ is given as

$$COV(Q) = \frac{\sigma(Q)}{\mu(Q)}. \quad (2.27)$$

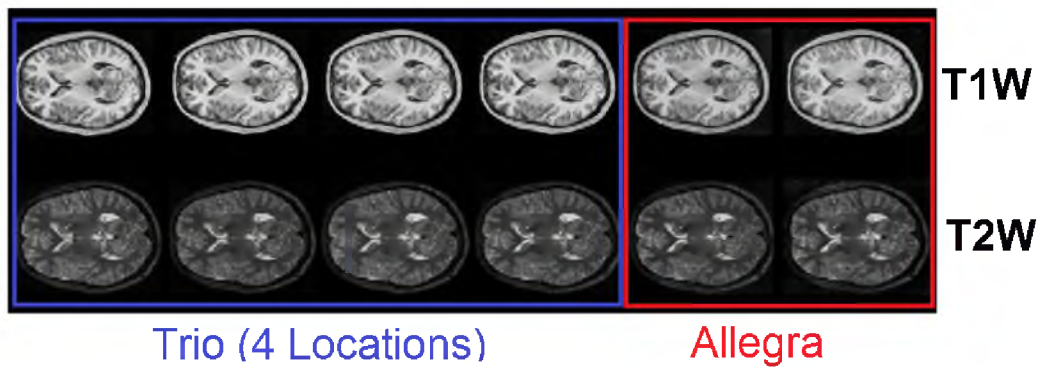


Figure 2.11: Seven T1W (top row) and T2W (bottom row) scans of Phantom 1, acquired across 2 scanner types and 4 locations. The scans are all co-registered - the five leftmost scans belong to the Trio scanner, while the two rightmost scans belong to the Allegra scanner.

Since WIVID is a regional measure, the COV of WIVID is computed for each cortical region. Let the set of images belonging to phantom p obtained under different scanning conditions be denoted by \mathbf{I}_p . The corresponding vector of WIVID values for a region R of the brain, from scans of modality m , is given by \mathbf{WIVID}_p^R . The COV of the WIVID for a region R can be written as

$$COV(\mathbf{WIVID}_{p,m}^R) = \frac{\sigma(\mathbf{WIVID}_{p,m}^R)}{\mu(\mathbf{WIVID}_{p,m}^R)}. \quad (2.28)$$

In comparison with WIVID, signal intensity measurements are computed in a voxel-wise manner, therefore leading to voxel-wise maps of COV. To obtain a regional estimate of the COV of signal intensity, we compute the mean COV of signal intensity averaged over all voxels in a distinct cortical region. As defined previously, for a set of images of scan modality m belonging to a phantom p , the voxel x has intensities denoted by $\mathbf{I}_{p,m}(x)$. The COV for the voxel x computed over the entire set of images is given by

$$COV(\mathbf{I}_{p,m}(x)) = \frac{\sigma(\mathbf{I}_{p,m}(x))}{\mu(\mathbf{I}_{p,m}(x))}. \quad (2.29)$$

From the above equation, the mean COV over all N_R voxels in a region R can be computed as

$$\overline{COV}(\mathbf{I}_{p,m}^R) = \frac{\sum_{x \in R} COV(\mathbf{I}_{p,m}(x))}{N_R}. \quad (2.30)$$

2.4.2 Visual Inspection of Image Quality

Figure 2.12 shows all the Trio scans obtained from Phantom 1. The Allegra scans belonging to Phantom 1 are displayed in Figure 2.13. To serve as a point of comparison for image quality, Allegra scans and Trio scans of Phantom 1 can be seen in Figure 2.11. From visual inspection, it is clear that the image quality of the Trio images is superior to those of Allegra, and it is apparent that Trio images display greater stability. This is particularly observed in the case of the T2W scans, in which the image quality is drastically improved in Trio scans compared with Allegra. In the sections to follow, we restrict our analysis to Trio scans alone and move to quantitative analysis.

2.4.3 Quantitative Analysis of WIVID

In our analysis, COV (Coefficient of Variation) is adopted as the tool of choice to study stability across measurements, along the lines of similar studies of stability for neuroimaging data [117]. Coefficient of variation can be defined as the ratio of standard deviation to overall mean. As described in the previous section, 2 phantoms were scanned repeatedly

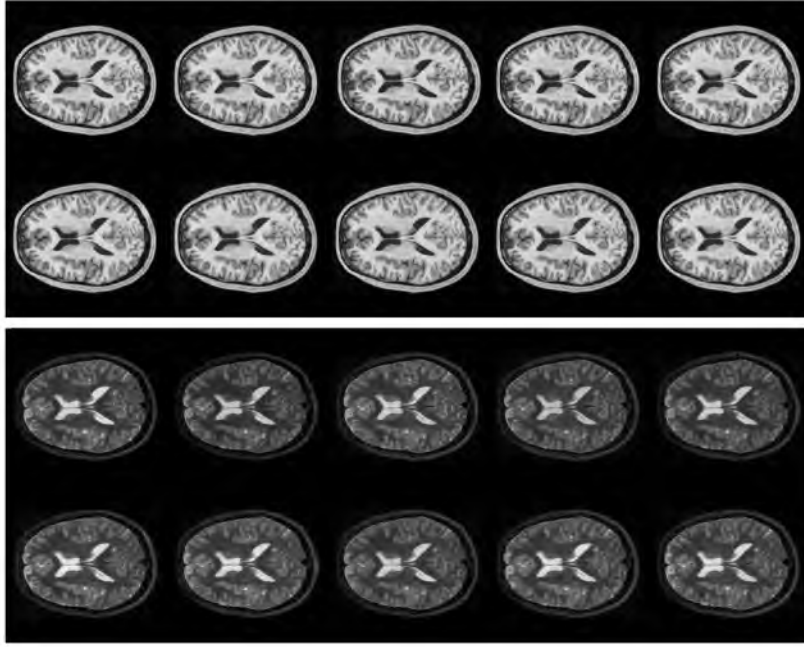


Figure 2.12: Bias corrected Trio scans of Phantom 1 from the 4 scan locations (top 2 rows, scans are repeated in 2 scanner in one of the locations), with T1W images (top 2 rows) and T2W images (bottom 2 rows).

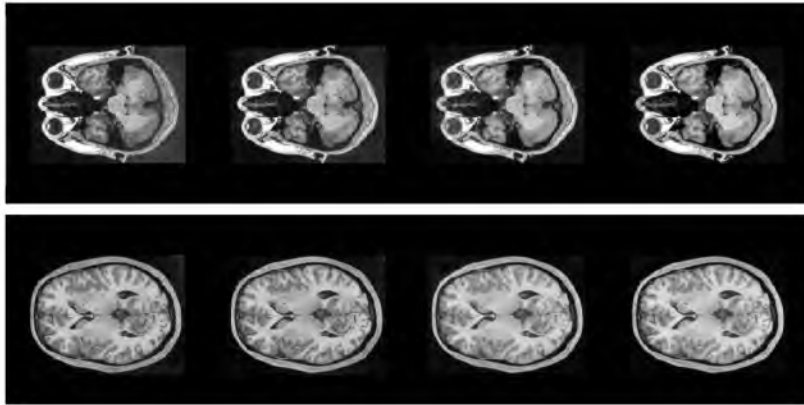


Figure 2.13: Bias-corrected T1W Allegra scans of Phantom 1 from the 4 channel coil (2 leftmost images), and volume coil (2 rightmost images).

using 2 scanners (Allegra and Trio) at 4 locations. We apply the processing pipeline and the registration framework described above to the scans belonging to each phantom. The Coefficient Of Variation (COV) for WIVID in each region R is computed. The mean COV for intensity in each region R is also computed, by averaging the voxel-wise COV over all voxels in the region R .

The results of the phantom experiment show that with changing external scanning

conditions and scanner type, the COV of WIVID is significantly lower than the mean COV of intensity values. This is illustrated in Figure 2.14, which shows the relative stability of WIVID when external factors such as scanner type, locations, and conditions of scan are varied. Since the COV is analyzed in a region-wise manner, it is shown for each of the 15 lobar parcellation regions. In most regions of the brain, the regional COV of WIVID lies between 2 and 5 percent. Therefore, any changes in WIVID across scans that lie in a much higher range can be predominantly attributed to actual changes in the appearance of images rather than to artifacts due to scanning conditions. In comparison, the mean COV of intensity for a region is much higher, ranging between 5 and 10 percent for T1W scans, and between 10 and 20 percent for T2W scans. Subsequently, lower COV values indicate lower standard deviation and better stability values. As a point of reference, it was shown previously that the COV value associated with volumetric analysis for the same dataset ranged from 0.5 and 10 percent for major cortical and subcortical structures [123]. From the above analysis of COV values, it can be concluded that the stability of the WIVID measure is superior to that of signal intensity, providing a strong motivation for adopting WIVID as the biomarker of choice for quantifying intertissue appearance variations in brain images.

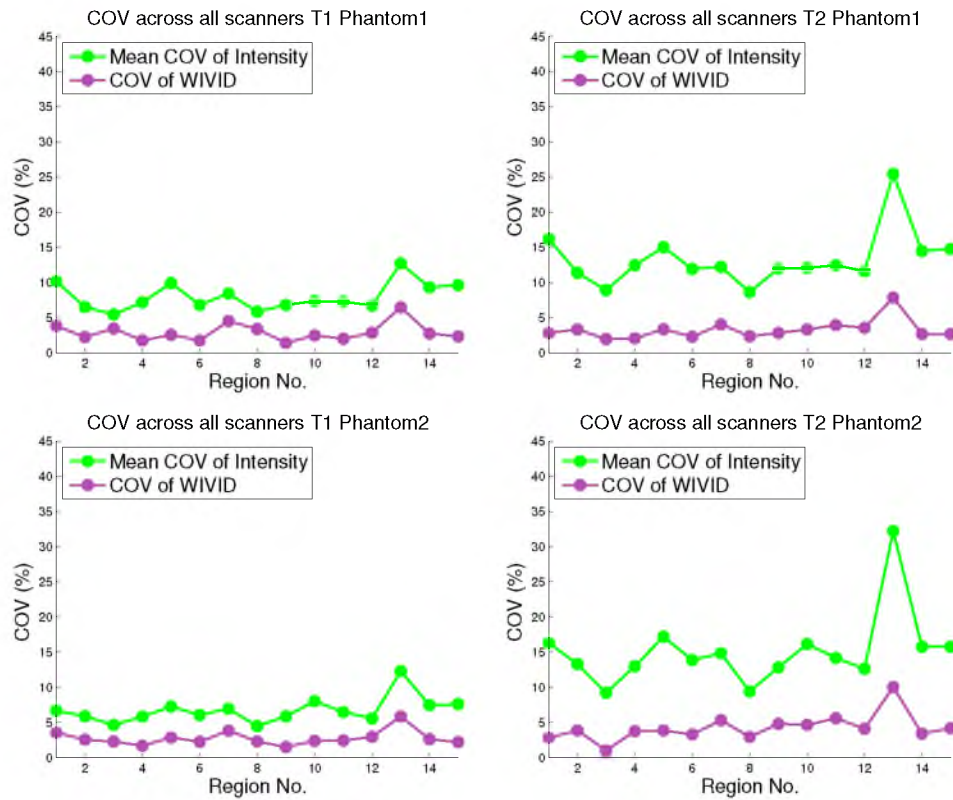


Figure 2.14: The COV (Coefficient of Variation) values of WIVID and Mean COV values of intensity, plotted for T1W (left column) and T2W (right column) scans belonging to Phantom 1 (top row) and Phantom 2 (bottom row) across 15 brain regions. Brain regions are indexed as follows: Occipital Right = 1, Temporal Right = 2, Subcortical Right = 3, Frontal Right = 4, Cerebellum R = 5, Parietal Right = 6, Occipital Left = 7, Subcortical Left = 8, Frontal Left = 9, Cerebellum Left = 10, Parietal Left = 11, Temporal L = 12, CSF Left = 13, Prefrontal Right = 14, Prefrontal Left = 15.

CHAPTER 3

STATISTICAL MODELING OF INTENSITY VARIABILITY IN DEVELOPING BRAIN MRI

3.1 Normative Models of Growth

Statistical modeling and analysis is possible with the development of quantitative markers. In the context of neuroimaging, these quantitative markers could range from volumetric, morphometric, and diffusion indices, to appearance indicators. Modeling of these markers across time can establish standardized trajectories of change characterizing normal development. These normative models of changes in imaging biomarkers are important for understanding crucial neurodevelopmental processes [60].

The origin of several neurobiological disorders can be traced back to early brain development [46,125], which can be attributed to plasticity of the early brain and its subsequent vulnerability to external influences. Understanding the origin and course of brain disorders even at early stages is possible with the creation of statistical models for use with neuroimaging studies. Normative models of change can be used to establish standards related to healthy brain development and assess the degree of deviation from these standards in the case of abnormal development. Procedures for early diagnosis of disorders based on normative growth information can then be formulated, facilitating the use of early interventional therapies that are effective when applied to the highly plastic infant brain. In addition, the normative statistical models developed can be used to evaluate interventional procedures and therapies following early diagnosis [60]. The importance of establishing normative models that map changes in neurodevelopmental biomarkers is outlined in Figure 3.1.

A major goal of this dissertation is to understand the nature of appearance changes in brain tissues across time, which can be accomplished by statistical modeling of the WIVID biomarker that was introduced in Chapter 2. Given WIVID data from large-scale neuroimaging datasets, a question at hand is, “What is the best method of modeling this data”? The neuroimaging data studied in this work is longitudinal in nature - i.e., it consists

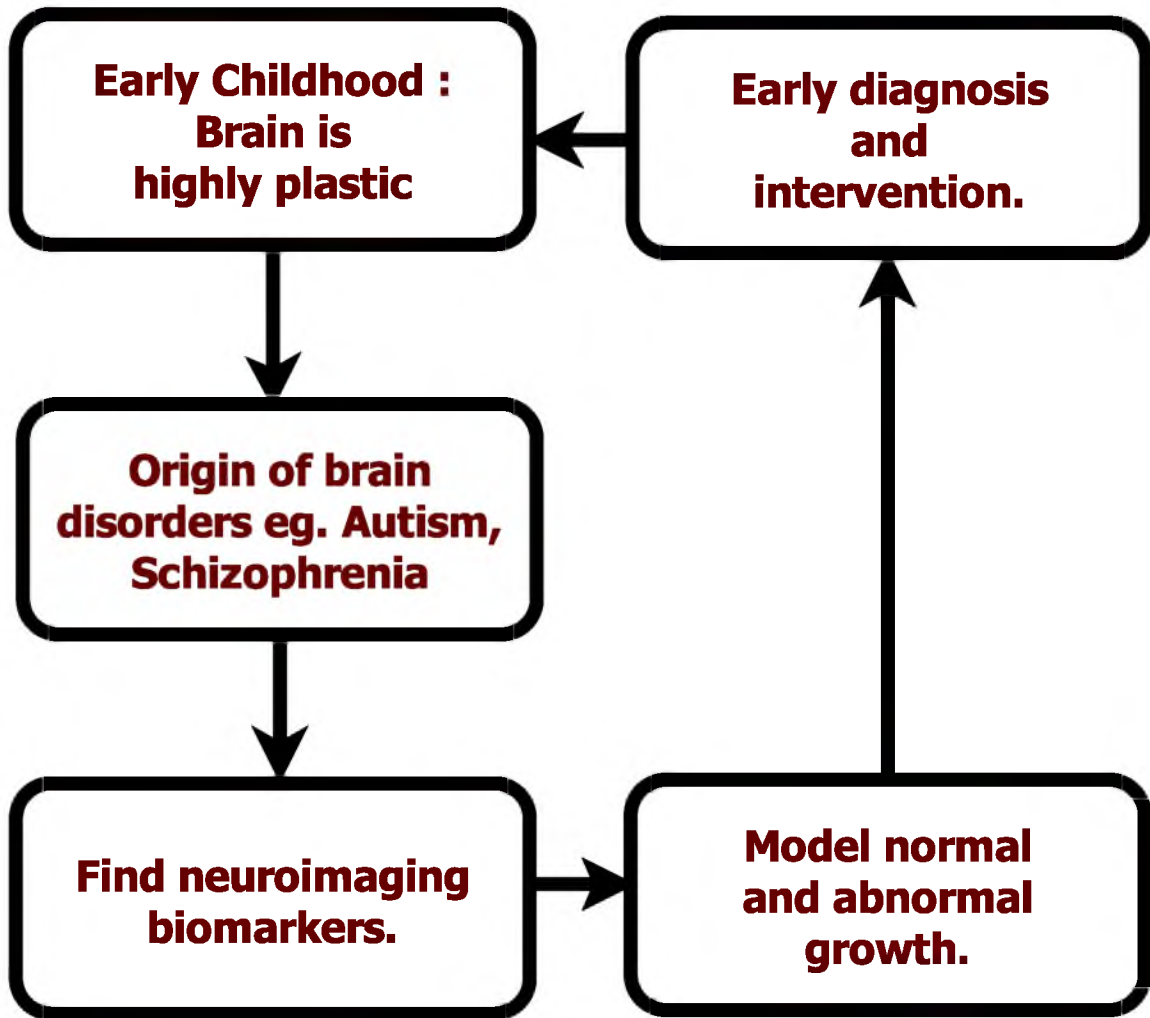


Figure 3.1: Importance of establishing normative models to quantify changes in neurodevelopmental biomarkers.

of repeated scans of the same subject being acquired over time. Modeling data of this nature using traditional cross-sectional modeling approaches could misrepresent the data, being insensitive to individual differences in development [54, 58]. Resulting changes that are reported using this approach could be distorted or might not be detected altogether [126]. Alternatively, methods for modeling of longitudinal data could be employed to present a more accurate picture of brain development [127]. Longitudinal modeling of data takes into account the correlations between repeated samples obtained from the same subject across time. In this chapter we introduce basic concepts in statistical modeling followed by discussions of comprehensive schemes for testing and modeling of longitudinal data. Multivariate analysis for jointly modeling changes in different modalities is also introduced here. These schemes for longitudinal and multivariate data modeling will be practically

applied to WIVID data and further explored in later chapters.

3.2 Modeling Repeated Measures Data

In this section, we discuss concepts in data modeling that are essential precursors to understanding longitudinal mixed effects modeling. The concepts covered include the nature of repeated measures data, quantitative tools for testing data models, the role of visual analysis in data modeling, parametric and nonparametric modeling, linear and nonlinear cross-sectional modeling, adding group-based covariates, and subject-specific modeling. While these concepts have already been discussed in the context of NLME (Nonlinear Mixed Effects) and LME (Linear Mixed Effects) modeling [17, 128, 129], we organize this information and present the pipeline involved with the specific application of modeling neuroimaging biomarkers in view. Figure 3.2 illustrates the options available for statistical modeling of longitudinal data.

To illustrate some of the concepts in basic data modeling and mixed effects modeling, we use a subset of the ACE-IBIS data that is analyzed in greater detail in Chapter 4. The subset used consists of scans from 92 subjects, obtained at three time points (6, 12, and 24 months), and categorized based on risk/diagnosis for Autism Spectrum Disorder (ASD). The grouped data points analyzed include WIVID contrast measures from 22 controls, 14 HR+ (High Risk Positive) subjects, and 56 HR- (High Risk Negative) subjects. To simplify the analysis for purposes of discussion, only positively valued WIVID measures are included in this example subset. Additionally, only WIVID contrast values of the right temporal lobe from T1W scans are studied here, for the sake of illustrating principles of statistical modeling.

3.2.1 Repeated Measures Data

“Repeated measures data” refers to data that is generated by observing a number of individuals repeatedly under differing experimental conditions [130]. In longitudinal data, the repeated measures observations could be obtained by ordering in time or space, resulting in serial correlations between data.

In the past, repeated measures data addressed measurements from experiments such as epidemiological studies, where a large number of individuals were observed from a small number of time points [129]. Since large-scale neuroimaging datasets consist of repeated scans of the same subject taken at different time points, this also falls within the repeated measures category. Traditional cross-sectional fits to the data might not be sufficient. Due to the highly correlated nature of data belonging to the same subject obtained at

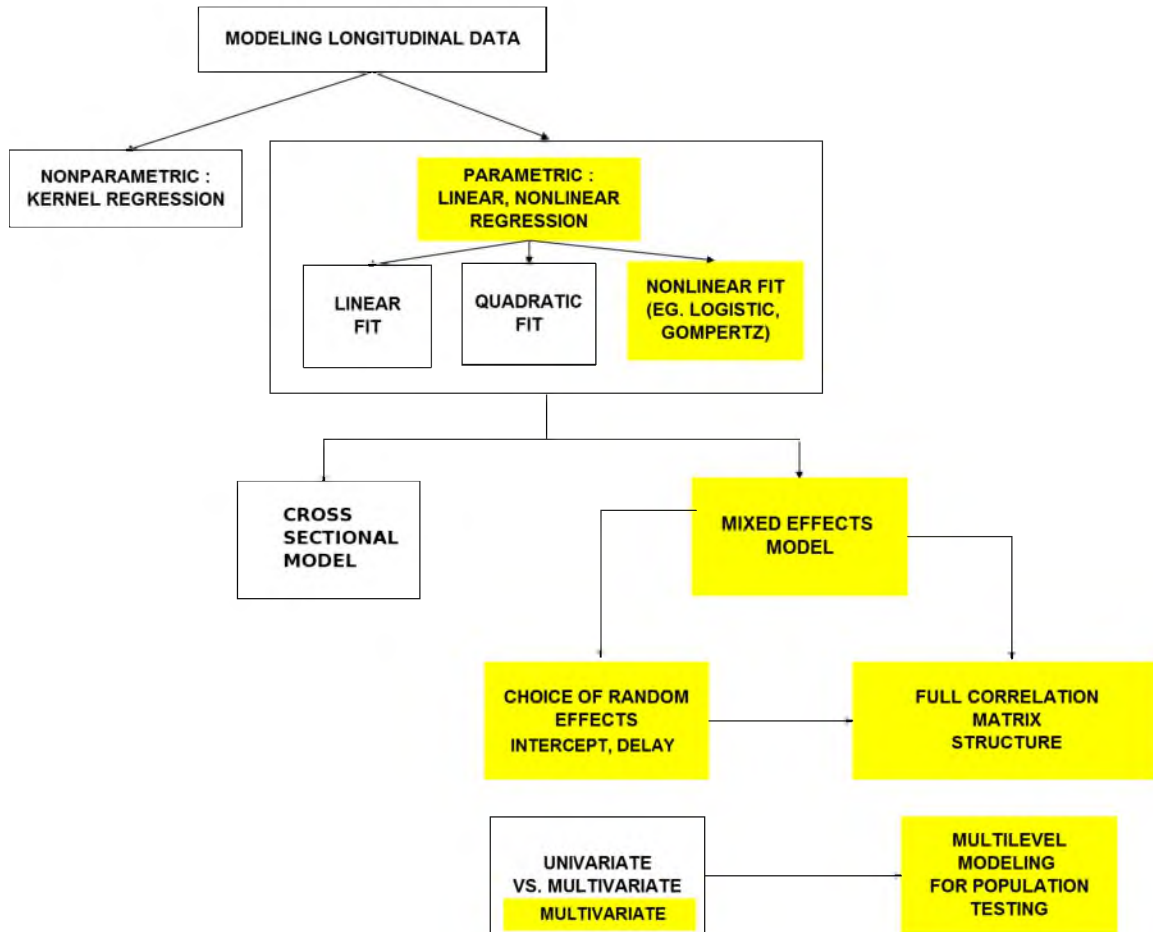


Figure 3.2: Description of process governing model-selection for longitudinal neuroimaging WIVID biomarker data from the ACE-IBIS dataset.

different time points, modeling the data in a cross-sectional manner would violate the necessary assumptions of independence. Another possibility for modeling repeated measures data is subject-specific modeling. However, this type of modeling restricts the estimated model to the subjects under study and may not be easily extendable to other subjects. The number of samples per subject required to estimate this model is also high, and due to limited samples the resulting per-subject fit might not be accurate. Therefore, the mixed effects model simultaneously takes into account the subject-specific (random effects) and population (fixed effects) trends. In addition, major advantages of using a mixed effects modeling approach to model longitudinal data include robustness to outliers and noise, as well as the ability to utilize incomplete datasets with missing and unbalanced data acquisition. In the following sections, nonparameteric and parametric models for characterizing neuroimaging biomarkers are discussed.

3.2.2 Quantitative Analysis of Data Fit

The major quantitative tools used to analyze the quality of a model fit include the following: 1) residual standard error, 2) residual plots, 3) R-squared error or R^2 , 4) ANOVA (Analysis of Variance) results, 5) standard deviations of estimates, 6) confidence intervals, and 7) AIC (Akaike Information Criterion). Of these, the AIC is a very important criterion that can be used for examining models. These tools can be used to compare different models for fitting data, examine grouping factors, and also to diagnose issues with these models. In the discussion that follows, y is assumed to be the outcome variable obtained at time instants t_k indexed by $k = 1, 2, \dots, K$. Predicted values of the outcome variable are denoted by \hat{y} . At a specific time instant t_k , the outcome variable is $y(t_k)$ and the predicted outcome variable is $\hat{y}(t_k)$. From this definition, the sum of squared error (SSE) can be defined as

$$SSE = \sum_{k=1}^K \epsilon(t_k)^2 = \sum_{k=1}^K (y(t_k) - \hat{y}(t_k))^2. \quad (3.1)$$

(1) The residual standard error (RSE) can be defined in terms of the SSE as

$$RSE = \sqrt{\frac{SSE}{K - dof}}, \text{ where, } dof \text{ is the number of degrees of freedom.} \quad (3.2)$$

(2) Residual plots consist of the residual $y(t_k) - \hat{y}(t_k)$ corresponding to each value $y(t_k)$ plotted on the y-axis, versus t_k values plotted on x-axis. If trends are seen in the values of the residual with respect to the covariate, it could be an indicator of an underlying problem with the fit. Residuals also verify assumptions made regarding the nature of statistical errors ϵ .

(3) The R-squared error or R^2 is the proportional reduction in squared error resulting from the model fit. Consider a model that does not use independent variables, for which prediction is based only on the mean outcome. In this case, the term total sum of squares (TSS) refers to the sum of squared errors for this model, given by

$$TSS = \sum_{k=1}^K (y(t_k) - \bar{y})^2. \quad (3.3)$$

Based on the TSS and SSE values, the R^2 measure is given by

$$R^2 = 1 - \frac{SSE}{TSS}. \quad (3.4)$$

The corrected R^2 error, which penalizes for extra independent variables p in the model, is given by

$$R^2 = 1 - \frac{\frac{SSE}{n-p-1}}{\frac{TSS}{n-1}}. \quad (3.5)$$

(4) The ANOVA (Analysis of variance) results are derived from definitions of TSS, SSE, and RegSS. RegSS or the regression sum of squares is given by $RegSS = \sum_{k=1}^K (\hat{y}(t_k) - \bar{y})^2$. The decomposition $TSS = SSE + RegSS$ constitutes the basis of ANOVA. ANOVA tests for statistical significance, mainly by differentiating between two sources of variation - intrinsic variation within a dataset from its mean and variation with respect to a predicted value. The significance of the factors involved in prediction can be tested using ANOVA.

(5) Standard deviations of estimated parameters can also be used for diagnosis of goodness of fit in the case of mixed effects models. While the formulation for standard deviation is outside the scope of this dissertation, it could be broadly defined as the deviation associated with the estimated parameter values of the fit. The standard deviation of each parameter is a result of the ML (Maximum Likelihood)-based estimation procedure, and must be within reasonable limits for the model fit to be considered.

(6) As a by-product of the ML-based estimation procedure, the mean and standard deviations of each parameter can also be used to produce bands of confidence intervals. Overlap in confidence intervals between different groups can indicate whether the grouping or categorization needs to be accounted for in the model.

(7) Akaike Information Criterion (AIC) is a critical diagnostic measure for analyzing fits from mixed effects models. Considering $logLik$ - the log-likelihood estimate from the ML procedure in mixed effects modeling, and the number of parameters to be estimated - n_{par} , the AIC value can be defined as

$$AIC = -2logLik + 2n_{par}. \quad (3.6)$$

A lower AIC value corresponds to a better model fit for the data.

3.2.3 Visual Analysis of Data Points

Prior to modeling the data, visual observation of the data is crucial. An important question to answer when applying longitudinal statistical modeling is “which function best fits the data?” In certain cases, visual analysis of the data can lead to an improved understanding of the inherent nature of variations in the data and mechanisms governing these variations. Based on this understanding, inspection of various models used for data fitting and verification of the best model fit serve as a crucial addition to quantitative checks such as the residual standard error and Akaike Information Criterion (AIC). Among other

factors, sparsely sampled data and missing data points might provide misleading results if only quantitative checks are considered for deciding the best fit to the data.

To illustrate this concept, WIVID data is modeled in a cross-sectional manner using both quadratic and nonlinear fits, as shown in Figure 3.3. From neurobiological evidence, tissue appearance differences, as quantified by the WIVID, will constantly increase for the age range under study in this dataset, until they stabilize at an asymptotic point. This asymptotic nature makes nonlinear growth functions, such as the Gompertz or Logistic functions (discussed extensively in the next section), an ideal fit to the data. However, on examination of the residual squared error values, the quadratic fit is only a marginally weaker fit compared to logistic and Gompertz fits (the residual standard error values are 0.06473, 0.06293, and 0.0767 for logistic, Gompertz, and quadratic fits respectively). In this example, the knowledge that tissue appearance difference does not decrease during this age range would be valuable for choosing the right function to model the data.

3.2.4 Nonparametric Models via Kernel Regression

Kernel regression can be used to get an initial estimate of spatiotemporal biomarker variation without making any assumptions based on parametric models. Considering that each subject i has been scanned at time indices k , denoted by $t_{i,k}$, we denote the biomarker of the subject i computed at the k th scan by $y(t_{i,k})$. The measure y obtained for a single subject at specific time instants given by $y(t_{i,k})$, is interpolated using kernel regression, which results in a continuous trajectory of change in the biomarker for the specific subject i , over all time points t , given by

$$y(t_i) = \frac{\sum_{t_{i,k}} K(t, t_{i,k}) y(t_{i,k})}{\sum_{t_{i,k}} K(t, t_{i,k})} . \quad (3.7)$$

A population trend is often essential for representing the overall information gained from several subject-specific trends. From the computation of the individual trend $y(t)$ using kernel regression, the mean value for the appearance measure at each time t for the entire population can be found. This value is found from the population trend, which is obtained by integrating the measure from each subject i over all subjects $i = 1, 2, \dots, N$ at

⁵The convention used here is a departure from that in Chapter 2. In Chapter 2, the biomarker corresponding to region R_l , modality m and subject i was denoted by $WIVID_{i,m}^{R_l}(t_k)$, where time is indexed by (t_k) . In this case, the common indexing of time points across all subjects is implicit since the term (t_k) is independent of the subject index i . This common indexing follows from the simplification made in Chapter 2, that the markers modeled are extracted from data acquired at the same time points across all subjects. Since the subject of the current chapter is exclusively data modeling, this assumption is relaxed - each subject i is assumed to be scanned at time points $t_{i,k}$, therefore adding the flexibility that these time points could vary between subjects.

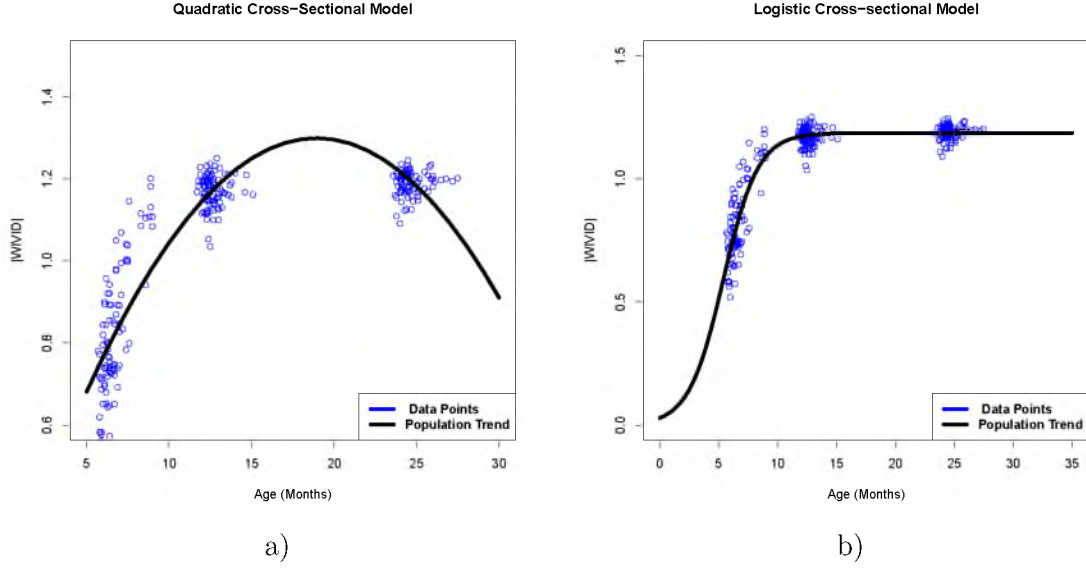


Figure 3.3: WIVID contrast measure changing with age are modeled by (Left) Quadratic cross-sectional fit, and (Right) Logistic cross-sectional fit.

each point t , and dividing this per-instant value by the total number of subjects, which is mathematically formulated as

$$y(t) = \frac{1}{N} \times y(t_i) . \quad (3.8)$$

Here, $K(\cdot, \cdot)$ denotes a kernel function that is chosen to be a Gaussian kernel in all our subsequent analysis. The main feature of this approach is that a parametric representation of change in data is eschewed in favor of a broader, nonparametric approach. Examples of the kernel regression fit performed for contrast change in a specific lobe of the brain are shown in Figure 3.4.

Averaging of subject-specific curves at each time point is done with the purpose of generating an average curve for the entire population, based on assumptions about the nature of distribution of values across all subjects at each time instant. In the sense in which KDE is implemented here, the results can be interpreted as subject specific. However, kernel regression can also be applied simultaneously to data points across *all* subjects. In this case, the correlated nature of repeated measurements from the same subject is ignored, and kernel regression will be cross-sectional in nature.

Kernel regression requires densely sampled data points to be effective, mainly due to the large number of values that are estimated. Practical application of the kernel regression method to our neuroimaging datasets is problematic, since it requires multiple scans of each subject at closely spaced time points for effective computation. The close clustering of time

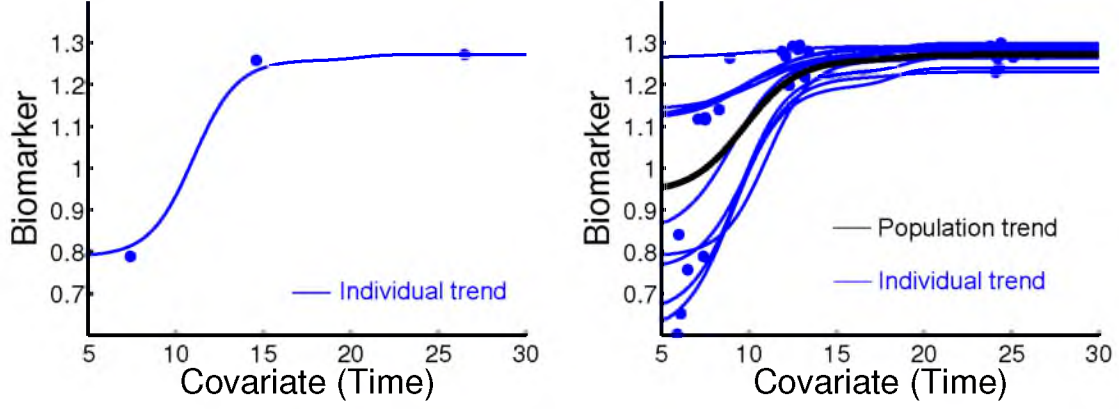


Figure 3.4: Kernel regression of biomarker data is performed with respect to the covariate time, resulting in a) individual subject-specific trend (blue), b) population trend (black) computed from individual trajectories of 10 subjects (blue).

points at which the different subjects were scanned and their imbalanced nature is another factor that leads to problems. Some problems that are observed in the resulting appearance change trajectories include artifacts such as stair-casing and excessive dependence of the fit on the kernel parameters that are chosen [131]. In addition, the use of kernel regression does not result in a set of parameters or values that can be easily used by clinicians to understand the behavior of the quantities under study. Hence, although kernel regression provides a fair initial idea of the nature of trajectories of biomarkers studied, we will investigate adopting parametric modeling techniques as suitable methods for spatiotemporal analysis on this dataset.

3.2.5 Parametric Regression Models

Prior to longitudinal modeling, it is beneficial to apply cross-sectional regression-based modeling to the biomarker data y of interest and explore the results. Cross-sectional modeling essentially consists of fitting a single model to data from the entire population. This approach assumes that all the data points being fit are independent, and hence it does not consider the correlation between repeated measurements from the same subject. By applying cross-sectional modeling, we can estimate and compare the efficacy of different functions in modeling our data, as an initial step. This approach can also be used to find initial estimates of growth parameters that are critical to implementation of more advanced models. The effectiveness of the models fitted can be quantified and compared

⁶Since we study data modeling with the view of applying it to neurodevelopmental biomarkers, time t is assumed by default to be the covariate of interest.

using diagnostic tools such as the residual standard error, R-squared errors, and ANOVA (Analysis Of Variance) values [130] that were defined earlier. In addition, the standard deviations of the estimated values in fits could be analyzed and checked for being within reasonable limits.

A linear cross-sectional approach is first explored for modeling the outcome variable $y(t)$ corresponding to the biomarker of interest, which changes with respect to the covariate time t^6 , which is described by the equation

$$y(t) = \phi_1 t + \phi_2, \text{ where } \phi_1 = \text{slope}, \phi_2 = \text{intercept}, t = \text{time/age}. \quad (3.9)$$

Results from this linear fit can be seen in Figure 3.5a.

An alternative cross-sectional regression fit to the data is the quadratic model, which consists of an additional quadratic term, given by the equation

$$y(t) = \phi_1 t^2 + \phi_2 t + \phi_3. \quad (3.10)$$

Linear and quadratic cross-sectional models are generally fit using least squares estimation, although alternative methods exist. Using this technique, the values of the coefficients in the model - for example, ϕ_1 and ϕ_2 , are estimated such that the sum of squares error between the predicted outcome and actual (i.e., empirical) outcome value is minimized. Considering the linear cross-sectional model shown above, and given that K experimental outcome values denoted by $y(t_k)$ are recorded at time instants t_k , the sum of squared error to be minimized can be given as

$$\sum_{k=1}^K \epsilon_{t_k}^2 = \sum_{k=1}^K (y(t_k) - \phi_1 t_k + \phi_2)^2. \quad (3.11)$$

Alternately, nonlinear growth functions can also be implemented using the cross-sectional model. Some major nonlinear growth functions and their equations, with parameters ϕ_1, ϕ_2, ϕ_3 , are listed below.

$$y(t) = \phi_1 \times e^{-\phi_2 \phi_3^t}, \dots \text{ Gompertz fit.} \quad (3.12)$$

$$y(t) = \phi_1 + \frac{\phi_2}{(1 + e^{\frac{\phi_3 - t}{\phi_4}})}, \dots \text{ Logistic fit.} \quad (3.13)$$

$$y(t) = \phi_1 + (\phi_2 - \phi_1)e^{-e^{\phi_3 t}}, \dots \text{ Asymptotic fit.} \quad (3.14)$$

The trajectories in Figure 3.5b model the entire sample data using the three nonlinear growth functions mentioned above. The parameters of the nonlinear growth function are estimated using nonlinear least squares methods.

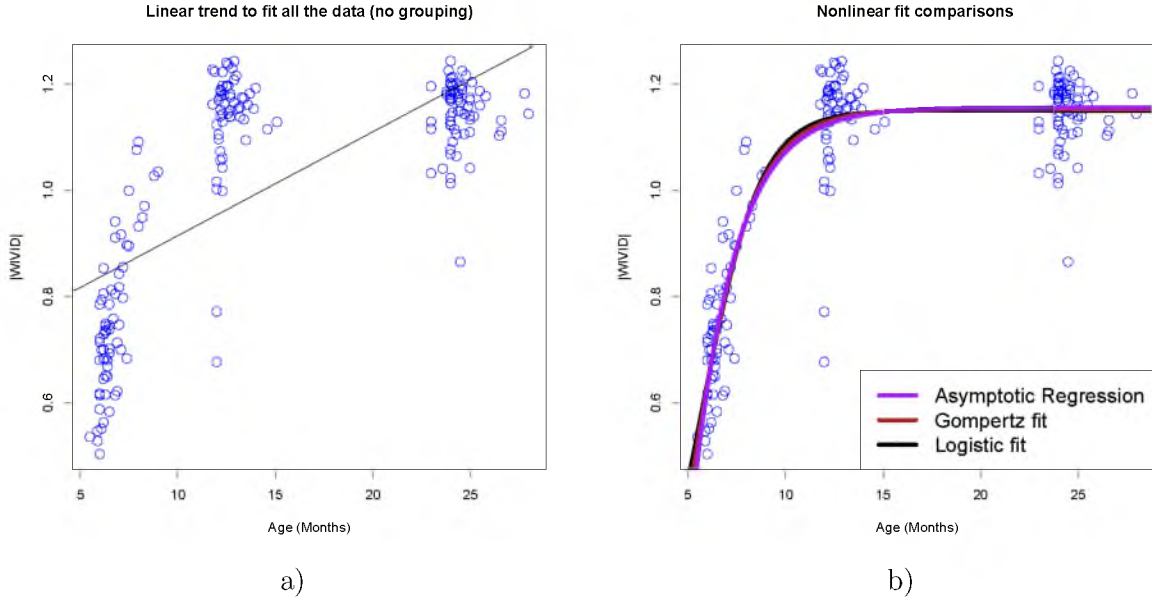


Figure 3.5: Examples of cross-sectional fits to data: (a) Linear cross-sectional fit to data independent of grouping, (b) Nonlinear cross-sectional fits to data (Logistic, Gompertz, and Asymptotic), independent of grouping.

Based on the asymptotic nature of brain appearance change trajectories for the age range under study, the quadratic fit is not implemented since it predicts an increase in the data followed by a decrease. While this trajectory might hold true for brain development curves from birth, the ACE-IBIS data, which ranges from 6 months to 2 years of age, generally displays only trends of increasing contrast. This asymptotic nature of contrast makes nonlinear functions such as the Gompertz, logistic, and asymptotic fits suitable for data modeling.

Based on the AIC values shown in Table 3.1, it could be observed that the logistic function provides the best cross-sectional fit, closely followed by Gompertz and asymptotic functions.

The AIC values of the linear fit were dramatically higher than for the nonlinear fits, indicating that nonlinear fits are much more suited for analysis of the dataset being studied. Further, all three nonlinear fits also displayed much lower values for residual standard error, compared with the linear fit. The logistic function also had the lowest value for the residual standard error, further establishing its suitability for modeling WIVID contrast data compared to other models. Upon studying the parameters and their associated standard error values, it was observed that the delay parameter for the Gompertz curve had the highest standard error of up to 25 percent of the value of the parameter itself, which

Table 3.1: Diagnostic Parameters for Assessment of Fit of Linear and Nonlinear Cross-Sectional Models.

Function	Linear	Gompertz	Logistic	Asymptotic Regression
AIC	-313.5890	-721.3247	-722.8527	-718.4570
Residual Standard error	0.1361	0.06491	0.06473	0.06525
Par1	0.799704	1.188303	1.185722	1.191400
StdErr(Par1)	0.018027	0.005474	0.005226	0.005759
Par2	0.017933	15.969526	5.402050	-7.311748
StdErr(Par2)	0.001092	4.323236	0.082397	1.857816
Par3	NA	0.564018	1.468603	-0.737506
StdErr(Par3)	NA	0.024335	0.111540	0.074316

indicates that the Gompertz fit implemented using cross-sectional modeling might be very unstable.

A comparison of the three nonlinear growth models shown in Figure 3.5b indicates that their resulting fits to the data are very similar. The main difference between them seems to be the speed at which they approach the asymptote. Although the logistic function sharply approaches the asymptote, the Gompertz function approaches the asymptote less sharply. Again, whereas the logistic function is constrained to be symmetric on both sides of the midpoint, this is not true for the Gompertz function. In comparison with the logistic and Gompertz functions, the asymptotic function approaches the asymptote the slowest.

In the linear and nonlinear models mentioned above, the only covariate used is age. Additional factors can be added to the analysis in such a way that they influence one or all of the model parameters. Examples of covariates that can be meaningful to clinical population analysis are sex, diagnostic results, and genetic factors. As an illustrative example, we change the linear cross-sectional analysis by adding a covariate to the intercept that we simply call “group.” Any number and combination of the estimated parameters can be associated with any of the groups. The mathematical equation for linear cross-sectional analysis can be rewritten as

$$y(t) = \phi_1 t + \phi_{2,group}, \quad (3.15)$$

where ϕ_1 refers to the common slope for all subjects and $\phi_{2,group}$ refers to the intercept for each group. In this case, the model fit to the data will be distinct for the distinct population groups of interest, although they will share the estimate for slope.

Extending this example to the case of both slope and intercept being grouped, the equation for linear cross-sectional analysis can be written as

$$y(t) = \phi_{1,group} t + \phi_{2,group}, \quad (3.16)$$

where $\phi_{1,group}$ refers to the slope based on group and $\phi_{2,group}$ refers to the intercept for each group. The resulting linear fit has grouping for both slope and intercept parameters.

The estimation of grouped covariates can also be extended to the case of nonlinear models. Some growth functions of interest can be reformulated as given below, assuming that all their parameters are dependent on the group covariates.

$$y(t) = \phi_{1,group} \times e^{-\phi_{2,group}\phi_{3,group}^t}, \dots \text{ Gompertz fit.} \quad (3.17)$$

$$y(t) = \frac{\phi_{1,group}}{(1 + e^{\frac{\phi_{2,group}-t}{\phi_{3,group}}})}, \dots \text{ Logistic fit.} \quad (3.18)$$

$$y(t) = \phi_{1,group} + (\phi_{2,group} - \phi_{1,group})e^{-e^{\phi_{3,group}t}}, \dots \text{ Asymptotic fit.} \quad (3.19)$$

Nonlinear logistic fitting applied to the entire dataset with grouping for all three parameters can be seen in the red, blue, and green curves in Figure 3.6. Whether or not grouping is necessary for an improved model fit needs to be assessed. Comparisons of residual standard error from Gompertz and logistic functions, when model fitting is done with and without grouping of the data, can be seen in Table 3.2. The data are grouped based on risk/diagnosis category. From this table, it can be inferred that grouping based on risk/diagnosis category reduces the residual standard error and is beneficial to the modeling.

The significance of each covariate associated with each parameter can be measured using ANOVA tests for the data. That is, if the p-value of the ANOVA tests is significant for the covariate-parameter combination of interest, it implies that the covariate is significant with respect to that parameter and should be included in the analysis. In this case, the group is the covariate of interest, and the p-value associated with this grouping, corresponding to each of the parameters (for example, ϕ_1 , ϕ_2), can determine if the covariates need to be included into the analysis.

To summarize, the results of applying linear and nonlinear cross-sectional models on the autism data introduced earlier can be seen in Figure 3.6. Results of applying the quadratic model fit on the same data are displayed. It was observed that cross-sectional nonlinear models were far superior to the cross-sectional linear model. In addition, it was seen that the best nonlinear fit to the data was provided by using the logistic growth function in the cross-sectional model. It should be mentioned that since WIVID contrast changes across all brain lobes roughly follow the same asymptotic pattern observed in the right temporal lobe, nonlinear functions might be the best fit for WIVID changes in these lobes as well.

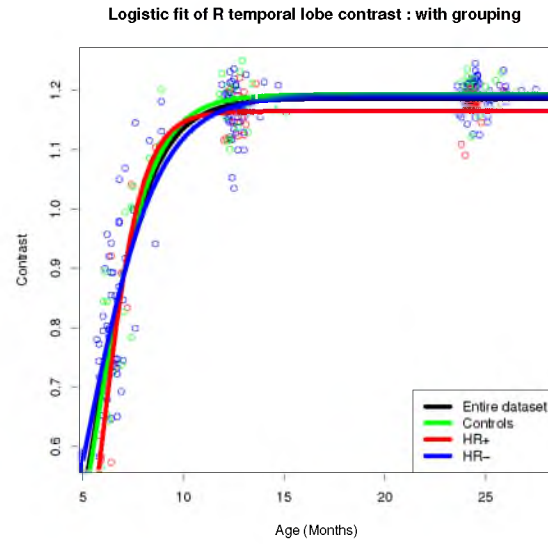


Figure 3.6: Logistic cross-sectional fit to the data with grouping applied to all three logistic parameters. The grouping is performed on the basis of ASD risk/diagnosis groups.

Table 3.2: Comparison of Gompertz and Logistic Cross-Sectional fits with and without grouping based on risk/diagnosis category.

Function	Gompertz (grouping)	Gompertz	Logistic (grouping)	Logistic
Residual Standard error	0.06293	0.06491	0.06271	0.06473

The defining feature of these cross-sectional models is that a single set of model parameters is estimated for the entire population or for each population group. In the following sections, the concept of repeated measurements will be introduced. In the case of repeated measures data, each subject might have a separate set of estimated model parameters. Advanced methods for effective modeling of repeated measures data will subsequently be introduced.

3.2.6 Subject-Specific Modeling of Data

A method of modeling repeated measures data is to perform subject-specific fits: that is, each subject is fit with an individual model. For example, we can redefine a subject-specific extension for the basic linear cross-sectional model defined earlier as

$$y(t) = \phi_{1,i}t + \phi_{2,i}, \text{ where } i \text{ refers to each individual subject.} \quad (3.20)$$

Each subject can be thought of as a “group,” which makes subject-specific analysis very similar to covariate-based analysis, with each subject being the covariate in question. In the equation mentioned above, both parameters, ϕ_1 and ϕ_2 , are assumed to be subject dependent. However, depending on the specific case under study, some of the parameters of the fit could be assumed as subject-specific, while others could be the same for the entire population. An example is given below in which the slope alone is assumed to be subject-specific, whereas the intercept is common for the population.

$$y(t) = \phi_{1,i}t + \phi_2. \quad (3.21)$$

Similar subject-specific fits could also be applied with nonlinear growth functions. A subject-specific fit to data for the logistic growth function is given below, and can be seen in Figure 3.7. The subject-specific model for the logistic function can be specified as follows

$$y_i(t) = \frac{\phi_{1,i}}{(1 + e^{\frac{\phi_{2,i}-t}{\phi_{3,i}}})}. \quad (3.22)$$

As mentioned previously, subject-specific fits cannot be easily extended to broader population data, because they are estimated specifically for the data under study. As a result, it is difficult to determine overall trends in data based on subject-specific fits alone. Subject-specific fits of data also require a large number of samples for proper estimates of values. These reasons, coupled with the issues with applying cross-sectional fits to repeated measures data, lead us to explore mixed effects modeling techniques in the following section.

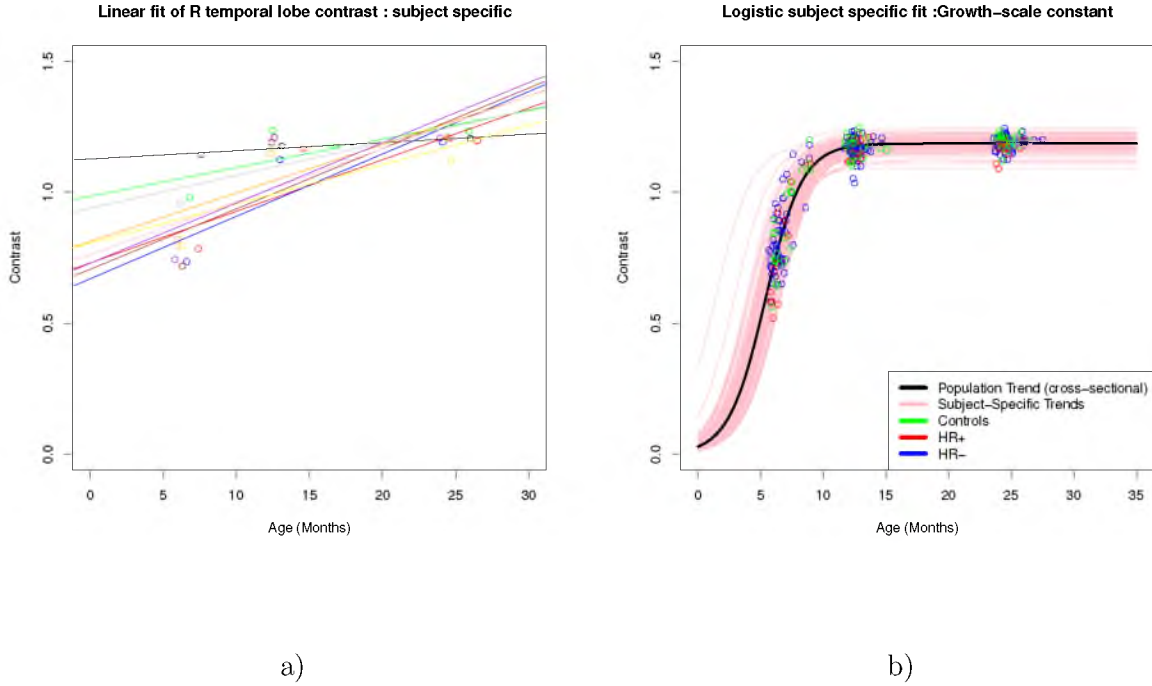


Figure 3.7: Subject-specific fits to WIVID data: (a) Linear function, subject-specific fit (each subject is encoded with a specific color), b) Logistic subject specific fit to the data keeping the rate parameter ϕ_4 constant (each risk/diagnosis group is encoded with a specific color).

3.3 Mixed Effects Modeling

Apart from being intuitively appealing, mixed effects models allow nonconstant correlations among observations and unbalanced data designs. Both linear and nonlinear functions can be implemented using mixed effects modeling. Here, since the superiority of the nonlinear fit for the WIVID data has already been established, only NLME modeling is discussed.

3.4 Nonlinear Mixed Effects Modeling

Nonlinear mixed effects modeling, also known as NLME, jointly estimates both population and individual trajectories of change. In our approach, we use the established framework of hierarchical NLME to model contrast change with time. This NLME model, proposed by Lindstrom and Bates [132], models the observation of the i th individual at the j th time point $t_{i,j}$ as

$$y_{ij} = f(\phi_i, t_{ij}) + e_{ij}. \quad (3.23)$$

We begin by only considering a univariate observation y_{ij} . Here $i = 1, \dots, N_{ind}$ refers to the different individuals indexed by values ranging from 1 to N_{ind} , and $j = 1, \dots, T_{ind}$ indicates the time points of scan, ranging from 1 to T_{ind} . The function f is the nonlinear growth function of choice that is used to model the contrast change trajectory, dependent on the covariate vector t_{ij} as well as the individual-specific parameter vector ϕ_i . The error term e_{ij} refers to the i.i.d error following the distribution $e_{ij} \sim N(0, \sigma^2)$. The parameter vector ϕ_i consists of both fixed and random effects components. This can be written as

$$\phi_i = \mathbf{A}_i \boldsymbol{\beta} + \mathbf{B}_i \mathbf{b}_i, \text{ where } \mathbf{b}_i \sim N(0, \boldsymbol{\psi}). \quad (3.24)$$

That is, the fixed effects and random effects design matrices are given by \mathbf{A}_i and \mathbf{B}_i for each subject, and the p-vector of fixed effects is given by $\boldsymbol{\beta}$ and the q-vector of random effects is given by \mathbf{b}_i . The random effects that contribute to parameter ϕ_i are assumed to be normally distributed with variance-covariance matrix $\boldsymbol{\psi}$ over all subjects.

Since we want to model the highly nonlinear trends seen in contrast change, and eventually extract parameters that describe maturational processes in terms of contrast change, we adopt a nonlinear mixed effects modeling approach using a growth function. We systematically evaluate three nonlinear growth functions - the Asymptotic, Logistic, and Gompertz functions. The question of which function shows the best fit overall can be answered by both comparing the Akaike Information Criterion (AIC) values and by examining the estimated random effects parameters, residual standard error, and residual plots of the NLME fit versus the actual observed values.

For the purpose of this discussion, the NLME models introduced correspond to the logistic growth function, although they can be extended to other functions as well. Parametric growth models such as the logistic and Gompertz growth function provide concise description of the data and greater flexibility compared with the standard linear mixed effects representation. The logistic function using the NLME format, in which the parameters of the growth function ϕ consist of both fixed and random effects, can be written as

$$f(\phi_i, t_{ij}) = \phi_{i1} + \frac{\phi_{i2}}{1 + \exp \frac{\phi_{i3} - t_{ij}}{\phi_{i4}}}. \quad (3.25)$$

Since our logistic fit is unique for each region, the response y_{ij}^R for a region R is

$$y_{ij}^R = \phi_{i1}^R + \frac{\phi_{i2}^R}{1 + \exp \frac{\phi_{i3}^R - t_{ij}}{\phi_{i4}^R}} + e_{ij}. \quad (3.26)$$

The parameters of nonlinear growth functions can also be interpreted in a manner that is intuitive. For example, the parameters of the logistic growth function can be interpreted

as follows: 1) ϕ_1 is the left asymptotic parameter for very small values of input, 2) ϕ_2 is the right asymptotic parameter for very large values of input, 3) ϕ_3 is the Inflection Point parameter also called Delay, which indicates the time taken to reach half the difference between left and right asymptotic values, and 4) ϕ_4 is the rate parameter, also called a scale parameter, denoting a scaling function on the time axis.

Crucial observations can be drawn regarding the nature of the NLME fit from Table 3.3. It can be seen that the logistic NLME models provide the best fit to the longitudinal data based on the values of the standard error parameter.

3.5 Extension of NLME to Multimodal Data

Since we want to model multimodal data, the j th response at time t_{ij} belonging to each subject i , modality m in a region R is written as:

$$y_{ij,m}^R = \beta_{1,m} + b_{i1,m} + \frac{\beta_{2,m} + b_{i2,m}}{1 + \exp^{\frac{\beta_{3,m} + b_{i3,m} - t_{ij}}{\beta_{4,m} + b_{i4,m}}}} + e_{ij}. \quad (3.27)$$

Therefore, for the entire set of multimodal images, the responses can be modeled as

$$\begin{bmatrix} y_{ij,m}^R \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} f(\phi_{i,m}, t_{ij}) \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} + e_{ij}. \quad (3.28)$$

In order to jointly study both variability within a modality (between individuals) and across modalities, the random effects belonging to all modalities are assumed to follow a multivariate normal distribution, the parameters of which are estimated. In this manner, the growth patterns of scans from different modalities are associated and jointly estimated rather than separately. Therefore, for both the modalities $[m = 1, 2, \dots, M]$ that we consider, the joint random effects associated with parameters b_{i2} and b_{i3} (corresponding to parameters ϕ_2 and ϕ_3 , i.e., asymptote and delay) are

$$\mathbf{b}_i = [b_{i2,1} \quad b_{i3,1} \quad \dots \quad b_{i2,M} \quad b_{i3,M}]^T \sim N(0, \boldsymbol{\psi}). \quad (3.29)$$

Inferences can be made based on the above statistical analysis by 1) studying the estimated mixed effects parameters and resulting growth trajectories, and 2) hypothesis testing to find significant differences in parameters belonging to different modalities. The details about computation of estimated parameters $(\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\sigma}^2)$ as well as hypothesis testing can be found in [58].

Table 3.3: AIC values for choosing optimal configuration of random Effects for NLME model. Abbreviations used: Parameter (Par), Standard Error (Std Error), and Variance (Var) . *NOTE: The quantity within the brackets in the case of the NLME fit refers to the index of the parameters (abbreviated as Par.), that were considered as random effects in the corresponding mixed effects model. Here, the random effects ϕ_1 , for example, is represented by the index 1.

Quantity(all logistic fits)*	Cross-sectional	NLME(12)	NLME(13)	NLME(1)
AIC	CS	-842.5382	-820.7867	-721.1057
Par 1 (Fixed effect)	1.57	1.180	1.182	1.1856
Par 2 (Fixed effect)	5.4020	5.857	5.699	5.3967
Par 3 (Fixed effect)	1.4686	0.855	1.1488	1.473
(Par 1) Std Error	0.005	0.003	0.003	0.0053
(Par 2) Std Error	0.082	0.076	0.0519	0.0821
(Par 3) Std Error	0.11	0.072	0.095	0.1110
Var Par 1 (Random effects)	NA	0.0212	0.017	0.0123
Var Par 2 (Random effects)	NA	0.4265	NA	NA
Var Par 3 (Random effects)	NA	NA	0.613	NA
Within group error	NA	0.0289	0.03099	0.0633

3.6 NLME Modeling on Synthetic Data

To better understand multivariate NLME analysis, bivariate synthetic data were generated. Trajectories of change of two variables were generated using logistic growth functions such that the random effects parameters underlying the growth function belong to a multivariate Gaussian distribution. Two experiments were conducted independently in which the random effects parameters of the growth functions were 1) strongly correlated and 2) uncorrelated, as can be seen in Figure 3.8. The fixed effects were the same for both experiments. The individual and subject-specific trends were then modeled using NLME-based mixed effects analysis. The NLME model fit was first done separately for each variable, followed by a joint modeling of both the variables using the multivariate version of NLME. Our most important inference from these phantom experiments is that when correlation is present between parameters of the two variables, a multivariate NLME fit results in a significantly lower AIC value compared with the summed AIC values of individual fits for each variable, as seen in Table 3.4. Thus the superiority of the multivariate fit is observed, which confirms the need for multivariate analysis, particularly when correlation is present between variables in a dataset. Results of hypothesis testing (not shown), which are consistent with our experimental design, as well as accuracy of estimated growth function parameter values, shown in Table 3.5, further confirm the efficacy of the multivariate NLME framework.

3.7 NLME: Analysis and Inferences

The inferences that can be made from the statistical analysis outlined above consists primarily of understanding patterns of brain growth and timing based on the parameters of the estimated growth function. Since the logistic fit was most stable and suitable to the data studied, we will discuss the parameters of the logistic growth function. The different parameters of the logistic function represent asymptote (final contrast value in the adult-like brain), delay (time taken to reach 1/2 of the final asymptotic value), and a rate parameter. A study of these parameters, particularly of the delay parameter, has the potential to illustrate the heterosynchronous nature of brain maturation that takes place at different rates in different regions of the brain.

Hypothesis testing is done on the parameters of the growth function to check for significant differences between distinct population groups. The exact procedure for this is outlined in the NLME literature [58, 128]. It should be noted that when parameters from multiple regions are compared, correction for multiple comparisons is employed.

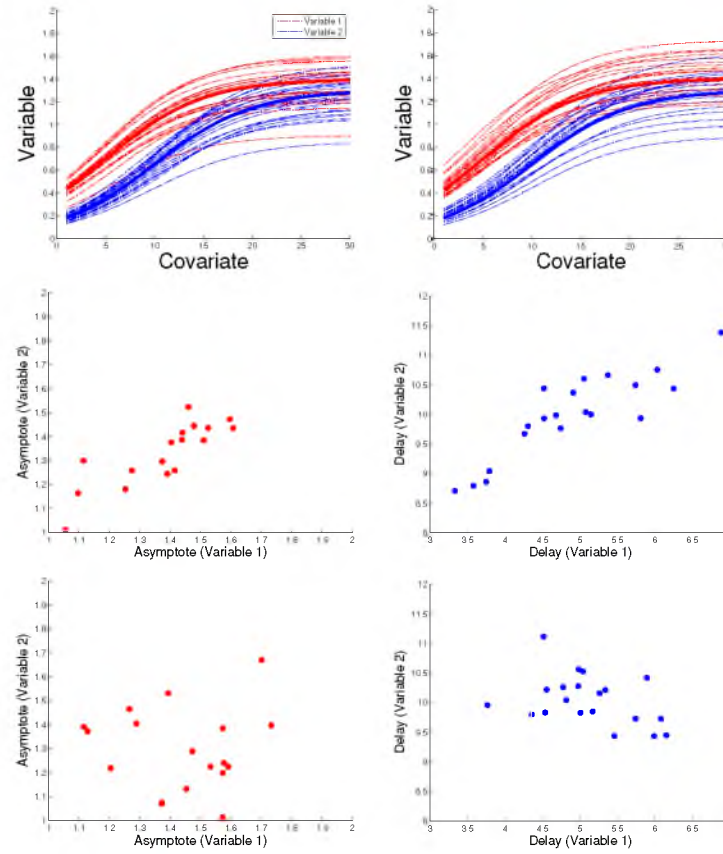


Figure 3.8: Top row: (Left) Strongly correlated phantom data, uncorrelated phantom data, Middle row: Trajectories of change in two variables with strongly correlated growth parameters, and Bottom row: Trajectories of change in two variables with uncorrelated (bottom) growth parameters, experimental values of (center) asymptote and (right) delay.

Table 3.4: AIC comparisons for separate and joint fits of multivariate synthetic data.

Relation Between Variables	AIC(Var.1) + AIC(Var.2)	AIC(Var.1 + Var.2)
Strong Correlation	-869.751	-1015.242
No Correlation	-866.006	-864.939

Table 3.5: Multivariate NLME: Estimation of growth parameters and hypothesis testing.

Parameter	Var.1(Truth)	Var.1(Estd)	Var.2(Truth)	Var.2(Estd)	p-value
Asymptote	1.4	1.408	1.3	1.308	< 0.001
Delay	5	5.029	10	10.290	< 0.001
Scale	5	4.983	5	5.008	0.089

In this work, we correct for multiple comparisons using the False Discovery Rate (FDR) method. In addition, since the different modalities have been modeled in a joint manner, it is possible to do hypothesis testing based on the estimated fixed effects and their sampling distributions. By comparison of the fixed effects parameters for a specific cortical region across all modalities, conclusions can be drawn as to the timing of specific components or phases that constitute the complete maturational process.

CHAPTER 4

TISSUE APPEARANCE ANALYSIS : EXPERIMENTS WITH IBIS (INFANT BRAIN IMAGING STUDY) DATA

4.1 Studies of Infant Brain Appearance Change

In this chapter, we study WM-GM intensity variations in longitudinal brain MR images from the ACE-IBIS (Autism Centers of Excellence - Infant Brain Imaging Study) project. The dataset analyzed consists of the brain MR images of children scanned at approximately 6 months, 12 months, and 24 months of age. Intensity variations between WM and GM tissues are quantified for this dataset by computation of the WIVID contrast measure, and changes in the WIVID measure over time are modeled to study neurodevelopmental processes. The purpose of studying intertissue appearance change in this infant brain dataset is two-fold - firstly, it serves to deepen our insight into mechanisms underlying normal brain development, and secondly, it enables detection of differences between the neurodevelopmental trajectories of distinct population groups.

The ACE-IBIS dataset consists of T1W (T1-Weighted) and T2W (T2-Weighted) brain MR scans. The multimodal nature of this dataset enables the study of distinct biological properties captured by each modality. In order to facilitate this, values of WIVID (White-gray Intensity Variation in Infant Development) are calculated independently for each modality included in this study.

The WIVID measures studied are also computed in a region-dependent manner. That is, a specific value of WIVID is extracted for each cortical region of the brain. The fact that brain growth is spatially heterosynchronous is the motivation for studying the WIVID measure independently for each brain region. It should be noted that a specific region being studied is further categorized based on the hemisphere to which it belongs. For example, in the computation and study of the WIVID measure, values from the left temporal lobe are treated as distinct from those of the right temporal lobe for the same subject.

In addition to its modality-specificity and region-specificity, the WIVID measure is also analyzed independently for distinct population groups. The subjects studied in the ACE-IBIS dataset belong to different population groups, and could be categorized based on sex and based on their risk/diagnosis category for ASD (Autism Spectrum Disorder). Similarly, all the subjects could be classified into either male or female population groups. Studies of differences in the spatiotemporal WIVID trajectories of distinct population groups are performed, with potential applications including early diagnosis of brain disorders and greater understanding of mechanisms underlying abnormal brain growth.

The study of intertissue image appearance in ACE-IBIS data is divided into quantification, modeling, and analysis components. The first component comprises stepwise processing of infant MR scans and extraction of WIVID contrast measures. Detailed discussion of the image processing pipeline employed for extraction of the WIVID measure has already been given in Chapter 2. In this chapter, we augment this discussion with actual examples drawn from processing of the ACE-IBIS dataset. These examples display intermediate results from key stages of the pipeline used for image processing and extraction of the WIVID biomarker. The second component concerns determining the best model fit for the WIVID data under study, following the steps described in Chapter 3. Initially, we visually analyze the WIVID data points to investigate which model would fit the data best. A comparison of data fits from cross-sectional and mixed effects models, linear and nonlinear models, and several distinct nonlinear growth functions has already been discussed in Chapter 3 with respect to a subset of the ACE-IBIS data. In this chapter, quantitative proof is provided for choice of model using AIC (Akaike Information Criterion) values resulting from fitting models to the entire dataset under study. Comparison of the nature of the resulting fit with prior neurobiological knowledge is also an integral part of this step. In the third component, we analyze the parameters resulting from fitting of the best model to the data. The parameters analyzed describe growth and are used to compare trajectories of changes in the WIVID contrast measure, across brain regions and modalities. Hypothesis testing is also done to predict significant differences in these growth parameters between distinct population groups under study. Differences based on sex and risk/diagnosis for ASD are explored.

Two interrelated questions are raised by modeling WIVID values based on ACE data: first, “How can the changing directionality of the WM-GM intensity gradient be accounted for in the WIVID contrast measure?”, and second, “Given prior biological knowledge that WM-GM intensity differences first decrease and then increase, how can WIVID values be

effectively modeled?”. These questions are linked, since the intensity variations quantified on the increasing and decreasing sections of this growth curve have WM-GM gradients oriented in opposite directions. As described in Chapter 2, methods to quantify WIVID direction have been developed, by means of which the WIVID measure can be assigned a “sign” based on the relative direction of the intensity gradient of gray matter with respect to white matter. Insufficient samples belonging to the decreasing portion of the WIVID change trajectory prevent usage of WIVID points with negative direction. Instead, only the increasing portion of the WIVID curve is modeled, thus explaining the restriction of analysis to WIVID points with positive direction. Therefore, for purposes of preliminary exploration, only positively valued WIVID data points are considered for the statistical analysis. For purposes of visualization, WIVID values with both positive and negative directionality are displayed in some figures. In these figures, only the magnitude of the WIVID measure is plotted, while the assigned direction is indicated by color of the displayed data points. Details of WIVID computation, visualization of data points, and choice of best model fit are discussed in the following sections, followed by analysis of normative and population-specific trends.

4.1.1 ACE-IBIS Study

The ACE-IBIS (Autism Centers of Excellence: Infant Brain-Imaging Study) study is a longitudinal study with image processing and analysis of structural MRI/DTI of infants at high risk for autism at approximately 6, 12, and 24 months of age. The goal of this multisite network study is to analyze infants at genetic risk for ASD based on having an older sibling diagnosed with autism³ [133], using both neuroimaging and behavioral data. This investigation is funded by the National Institutes of Health and is informally called the Infant Brain Imaging Study (IBIS). The ACE-IBIS network includes four clinical data collection sites (University of North Carolina at Chapel Hill, University of Washington, Children’s Hospital of Philadelphia, Washington University in St. Louis), a data coordinating center (Montreal Neurological Institute at McGill University), and two image processing sites (University of Utah, University of North Carolina).

The study enrolls 1) 6-month-old high-risk infants (with an older sibling who is diagnosed with ASD) who are seen for follow-up assessments at 12 and 24 months of age, 2) 12-

³To quote the details of the ACE study from a recent paper by Hazlett et al. [133], “Subjects were characterized as having high risk if they had an older sibling with a diagnosis of an ASD that was documented in a clinical diagnostic report and confirmed by the Autism Diagnostic Interview-Revised administered at enrollment. Subjects were enrolled in the low-risk group if they had an older sibling without evidence of ASD and no family history of a first- or second-degree relative with an ASD.”

month-old high-risk infants who are followed up at 24 months, and 3) a comparison group of typically developing controls - infants considered to be at low risk for autism (i.e., older siblings are developing typically), who are seen at 6, 12, and 24 months. The test detecting diagnosis of autism is administered at 24 months of age (and again at 36 months of age if the timeline of the project allows). Due to subject availability, data acquisition, and processing issues, it must be mentioned that not all subjects have scans available from the required follow ups - for example, some 6-month-old high-risk infants might not have scans available at 12 months of age, and might only have follow-up scans available at 24 months of age.

Based on combined risk and diagnosis information, the subjects in the dataset can be divided into four groups for the purpose of our analysis. The first group, HR+ (High Risk Positive), consists of subjects born with a high genetic risk for autism, who were diagnosed positive for ASD. The second group, HR- (High Risk Negative), consists of subjects who are born with a high genetic risk for autism, diagnosed as negative for ASD. The third group consists of healthy controls (HC), subjects without high genetic risk for autism, who were also diagnosed as negative for ASD. The fourth group consists of controls positive (C+), subjects without high genetic risk for autism who were diagnosed as positive for ASD. Subjects from the C+ group were limited in number and have hence been mostly excluded from the analysis presented. Additional scans from two subjects diagnosed with Downs Syndrome (DS) have also been included. Both male and female subjects were included in the analysis.

4.1.2 MR Image Acquisition

A 3-T Siemens Tim Trio scanner with a 12-channel head coil was used for MR acquisition at all clinical sites. The scans were acquired without any sedation, when the infants were asleep. Steps such as feeding, swaddling, and protecting the ears of the infants were undertaken, along with conditioning with scanner sounds from a compact disc to ensure that the infant is able to sleep and can be successfully scanned.

The details of MR acquisition provided below have been reproduced from a paper by Hazlett [133]. The MR acquisition protocol includes 1) a localizer scan ; 2) a three-dimensional T1 magnetization-prepared rapid acquisition gradient-echo (MPRAGE) scan: TR (Repetition Time) =2,400 ms, TE (Time to Echo) =3.16 ms, 160 sagittal slices, field of view (FOV)=256 mm, voxel size=1 mm³ ; 3) a three-dimensional T2 fast spin echo (FSE) scan: TR=3,200 ms, TE=499 ms, 160 sagittal slices, FOV=256 mm, voxel size=1 mm³ ; and 4) a 25-direction diffusion tensor imaging scan: TR=12,800 ms, TE=102 ms, slice thickness=2 mm isotropic, variable b value=maximum of 1,000 s/mm², FOV=190

mm. Regional distortions were corrected and image quality was assessed with the help of a LEGO phantom that was scanned each month at all the acquisition sites. Additionally, two human phantoms were scanned once each year per scanner. Scanner stability and reliability were tested across all sites and scanners using the LEGO phantom and human phantom data.

4.1.3 Sample Used in Analysis

For the analysis we perform, a subset of the data obtained in the ACE-IBIS study was included. The total number of subjects whose scans were used was 306. Of these, 185 subjects scanned at all three time points, 71 subjects scanned at 12 and 24 months, 41 subjects scanned at 6 and 24 months, and 14 subjects scanned at 24 months alone. There were 70 controls, 56 HR+ (High Risk Positive) subjects, and 174 HR- (High Risk Negative) subjects included in the study. Categorized based on sex, 193 subjects were male and 113 subjects were female. At each time point, two scans belonging to T1W and T2W modalities were obtained. Therefore, a total of 1556 scans were obtained from this study.

Properties of the data subset used are listed below as follows - 1) Imbalanced in Number of Samples : The number of samples obtained per subject was imbalanced, that is, each subject was scanned at one to three time points ; 2) Imbalanced in Covariates : The values of the covariates were imbalanced - that is, the subjects were not scanned exactly at the same time points ; and 3) Unequal sample sizes : The sample sizes of the data was unequal - that is, the number of subjects belonging to each population group was unequal.

In the following section, we illustrate a practical example of the pipeline for computation of the WIVID contrast measure, implemented on a healthy controls subject as part of a population study.

4.2 Computation of the WIVID Measure

Computation of the WIVID measure follows the procedure outlined in Chapter 2, which is based on calculation of the HD (Hellinger Distance) between the nonparametric intensity distributions for gray and white matter tissue classes. Initially, steps such as alignment to a template, bias correction, intrasubject registration and intersubject registration are performed. The original series of images belonging to a control subject from the ACE-IBIS study is displayed in Figure 4.1.

From these scans, it can be observed that the T1W images displayed have very low WM-GM contrast at the age of 6 months, and increase in contrast with time. At the age of 24 months they display excellent contrast. In the case of T2W images, the contrast is

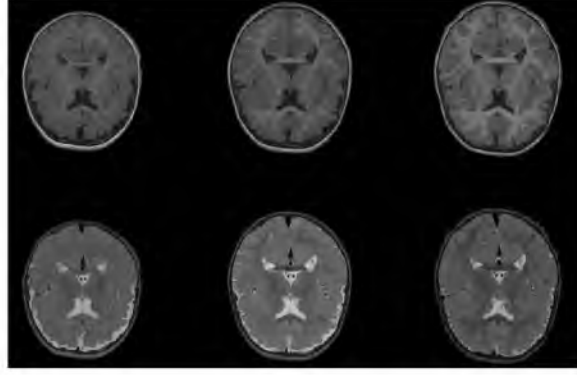


Figure 4.1: Multimodal, repeated measures dataset : Series of longitudinal images from a single subject taken at 6 months (left), 12 months (center), and 24 months (right) of age.

close to zero at 6 months, and increases slowly over time up to 24 months of age. It can be observed that T1W images “lead” T2W images in terms of WM-GM contrast change

Coregistered scans belonging to the same subject have been shown in Figure 4.2. These scans are the result of intrasubject registration. Coregistration is an essential precursor to proper segmentation of early time point images into the major tissue classes. This is mainly since, after coregistration, probability maps from later time point images can be employed as priors for improving the segmentation of early time point images that otherwise suffer from lack of contrast.

The multimodal set of scans at each time point is segmented using the longitudinal method outlined above. Results from segmentation can be seen in Figure 4.3, which shows WM and GM binary label maps corresponding to scan sets from each time point. Segmentation of these scan sets was performed using longitudinal techniques and resulted in classification of the brain into major tissue classes.

Binary label maps as priors corresponding to the multimodal set of scans obtained at each time point, are shown here. The displayed scans and their label maps lie in their original coordinate space (images corresponding to different time points are not coregistered). In this figure, it should be noted that the parcellation maps that have been coregistered to the set of scans at each time point can be seen in the third column, dividing the brain into major cortical regions. These scans and their parcellation maps lie in their original coordinate space, that is, images corresponding to different time points are not coregistered. It can be seen that the use of a probabilistic prior for segmentation of early time points yields satisfactory results, especially considering the poor intertissue contrast during early stages of brain development. Consistency in the segmented label maps across different time points is also ensured by this process, as seen in the segmentation results.

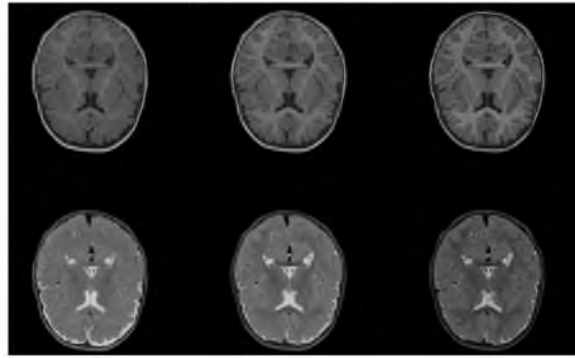


Figure 4.2: Intrasubject image registration: Series of coregistered longitudinal images from the same subject, taken at 6 months (left), 12 months (center), and 24 months (right) of age. The coregistration was performed using nonlinear deformation techniques.

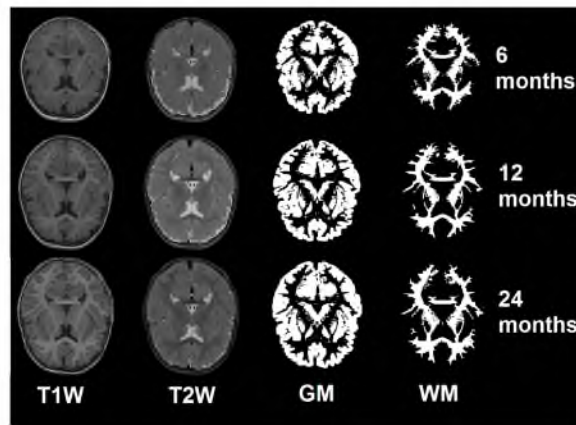


Figure 4.3: Image segmentation: Results from segmentation of a longitudinal series of images from the same subject, obtained at 6 months (first row), 12 months (second row), and 24 months (third row) of age. The multimodal scan set that is segmented consists of T1W (first column) and T2W (second column) scans.

After segmentation, parcellation of the image set at each time point into major brain regions is shown in Figure 4.4. Parcellation is performed by deformation of an anatomical atlas with delineated cortical regions to the coordinate space of scans under study. Exploration of changing WM-GM intensity variation on a per-lobe basis is possible due to parcellation of brain scans.

As a result of segmentation of the MR scans into major tissue classes and parcellation into cortical regions, WM and GM intensity distributions from important brain regions can be extracted. Sample WM and GM distributions from the left frontal lobe and left occipital lobe are displayed in Figure 4.5. It should be noted that in this figure, each distribution is normalized such that the area under it sums to unity. For example, the area under the WM distribution for the left temporal lobe in the T1W scan sums to unity. As a result,

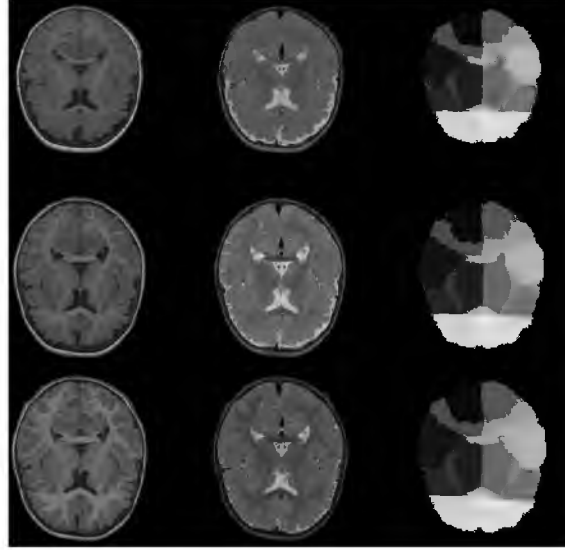


Figure 4.4: Image parcellation: Results from parcellation of a longitudinal series of images from the same subject into major cortical regions. The scans in this series were obtained at 6 months (first row), 12 months (second row), and 24 months (third row) of age. The multimodal scan set which is segmented consists of T1W (first column) and T2W (second column) scans, and uses the binary label maps for GM (third column) and WM (fourth column).

the shapes of distributions do not reflect the number of points constituting each tissue class. From these plots, it can be concluded that in general, the distance of separation between WM and GM intensity distributions increases with time. Lobar differences in the configuration of WM and GM intensity distributions are also evident from this figure, particularly for values from T1W scans. The left occipital acquires considerable WM-GM contrast for the T1W scan at 6 months of age compared with the left frontal lobe - an observation that is consistent with theories of heterosynchronicity in brain development.

WIVID contrast values corresponding to the distributions displayed are shown in Figure 4.6. As a departure from the convention adopted in this chapter of only displaying the magnitude of WIVID, the WIVID values plotted in this figure include their directional attribute. The direction is conveyed by their sign, with these signs provided by the direction of gradient as defined in Chapter 2. The general trend of WIVID values increasing with age can be observed, although the left occipital lobe shows a deviation at 24 months.

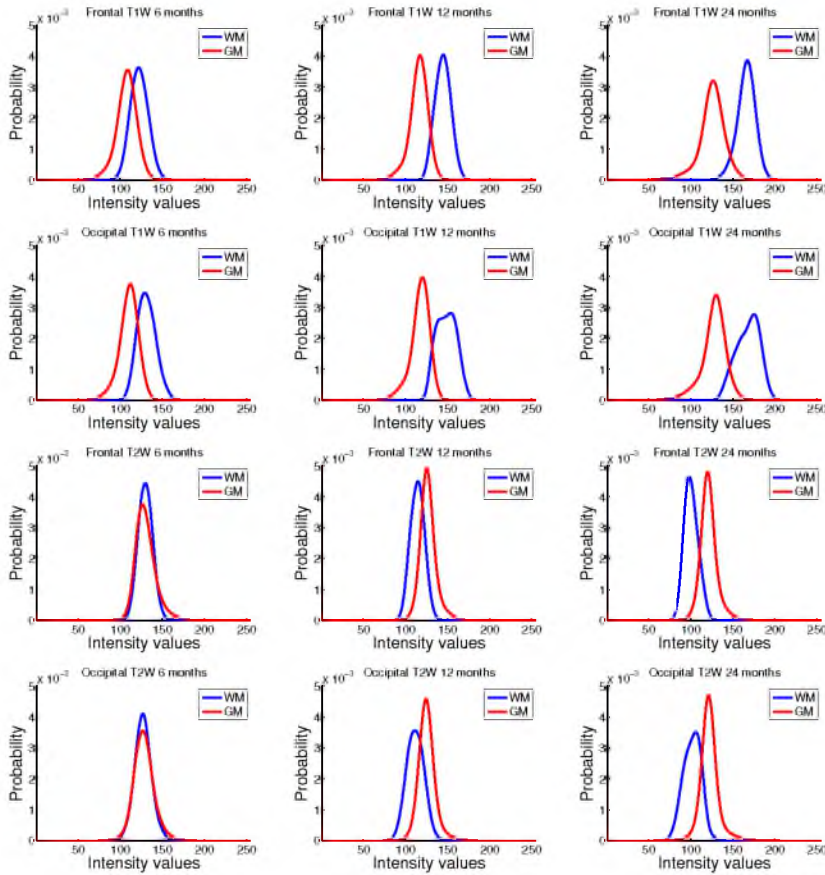


Figure 4.5: White matter (blue), and GM (red) intensity distributions are displayed for frontal and occipital lobes of the brain. The distributions that are displayed correspond to scans acquired at 6 months (first column), 12 months (second column), and 24 months (third column) of age. The distributions correspond to cortical regions in the left hemisphere.

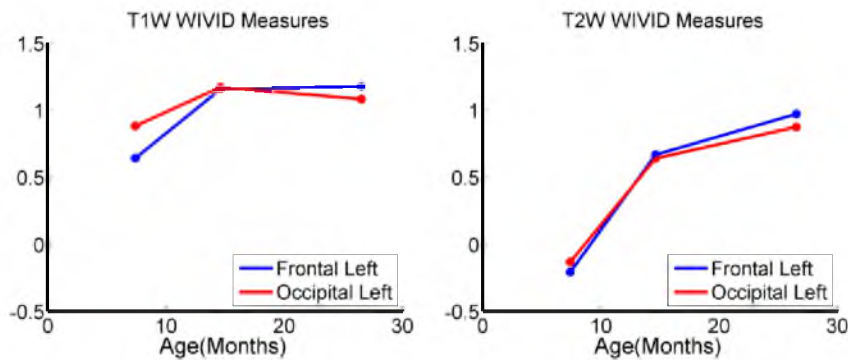


Figure 4.6: Signed WIVID contrast values extracted from two major cortical lobes of the brain are displayed for scans from different time points. The sign for the contrast values is derived from the directionality attribute.

From this plot of absolute WIVID values, intermodal and interlobe differences in trajectories of change in the WIVID contrast measure can be clearly seen. In particular, the early growth of WIVID values in the left occipital lobe when compared with the left frontal lobe is observed. Statistical evidence for these interlobe and intermodal differences in the WIVID contrast measure will be analyzed in later sections.

4.3 Experiments for Finding the Best Model Fit

In Chapter 3, a subset of the ACE-IBIS under study was analyzed, and the best methods for statistical modeling of this subset were determined. Nonparametric methods such as kernel regression were explored empirically and considered unsuitable for modeling WIVID across this dataset. Further, it was established via cross-sectional modeling that nonlinear growth functions were superior to linear fits of the data. Additionally, the superiority of mixed-effects models when compared with cross-sectional models for fitting the longitudinal data was confirmed. In this section, we examine some conclusions that were made related to the sample subset of the data studied (choice of parametric fit over nonparametric fit, nonlinear fit over linear fit, mixed effects fit over cross-sectional fit) and check if they remain true for the entire population. In addition, we confirm the superiority of nonlinear vs. linear fits, and mixed-effects vs. cross-sectional fits by modeling the entire population of healthy controls using each of these methods and comparing the results quantitatively using statistical tools. Further, once a particular fit is chosen, other options are explored in depth. These could include the type of nonlinear model function used, type of grouping employed to best model the data, univariate vs. multivariate modeling for multimodal data, and choice of fixed and random effects parameters. The population of healthy controls was used to analyze choices for the best model fit.

4.3.1 Choice of Parametric Fits for Modeling Data

Two major reasons can be cited regarding the choice of parametric fit for modeling the data, instead of the nonparametric kernel regression fit. The first of these is concerned with the fact that MR scans in the ACE-IBIS dataset are sparsely sampled across time. Kernel regression involves estimation of fit at several intermediate points between samples; therefore, a maximum of three time points per subject is too sparse for proper estimation of the curve using kernel regression. In addition, data in the current sample is concentrated around the 6 months, 12 months, and 24 months time points, making the samples imbalanced and difficult to model via kernel regression. The second major reason is that parametric models for the data result in concise parameters by means of which behavior

of the data can be described. This succinct representation of change in data is highly valuable in clinical studies such as the ACE-IBIS. Concise parametric representation allows for quick translation of statistical results into easily interpretable parameters. Finally, kernel regression is highly dependent on choice of bandwidth and parameters of the kernel. Results from regression might be variable based on choice of these parameters. Since stability and consistency are central goals of our modeling exercise, we adopt relatively stable parametric models for analyzing our data.

4.3.2 Cross-Sectional Model for Fitting data

In this section, we perform a comparison of various cross-sectional fits to the data by studying their AIC values. Cross-sectional fitting can provide an accurate initial estimate of the best methods for modeling the data. Linear and quadratic fits, discussed in detail in Chapter 3, are applied to the data. Additionally, we also apply three major nonlinear growth functions - the logistic function, Gompertz function, and asymptotic function. These three functions model change that is highly nonlinear and asymptotic in nature. Considering the response variable y being studied at time instants t , the equations describing these functions, along with parameters ϕ_1, ϕ_2, ϕ_3 , are listed below.

$$y(t) = \frac{\phi_1}{(1 + e^{\frac{\phi_2 - t}{\phi_3}})}, \dots \text{Logistic fit.} \quad (4.1)$$

$$y(t) = \phi_1 \times e^{-\phi_2 \phi_3^t}, \dots \text{Gompertz fit.} \quad (4.2)$$

$$y(t) = \phi_1 + (\phi_2 - \phi_1)e^{-e^{\phi_3 t}}, \dots \text{Asymptotic fit.} \quad (4.3)$$

The results from model fits of all the above functions to WIVID contrast data from T1W and T2W scans can be seen in Table 4.1 and Table 4.2. Based on the vastly lower AIC values for the nonlinear models, it can be concluded that nonlinear models are better fits for the data compared to linear models. Amongst nonlinear fits for the data, the AIC values are extremely close, and hence it is not conclusive as to which nonlinear growth function is the best fit at this stage of the modeling procedure.

4.3.3 Cross-Sectional vs. Mixed Effects Models

Earlier, it was determined that fitting the ACE-IBIS data with nonlinear functions is superior compared with fits using linear functions when implemented using a cross-sectional approach. As a continuation of the discussion on choice of best model fit to the data, we now implement nonlinear growth functions using mixed effects methods.

Table 4.1: AIC values from results of fitting various cross-sectional models to WIVID contrast values for T1W scans changing with time. The AIC value is substituted by CC (cannot converge) when the model fit estimation does not converge. The abbreviations used for the cross-sectional models are as follows: Lin. = Linear, Quad. = Quadratic, Log. = Logistic, Gomp. = Gompertz, and Asymp. = Asymptotic. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Lin.	Quad.	Log.	Gomp.	Asymp.
Fron L	-204.0227	-417.8263	-481.9439	-482.9996	-483.1256
Temp L	-192.0841	-418.8859	-474.7366	-472.4878	-468.6484
Pari L	-364.3062	-500.488	-542.3359	CC	-543.6835
Occi L	-226.6109	-407.2369	-473.6326	CC	-472.5627
Fron R	-206.2182	-418.7146	-500.2523	-501.5194	-502.0333
Temp R	-177.4749	-403.8914	-460.5110	-457.8785	-453.3481
Pari R	-374.7357	-518.5823	-547.1465	-547.2528	-547.2614
Occi R	-218.7208	-404.7162	-467.2036	-466.5585	-465.2907

Table 4.2: AIC values from results of fitting various cross-sectional models to WIVID contrast values for T2W scans changing with time. The abbreviations used for the cross-sectional models are as follows: Lin. = Linear, Quad. = Quadratic, Log. = Logistic, Gomp. = Gompertz, and Asymp. = Asymptotic. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Lin.	Quad.	Log.	Gomp.	Asymp.
Fron L	-216.3274	-385.1637	-392.7746	-394.7163	-389.4201
Temp L	-233.6071	-350.7108	-354.3747	-355.6372	-353.3851
Pari L	-239.038	-385.6701	-389.1405	-395.6181	-401.2010
Occi L	-171.4767	-376.7392	-381.1361	-389.1581	-390.0586
Fron R	-225.4172	-383.119	-389.4781	-502.0333	-386.9869
Temp R	-229.2674	-351.0523	-353.7357	-483.3481	-354.1122
Pari R	-219.4007	-366.2591	-371.0914	-547.2614	-376.7596
Occi R	-168.2874	-370.7254	-373.9839	-465.2907	-385.0810

The differences between the cross-sectional model and mixed effects model have been discussed in detail in Chapter 3. Essentially, the mixed effects model consists both of a fixed component denoting values estimated for the entire population under study, and a random component that estimates the deviations of individual subject trajectories from the population curve. For example, the logistic function introduced earlier can be extended into a mixed-effects model. The logistic function implemented using the NLME format, in which the parameters of the growth function ϕ consist of both fixed and random effects, can be written as

$$f(\phi_i, t_{ij}) = \phi_{i1} + \frac{\phi_{i2}}{1 + \exp^{-\frac{\phi_{i3} - t_{ij}}{\phi_{i4}}}}. \quad (4.4)$$

Here, i is the subject being studied at time points t_{ij} . The parameter vector ϕ_i in this equation consists of both fixed and random effects components. This can be written as

$$\phi_i = \mathbf{A}_i \boldsymbol{\beta} + \mathbf{B}_i \mathbf{b}_i, \text{ where } \mathbf{b}_i \sim N(0, \boldsymbol{\psi}). \quad (4.5)$$

Restating the definition given in Chapter 3, the fixed effects and random effects design matrices are given by \mathbf{A}_i and \mathbf{B}_i for each subject, and the p-vector of fixed effects is given by $\boldsymbol{\beta}$ and the q-vector of random effects is given by \mathbf{b}_i . Of all the parameters describing a function, some can be chosen to have purely fixed effects components, and some can be chosen to have random effects components also. Based on the diagnostic tools analyzed in Chapter 3, it was concluded that the mixed effects model was clearly superior to cross-sectional models for fitting longitudinal data. Here, we assume the same is true for the case of the entire population of healthy controls, particularly since we know that repeated measures data cannot be correctly modeled using the population-level cross-sectional approach.

4.3.4 NLME Models for ACE-IBIS Data

The three nonlinear growth functions mentioned earlier were implemented, using both cross-sectional and mixed effects modeling approaches. The NLME fits assumed ϕ_1 and ϕ_2 to be random effects parameters, whereas ϕ_3 was assumed to be purely a fixed effects parameter. This choice was because of the fact that in the nonlinear functions studied, ϕ_3 happens to be quite unstable compared with the other parameters, primarily because it refers to a scaling factor that is relatively complex to estimate. It was observed from results using only ϕ_1 as random effects that the AIC values of the resulting fits (not displayed here) was not lower than for the case presented, in which ϕ_1 and ϕ_2 were both used as random effects parameters. When ϕ_1 and ϕ_3 were assumed to be the random effects parameters,

with ϕ_2 assumed to be a fixed effect parameter, it was concluded that the fit was highly unstable and did not converge in several cases. Additionally, the standard deviations of estimated values were very high for these fits, often ranging till up to 20 percent of the estimated value itself.

Based on the above discussion, the ACE-IBIS data were fit with the logistic, Gompertz, and asymptotic nonlinear growth functions using mixed effects methods, maintaining ϕ_1 and ϕ_2 as random effects parameters. AIC values resulting from these fits can be compared in Table 4.3 and Table 4.4. From these values, it can be concluded that the logistic function is the best and most stable for fitting WIVID contrast values from T1W scans, given that the NLME models implemented using Gompertz and asymptotic cannot converge to a stable estimate in several cases. In the case of WIVID contrast values from T2W scans, the logistic function proved to be the best fit according to AIC values measured in different regions. In the left and right occipital lobes, although the logistic function did not have the lowest AIC value compared to the Gompertz and asymptotic, it still had acceptable AIC values. Finally, the correlation matrices and standard deviations of estimated parameters in the logistic model fit remained within acceptable bounds compared to these values from Gompertz and asymptotic fits.

4.3.5 Multivariate Versus Univariate Modeling

In the final stage, we investigate whether joint multivariate to WIVID contrast values is preferable to fitting each modality separately. In order to do this, we compare AIC values of the joint multivariate fit to the summed AIC values resulting from fitting to each modality separately. The results of this comparison can be assessed in Table 4.5. From the resulting AIC values, it can be concluded that the joint fit is superior to the individual fit in all lobes except for the frontal lobe. This superiority of the joint fit could be attributed to the fact that WIVID contrast change trajectories have a non-zero correlation between modalities. This intermodal correlation in tissue appearance change in turn contributes to estimating a better fit to the data. Additionally, when T1W and T2W data are jointly studied, the number of samples available to estimate a particular trajectory is also increased (although the number of variables needed for determining this trajectory also increase due to nonzero correlation between modalities).

4.3.6 Other Factors to Consider in NLME Modeling

Additional factors to consider in NLME modeling include the structure of the random effects variance-covariance matrix, and heteroscedasticity of random effects.

Table 4.3: AIC values from results of fitting various nonlinear models using NLME methods to WIVID contrast values for T1W scans changing with time. The AIC value is substituted by CC (cannot converge) when the model fit estimation does not converge. The abbreviations used for the cross-sectional models are as follows: Log. = Logistic, Gomp. = Gompertz, and Asymp. = Asymptotic, NLME = Non linear mixed effects. The NLME fits assumed ϕ_1 and ϕ_2 to be random effects parameters. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Log.(ME)	Gomp.(ME)	Asymp.(ME)
Fron L	-539.9489	-489.004	CC
Temp L	-510.3468	-487.9058	CC
Pari L	-521.7956	CC	CC
Occi L	-568.9369	CC	CC
Fron R	-560.7123	CC	CC
Temp R	-495.3094	CC	-474.2473
Pari R	-521.9392	CC	CC
Occi R	-578.9379	CC	CC

Table 4.4: AIC values from results of fitting various nonlinear models using NLME methods to WIVID contrast values for T2W scans changing with time. The AIC value is substituted by CC (cannot converge) when the model fit estimation does not converge. The abbreviations used for the cross-sectional models are as follows: Log. = Logistic, Gomp. = Gompertz, and Asymp. = Asymptotic, NLME = Non-Linear Mixed Effects. The NLME fits assumed ϕ_1 and ϕ_2 to be random effects parameters. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Log.(ME)	Gomp.(ME)	Asymp.(ME)
Fron L	-401.4559	-399.5792	-389.4201
Temp L	-371.2675	-353.3851	-363.1307
Pari L	-419.958	-401.2010	-416.7448
Occi L	-428.7865	-390.0586	-430.8633
Fron R	-401.5341	-386.9869	-388.2832
Temp R	-366.3393	-354.1122	-364.9019
Pari R	-419.5685	-376.7596	-414.1612
Occi R	-373.9839	-420.5339	-420.1369

Table 4.5: AIC values from results of NLME-based fitting of the logistic function to WIVID contrast values for T1W and T2W scans changing with time, using (first column) joint multivariate modeling, and (second column) the summed AIC value from univariate fits to T1W and T2W scans. The logistic NLME fits assumed ϕ_1 and ϕ_2 to be random effects parameters. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	AIC(T1+T2)	AIC(T1) +AIC(T2)
Fron L	-938.7905	-941.4048
Temp L	-911.3041	-881.6143
Pari L	-983.4265	-941.7536
Occi L	-1015.866	-997.7234
Fron R	-944.5014	-962.2464
Temp R	-911.3038	-861.6487
Pari R	-986.0895	-941.5077
Occi R	-1034.893	-999.4718

Grouping due to additional factors such as sex, and ways to model “signed” WIVID values (that is, WIVID contrast values that have been assigned a directionality) must also be considered. In the experiments outlined above, the random effects variance-covariance matrix estimated in the NLME procedure was considered to be a full matrix. However, assumptions could be made regarding the variance-covariance matrix that reduce the number of parameters estimated in the matrix and enforce a certain pattern of change.

For example, the variance-covariance matrix could be assumed to be a diagonal matrix, thereby constraining the covariance between different mixed effects parameters to be zero. In practical terms, this could imply that the asymptote and delay are uncorrelated in the logistic fit performed above. Since the WIVID contrast measure is novel and we have limited knowledge of its changing behavior with time, we assume a full variance-covariance matrix in this chapter, for modeling purposes.

A closely related consideration is that the random effects values estimated for each subject are currently assumed to be Gaussian in nature in each group. That is, these random effects form a single cluster per group, and the shape of this cluster is estimated by the random effects variance-covariance matrix. Experiments with heteroscedasticity can be used to further subdivide this cluster into subclusters, which is not to be confused with grouping factors that naturally divides the random effects into group-based clusters. Each cluster of random effects parameters will then assumed to be Gaussian. Although

this increases the number of parameters to be estimated, heteroscedasticity could be useful when the population under study does not conform to one single pattern of deviation. For example, in the logistic model under study, the asymptote parameters of the healthy controls data could be divided into two clusters if heteroscedastic. However, the implications of heteroscedasticity are not completely understood, particularly in the context of testing for significance between parameters of the fit. Since hypothesis testing forms a core part of the inferential process we use to understand intertissue appearance changes, and due to lack of evidence that random effects for the data studied are heteroscedastic, we assume homoscedasticity over heteroscedasticity for modeling the WIVID contrast data.

In order to consider variability due to grouping factors such as sex, we perform the hierarchical or multilevel modeling as described in Chapter 3. The mixed effects model is inherently hierarchical in nature since it consists of a population-level model and subject-level model.

We have understood from the discussion in the last section that joint fits to multimodal data are superior to univariate fits to data from each modality. In addition, consider the case that grouped data are being modeled - for example, data consisting of both male and female subjects are analyzed. In this case, we have a multilevel model, with the first level being modality, the second level being based on grouping factors such as sex, and the third level being subject-level variations. If data are studied in this hierarchical manner, comparisons between different groups can also be easily made based on p-values resulting from the multilevel model estimation.

Finally, the question of modeling the directionality attribute of WIVID contrast values is examined. Based on the discussion of WIVID contrast direction in Chapter 2, it could be concluded that WIVID contrast is assigned a directionality attribute or sign, based on the configuration of WM and GM intensity distributions at a point, with respect to the configurations across time. As a result, the WIVID computations we made have also included a sign or direction component. The behavior of WIVID contrast across infant brain scans from birth to 2 years of age generally involves a trajectory of decreasing WIVID followed by increasing WIVID. The decreasing section of the WIVID contrast curve is assigned negative direction, and the increasing section of the WIVID contrast curve is assigned a positive direction. Figure 4.7 illustrates the directionality concept in the left parietal lobe for the ACE-IBIS data under study. Since T1W WIVID values for the age range under study are always positive, our attention will be mainly on the T2W data.

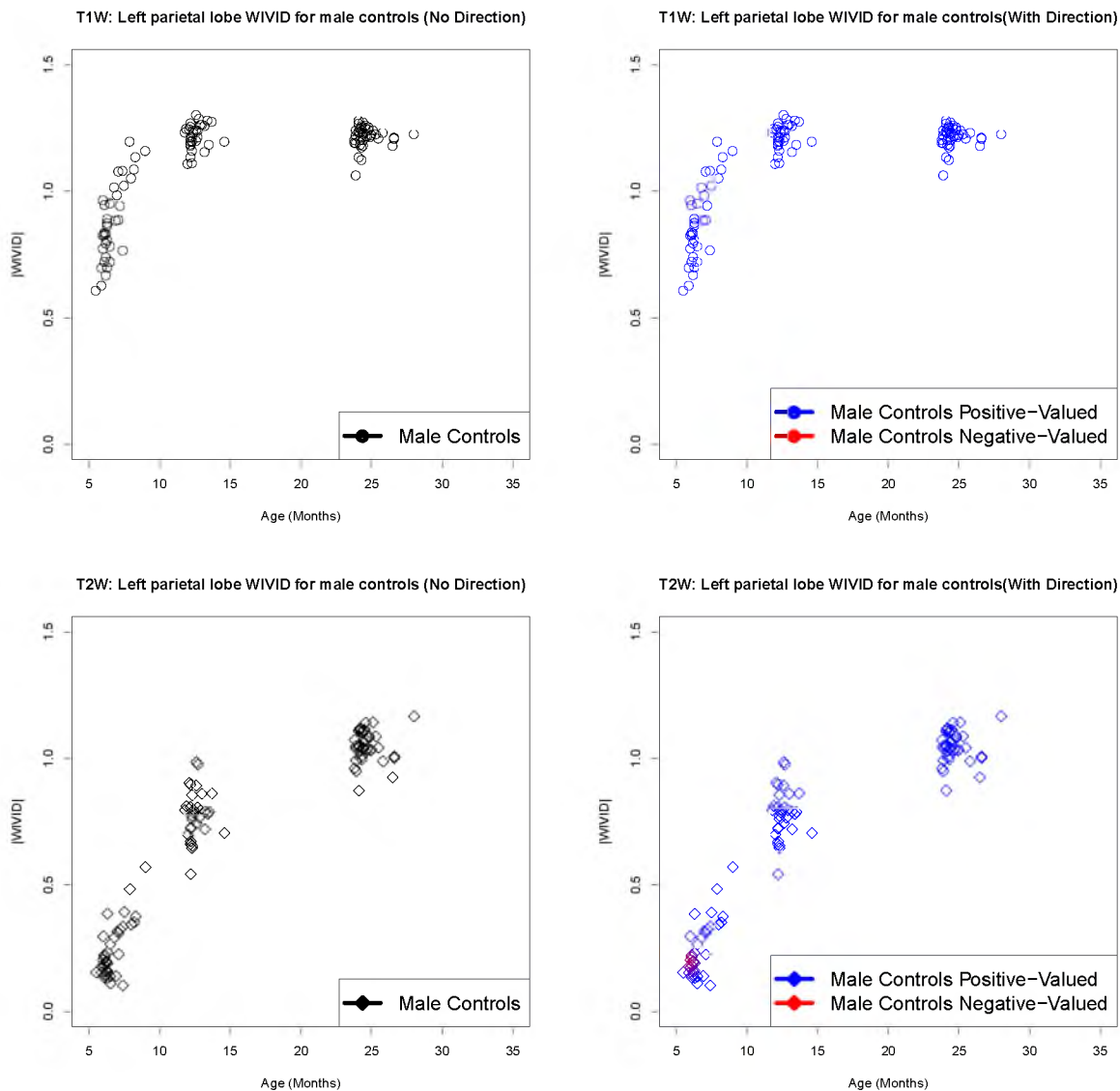


Figure 4.7: Concept of (left) contrast without direction, and (right) contrast with direction, is illustrated for the case of the left parietal lobe. WIVID values associated with (top row) T1W scans, and (bottom row) T2W scans, are displayed. Since T1W scans usually do not have negative-valued WIVID data, the positive and negative curves in the data cannot be seen in T1W scans. Although data points have directional attributes, only the absolute values of WIVID are displayed, hence ensuring all data points fall in the positive range.

In most lobes, only a small fraction of the total number of T2W WIVID contrast points at the first time point has negative direction assigned. Data from other time points (at 12 and 24 months) are entirely positively valued. As a result, the number of data points that can be used to define the negative or decreasing portion of the curve is extremely limited. Data symmetry or extrapolation assumptions will have to be made if the two-sided data is to be modeled. For the purposes of this work, we limit our analysis to only positively valued WIVID contrast points. As a result, this work models only the increasing portion of the WIVID contrast curve.

4.3.7 Experiments for Modeling ACE-IBIS Data

To summarize results from the experiments shown above, the following conclusions can be made: parametric modeling was chosen instead of kernel regression, nonlinear models provided superior data fits compared with linear models, and the logistic fit was chosen to be the best nonlinear models for fitting the data using mixed effects methods. Additionally, the logistic mixed effects fit implemented modeled asymptote and delay parameters as random effects, while the rate parameter was considered to be purely fixed effects. Extending this model to multiple modality data, the joint multivariate fit was favored for modeling multimodal data over separate univariate fits for each modality. Further, the analysis considered the variance-covariance matrix for random effects to be a full matrix without any restrictions such as zero nondiagonal values being applied. The random effects were also assumed to be homoscedastic. Comparison between different population groups was facilitated by employing a hierarchical multilevel fit to the grouped data. Finally, for the analysis performed, only positively valued WIVID contrast data points were considered.

Theoretical reasoning also supports these claims. Neurobiological processes underlying WM-GM contrast change in the infant brain are highly nonlinear in nature, justifying the use of nonlinear growth functions [58]. Cross sectional modeling of longitudinal data fails to account for the correlations between repeated measurements, affirming the need for an approach such as the mixed effects model that estimates inherent correlations in this data. Other approaches such as kernel regression that are nonparametric in nature cannot properly model this dataset since brain scans are sparsely sampled across time, i.e., they have been acquired only at three time points. Moreover, parametric nonlinear growth models provide growth parameters that describe growth effectively, are intuitively appealing, and have potential as clinical indicators. On the basis of the reasoning outlined above, parametric nonlinear growth functions implemented using a mixed effects approach would provide the ideal solution to the question of modeling WIVID contrast values for

ACE-IBIS data. The choice of multivariate data modeling is also justified, particularly since specific synchrony and patterns can be seen in WM-GM appearance changes across different modalities [29,36]. These intermodal patterns justify the inclusion of the correlation between growth parameters across modalities in our joint, multivariate model.

4.4 WIVID Contrast Analysis for ACE-IBIS Data

4.4.1 Visual Analysis of WIVID Data Points

By visual observation of WIVID data that has positive directionality, as shown in Figures 4.8 and 4.9, we can make certain qualitative conclusions. A crucial observation is that WIVID increases monotonically in a nonlinear manner for the age range under study. The left and right temporal lobes, frontal lobes, and parietal lobes look quite similar in terms of WIVID values for T1W and T2W modalities. The T1W WIVID value of these lobes ranges approximately between 0.5 and 1.3 units. The occipital lobes, which are the earliest to mature, have T1W WIVID values that are close to saturation even as early as 6 months of age. Frontal, temporal, and parietal lobes have relatively low WIVID values for T1W scans at 6 month but quickly reach saturation around 12 months of age. The occipital lobe is the earliest to mature, as indicated by WIVID values in T2W scans. The frontal lobes are the latest to mature in terms of WIVID values for T2W scans. The T2W parietal lobe WIVID value at 6 months looks higher than that of the frontal and temporal lobes in terms of growth. Maturing at a relatively slower pace compared with the parietal lobe, the T2W temporal lobe contrast still seems to undergo growth earlier compared with the frontal lobe.

In terms of intermodal variation in appearance, it could be seen that T1W WIVID values are much higher than T2W WIVID values, particularly at 6 and 12 months of age. T2W WIVID values show a clear lag behind T1W WIVID values at every time point. Whereas T2W WIVID values continue growing up to 24 months of age, T1W WIVID values begin to saturate at 12 months of age. These qualitative observations based on WIVID confirm observations found in neuroimaging literature - mainly confirming that T1W scans develop WM-GM contrast at a much earlier age compared with T2W scan, and although T1W scans predominantly undergo appearance change up to the age of 6 months, T2W scans undergo appearance change from 6 months up to 2 years of age [29,36]. In the following sections, these qualitative observations will be studied in greater detail and quantified using statistical parameters of interest. In addition to obvious interlobe and intermodality variations, other factors such as sex and risk/diagnosis group could further affect WIVID values. Since the

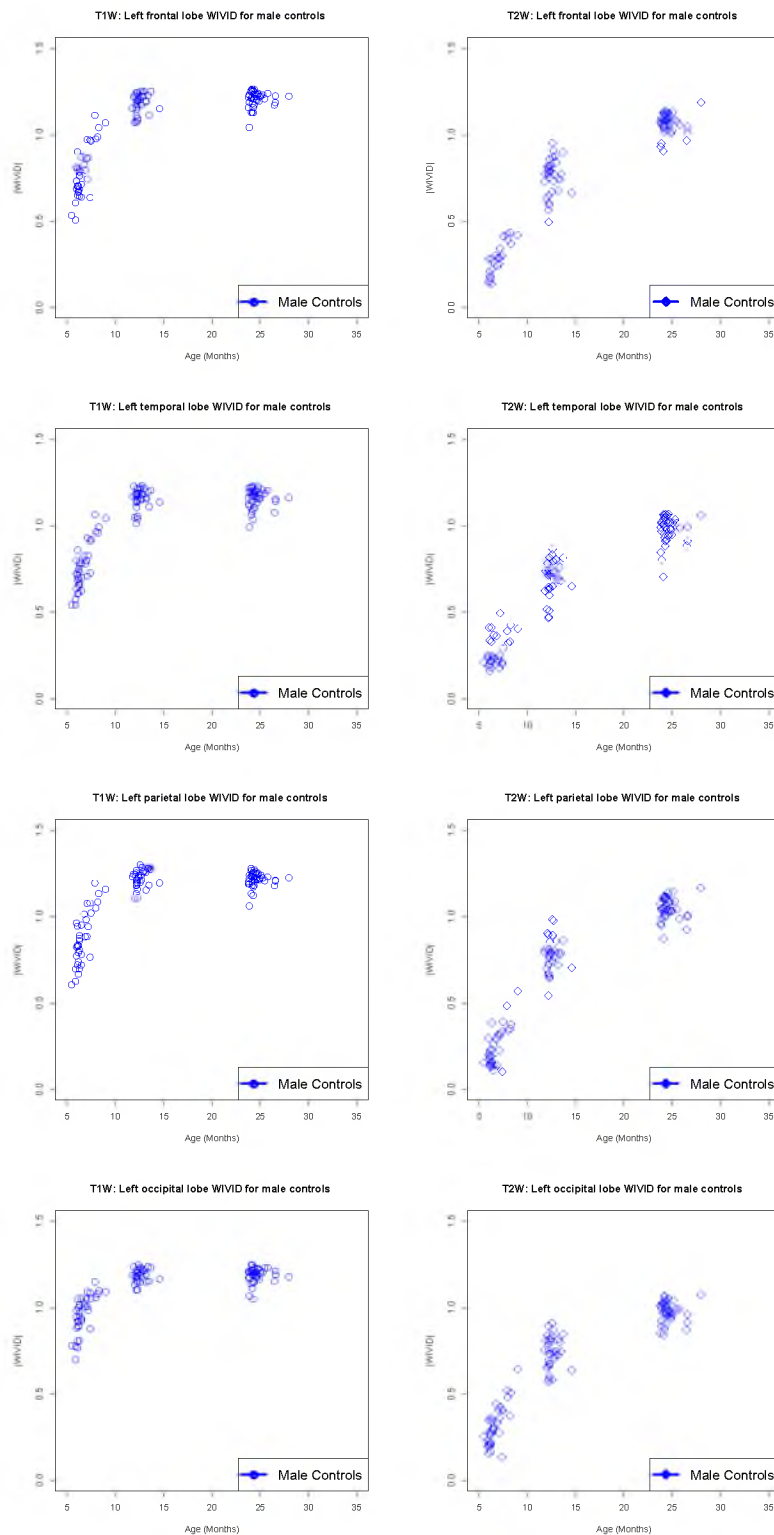


Figure 4.8: WIVID contrast values displayed for T1W (left column) and T2W (right column) scans, for major lobes in the left cortical hemisphere.

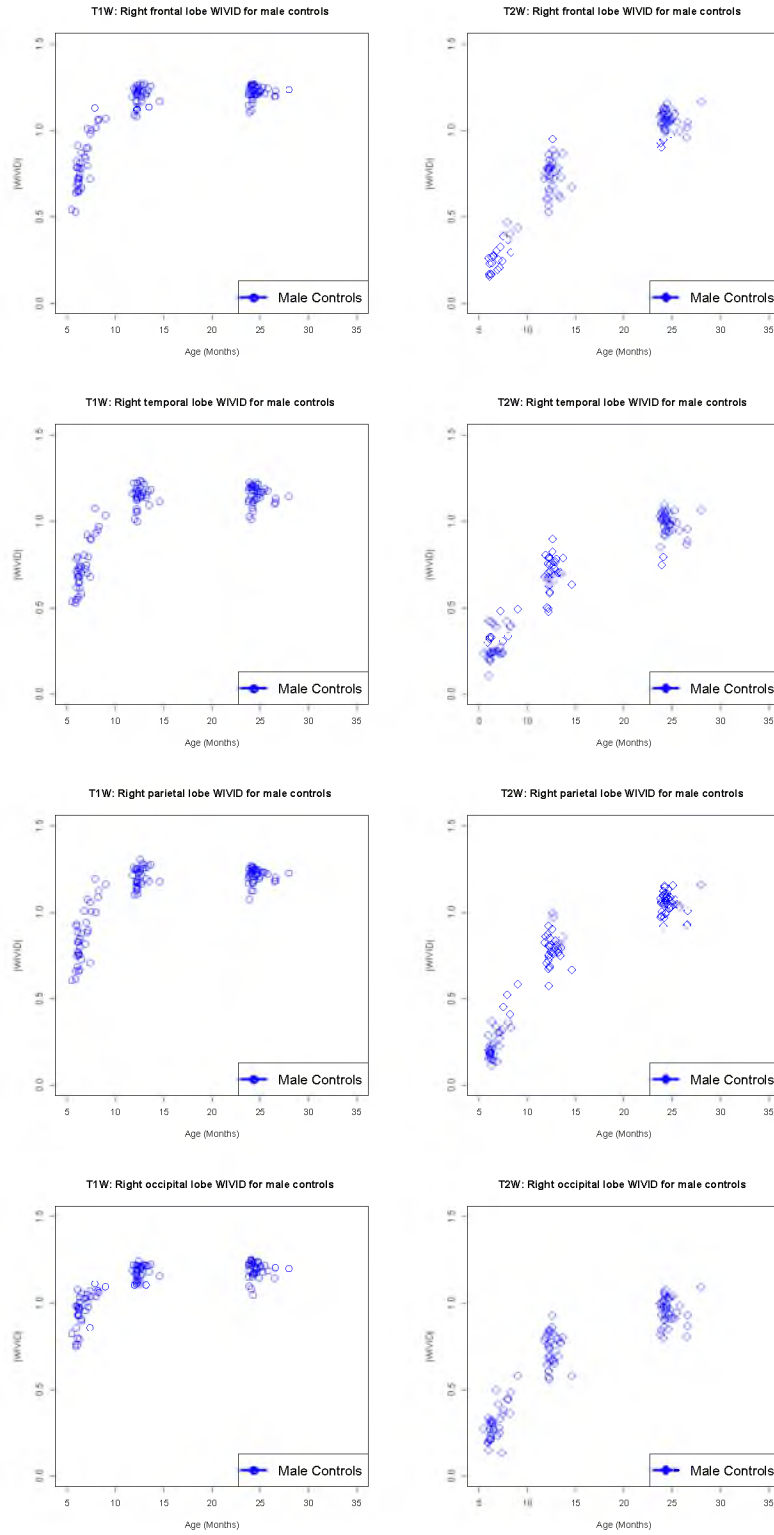


Figure 4.9: WIVID contrast values displayed for T1W (left column) and T2W (right column) scans, for major lobes in the right cortical hemisphere.

variations seen due to these factors are much more subtle, any visual-observation-based discussion would likely be inconclusive and speculative. However, data grouped by these factors are displayed and analyzed in the following sections.

4.4.2 WIVID Contrast Change: Intermodal Variations

In order to find intermodal differences in normative contrast trajectories, healthy male controls were studied. Nonlinear logistic growth functions were fit to the entire dataset of healthy controls using a hierarchical mixed effects scheme based on grouping of the data into male and female subjects. For the purpose of better understanding intermodal behavior, only scans from male controls were examined. This was mainly for the purposes of establishing normative trends, since the number of male controls was larger than the number of female subjects. In addition, later sections study the behavior of male subjects grouped by ASD risk/diagnosis, further establishing the need for study of normative trajectories of change in WIVID for male subjects in the controls group. In later sections, variations in growth parameters due to region and sex will be discussed, based on the same data.

Plots resulting from fitting logistic NLME functions to T1W and T2W WIVID contrast values can be seen in Figure 4.10. From these plots, clear differences in the growth trajectories based on modality can be seen. It can be clearly observed that T2W intertissue contrast develops much later than T1W intertissue contrast, resulting in a lag or time-shift between the two signals. While T1W contrast increases rapidly from 6 months of age and reaches saturation at around 12 to 15 months of age, T2W contrast increases at a slower pace initially and continues to vary up to 2 years of age. Essentially, MR image appearance becomes close to adult-like in terms of relative intensity of white and gray matter at around 12 months of age in T1W scans. In comparison, the intensity variation between white matter and gray matter keeps increasing until 2 years of age in T2W scans. Since myelination is known to be one of the key processes contributing to contrast in T1W and T2W images, this pattern is in conformity with the well-established knowledge that myelin appears earlier and proceeds faster on T1W images than it does on conventional T2W images [29].

Intermodal variations in growth parameters resulting from the logistic fit can be analyzed based on Table 4.6. These growth parameters are also plotted in Figure 4.11 for quick visual comparison. The observations made from the plots of T1W and T2W WIVID values changing are further confirmed by analysis of the parameter values of the fits. It is seen that the asymptote value of T1W scans is much higher than that of T2W scans, which is justified by neurobiology, since T2W intertissue contrast can change up to 3 years.

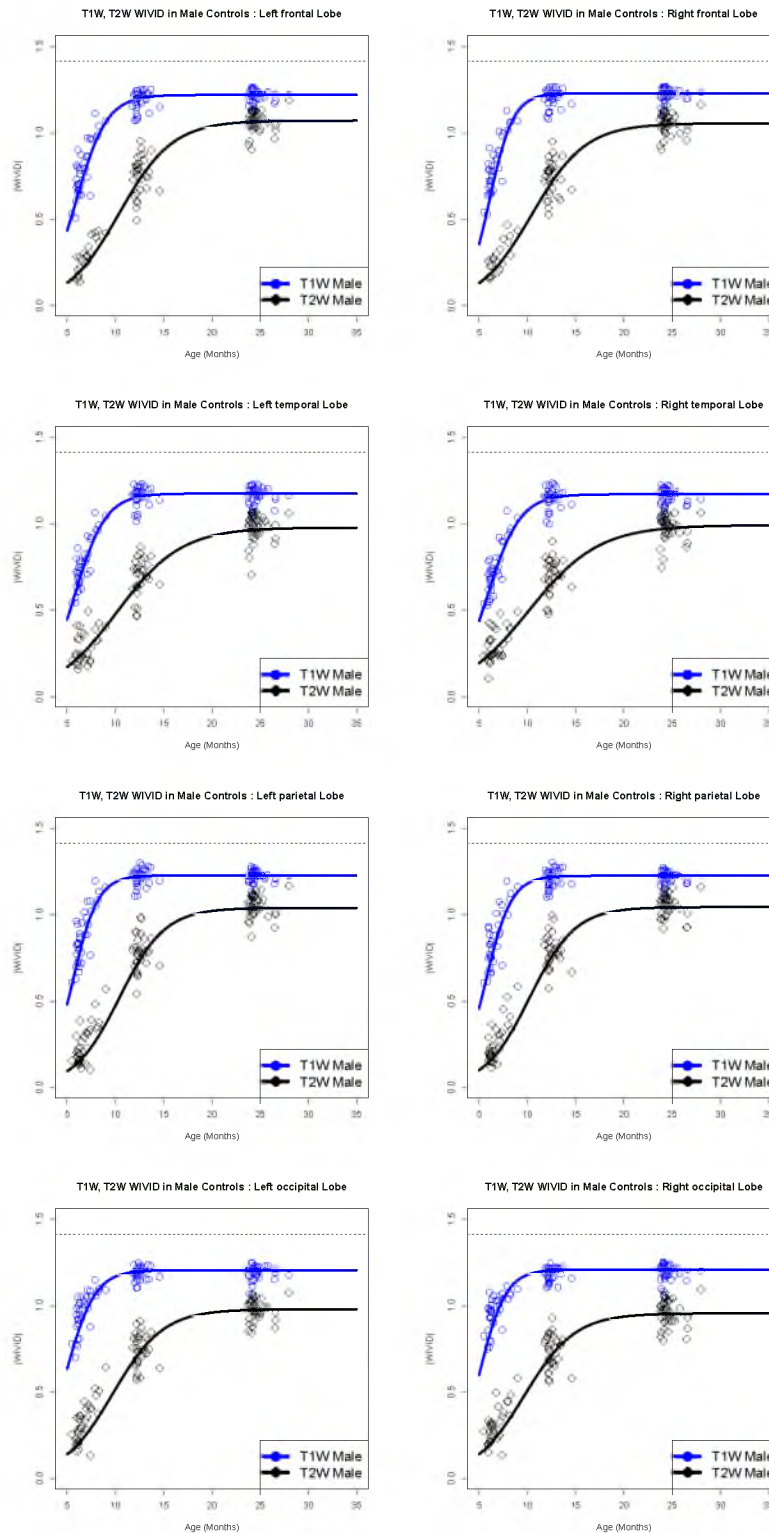


Figure 4.10: Multilevel NLME modeling of WIVID contrast change in healthy male controls: Plots of WIVID trajectories for T1W scans (blue) and T2W scans (black) population groups. The lobes studied are Frontal (first row), Temporal (second row), Parietal (third row), and Occipital (fourth row), with left hemisphere (first column) and right hemisphere (second column) regions in each of these lobes analyzed.

Table 4.6: Multilevel NLME Logistic Modeling for T1W and T2W WIVID values based on healthy male controls: Parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from logistic fit to model WIVID contrast values changing with time for T1W and T2W scans, with these scans are denoted by T1W and T2W, respectively. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp (T1)	Asymp (T2)	Delay (T1)	Delay (T2)	Rate (T1)	Rate (T2)
Fron L	1.2172	1.0693	5.8880	10.3819	1.6242	2.7084
Temp L	1.1664	0.9864	5.6558	9.7758	1.4959	3.1484
Pari L	1.2192	1.0502	5.4313	9.8531	1.1737	2.4862
Occi L	1.1917	0.9785	4.6548	9.0184	1.3153	2.9338
Fron R	1.2166	1.0603	5.6780	10.2035	1.3482	2.8156
Temp R	1.1587	0.9897	5.7826	9.5976	1.4964	3.2866
Pari R	1.2150	1.0542	5.4770	9.7989	1.2558	2.5138
Occi R	1.1932	0.9585	4.4946	9.0619	1.4483	2.7809

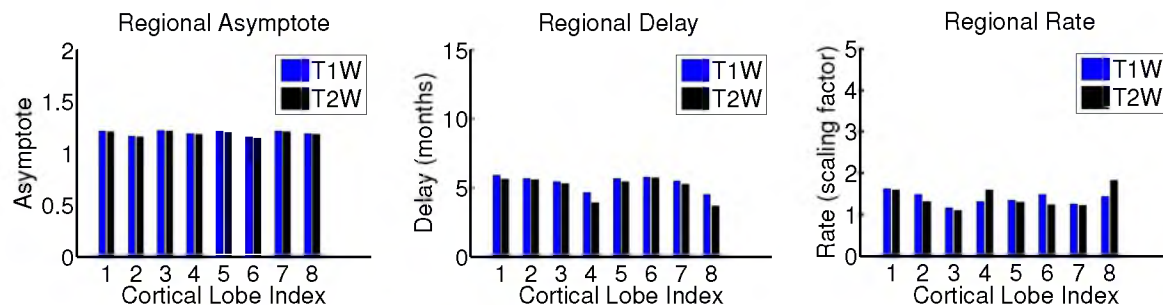


Figure 4.11: Multilevel NLME modeling of WIVID differences in T1W and T2W scans for healthy male controls: Plots of parameters asymptote (first column), delay (second column), and rate (third column) that result from fitting a logistic model to WIVID contrast values for T1W scans (blue) and T2W scans (black). The x-axis of these plots represents the indices corresponding to the lobes studied - 1 = frontal left hemisphere, 2 = temporal left hemisphere, 3 = parietal left hemisphere, 4 = occipital left hemisphere, 5 = frontal right hemisphere, 6 = temporal right hemisphere, 7 = parietal right hemisphere, 8 = occipital right Hemisphere.

Vast differences are also seen in the timing of the WIVID contrast change curves, as reflected in the delay and rate parameters. The delay parameter shows a difference of 4 to 5 months for T1W and T2W scans, which implies a correspondence to the timing difference in the maturational processes captured by each of these modalities. It is also very interesting to note that the difference in delay parameters follows a lobar pattern. The difference between delays belonging to each modality is highest in the left and right frontal lobes. The rate parameter is highly correlated with the delay parameter. On examining the rate parameter, it can be seen that the rate parameter is much higher for the T2W WIVID values compared with the T1W WIVID values. This difference in the rate reflects growth after the midpoint of the logistic curve, and indicates that while the T2W WIVID value continues increasing after this midpoint, the T1W WIVID saturates and does not increase further. Further, upon examination of the results from hypothesis testing for significant differences between modalities across all three parameters, it was concluded that *all* parameters had significant differences. These significant differences were also consistently present across *all* brain regions.

4.4.3 WIVID Contrast Modeling: Interlobe Variations

As a result of differential spatial patterns of maturation, it can be seen that intertissue contrast varies across different spatial regions [29, 36, 69]. To quantify this phenomenon, regional-specificity was incorporated as a key feature of the WIVID measure. In this section we examine interlobe variations in the WIVID contrast measure. Since healthy male controls form a sufficiently large population group in our study we study interlobe variations within their group. Lobar patterns in healthy male controls can be observed from Table 4.6.

A key finding here is that the delay parameters are consistent with established posterior-to-anterior and inferior-to-superior patterns of maturation found in neuroimaging literature [29]. Delay parameters are highest in frontal and temporal lobe and lowest in the occipital lobe by a large margin.

This evident difference in delay parameters across brain lobes can serve as indicators of heterosynchronous maturational patterns that are seen across different regions of the brain. It is also interesting to note that this difference between delay parameters is also consistent across both T1W and T2W modalities.

4.4.4 WIVID Contrast Variations Based on Sex

As a result of multilevel modeling performed by considering groups based on sex, comparisons between male and female populations could be made. The NLME logistic fits

to the WIVID data from T1W and T2W scans are shown in Figure 4.12 and Figure 4.13, respectively. Graphical comparison of the resulting growth parameters are also shown in Figure 4.14.

An interesting phenomenon is that the delay values associated with T1W scans are higher for male subjects compared with female subjects, again observed in all regions studied. These results, which suggest that male subjects experience delayed maturation compared with females, could have interesting implications in the study of sexual dimorphism in early brain development. Observing the rate parameter, it could be concluded that with the exception of the left and right occipital lobes, the growth rate as visible on T1W scans is higher for male subjects compared with female subjects. The growth rate in the logistic function approximately tracks the time taken to travel from 50 percent of the final value to 75 percent of the final value. In this case, since male subjects initially experience a delay in development, a higher growth rate could also be intrinsically correlated with this delay.

It is interesting to see that T2W WIVID contrast trajectories follow slightly different patterns of sex differences compared with trajectories associated with T1W scans. Except for the right frontal lobe, all brain regions show a higher value of asymptote for female subjects compared with male subjects in T1W scans. This observation is consistent with the hypothesis that the T2W WIVID contrast undergoes delayed growth compared with the T1W WIVID contrast; therefore, the T2W WIVID value at any point is a time-shifted version of the T1W WIVID value at an earlier point. Initially, maturation is delayed in males compared with females, a fact that is indicated by T1W delay values. Over time, the maturation of males increases to the same level as females, as reflected in T1W asymptote values. However, the asymptote value of the T2W curve corresponds to the early phase of the T1W curve, which explains the seemingly opposing observations in T1W and T2W WIVID contrast values with respect to the asymptote parameters of their logistic fits.

The delay parameter associated with T2W WIVID contrast trajectories conforms to the patterns seen in the T1W WIVID contrast curves: male subjects display higher delay compared with female subjects. This difference in delay between male and female contrast curves can be clearly observed across all brain regions without exception. In comparison, analysis of the rate parameter for male and female groups showed no clear patterns of difference.

The resulting logistic growth parameters corresponding to male and female controls are displayed in Table 4.7 and Table 4.8 for T1W and T2W contrast values, respectively. From plots of WIVID contrast change trajectories for T1W scans, a subtle difference in the

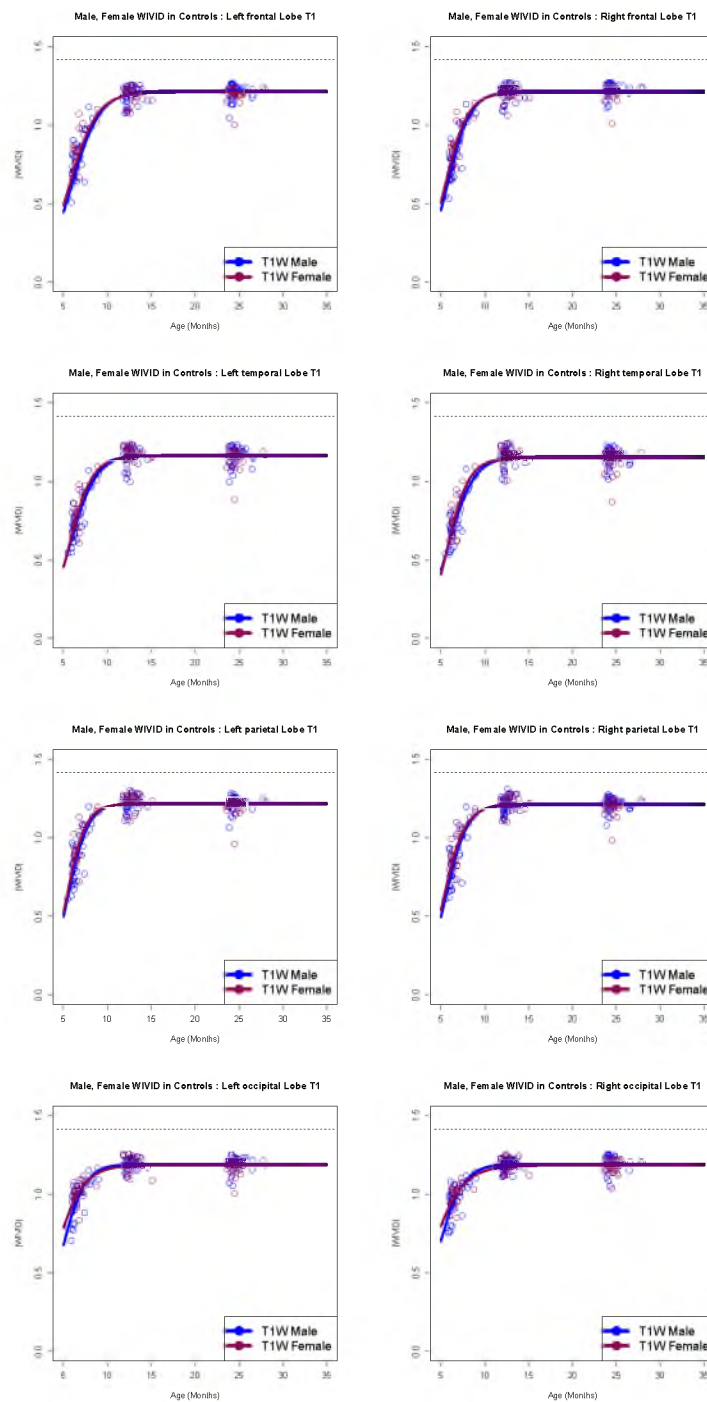


Figure 4.12: Multilevel NLME Modeling of WIVID Differences Based on Male/Female Categorization: Plots of WIVID trajectories for T1W scans are displayed for male (blue) and female (dark pink) population groups. The lobes studied are frontal (first row), temporal (second row), parietal (third row), and occipital (fourth row), with left hemisphere (first column) and right hemisphere (second column) regions in each of these lobes analyzed.

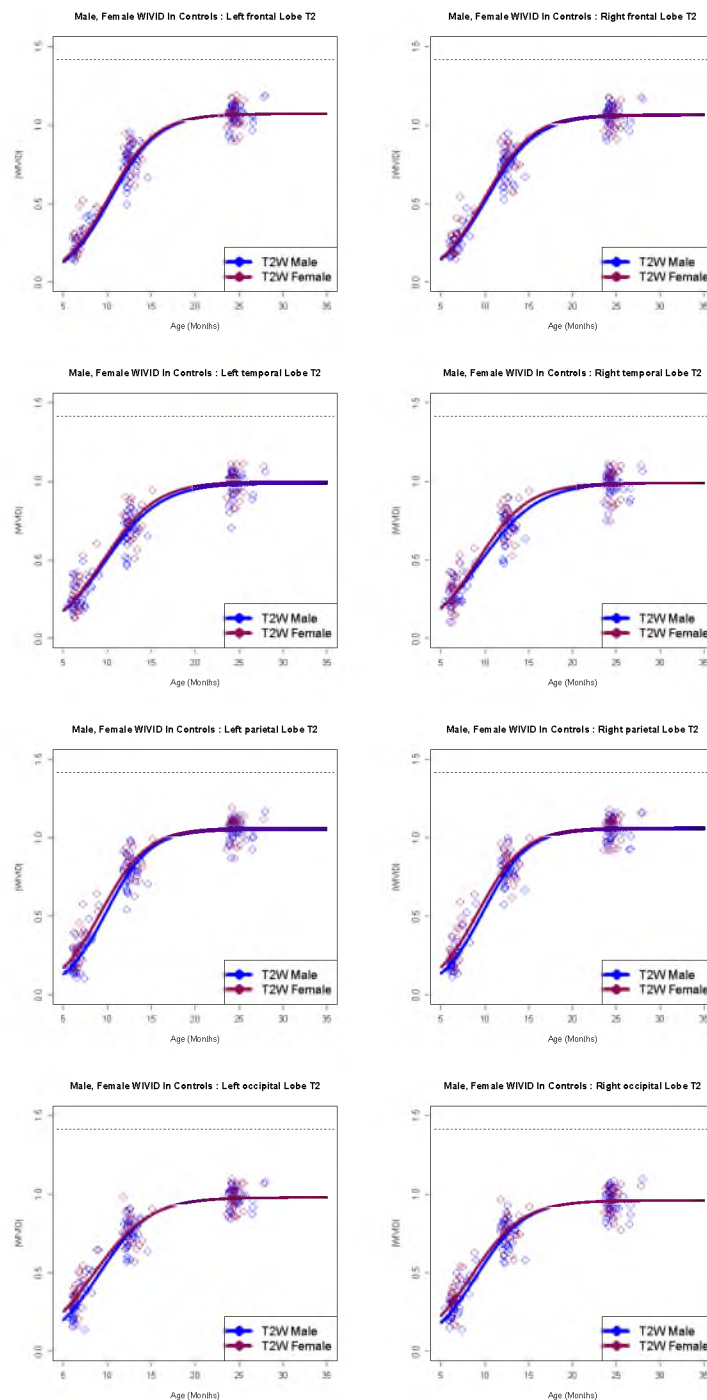


Figure 4.13: Multilevel NLME Modeling of WIVID Differences Based on Male/Female Categorization: Plots of WIVID trajectories for T2W scans are displayed for male (blue) and female (dark pink) population groups. The lobes studied are frontal (first row), temporal (second row), parietal (third row), and occipital (fourth row), with left hemisphere (first column) and right hemisphere (second column) regions in each of these lobes analyzed.

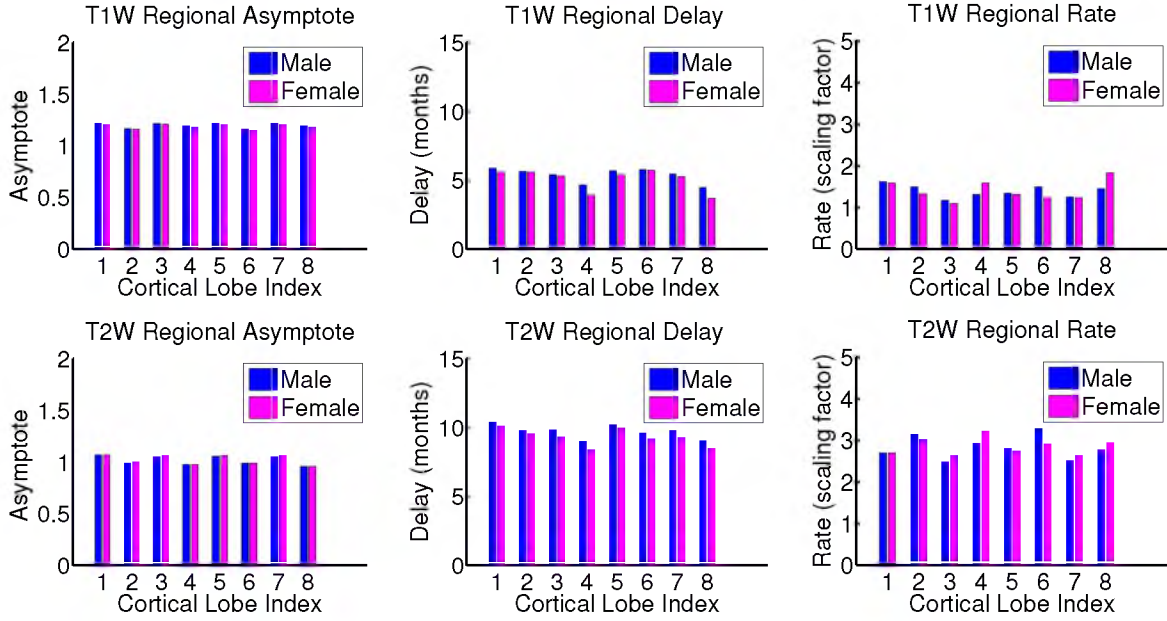


Figure 4.14: Multilevel NLME Modeling of WIVID Differences Based on Sex: Plots of parameters asymptote (first column), delay (second column), and rate (third column) that result from fitting a logistic model to WIVID contrast values for T1W scans (top row) and T2W scans (bottom row). Parameter values are displayed for male (blue) and female (dark pink) population groups. The x-axis of these plots represent the indices corresponding to the lobes studied - 1 = frontal left hemisphere, 2 = temporal left hemisphere, 3 = parietal left hemisphere, 4 = occipital left hemisphere, 5 = frontal right hemisphere, 6 = temporal right hemisphere, 7 = parietal right hemisphere, 8 = occipital right hemisphere.

asymptote between male and female subjects can be observed. This difference in asymptote can be confirmed from Table 4.7, in which it is seen that the asymptote is higher for male subjects compared with female subjects.

This difference in asymptote is consistently observed across all regions. Without correction for multiple comparisons, the P-values for T2W WIVID data show significance for parietal and occipital lobes in both hemispheres, at a significance level of 0.05. After correction for multiple comparisons, these P-values for parietal and occipital lobes are slightly above the significance-level, at around 0.06. These results from hypothesis testing indicate that maturational trajectories indeed follow different patterns in male and female subjects. It is also of interest that P-values corresponding to delay parameters are low for the parietal and occipital lobes that mature relatively early.

Table 4.7: Multilevel NLME Logistic Modeling for T1W WIVID values from male, female groups: Parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from logistic fit to model WIVID contrast values changing with time for male and female groups, with these groups are denoted by M and F respectively. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp (M)	Asymp (F)	Delay (M)	Delay (F)	Rate (M)	Rate (F)
Fron L	1.2172	1.2070	5.8880	5.6080	1.6242	1.5969
Temp L	1.1664	1.1600	5.6558	5.5864	1.4959	1.3239
Pari L	1.2192	1.2138	5.4313	5.3063	1.1737	1.1004
Occi L	1.1917	1.1830	4.6548	3.9179	1.3153	1.5969
Fron R	1.2166	1.2055	5.6780	5.4260	1.3482	1.3128
Temp R	1.1587	1.1480	5.7826	5.7403	1.4964	1.2401
Pari R	1.2150	1.2068	5.4770	5.2679	1.2558	1.2364
Occi R	1.1932	1.1826	4.4946	3.6760	1.4483	1.8299

Table 4.8: Multilevel NLME Logistic Modeling for T2W WIVID values from Male, Female groups: Parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from logistic fit to model WIVID contrast values changing with time for male and female groups , with these groups are denoted by M and F respectively. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp (M)	Asymp (F)	Delay (M)	Delay (F)	Rate (M)	Rate (F)
Fron L	1.0693	1.0702	10.3819	10.1066	2.7084	2.7080
Temp L	0.9864	0.9985	9.7758	9.5575	3.1484	3.0195
Pari L	1.0502	1.0619	9.8531	9.3146	2.4862	2.6350
Occi L	0.9785	0.9786	9.0184	8.3888	2.9338	3.2221
Fron R	1.0603	1.0654	10.2035	9.9642	2.8156	2.7532
Temp R	0.9897	0.9895	9.5976	9.1731	3.2866	2.9132
Pari R	1.0542	1.0636	9.7989	9.2930	2.5138	2.6430
Occi R	0.9585	0.9611	9.0619	8.4870	2.7809	2.9438

This could suggest that sex differences associated with brain appearance change are specific to only specific certain brain regions. The extent of sex differences might follow a pattern that is closely related to patterns of regional brain maturation. P-values resulting from hypothesis testing between male and female WIVID curves are enumerated in Table 4.9 and Table 4.10. P-values associated with T1W data are quite high and do not show significance. However, P-values corresponding to T2W data points are much lower, particularly for the delay parameter.

4.4.5 Effect of ASD Risk/Diagnosis Category on WIVID

In this section, we compare trajectories of the WIVID measure changing with time across population groups that are categorized based on their genetic risk and diagnosis for ASD. As mentioned earlier, subjects were classified on the basis of ASD risk/diagnosis into HR+ (High Risk Positive), HR- (High Risk Negative), HC (Healthy Controls), and C+ (Controls Positive). Here, we use hierarchical multilevel NLME modeling to model and compare growth trajectories between different population groups. Our focus in this section is on HR+ and HR- groups, with 44 subjects from the HR+ category and 105 subjects from the HR- category included in our analysis.

Only male subjects are included in this comparison, which is important since sex differences that were discussed in the last section could otherwise be a confounding factor in the analysis. Further, the HR+ group consists of a low ratio of only 12 female subjects to 44 male subjects, but the HR- group consists of a much better ratio of 69 female subjects to 105 male subjects. As a result, if female subjects were included in the comparison of HR+ and HR- population groups, the vast differences in ratio of number of female to male samples between these groups might pose an additional problem.

On examining the trajectories of WIVID contrast change with time shown in Figures 4.15 and 4.16, it can be seen that curves belonging to the HR+ group saturate at a lower value when compared with curves belonging to the HR- group. This difference in the saturation point, quantified by the asymptote parameter in the logistic function, is especially evident in WIVID contrast values associated with T2W scans.

The parameters of the logistic growth curves belonging to HR+ and HR- population groups are shown in Figure 4.17, and enumerated in Table 4.11 and Table 4.12, for WIVID contrast change curves obtained from T1W and T2W scans, respectively. As a result of studying the growth parameters in these tables, it can be concluded that asymptote values from T1W and T2W scans are higher for the HR- group when compared with the HR+ group.

Table 4.9: Multilevel NLME Logistic Modeling for T1W WIVID values of Male, Female groups: P values corresponding to parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from fitting WIVID contrast values with the logistic growth function. P-values that are corrected for multiple comparisons are denoted by p* (correction was done using the False Discovery Rate (FDR) method). The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp p	Delay (p)	Rate (p)	Asym (p*)	Delay (p*)	Rate (p*)
Fron L	0.3822	0.1592	0.9057	0.6205	0.4508	0.9289
Temp L	0.6435	0.705	0.4717	0.6435	0.8007	0.9289
Pari L	0.6264	0.5853	0.7243	0.6435	0.7804	0.9289
Occi L	0.3999	0.2254	0.4699	0.6205	0.4508	0.9289
Fron R	0.2996	0.1839	0.8672	0.6205	0.4508	0.9289
Temp R	0.4555	0.8007	0.2714	0.6205	0.8007	0.9289
Pari R	0.4654	0.3565	0.9289	0.6205	0.5704	0.9289
Occi R	0.3006	0.1712	0.3382	0.6205	0.4508	0.9289

Table 4.10: Multilevel NLME Logistic Modeling for T2W WIVID values of Male, Female groups: P values corresponding to parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from fitting WIVID contrast values with the logistic growth function. P-values that are corrected for multiple comparisons are denoted by p* (correction was done using the False Discovery Rate method). The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp p	Delay (p)	Rate (p)	Asym (p*)	Delay (p*)	Rate (p*)
Fron L	0.9582	0.2418	0.9977	0.9942	0.3224	0.9977
Temp L	0.5771	0.4795	0.4761	0.9942	0.4795	0.6348
Pari L	0.5067	0.0183	0.2169	0.9942	0.0618	0.4395
Occi L	0.9942	0.0172	0.102	0.9942	0.0618	0.4080
Fron R	0.7615	0.3498	0.6361	0.9942	0.3997	0.7269
Temp R	0.9942	0.1522	0.0473	0.9942	0.2435	0.3784
Pari R	0.5983	0.0338	0.2672	0.9942	0.0676	0.4395
Occi R	0.8951	0.0232	0.2747	0.9942	0.0618	0.4395

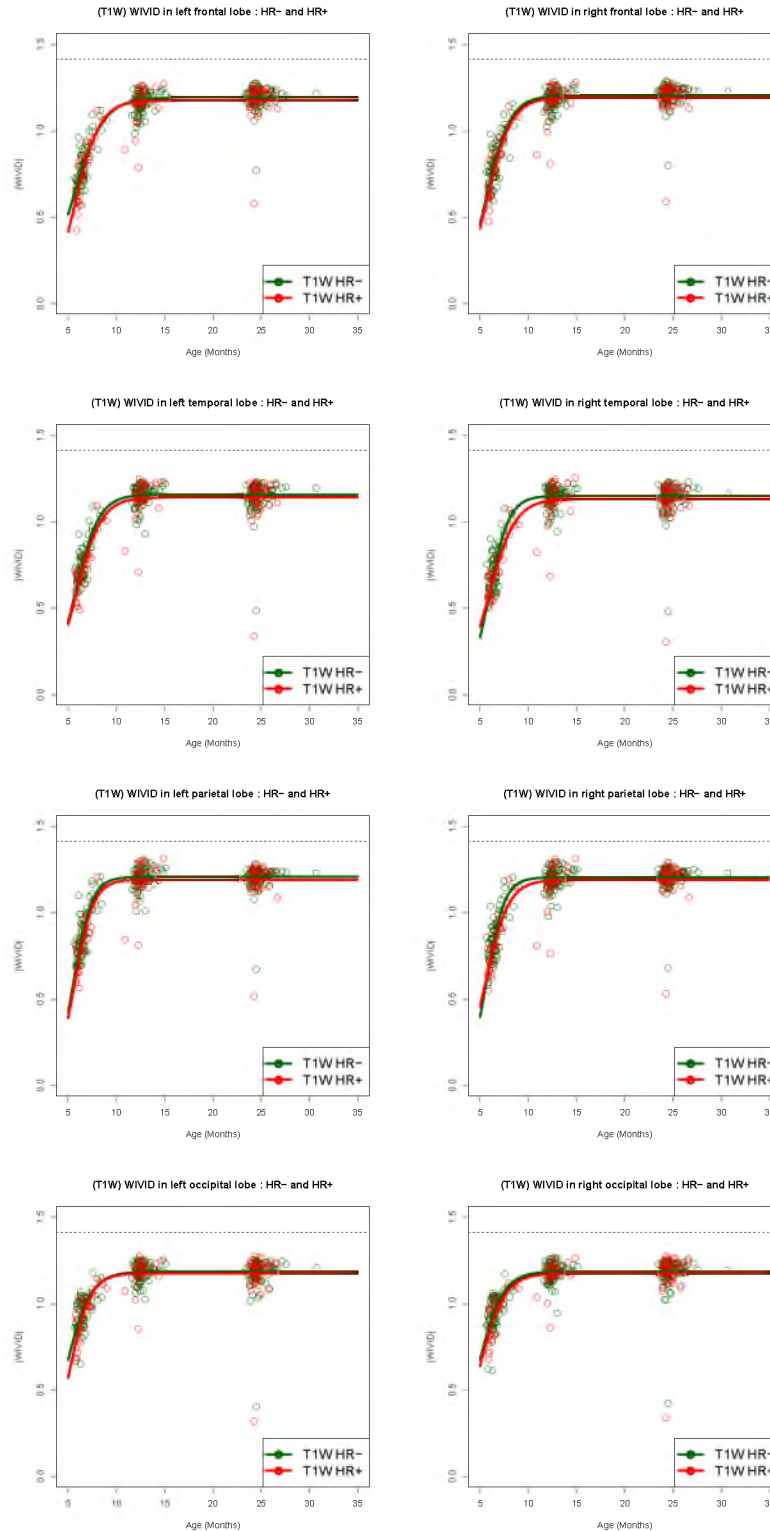


Figure 4.15: Multilevel NLME modeling of WIVID differences based on risk/diagnosis group for ASD: Plots of WIVID trajectories for T1W scans are displayed for HR+ or High Risk Positive (red) and HR- or High Risk Negative (green) groups.

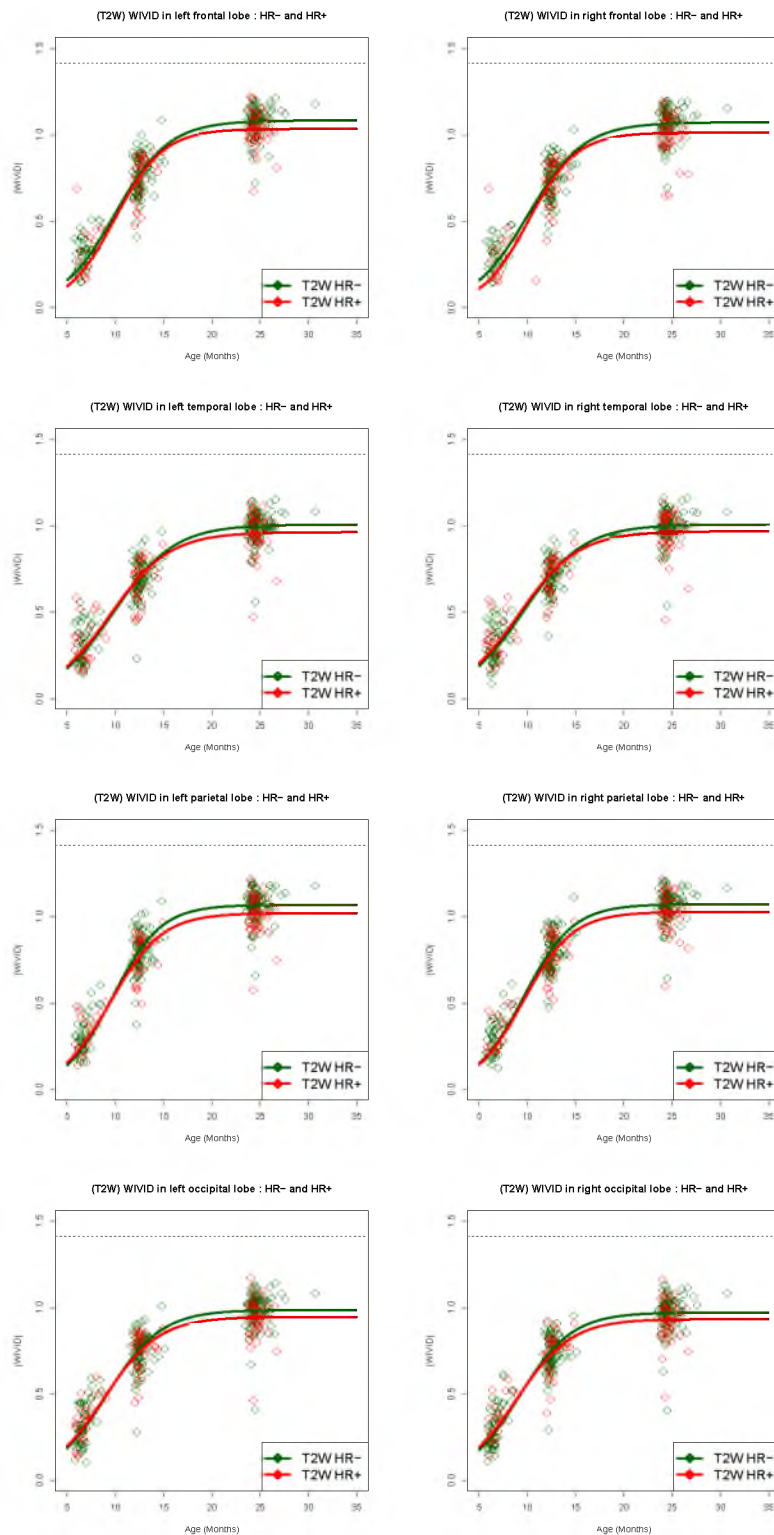


Figure 4.16: Multilevel NLME modeling of WIVID differences based on the basis of risk/diagnosis group for ASD: Plots of WIVID trajectories for T2W scans are displayed for HR+ or High Risk Positive (red) and HR- or High Risk Negative (green) groups.

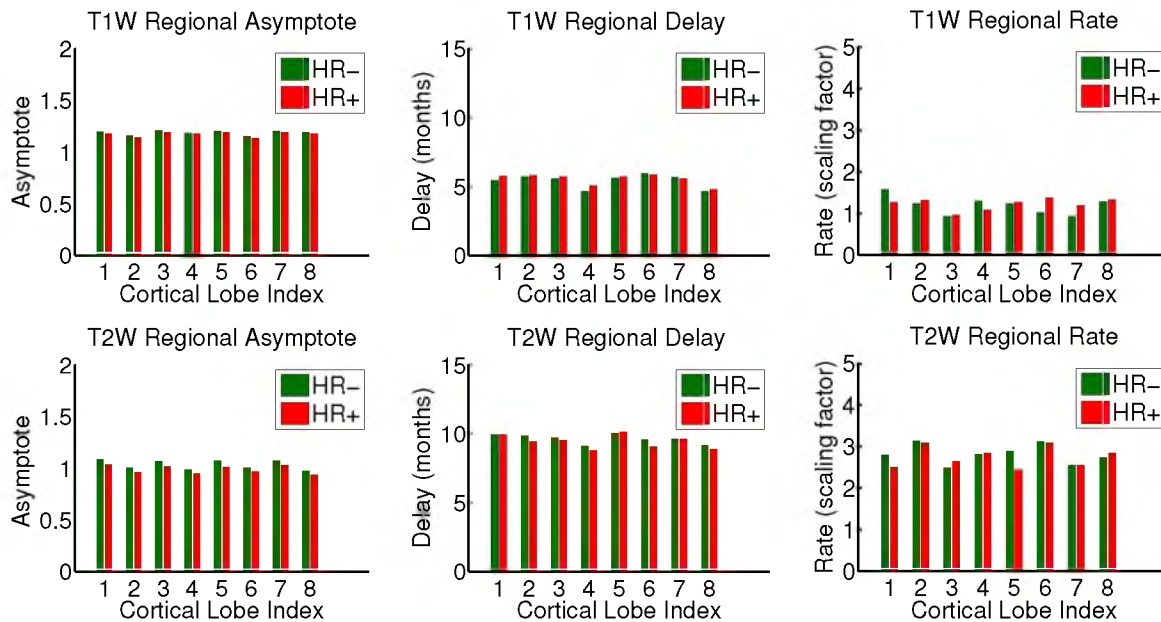


Figure 4.17: Multilevel NLME modeling of WIVID differences based on the basis of risk/diagnosis group for ASD: Plots of parameters asymptote (first column), delay (second column), and rate (third column) that result from fitting a logistic model to WIVID contrast values for T1W scans (top row) and T2W scans (bottom row). Parameter values are displayed for HR+ or High Risk Positive (red) and HR- or High Risk Negative (green) population groups. The x-axis of these plots represent the indices corresponding to the lobes studied - 1 = Frontal left hemisphere, 2 = Temporal left hemisphere, 3 = Parietal left hemisphere, 4 = Occipital left hemisphere, 5 = Frontal right hemisphere, 6 = Temporal right hemisphere, 7 = Parietal right hemisphere, 8 = Occipital Right Hemisphere.

Table 4.11: Multilevel NLME Logistic Modeling for T1W WIVID values from HR-, HR+ groups: Parameters (Asymptote - abbreviated as Asym, Del, and Rate) from logistic fit to model WIVID contrast values changing with time. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asym(HR-)	Asym(HR+)	Del(HR-)	Del(HR+)	Rate(HR-)	Rate(HR+)
Fron L	1.1936	1.1759	5.4351	5.7705	0.0187	1.5795
Temp L	1.1575	1.1413	5.6999	5.7878	0.5573	1.2419
Pari L	1.2064	1.1903	5.5791	5.6946	0.4516	0.9396
Occi L	1.1845	1.1736	4.6403	5.0636	0.2089	1.2955
Fron R	1.2032	1.1911	5.6013	5.7000	0.5119	1.2428
Temp R	1.1494	1.1316	5.9363	5.8670	0.6201	1.0294
Pari R	1.2037	1.1884	5.6670	5.5677	0.9351	1.1962
Occi R	1.1862	1.1740	4.6550	4.7707	1.2912	1.3265

Table 4.12: Multilevel NLME Logistic Modeling for T2W WIVID values from HR-, HR+ groups: Parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from logistic fit to model WIVID contrast values changing with time. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asym(HR-)	Asym(HR+)	Del(HR-)	Del(HR+)	Rate(HR-)	Rate(HR+)
Fron L	1.0821	1.0333	9.9365	9.9678	2.8066	2.4997
Temp L	1.0061	0.9623	9.8565	9.4127	3.1327	3.0848
Pari L	1.0680	1.0197	9.7032	9.5284	2.4933	2.6490
Occi L	0.9861	0.9466	9.0991	8.7971	2.8131	2.8523
Fron R	1.0711	1.0121	10.0627	10.1413	2.8858	2.4502
Temp R	1.0067	0.9672	9.5810	9.0708	3.1222	3.0844
Pari R	1.0719	1.02706	9.6308	9.6034	2.5509	2.5538
Occi R	0.9714	0.9335	9.1507	8.8702	2.7437	2.8437

Some differences were also seen in delay and rate parameters. It was interesting to note that except in the case of the frontal lobes, a delay associated with WIVID curves based on T2W scans is consistently *higher* for the HR- group when compared with the HR+ group. Upon analyzing the p-values shown in Table 4.13 and Table 4.14, it can be concluded that asymptote values associated with T2W contrast values are significant even after applying multiple comparison corrections. The right frontal lobe also showed significant differences in the rate parameter associated with WIVID curves based on T2W scans. The rate parameter for this brain region is higher for the HR- group than the HR+ group.

These findings confirm the hypothesis that WIVID contrast values could be significant biomarkers in the analysis of neurodevelopment. It is especially interesting to note that asymptote values are significant across all brain regions. The significant differences seen in the rate parameter for the right frontal lobe might also hold potential for research related to an early diagnosis of autism. The analysis shown, which finds statistical evidence for links between intertissue appearance and infant brain disorders is a major contribution of this work.

Table 4.13: Multilevel NLME Logistic Modeling for T1W WIVID values of HR-, HR+ groups: P values corresponding to parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from fitting the WIVID contrast values with the logistic growth function. P-values that are corrected for multiple comparisons are denoted by p*. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp p	Delay (p)	Rate (p)	Asym (p*)	Delay (p*)	Rate (p*)
Fron L	0.1531	0.0187	0.1027	0.4133	0.7086	0.3549
Temp L	0.2817	0.5573	0.7228	0.4133	0.7086	0.8921
Pari L	0.2131	0.4516	0.8533	0.41333	0.7086	0.8921
Occi L	0.4419	0.2089	0.4292	0.4419	0.7086	0.8584
Fron R	0.31	0.5119	0.8782	0.4133	0.7086	0.8921
Temp R	0.2428	0.6201	0.0925	0.41333	0.7086	0.3549
Pari R	0.239	0.5167	0.1331	0.4133	0.7086	0.3549
Occi R	0.3904	0.7287	0.8921	0.4419	0.7287	0.8921

Table 4.14: Multilevel NLME Logistic Modeling for T2W WIVID values of HR-, HR+ groups: P values corresponding to parameters (Asymptote - abbreviated as Asymp, Delay, and Rate) from fitting WIVID contrast values with the logistic growth function. P-values that are corrected for multiple comparisons are denoted by p*, with significant p* values highlighted in red for significance levels of 0.05. The abbreviations used to represent cortical regions are as follows: Fron = Frontal, Temp = Temporal, Pari = Parietal, and Occi = Occipital, L = Left hemisphere, and R = Right hemisphere.

Lobe	Asymp p	Delay (p)	Rate (p)	Asym (p*)	Delay (p*)	Rate (p*)
Fron L	0.0032	0.9001	0.0317	0.0128	0.9001	0.1268
Temp L	0.0173	0.1577	0.8134	0.02768	0.5856	0.9641
Pari L	0.0061	0.4393	0.2366	0.01626	0.7028	0.6309
Occi L	0.0257	0.2627	0.8224	0.0342	0.5856	0.9641
Fron R	0.0011	0.7561	0.0043	0.0088	0.9001	0.0344
Temp R	0.0328	0.0877	0.8436	0.0369	0.5856	0.9641
Pari R	0.0092	0.8996	0.9816	0.0184	0.9001	0.9816
Occi R	0.0369	0.2928	0.5639	0.0369	0.5856	0.9641

CHAPTER 5

STUDY OF BRAIN TISSUE APPEARANCE IN INFANTS

5.1 Brain Appearance in the First Year After Birth

In this chapter, we study WM-GM intensity variations in longitudinal brain MR images from birth to 1 year of age. The dataset analyzed consists of multimodal brain MR images of children scanned at five time points: approximately at 2 weeks, 3 months, 6 months, 9 months, and 12 months of age. The WIVID measure is used to quantify and analyze intensity variations for this dataset by computation of intertissue contrast at each of the five time points scanned. Images obtained during the first year after birth are of special significance in pediatric neuroimaging studies, particularly since this stage of infant brain development is a critical period of rapid growth and contrast change [29, 36, 44]. This early stage of infant neurodevelopment is also a time when the brain is highly plastic and vulnerable to external influences, and hence contains key answers to questions concerning the origins of psychiatric and neurological disorders. In addition, images from this age range can distinctly capture the process of contrast-inversion that takes place in the early brain.

Three brain tissue appearance patterns can be discerned during neurodevelopment - infantile (at birth), iso-intense, and adult-like [134]. During early brain development, as captured by neuroimaging studies, the intensity gradient between WM and GM reduces from birth up to a point of iso-intensity. After this point, a reversal of the direction of the intensity gradient takes place, while the intertissue intensity variation continues increasing from this point up to 2 years of age. This pattern is clearly seen in both T1W and T2W modality scans. For example, in T1W scans, the intensity of WM is initially much lower than that of GM, and this pattern undergoes a reversal between 3 months and 6 months of age, after which point WM is of higher intensity than GM. The reverse sequence can be observed at a later time point, between 6 months and 8 months of age, on T2W scans. Contrast inversion takes place at different time points for each subject based on the degree of tissue maturation. Hence, the point of contrast reversal could be a

useful temporal neuroimaging marker of brain appearance change, which in turn indicates progression of underlying neurodevelopmental processes. This process of contrast flip can be clearly observed in the T1W scans in Figure 5.1. In comparison, T2W scans display a lag in this process, and contrast flip takes place around 2-3 months later in these scans.

Courtesy of Dr. Weili Lin and Dr. Dinggang Shen, UNC Chapel Hill, we were given access to a unique five time-point early infant MRI dataset. The five time-point serial dataset analyzed consists of T1W (T1-Weighted) and T2W (T2-Weighted) brain MR scans. The scans were previously segmented, utilizing longitudinal techniques to enable improved tissue segmentation, particularly during the phase of iso-intensity [106, 134, 135]. Specific challenges unique to this dataset include computing accurate deformations for intrasubject image registration, especially given the low contrast and noisy quality of infant brain images.

Additionally, appropriate nonlinear methods for modeling WIVID contrast change trajectories need to be adopted to account for nonmonotonic changes in intertissue contrast. Essentially, since WIVID contrast undergoes a decrease followed by an increase, it assumes the form of a nonmonotonic function. A limited number of time points at which scans were acquired further complicates modeling of this curve. As described in Chapter 3, we adopt a nonparametric kernel-regression-based modeling scheme to characterize this curve of temporally varying WIVID.

5.1.1 Five Time Point Data Under Study

The current dataset under study consists of T1W and T2W MR data acquired at five distinct time points from 10 subjects. Two scans were discarded due to low image quality. MR images that are used were acquired using a Siemens head-only 3T Allegra scanner with a circular polarized head coil [134]. T1 images were acquired from a 3T head-only MR scanner consisting of 144 slices at a resolution of $1 \times 1 \times 1 \text{ mm}^3$, and the following parameters of scan: $\frac{TR}{TE} = 19004.38ms$, and flip angle of 7. T2 images were obtained at the resolution $1.25 \times 1.25 \times 1.95 \text{ mm}^3$, and the following parameters of scan: $\frac{TR}{TE} = 7380119ms$, and flip angle of 150. Preprocessing steps included skull stripping, bias correction, and tissue segmentation.

5.1.2 Intrasubject Image Registration

The sample image set belonging to a single subject is displayed in Figure 5.2. The data were presegmented using advanced longitudinal methods [134], hence removing the need for additional tissue segmentation. Further, it must be noted that the image series displayed in Figure 5.2 have already undergone intersubject registration. As described

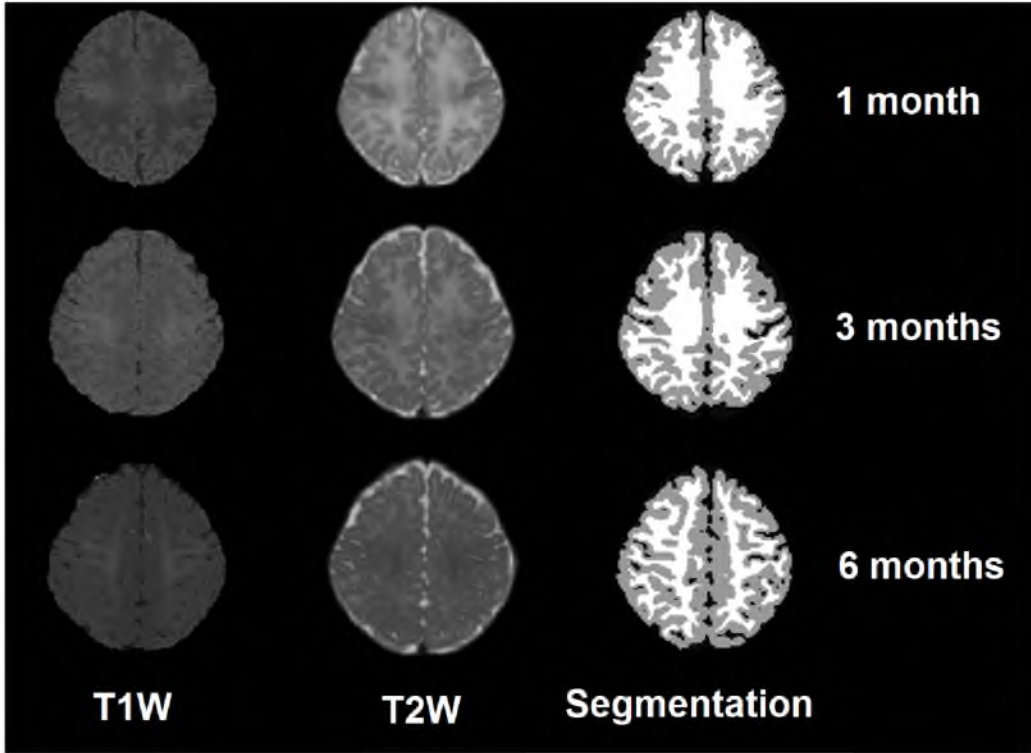


Figure 5.1: Sample scans from serial, multimodal infant brain images along with their tissue segmentations. Note: These scans are yet to be co-registered.

in depth in Chapter 2, ANTS was implemented for diffeomorphic registration of images using CC (Cross-Correlation) as the metric of choice [94]. Since five time points have been scanned for each subject, several options exist for multimodal intrasubject registration such that all scans belonging to a single subject are coregistered. Upon visual analysis of resulting registrations and quantitative comparisons using CC as the image matching metric, it was found that using both T1W and T2W scans for registration, along with segmentation results as a third channel, results in optimal registration. It was also concluded that intrasubject registration is best accomplished by nonlinear deformation of all the scans to the latest time point scans.

The primary purpose of intrasubject registration in this analysis lies in performing atlas-based parcellation of the MR brain scans. Once intrasubject registration is complete, the parcellation atlas, which delineates cortical regions of interest, can be accurately co-registered to the scans belonging to each subject, particularly in the case of low-contrast or iso-intense scans that have insufficient structure for nonlinear registration. Following parcellation, the MR scans are segmented into the major cortical lobes of interest, as displayed in Figure 5.3.

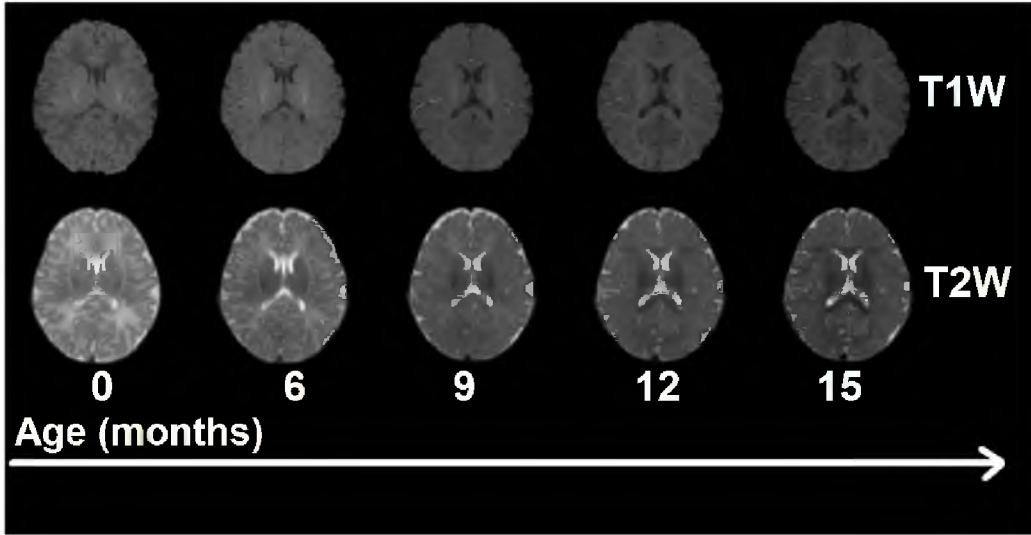


Figure 5.2: Multimodal, repeated measures dataset: Co-registered series of longitudinal images from a single subject taken at five time points.

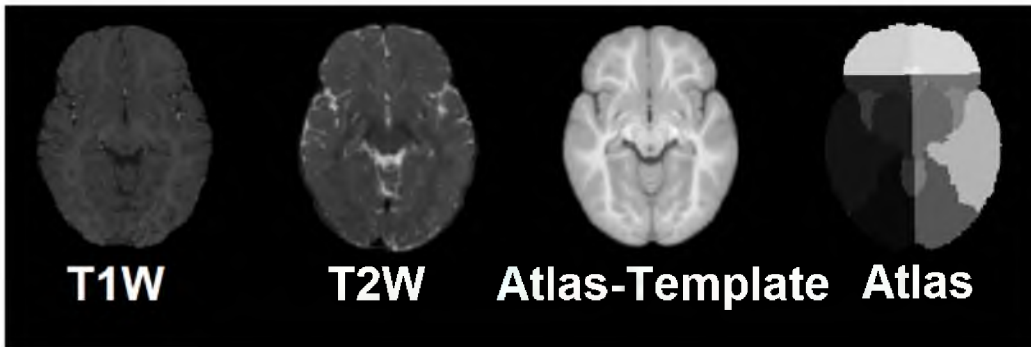


Figure 5.3: Atlas-based parcellation of multimodal data.

5.2 Computation of the WIVID Measure

Once scans are co-registered and parcellated into major anatomical regions, WIVID is computed for the five time points using the methodology delineated in Chapter 2. Modeling of the WIVID measure changing with time is challenging, primarily because the resulting WIVID function is not monotonic, hence eliminating the use of several linear and nonlinear parametric functions. To obtain a preliminary estimate of WIVID contrast change trajectories with time, we use kernel regression. The procedure for kernel regression has been outlined in Chapter 3, and basically involves estimating change in a quantity via adoption of a Gaussian-kernel-weighted smoothing procedure. Kernel regression consists of estimating a subject-specific trajectory for each subject, followed by averaging this across all subjects to generate a population-averaged trajectory. Subject-specific individual tra-

jectories displaying change in WIVID are displayed in Figure 5.4. Results from estimating a population trajectory for the right frontal lobe WIVID data for T1W and T2W scans using kernel regression are displayed in Figure 5.5.

Optimal kernel size was fixed at a standard deviation of 1.3 months following heuristic analysis. Finally, population trajectories of changing WIVID values can be compared across the major brain lobes as well. Results comparing WIVID change trajectories for the four major lobes in the right cortical hemisphere are displayed in Figure 5.6 for both T1W and T2W scans.

5.3 Results and Discussion

As a result of spatiotemporal modeling and analysis of WIVID change during the first year after birth, it is possible to discern interesting patterns of tissue appearance changes in the earliest stages of brain development. Studying these patterns has potential clinical utility, particularly since maturational patterns of the early developing brain contain answers to several important neurobiological questions [29]. Additionally, the time frame under study enables characterization of the pattern of decrease in contrast, followed by an increase after the point of contrast reversal. In Figure 5.1, this pattern is visible on the T1W MRI scans prior to quantification of appearance. Following quantification of appearance by means of WIVID computation, it is possible to map spatiotemporal changes in the WIVID values across the first year after birth. The timing and sequence of this nonmonotonic trajectory of contrast change might hold considerable value in the study of early brain maturation. It is possible that in the future, the points of contrast inversion might also be extracted following this spatiotemporal trajectory mapping.

Upon analysis of both the subject-specific trends in WIVID across modalities in Figure 5.4, and the population trends in WIVID in Figure 5.5, distinct patterns of WIVID change can be discerned in the T1W and T2W modalities. It can be clearly seen that patterns of WIVID change in T2W scans have a temporal “lag” with respect to those in T1W scans. While contrast flip can be estimated to take place at around roughly 2-5 months of age for T1W scans, the contrast flip was estimated to take place at 5-8 months of age in T2W scans. Moreover, this finding is corroborated by established literature in the area of early brain imaging [12, 29, 36], which emphasized a similar lag in the contrast reversal process between T1W and T2W modalities. It must be noted here that the exact time of contrast reversal is dependent on the pulse sequence used, and hence it cannot be compared across imaging datasets with dissimilar scanning protocols. Patterns of contrast

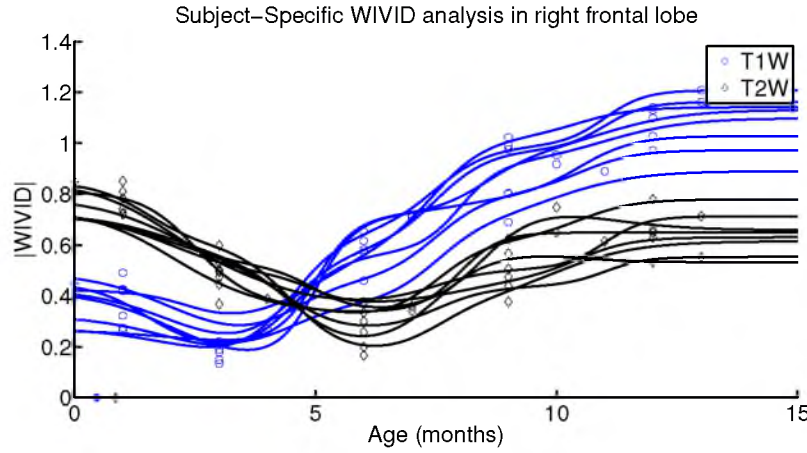


Figure 5.4: Temporal modeling of WIVID values from birth to 2 years of age in the right frontal lobe: Individual subject-specific trajectories.

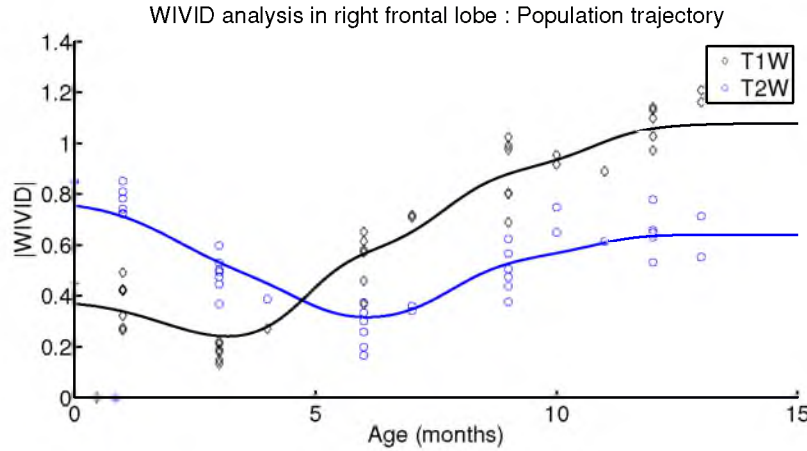


Figure 5.5: Temporal modeling of WIVID values from birth to 2 years of age in the right frontal lobe: Average population trajectory.

change in different brain regions can also be analyzed, although the limited nature of the dataset prevents drawing of clear conclusions.

In conclusion, it can be established that spatiotemporal modeling of WIVID during the first year after birth enables the study of characteristic patterns of contrast change seen in the early developing brain. The spatiotemporal trajectory of WIVID change distinctly captures three phases of brain appearance change that have been described qualitatively in previous imaging literature - contrast decrease, contrast iso-intensity followed by contrast flip or reversal, and finally increasing contrast up to adulthood. Finally, patterns of WIVID change across T1W and T2W modalities confirm qualitative findings in neuroimaging literature. It is observed that contrast reversal in T2W scans lags behind this point on

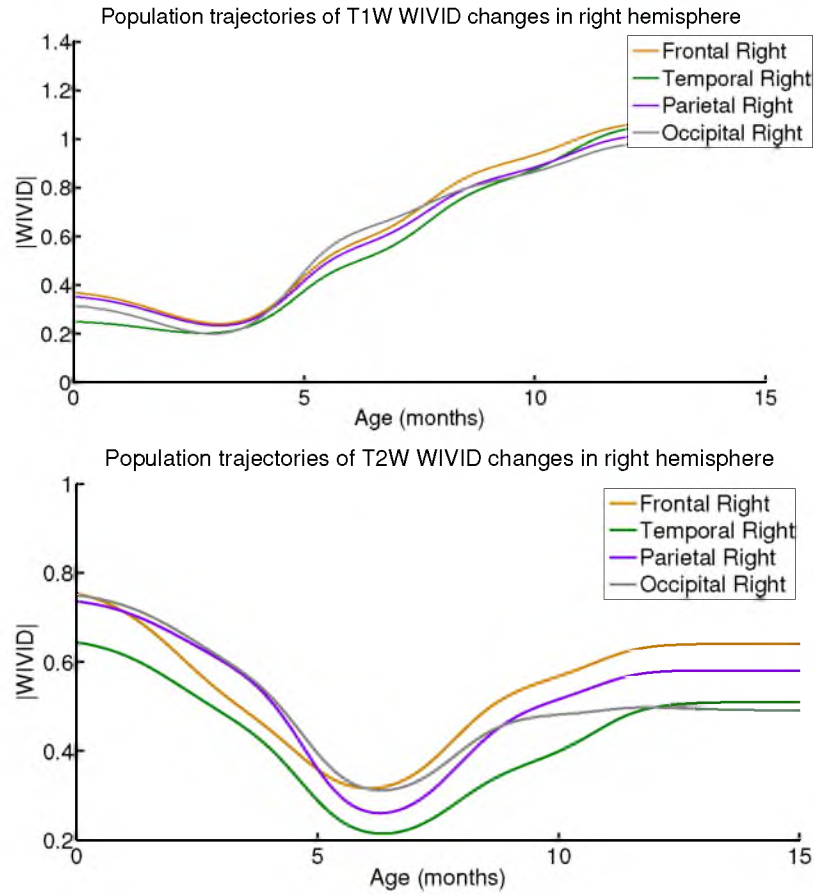


Figure 5.6: Temporal modeling of T1W (top), and T2W (bottom) WIVID values from birth to 2 years of age in the right frontal lobe : Average population trajectories across major brain lobes.

T1W scans. Further analysis of temporal biomarkers extracted from these WIVID change trajectories could be of potential utility in clinical studies of neurodevelopment. These temporal biomarkers, such as the point of contrast reversal, could provide insight into delayed trajectories of neurodevelopment in the case of subjects diagnosed with psychiatric and neurological disorders.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

6.1 Major Contributions

The rapidly changing appearance of MR scans observed during the course of neurodevelopment can be attributed in large part to the intensity variations between WM (White Matter) and GM (Gray Matter) tissue classes. A primary goal achieved by this dissertation has been the development of a method for quantification of WM-GM intensity variation during infant brain development. We refer to this neuroimaging biomarker as the WIVID (White-Gray Intensity Variation in Infant Development) measure. WIVID characterizes the intensity difference between WM and GM tissue classes by using the Hellinger Distance to measure the separation between WM and GM intensity distributions. An important feature of WIVID is its relative stability with respect to changing external conditions of scan. Given that MR images are prone to interscan variations resulting from instrument-related and subject-related factors [117, 120], the stability of any appearance measure developed becomes critical. To our knowledge, this is the only work that captures intertissue contrast in the infant brain by measuring the distance between underlying WM and GM distributions. Moreover, this computation was done in such a manner that reliance on intensity normalization techniques is reduced.

Our second contribution is a systematic implementation of statistical techniques for estimating trajectories of WIVID changing with time, resulting in a normative model of WM-GM contrast changes during infant brain development. Since the data to be modeled is longitudinal, additional challenges are posed by the repeated nature of measurements acquired from each subject. Experiments were conducted using several modeling methods - ranging from nonparametric kernel-regression-based techniques and parametric cross-sectional fits, to linear and Non-Linear Mixed Effects (NLME) methods. Following experimentation with these methods and assessment of quality of fit using diagnostic tools, the logistic growth function implemented using NLME was determined to be the best fit for modeling WIVID contrast change. As an extension to this fit, a multivariate version of

the NLME methodology was implemented, which jointly modeled WIVID contrast changes across multiple modalities. As a result of this comprehensive statistical modeling procedure, normative trends characterizing changes in tissue appearance during early development were established. Parameters of the underlying nonlinear growth model were also shown to be descriptive markers of MR image appearance properties. These parameters enabled quantification of differences observed in the normative WIVID contrast trajectories belonging to distinct modalities and brain lobes - observations that were so far conveyed mainly in qualitative terms.

Our third novel contribution is the assessment of population differences in the evolution of intertissue appearance during early brain development. Testing for population differences was implemented as an extension to the methodology implemented for statistical modeling of WIVID contrast changes with time. An analysis of differences between male and female groups led to several interesting inferences related to neurodevelopment. It was shown that brain maturation measured in terms of WM-GM contrast takes places earlier in female subjects than in male subjects. Comparisons between subjects classified as High Risk Positive (HR+) and High Risk Negative (HR-) for ASD (Autism Spectrum Disorder) were also drawn. Significant differences in asymptote parameters for T2W scans were found between HR+ and HR- groups. These significant differences in asymptote were seen across all brain lobes. Significant difference was also observed in the rate parameter in the right frontal lobe, which has been a region of special interest in several autism studies [136, 137]. These observations can provide exciting clinical insights into the origin of brain disorders. The observations concerning the rate parameter also demonstrate that early diagnosis for ASD might be a practical possibility. Observed differences between WIVID contrast change trajectories belonging to distinct populations also indicate the potential of the WIVID contrast measure as a meaningful biomarker for assessing brain development. Additionally, WIVID-based tissue appearance change was analyzed from birth to 1 year of age. WIVID analysis enabled plotting of the trajectory of WM-GM contrast change during this age span, which is crucial for neurodevelopmental studies. Studies of WIVID change for this early stage of brain maturation characterize the three distinct stages of intertissue contrast observed during infancy. These stages include a decrease in contrast following birth, reversal of direction of contrast between WM and GM, and finally an increase in contrast up to early adulthood. Clear differences in the timing and sequence of WIVID change trajectories were also observed between T1W and T2W modalities, with a lag of the T2W WIVID change curve with respect to that of the T1W WIVID curve.

6.1.1 Development of WIVID Measure

Brain maturation is comprised of several sequential biophysical and chemical processes, which manifest as changes in the relative contrast between WM and GM tissues in MR scans. Although some recent studies have analyzed intensity changes over time during early brain development [68, 69], a major limitation to progress in this field has been the lack of signal intensity calibration in MR scans. Some studies of the adult brain overcome this issue by analyzing the ratio, or difference, between signal intensity measurements from WM and GM tissues, rather than measuring raw signal intensity [76]. The ratio, or difference, between intensity values from two brain regions is a relative measure, and hence more stable compared to raw signal intensity. However, these methods involve complex operations that cannot be performed accurately on infants, such as estimation of cortical surfaces.

As a solution to this problem, we developed a technique by which the intensity distributions belonging to WM and GM tissue classes are extracted, and the intensity variation between the two tissue classes is captured by measuring the Hellinger Distance between them. This quantity extracted is referred to as WIVID, as mentioned previously. A major advantage offered by this technique is the lack of dependence on complex techniques such as intensity normalization or cortical surface extraction, which cannot be performed easily on infant brain images. A second crucial advantage of this technique is that the WIVID contrast measure is invariant to affine transformations of the underlying intensity distributions. For example, if WM and GM intensity distributions belonging to the same brain region undergo the same scaling and/or shifting transformations, the WIVID contrast measure will not change as a result. This affine invariance property stems from the corresponding property of the Hellinger distance that is used to compute the WIVID contrast measure. Other favorable properties of the WIVID contrast measure such as boundedness, symmetry, and identity, are discussed in Chapter 2. Importantly, the WIVID contrast measure can capture *shape differences between intensity distributions*. Therefore, it can capture subtle intensity variations that cannot be quantified using functions based on summary values such as the mean or median of WM and GM tissue intensities.

The WIVID contrast measure is computed in a region-specific and modality-specific manner. Region-specificity is critical to studies of the infant brain, mainly since large spatial variations in intertissue appearance are observable in these brain images. By computing WIVID on a per-region or per-lobe basis, appearance changes that are distinct for different brain regions can be analyzed. WIVID contrast is also computed separately for each modality, since T1W and T2W scans capture entirely different tissue properties [36].

Modality-specific computation of WIVID provides the opportunity to study the distinct neurobiological processes underlying appearance changes seen in each modality.

Practical computation of WM-GM intensity differences for a series of infant brain images presents several issues. High SNR and low intertissue contrast in the infant brain pose challenges to accurate segmentation and parcellation of the brain. In Chapter 2, we describe an optimal image processing pipeline that utilizes the longitudinal nature of the dataset studied to produce accurate segmentation and parcellation results. Intensity distributions of WM and GM are then extracted using nonparametric kernel density estimation techniques.

Chapter 2 demonstrates the behavior of the WIVID contrast measure for a series of phantom images, illustrating its sensitivity to appearance variations. Besides the magnitude or extent of intertissue intensity variations, a directional attribute is also attached to the WIVID contrast measure. The WIVID contrast measure undergoes first decreasing, and then increasing, trajectories of change during early brain development. The directional attribute associated with the WIVID contrast measure provides information related to the position of the current configuration of WM and GM intensity distributions on this nonmonotonic change curve. That is, the direction indicates whether the WIVID contrast measure at a particular time point belongs to the decreasing or increasing portion of the trajectory. A WIVID point on the decreasing portion of the contrast change curve is assigned a negative direction, whereas a WIVID point on the increasing portion is assigned a positive direction. Without the directional attribute, intertissue contrast values with the same WIVID magnitude, but opposite direction of WM-GM intensity gradient, would be considered identical. Given that the infant brain undergoes contrast reversal, characterization of direction is a crucial addition to complete WIVID-based contrast analysis, since it enables correct estimation of appearance change during infant brain development.

A traveling phantom study was used for analysis of stability of the WIVID contrast measure with respect to signal intensity. This study consisted of repeated scans obtained from two human phantoms across multiple scanners and multiple locations. The COV (Coefficient Of Variation) was calculated for capturing variations in WIVID contrast and signal intensity measures, *across* scans of the same subject. It was observed from this study that the COV of intensity for every region studied is higher than the corresponding COV of the WIVID contrast measure. This conclusion was valid even after intensity normalization, highlighting the superior stability of the WIVID contrast measure with respect to interscan variations, compared with the stability of signal intensity measurements.

6.1.2 Normative Trajectories of Appearance Change in Infants

In Chapters 3 and 4, a systematic evaluation of several statistical methodologies for modeling WIVID contrast across ACE-IBIS data was undertaken. Quantitative diagnostic tools such as the AIC values were employed to confirm the best model fit. A sequential procedure was adopted for assessing nonparametric and parametric models, linear and nonlinear models, and cross-sectional and nonlinear mixed-effects (NLME) models. Additionally, neurobiological factors concerning the WIVID contrast measure, which aids in the choice of model fit, were discussed. In Chapter 3, based on experiments with modeling a sample subset from the ACE-IBIS data, it was determined that the multivariate logistic growth function implemented using a nonlinear mixed-effects approach was the best choice for modeling WIVID contrast changes with time. In Chapter 4, some of these experiments were extended to the entire dataset of healthy male controls, and the choice of model was confirmed for studies of the entire population.

A significant contribution in this area is the application of joint multivariate NLME for modeling WIVID contrast from multimodal datasets. Assuming that correlation exists between the modalities being modeled, it was demonstrated by means of phantom experiments in Chapter 3 that the joint multivariate fit is superior to separate univariate fits to the data. This concept was also extended to the multimodal ACE-IBIS data in Chapter 4. It was confirmed using AIC values that joint multivariate modeling of WIVID contrast values from T1W and T2W scans results in a better fit to the data compared with individual fits for each modality. This multivariate NLME model sets up a convenient platform for analysis and comparison of appearance change trends across major cortical regions and distinct modalities.

Trajectories of WIVID contrast change, resulting from application of NLME models to normative ACE-IBIS data, were displayed in Chapter 4. Healthy male controls were used as the population subset of choice for establishing these normative models. In addition to plots of WIVID changing with time, parameters that provided a concise description of these trajectories were summarized in tables. Visualizations of these parameters were also given, enabling drawing conclusions regarding the nature of contrast change. As a result of these tools, evident patterns of normative appearance change could be quantified.

Two major conclusions could be made from the normative trends of WIVID contrast change. Firstly, significant intermodality differences were observed, reinforcing primarily qualitative observations that have been published in established neuroimaging literature [29, 36]. Appearance change trajectories between T1W and T2W scans exhibited

major differences in their timing and shape, as reflected by delay and rate parameters in their corresponding logistic models. Trajectories of WM-GM contrast in T2W scans were delayed in time compared with T1W scans. This crucial observation essentially quantifies the fact that T1W and T2W scans offer snapshots of different stages of neurodevelopment. Therefore, parameters enumerating differences in intermodal appearance change could serve as biomarkers for distinct neurobiological processes that comprise brain maturation.

The second significant conclusion that could be drawn from normative trajectories of WIVID contrast change was related to the spatially heterosynchronous nature of brain development. Pediatric brain imaging literature has consistently supported the hypothesis that brain maturation occurs differently across distinct brain regions. A sequential pattern of brain development across posterior-to-anterior and inferior-to-superior cortical regions was observed. This trend was confirmed upon analysis of parameter values from WIVID contrast change trajectories in Chapter 4. It was seen that the occipital lobe is the earliest to undergo appearance change, followed by the parietal and temporal lobes, while the frontal lobe was the last to mature. This pattern was consistent in both hemispheres, and could also be seen across T1W and T2W modalities.

In short, statistical modeling of normative WIVID trends was performed using a systematic procedure that adopted diagnostic tools to select the best fit to the data. Normative trajectories of WIVID contrast change resulted in interesting conclusions related to intermodal differences and interregion differences in brain development. Importantly, these differences followed patterns that confirmed observations made in the established neuroimaging literature. These findings further confirm the potential of the WIVID contrast measure for capturing appearance changes, and they validate the spatiotemporal modeling procedure employed as a tool for drawing inferences regarding differences between distinct WIVID contrast change trajectories.

In Chapter 5, a similar task of modeling WIVID trajectories was undertaken for an earlier age span, i.e., for subjects ranging from birth to 1 year of age. During this crucial phase of brain maturation, a trend of decreasing WM-GM contrast, followed by a period of iso-intensity or zero contrast, finally followed by a trend of contrast increasing up to 1 year, was characterized. Intermodality differences and a lag in T2W contrast change with respect to T1W contrast change could be established. This would quantify the phase shift observed in the growth characteristics of T1W and T2W scan modalities.

6.1.3 Population-wise Differences in Appearance Change

Population differences in WIVID contrast change curves were detected using a multilevel extension to the multivariate approach used for modeling multimodal data. As a result of implementing the hierarchical NLME scheme on WIVID values from ACE-IBIS data, it was possible to compare intertissue appearance change trajectories belonging to different brain regions. Crucial inferences could be made regarding the differential WIVID contrast change trajectories belonging to male and female groups, as well as between high-risk groups that were diagnosed positive (HR+) and negative (HR-) for ASD. Notable differences between male and female groups from the population of healthy controls were seen in the timing of appearance change trajectories. These timing differences, as reflected in values of the delay and rate parameters of the logistic function, indicated that female subjects undergo earlier maturation compared with male subjects. Delay parameters between male and female subjects for T2W data displayed significance before corrections were applied for multiple comparison using the FDR (False Discovery Rate) method. Although these significant differences were no longer observed after correcting for multiple comparisons (for a significance level of 0.05), the p-values associated with the delay parameters of the occipital and parietal lobes still displayed very low values (around 0.06) after correcting for multiple comparison. These values indicate that WIVID contrast change trajectories, particularly those associated with T2W scans, are indeed governed by distinct timing mechanisms based on sex. Further, the strong differences seen in the delay parameters *only* in the occipital and parietal lobes have interesting implications for the spatially dependent nature of trends in sexual dimorphism.

A second major contribution made in this direction is related to differences in contrast change trajectories between HR+ and HR- population groups diagnosed for ASD. It is of great clinical interest to note that significant differences were observed between the logistic parameters used for modeling HR+ and HR- WIVID contrast curves. The asymptote values associated with T2W scans displayed significance across *all* brain lobes that were studied. It was observed that the HR+ group, which was diagnosed with ASD, displayed lower values of asymptote compared with the HR- group. The rate parameter associated with T2W scans, measured in the right frontal lobe, also showed significant differences between the two groups. This is particularly interesting, since the frontal lobe has been linked with autism in several studies. These p-values remained significant even after correction was applied for multiple comparisons, establishing conclusively that maturational patterns associated with ASD are distinct from patterns of at-risk but healthy subjects. The significant differences

seen in the rate parameters are especially critical, since they indicate the possibility that early diagnosis of ASD could be made by analysis of WIVID contrast change.

To the best of our knowledge, these experiments are the first to study early brain development, sex differences, and differences based on ASD risk/diagnosis, in terms of WM-GM appearance changes. In addition to providing vital clinical clues regarding the origins of sexual dimorphism and brain disorders, the above experiments also demonstrate that the WIVID contrast measure is sensitive to variations in maturational trends across population groups. The potential of the WIVID contrast measure to characterize and capture trends in brain appearance is illustrated by these studies, laying the foundation for application of WIVID-based appearance analysis to other kinds of neuroimaging data.

6.2 Limitations

Several limitations faced in studies of WIVID contrast can be attributed to the nature of infant brain data. Infant brain scans are inherently prone to several image processing issues, including low tissue contrast, motion artifacts, low signal to noise ratios, large variability in size and shape, and rapid age-related changes [55]. Due to the risk of radiation exposure, infants cannot be scanned by sophisticated techniques that eliminate these issues. Infant brain datasets also consist of a limited number of temporal samples - for example, two to five scans are generally acquired from each subject over the period of birth to 2 years. Also, WIVID contrast is computed in a manner that mostly measures appearance change, but does not explore the *causes* for these changes. This gap between knowledge of effect and cause, coupled with the nonspecific nature of MR tissue contrast, could affect interpretation of results from WIVID-based studies. In the following paragraphs, we will discuss how these broad factors limit the current work.

6.2.1 Inaccuracies in Image Processing Procedures

Low tissue contrast is commonly observed in MR scans of infants along with other issues, such as the presence of noise and motion artifacts [55]. These factors adversely affect image processing procedures such as nonlinear registration, image segmentation into tissue classes, and parcellation of the brain into major cortical regions. Since these steps are critical for proper extraction of WM and GM intensity distributions and subsequent computation of WIVID, inaccuracies in these procedures could result in erroneous WIVID measurements. To make the pipeline optimal for infant brain scans, a longitudinal method was adopted by which the repeated nature of scans and consistency of brain anatomy across age are utilized for registration, segmentation, and parcellation. This pipeline is particularly effective in the

case of segmentation, where probabilistic segmentation maps from late time point images are used as priors for improving segmentation of early time points as described in Chapter 2 [106]. Further, an extension of the current method for WIVID contrast computation is outlined in Section 6.3, by which the effects of segmentation errors could be mitigated.

6.2.2 Intensity Inhomogeneities

Intensity inhomogeneities that cannot be attributed to changes in tissue characteristics can be a source of irregularities in WIVID contrast modeling. Whereas affine transformations that uniformly impacts *both* WM and GM intensity distributions from the region being studied do not affect the WIVID contrast measure, nonlinear transformations of the intensity distributions can adversely impact the results. Similarly, even affine transformations that affect WM and GM distributions from the same regions by varying degrees could result in inconsistent WIVID calculations. A primary cause for MR intensity inhomogeneity is the presence of the bias field, which can be attributed to nonideal scanning and acquisition conditions. We have sufficient evidence to believe that inhomogeneities caused by the bias field are multiplicative in nature, hence falling within the category of affine transformations. We also assume in this work that transformations such as the bias field will affect both WM and GM distributions from a single region uniformly. This assumption could be generally made due to the smoothly varying nature of the bias field, but it might not be applicable to all cases. It is not yet completely clear how the problem of a bias inhomogeneity that affects only one portion of a brain region can be solved in terms of effective WIVID contrast computation. We take an image processing approach to this problem by applying algorithms that have been proven effective in removing intensity inhomogeneities induced by the bias field to the scans studied [103]. However, this method must be approached with caution, especially since intensity inhomogeneities due to tissue characteristics might be mistaken for undesirable bias inhomogeneities by these algorithms, and they could be subsequently eliminated in the process. Intensity inhomogeneities might also be present due to reasons other than the MR bias field. Root causes of any major nonbias intensity inhomogeneities must be carefully examined, such that the adverse effects of these inhomogeneities on image analysis are clearly understood.

6.2.3 Limitations: Longitudinal Data

Several challenges in modeling the ACE-IBIS data stem from the presence of missing time points, unevenly spaced data (measurements are clustered around 6-, 12-, and 24-month time points), and imbalanced data (measurements from each subject are variable in

number). Due to the limited number of scans that are generally available in infant brain studies, appropriate statistical techniques must be used for proper modeling. For example, as discussed in Chapters 3 and 4, kernel-regression-based techniques cannot be used for data that are sparsely sampled over time. However, the NLME framework can be implemented for data that have all the limitations listed above. The robust nature of NLME modeling has been established in several studies [58, 128]. By using a hierarchical structure that estimates a population curve and individual deviations from this curve, the NLME model is relatively insensitive to noise and outliers.

In the NLME model we implemented, the random effects parameters of the nonlinear function used were assumed to follow a Gaussian distribution. In general this is a reasonable assumption to make, but it can be further verified by plotting the histogram of parameters from individual subject-specific fits if sufficiently large samples of data are available. If these histograms conform to the Gaussian distribution, verifiable also by the Chi-squared test and QQ-Plots, the assumption made in the NLME estimation will be valid. Deviations from this assumption can be accommodated by modeling non Gaussian distributions using heteroscedasticity of random effects parameters.

6.2.4 Limitations: Variable Scanning Protocols

As mentioned in Chapter 1, MR image appearance is dependent on signal measurements, which are in turn based on scanning parameters and the scan protocol used. MR scans acquired from the same subject can appear completely dissimilar if the scanning parameters and protocols used are different. Hence, it would only be reasonable to compare WIVID contrast values for scans acquired using common scanning protocols and scan parameters. Important parameters of scan include flip angle, TE (Time to Excitation), and TR (Repetition Time), and magnitude of the main magnetic field. As a result of this limitation, caution must be taken, particularly when values from studies using different scanning protocols are interpreted together.

6.2.5 Nonspecificity of MR Image Appearance

Changes in MR tissue appearance can be attributed to several causes, with precise relationships between neurobiological processes and changes in MR signal remaining somewhat unknown [138]. Not only are the T1W and T2W MR signals sensitive to myelinational processes, but they are also affected by changes to several factors including water content, compartmentalization related to changes in axonal structure, and iron content [139]. As a result, it might not be valid to always link changes in MR image appearance to increased

myelination alone [140]. This limitation can be overcome by more detailed studies of processes underlying MR signal change, with a focus on the infant age range.

6.3 Future Work

The WIVID contrast measure as specified in this dissertation can be further developed through several types of future work. Experiments that have been performed so far on the WIVID measure have resulted in demonstration of its sensitivity to maturational differences and stability with respect to interscan variations. These properties of the WIVID contrast measure, along with its potential as a clinical tool for examining population differences, reaffirm the need to carry this research forward, which could be done in several different ways, as outlined below.

The computation of the WIVID method could be modified such that the contrast measure developed is more robust to segmentation errors. Also, the current WIVID measure captures the intensity variation between tissues. An adaptation of this WIVID measure could be implemented to capture intensity variations between regions. From the point of view of behavioral analysis and structural-functional links, correlations between changing tissue appearance and behavioral scores could be investigated. From a neurobiological perspective, the biophysical and chemical processes underlying appearance change could be closely examined to establish more concrete links between neurodevelopmental processes and their manifestation as changes in MR image appearance. Further, it is even possible to use the contrast measure we define for assessing intertissue appearance in adult brain images as a biomarker to evaluate aging and progression of diseases. Finally, although the WIVID measure was developed with the view to capture intensity variations in the brain, it could also be extended to nonbrain applications. Details of all these possibilities for future work are discussed below.

6.3.1 Robustness to Segmentation Errors

In Chapter 2, the development of a longitudinal segmentation pipeline was outlined for achieving accurate brain tissue classification. However, the infant brain suffers from several issues, including lack of sufficient contrast, high noise levels, and motion artifacts, making even the best possible segmentations somewhat suboptimal. The current WIVID contrast measure is based on WM and GM intensity distributions that have been derived from WM and GM voxels that are determined by binary membership to their respective tissue classes. However, it is possible to extend this concept to the case of voxels with fuzzy membership to tissue classes. Essentially, the KDE (Kernel Density Estimation) procedure that is used for

generating tissue intensity distributions can be extended for use with voxels that have fuzzy memberships to the tissue classes being studied. For instance, to generate the WM intensity distribution, rather than using only the voxels that have been classified on a binary basis as belonging to the WM tissue class, all voxels with nonzero values of fuzzy membership to the WM tissue class can be considered. In this case, the KDE equation can be extended very easily by weighting the contribution of each voxel to the distribution. The weight for each voxel is determined by the value of its fuzzy membership to the tissue class being studied. By converting the previous dependencies of this technique on binary label maps to dependencies on fuzzy segmentation maps, the effects caused by errors in segmentation might be reduced. This method poses several challenges such as potential instability of intensity distributions generated and the bimodal nature of the resulting distributions. These issues might decrease the effectiveness of the Hellinger Distance in computing the distance between WM and GM distributions. However, by careful exploration of this idea, the dependence of the current WIVID contrast measure on accurate segmentation maps might eventually be minimized.

6.3.2 Regional WIVID Measure

The current WIVID measure has been developed to measure intensity variation between WM and GM tissue classes. While extending this to any other tissue classes of interest is straightforward, it might be more complex to analyze intensity differences between distinct *regions* using this method. Measurement of region contrast differences was explored in the early stages of this work [141]. In several neurological conditions of interest, ranging from Gadolinium-accumulation to Alzheimer’s Disease, interesting results were found upon comparing signal intensity measures from different regions of the brain [142]. By measuring the overlap in intensity distributions belonging to the two regions of interest, some subtle discoveries can be made that might be outside the potential of straightforward techniques, such as ratios of mean intensities of the two regions or cross-correlation between their overall distributions. Contrast change trajectories in two or more regions can also be analyzed jointly in terms of relative growth rate and other similar parameters.

6.3.3 Brain-Behavior Studies using WIVID

The ACE-IBIS and other infant brain imaging datasets are also rich repositories of complex behavioral information. For instance, several behavioral scores are tracked in the ACE-IBIS study, including gaze-tracking indices, scores on learning tests, and measurements of social skills [133]. Since intertissue appearance changes are biomarkers that indicate

underlying brain patterns of brain maturation, it would be interesting to investigate links between WIVID contrast measures and associated subject-specific behavioral scores. A primary process linked with WM-GM contrast change is myelination. Myelination of a brain region, in turn, is associated with the functional maturity of that region. As a result, the links between WM-GM contrast change and functional or behavioral scores are of considerable scientific potential. Since the WIVID contrast measures we study are region-specific, such an investigation could deepen our insight into relationships between brain structure and function.

6.3.4 Neurobiology of WM-GM Contrast

Although histological and neurological studies have provided proof for the concept that neurodevelopmental processes result in changes in WM-GM contrast, precise relationships to describe these connections have not been found. For example, it is hypothesized that premyelination stages and tightening of the myelin spiral around the axon contribute to the appearance changes seen in T1W and T2W scans. However, precise observations of changes in MR relaxation parameters and signal with the development of *each* of these processes have not been made. If the neurobiological underpinnings related to specific phases of MR signal intensity change could be established, further clues would result as to the precise origin of brain disorders.

6.3.5 WIVID Across Multiple Scanning Protocols

It is a common fact that MR signal intensity and intertissue appearance are heavily dependent on the scanning protocols used for image generation [29]. It was noted that the timing sequences associated with appearance change trajectories from MR scans using vastly dissimilar scanning protocols might be completely different [29]. This is reasonable, primarily since the nature of the measured MR signal and the tissue properties that are captured are entirely dependent on the scanning protocol used, as described in Chapter 1. It would hence be useful if the WIVID contrast measure could be explored based on MR scans generated using different scanning protocols. If conclusive results could be extracted by using several different scanning protocols, eventually it might be possible to compile an entire database of tissue appearance across different types of MR scans, which might also facilitate interstudy analysis. That is, brain imaging studies that used different scanning protocols could be examined in a combined manner.

6.3.6 WIVID for Other Neurobiological Processes

In this work, we have restricted our usage of WIVID contrast to infant brain imaging studies. However, WM-GM contrast also plays a crucial role in other neurobiological processes such as aging and progression of diseases such as MS (Multiple Sclerosis) [76, 79, 143, 144]. Application of the WIVID measure to studying these processes could be extremely valuable, particularly since it might be sensitive to variations that cannot be fully captured using summary measures such as the mean intensity of a region. WIVID contrast in these studies might also not face the same limitations as infant studies - for example, segmentation and parcellation of the adult brain would be superior to that of the infant. However, other challenges in these datasets, such as the presence of plaques in aging brain studies and lesions in MS studies, need to be overcome. In spite of these limitations, it might be worthwhile to explore such datasets, especially given the demonstrated potential of the WIVID measure for quantifying appearance changes.

6.3.7 Application of WIVID to Nonbrain Data

Besides neuroimaging studies, several other types of medical imaging studies utilize the concept of contrast. For example, contrast is used in LGE (Late Gadolinium Enhanced) imaging, in which a contrast agent is injected for enhancing intensity differences between tissues. Intertissue contrast is often measured on these datasets by computing a simplistic difference of mean intensities of the two regions studied. Other studies in which the concept of contrast and intensity is directly or indirectly used include detection of tumors [145], examining appearance of the myocardium [146], and studies of the uterus [147]. The work we discussed in this dissertation can be applied to other types of medical imaging studies outside of neuroimaging, by extension of the current concept of WM-GM contrast as measured by WIVID to contrast between any two tissues of interest. Although some of these modalities such as CT might not be affected by lack of calibration, the concept of using the Hellinger Distance between distributions to capture contrast might still be relevant. Finally, contrast is often measured visually to confirm optimal settings of scanner parameters with the final goal of properly capturing tissue structures in modalities such as MR. That is, the parameters of scan acquisition might be varied, and resulting images will be studied until optimal tissue structure is visible. A possible application of the WIVID contrast measure would be to quantify this process in such a way that setting parameters of scan becomes accurate and standardized.

APPENDIX

INTENSITY NORMALIZATION FOR INFANT MRI

Intensity normalization is defined as correction of intrainage signal inhomogeneities as well as standardization of image intensities to a required range [35]. A variety of sophisticated intensity normalization techniques are available, many of which have been established as extremely effective in the normalization of adult brain images [148]. These techniques essentially transform MR image intensities such that the image intensities for a certain tissue class across scans of different subjects are “normalized”, or made to occupy a certain standard intensity range.

In the infant age-range, intensity normalization techniques face several challenges, including inherent variability of signal intensity due to developmental processes and lower signal to noise ratios. Difficulty discerning intrinsic intensity differences due to neurodevelopment from those differences caused by variable scanning conditions is a great challenge to intensity normalization for infant datasets. In addition, the tissue intensity of different regions is often vastly different in the infant brain of a single subject. This could be attributed to the fact that each cortical region might be at a different stage of maturation and hence might have differential tissue characteristics, contributing to natural variations in tissue intensities, particularly for WM.

Therefore, intensity normalization in infant datasets consists of a relatively simple approach. This approach consists of division of intensities throughout the brain uniformly by a single constant. The constant factor for division is determined by the intensity distribution of regions of fatty tissue and ventricular CSF in T1W and T2W images, respectively [67]. The normalization constant is computed as the mean, median, or a percentile (e.g. 90th percentile) of the intensity distributions of these regions. We now mathematically define normalization in the case of the constant factor being computed using the mean operation. Considering the original T1W image $ImT1$, the normalized version of this image $ImT1_{normalized}$, fatty tissue region consisting of total of $N_{FattyTissue}$

points with each such point denoted by $x \in FattyTissue$, normalization can be expressed mathematically for T1 images as

$$Int(ImT1_{normalized}) = \frac{Int(ImT1)}{\left(\frac{\sum_{x \in FattyTissue} ImT1(x)}{N_{FattyTissue}}\right)} \quad (A.1)$$

Similarly, for the T2W image $ImT2$, with ventricular CSF regions analogous to the fatty tissue regions defined for T1W images above, the corresponding normalized image $ImT2_{normalized}$ can be computed as

$$Int(ImT2_{normalized}) = \frac{Int(ImT2)}{\left(\frac{\sum_{x \in VentricularCSF} ImT2(x)}{N_{VentricularCSF}}\right)} \quad (A.2)$$

Since these regions (fatty tissue in T1W images and ventricular CSF in T2W images) are assumed to have a constant intensity throughout stages of neurodevelopment, by using the mean or median intensity of these regions as a normalizing factor, the intensity ranges of individual scans can be brought to occupy a standard intensity range.

REFERENCES

- [1] J. L. Prince and J. M. Links, *Medical Imaging Signals and Systems*. Upper Saddle River, N.J. : WB Saunders Co, 2006.
- [2] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential." *Comput Med Imaging Graph*, vol. 31, no. 4-5, pp. 198–211, 2007.
- [3] C. Jack, Jr, R. C. Petersen, P. C. O'Brien, and E. G. Tangalos, "Mr-based hippocampal volumetry in the diagnosis of alzheimer's disease," *Neurology*, vol. 42, no. 1, pp. 183–188, Jan, 1992.
- [4] M. A. Walter, R. A. Melzer, C. Schindler, J. Müller-Brand, A. Tyndall, and E. U. Nitzsche, "The value of [18f] fdg-pet in the diagnosis of large-vessel vasculitis and the assessment of activity and extent of disease," *EJNMMI*, vol. 32, no. 6, pp. 674–681, 2005.
- [5] J. Meller, F. Strutz, U. Siefker, A. Scheel, C. Sahlmann, K. Lehmann, M. Conrad, and R. Vosschenrich, "Early diagnosis and follow-up of aortitis with [18f] fdg pet and mri," *EJNMMI*, vol. 30, no. 5, pp. 730–736, 2003.
- [6] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Medical Imaging 2001*. International Society for Optics and Photonics, 2001, pp. 236–248.
- [7] "Neuroimaging: Visualizing brain structure and function," <http://ccnmtl.columbia.edu/projects/neuroethics/module1/foundationtext/>.
- [8] A. G. Filler, "The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: Ct, mri, and dti," *Nature Proceedings*, vol. 7, no. 1, pp. 1–69, 2009.
- [9] "Introduction to neuroimaging," <http://www.neuroradiologija.com/>.
- [10] M. T. Abou-Saleh, "Neuroimaging in psychiatry: An update." *J. Psychosom. Res.*, vol. 61, no. 3, pp. 289–293, Sep 2006.
- [11] P. I. Yakovlev and A. R. Lecours, "The myelogenetic cycles of regional maturation of the brain," *Regional Development of the Brain in Early Life*, pp. 3–70, 1967.
- [12] M. Van der Knaap and J. Valk, "Mr imaging of the various stages of normal myelination during the first year of life," *Neuroradiology*, vol. 31, no. 6, pp. 459–470, 1990.
- [13] P. R. Huttenlocher and A. S. Dabholkar, "Regional differences in synaptogenesis in human cerebral cortex," *JCN*, vol. 387, no. 2, pp. 167–178, 1997.

- [14] L. A. Glantz, J. H. Gilmore, R. M. Hamer, J. A. Lieberman, and L. F. Jarskog, "Synaptophysin and postsynaptic density protein 95 in the human prefrontal cortex from mid-gestation into early adulthood," *Neuroscience*, vol. 149, no. 3, pp. 582–591, 2007.
- [15] D. G. Nishimura, *Principles of magnetic resonance imaging*. Stanford University, 1996.
- [16] "Acrasnrsp practice parameter for the performance and interpretation of magnetic resonance imaging (mri) of the brain," <http://www.acr.org/ /media/E712C7C597AE44D8B1765D5ED5086086.pdf>, 2014.
- [17] N. Sadeghi, "Modeling and analysis of longitudinal multimodal magnetic resonance imaging: Application to early brain development," Ph.D. dissertation, Department of Bioengineering, University of Utah, December 2013.
- [18] W.-D. Heiss, "The potential of pet/mr for brain imaging," *EJNMMI*, vol. 36, no. 1, pp. 105–112, 2009.
- [19] G. Malinger, L. Ben-Sira, D. Lev, Z. Ben-Aroya, D. Kidron, and T. Lerman-Sagie, "Fetal brain imaging: a comparison between magnetic resonance imaging and dedicated neurosonography," *UOG*, vol. 23, no. 4, pp. 333–340, 2004.
- [20] R. Leblanc, A. Olivier, R. Pokrupa, and D. L. Arnold, "Accurate, noninvasive diagnosis of human brain tumors by using proton magnetic resonance spectroscopy," *Nat. Med.*, vol. 2, no. 3, 1996.
- [21] F. Howe, S. Barton, S. Cudlip, M. Stubbs, D. Saunders, M. Murphy, P. Wilkins, K. Opstad, V. Doyle, M. McLean *et al.*, "Metabolic profiles of human brain tumors using quantitative in vivo 1h magnetic resonance spectroscopy," *MRM*, vol. 49, no. 2, pp. 223–232, 2003.
- [22] P. J. Modrego, N. Fayed, and M. A. Pina, "Conversion from mild cognitive impairment to probable alzheimers disease predicted by brain magnetic resonance spectroscopy," *AJP*, 2014.
- [23] A. Villringer, J. Planck, C. Hock, L. Schleinkofer, and U. Dirnagl, "Near infrared spectroscopy (nirs): a new tool to study hemodynamic changes during activation of brain function in human adults," *Neurosci. Lett.*, vol. 154, no. 1, pp. 101–104, 1993.
- [24] A. M. Dale and E. Halgren, "Spatiotemporal mapping of brain activity by integration of multiple imaging modalities," *Curr. Opin. Neurobiol.*, vol. 11, no. 2, pp. 202–208, 2001.
- [25] F. W. Wehrli, J. R. MacFall, D. Shutts, R. Breger, and R. J. Herfkens, "Mechanisms of contrast in nmr imaging," *Jcat*, vol. 8, no. 3, pp. 369–380, 1984.
- [26] W. Nitz and P. Reimer, "Contrast mechanisms in mr imaging," *ER*, vol. 9, no. 6, pp. 1032–1046, 1999.
- [27] P. Baierl, C. Förster, H. Fendel, M. Naegele, U. Fink, and W. Kenn, "Magnetic resonance imaging of normal and pathological white matter maturation," *Pediatric radiology*, vol. 18, no. 3, pp. 183–189, 1988.

- [28] D. B. Plewes, "The aapm/rsna physics tutorial for residents. contrast mechanisms in spin-echo mr imaging," *Radiographics*, vol. 14, no. 6, pp. 1389–1404, 1994.
- [29] M. Rutherford, *MRI of the Neonatal Brain*. WB Saunders Co, 2002.
- [30] T. HE, "Normal maturation of the neonatal and infant brain: Mr imaging at 1.5 t," *Radiology*, vol. 166, pp. 173–180, 1988.
- [31] P. I. Yakovlev and A.-R. Lecours, "The myelogenetic cycles of regional maturation of the brain," *Regional Development of the Brain in Early Life*, pp. 3–70, 1967.
- [32] B. Holland, D. Haas, D. Norman, M. Brant-Zawadzki, and T. Newton, "Mri of normal brain maturation," *AJNR*, vol. 7, no. 2, pp. 201–208, 1986.
- [33] C. F. Hagmann, E. De Vita, A. Bainbridge, R. Gunny, A. B. Kapetanakis, W. K. Chong, E. B. Cady, D. G. Gadian, and N. J. Robertson, "T2 at mr imaging is an objective quantitative measure of cerebral white matter signal intensity abnormality in preterm infants at term-equivalent age 1," *Radiology*, vol. 252, no. 1, pp. 209–217, 2009.
- [34] L. Tzarouchi, L. Astrakas, V. Xydis, A. Zikou, P. Kosta, A. Drougia, S. Andronikou, and M. Argyropoulou, "Age-related grey matter changes in preterm infants: an mri study," *Neuroimage*, vol. 47, no. 4, pp. 1148–1153, 2009.
- [35] F. Jäger, *Normalization of Magnetic Resonance Images and its Application to the Diagnosis of the Scoliotic Spine*. Berlin : Logos Verlag, 2011, vol. 34.
- [36] A. J. Barkovich and C. Raybaud, *Pediatric Neuroimaging*. Lippincott Williams & Wilkins, 2012.
- [37] A. Barkovich, B. Kjos, D. Jackson Jr, and D. Norman, "Normal maturation of the neonatal and infant brain: Mr imaging at 1.5 t." *Radiology*, vol. 166, no. 1, pp. 173–180, 1988.
- [38] E. Courchesne, H. Chisum, J. Townsend, A. Cowles, J. Covington, B. Egaas, M. Harwood, S. Hinds, and G. Press, "Normal brain development and aging: quantitative analysis at in vivo mr imaging in healthy volunteers1," *Radiology*, vol. 216, no. 3, pp. 672–682, 2000.
- [39] S. Blinkov and I. Glezer, *The Human Brain in Figures and Tables: a Quantitative Handbook*. New York : Basic Books, 1968.
- [40] P. Huttenlocher and P. Huttenlocher, *Neural plasticity: The effects of environment on the development of the cerebral cortex*. Harvard University Press, 2002.
- [41] H. C. Kinney, B. A. Brody, A. S. Kloman, and F. H. Gilles, "Sequence of central nervous system myelination in human infancy. ii. patterns of myelination in autopsied infants." *J. Neuropathol. Exp. Neurol.*, vol. 47, no. 3, pp. 217–234, May 1988.
- [42] S. R. Quartz and T. J. Sejnowski, "The neural basis of cognitive development: a constructivist manifesto," *Behav. Brain. Sci.*, vol. 20, no. 4, pp. 537–56; discussion 556–96, Dec 1997.
- [43] E. Courchesne, "Brain development in autism: early overgrowth followed by premature arrest of growth," *Ment. Retard. Dev. Disabil. Res. Rev.*, vol. 10, no. 2, pp. 106–111, 2004.

- [44] M. S. Van Der Knaap, J. Valk, and F. Barkhof, *Magnetic Resonance of Myelination and Myelin Disorders*. New York : Springer, 2005.
- [45] J. H. Gilmore, C. Kang, D. D. Evans, H. M. Wolfe, J. K. Smith, J. A. Lieberman, W. Lin, R. M. Hamer, M. Styner, and G. Gerig, “Prenatal and neonatal brain structure and white matter maturation in children at high risk for schizophrenia,” *AJP*, 2014.
- [46] P. S. Hüppi, “Neuroimaging of brain development-discovering the origins of neuropsychiatric disorders?” *Pediatr. Res.*, vol. 64, no. 4, pp. 325–325, 2008.
- [47] J. N. Giedd, J. Blumenthal, N. O. Jeffries, F. X. Castellanos, H. Liu, A. Zijdenbos, T. Paus, A. C. Evans, J. L. Rapoport *et al.*, “Brain development during childhood and adolescence: a longitudinal mri study,” *Nature Neuroscience*, vol. 2, pp. 861–862, 1999.
- [48] P. Aljabar, K. Bhatia, M. Murgasova, J. Hajnal, J. Boardman, L. Srinivasan, M. Rutherford, L. Dyet, A. Edwards, and D. Rueckert, “Assessment of brain growth in early childhood using deformation-based morphometry,” *Neuroimage*, vol. 39, no. 1, pp. 348–358, 2008.
- [49] R. C. Knickmeyer, S. Gouttard, C. Kang, D. Evans, K. Wilber, J. K. Smith, R. M. Hamer, W. Lin, G. Gerig, and J. H. Gilmore, “A structural mri study of human brain development from birth to 2 years,” *J. Neurosci.*, vol. 28, no. 47, pp. 12 176–12 182, 2008.
- [50] J. H. Gilmore, W. Lin, M. W. Prastawa, C. B. Looney, Y. S. K. Vetsa, R. C. Knickmeyer, D. D. Evans, J. K. Smith, R. M. Hamer, J. A. Lieberman *et al.*, “Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain,” *J. Neurosci.*, vol. 27, no. 6, pp. 1255–1260, 2007.
- [51] P. S. Hüppi, S. Warfield, R. Kikinis, P. D. Barnes, G. P. Zientara, F. A. Jolesz, M. K. Tsuji, and J. J. Volpe, “Quantitative magnetic resonance imaging of brain development in premature and mature newborns,” *Ann. Neurol.*, vol. 43, no. 2, pp. 224–235, 1998.
- [52] B. S. Peterson, A. W. Anderson, R. Ehrenkranz, L. H. Staib, M. Tageldin, E. Colson, J. C. Gore, C. C. Duncan, R. Makuch, and L. R. Ment, “Regional brain volumes and their later neurodevelopmental correlates in term and preterm infants,” *Pediatrics*, vol. 111, no. 5, pp. 939–948, 2003.
- [53] J. Matsuzawa, M. Matsui, T. Konishi, K. Noguchi, R. C. Gur, W. Bilker, and T. Miyawaki, “Age-related volumetric changes of brain gray and white matter in healthy infants and children,” *Cerebr. Cortex*, vol. 11, no. 4, pp. 335–342, 2001.
- [54] N. Gogtay, J. N. Giedd, L. Lusk, K. M. Hayashi, D. Greenstein, A. C. Vaituzis, T. F. Nugent, D. H. Herman, L. S. Clasen, A. W. Toga *et al.*, “Dynamic mapping of human cortical development during childhood through early adulthood,” *Proc. Natl. Acad. Sci. USA*, vol. 101, no. 21, pp. 8174–8179, 2004.
- [55] M. Prastawa, J. H. Gilmore, W. Lin, and G. Gerig, “Automatic segmentation of mr images of the developing newborn brain,” *MEDIA*, vol. 9, no. 5, pp. 457–466, 2005.

- [56] W. Gao, W. Lin, Y. Chen, G. Gerig, J. Smith, V. Jewells, and J. Gilmore, "Temporal and spatial development of axonal maturation and myelination of white matter in the developing brain," *AJNR*, vol. 30, no. 2, pp. 290–296, 2009.
- [57] X. Geng, S. Gouttard, A. Sharma, H. Gu, M. Styner, W. Lin, G. Gerig, and J. H. Gilmore, "Quantitative tract-based white matter development from birth to age two years," *NeuroImage*, 2012.
- [58] N. Sadeghi, M. Prastawa, P. T. Fletcher, J. Wolff, J. H. Gilmore, and G. Gerig, "Regional characterization of longitudinal dt-mri to study white matter maturation of the early developing brain," *NeuroImage*, 2012.
- [59] P. Mukherjee, J. H. Miller, J. S. Shimony, T. E. Conturo, B. C. Lee, C. R. Almlil, and R. C. McKinstry, "Normal brain maturation during childhood: Developmental trends characterized with diffusion-tensor mr imaging 1," *Radiology*, vol. 221, no. 2, pp. 349–358, 2001.
- [60] R. Marsh, A. J. Gerber, and B. S. Peterson, "Neuroimaging studies of normal brain development and their relevance for understanding childhood neuropsychiatric disorders," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 47, no. 11, pp. 1233–1251, 2008.
- [61] S. C. Deoni, E. Mercure, A. Blasi, D. Gasston, A. Thomson, M. Johnson, S. C. Williams, and D. G. Murphy, "Mapping infant brain myelination with magnetic resonance imaging," *J. Neurosci.*, vol. 31, pp. 784–791, Jan 2011.
- [62] R. Kreis, T. Ernst, and B. D. Ross, "Development of the human brain: in vivo quantification of metabolite and water content with proton magnetic resonance spectroscopy," *MRM*, vol. 30, no. 4, pp. 424–437, 1993.
- [63] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike, "Mri-based myelin water imaging: A technical review," *MRM*, vol. 73, no. 1, pp. 70–81, 2015.
- [64] R. Henkelman, G. Stanisz, and S. Graham, "Magnetization transfer in mri: a review," *NMR in Biomedicine*, vol. 14, no. 2, pp. 57–64, 2001.
- [65] I. R. Leppert, C. R. Almlil, R. C. McKinstry, R. V. Mulkern, C. Pierpaoli, M. J. Rivkin, and G. B. Pike, "T2 relaxometry of normal pediatric brain development," *JMRI*, vol. 29, no. 2, pp. 258–267, 2009.
- [66] E. A. Krupinski, "Current perspectives in medical image perception," *Atten, Percept, & Psychophys*, vol. 72, no. 5, pp. 1205–1217, 2010.
- [67] J. Gilmore, W. Lin, I. Corouge, Y. Vetsa, J. K. Smith, C. Kang, H. Gu, R. Hamer, J. Lieberman, and G. Gerig, "Early postnatal development of corpus callosum and corticospinal white matter assessed with quantitative tractography," *AJNR*, vol. 28, no. 9, pp. 1789–1795, 2007.
- [68] N. Sadeghi, M. Prastawa, J. H. Gilmore, W. Lin, and G. Gerig, "Towards analysis of growth trajectory through multi-modal longitudinal mr imaging," in *Proc. of SPIE Medical Imaging*. Society for Optics and Photonics, 2010.

- [69] A. Serag, P. Aljabar, S. Counsell, J. Boardman, J. Hajnal, and D. Rueckert, "Tracking developmental changes in subcortical structures of the preterm brain using multi-modal mri," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 2011, pp. 349–352.
- [70] M. Prastawa, N. Sadeghi, J. H. Gilmore, W. Lin, and G. Gerig, "A new framework for analyzing white matter maturation in early brain development," in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 2010, pp. 97–100.
- [71] N. Sadeghi, M. Prastawa, J. H. Gilmore, W. Lin, and G. Gerig, "Spatio-temporal analysis of early brain development," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2010, pp. 777–781.
- [72] P. Aljabar, R. Wolz, L. Srinivasan, S. J. Counsell, M. A. Rutherford, A. D. Edwards, J. V. Hajnal, and D. Rueckert, "A combined manifold learning analysis of shape and appearance to characterize neonatal brain development," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 12, pp. 2072–2086, 2011.
- [73] I. Csapo, B. Davis, Y. Shi, M. Sanchez, M. Styner, and M. Niethammer, "Temporally-dependent image similarity measure for longitudinal analysis," in *Biomedical Image Registration*. New York : Springer, 2012, pp. 99–109.
- [74] —, "Longitudinal image registration with temporally-dependent image similarity measure," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 10, pp. 1939–1951, 2013.
- [75] S. H. Kim, V. S. Fonov, C. Dietrich, C. Vachet, H. C. Hazlett, R. G. Smith, M. M. Graves, J. Piven, J. H. Gilmore, S. R. Dager *et al.*, "Adaptive prior probability and spatial temporal intensity change estimation for segmentation of the one-year-old human brain," *J. Neurosci. methods*, vol. 212, no. 1, pp. 43–55, 2013.
- [76] D. Salat, S. Lee, A. van der Kouwe, D. Greve, B. Fischl, and H. Rosas, "Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast," *Neuroimage*, vol. 48, no. 1, pp. 21–28, 2009.
- [77] C. Davatzikos and S. M. Resnick, "Degenerative age changes in white matter connectivity visualized in vivo using magnetic resonance imaging," *Cerebr. cortex*, vol. 12, no. 7, pp. 767–771, 2002.
- [78] D. Salat, J. Chen, A. van der Kouwe, D. Greve, B. Fischl, and H. Rosas, "Hippocampal degeneration is associated with temporal and limbic gray matter/white matter tissue contrast in alzheimer's disease," *Neuroimage*, vol. 54, no. 3, pp. 1795–1802, 2011.
- [79] L. T. Westlye, K. B. Walhovd, A. M. Dale, T. Espeseth, I. Reinvang, N. Raz, I. Agartz, D. N. Greve, B. Fischl, and A. M. Fjell, "Increased sensitivity to effects of normal aging and alzheimers disease on cortical thickness by adjustment for local variability in gray/white contrast: a multi-sample mri study," *Neuroimage*, vol. 47, no. 4, p. 1545, 2009.
- [80] S. Magnaldi, M. Ukmar, A. Vasciaveo, R. Longo, and R. Pozzi-Mucelli, "Contrast between white and grey matter: Mri appearance with ageing," *ER*, vol. 3, no. 6, pp. 513–519, 1993.

- [81] M. S. Panizzon, C. Fennema-Notestine, T. S. Kubarych, C.-H. Chen, L. T. Eyler, B. Fischl, C. E. Franz, M. D. Grant, S. Hamza, A. Jak *et al.*, “Genetic and environmental influences of white and gray matter signal contrast: A new phenotype for imaging genetics?” *NeuroImage*, 2012.
- [82] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *IJCV*, vol. 1, no. 4, pp. 321–331, 1988.
- [83] J. Park, D. Metaxas, A. A. Young, and L. Axel, “Deformable models with parameter functions for cardiac motion analysis from tagged mri data,” *Medical Imaging, IEEE Transactions on*, vol. 15, no. 3, pp. 278–289, 1996.
- [84] L. H. Staib and J. S. Duncan, “Boundary finding with parametrically deformable models,” *IEEE TPAMI*, vol. 14, no. 11, pp. 1061–1075, 1992.
- [85] S. C. Joshi, M. I. Miller, and U. Grenander, “On the geometry and shape of brain sub-manifolds,” *IJPRAI*, vol. 11, no. 08, pp. 1317–1343, 1997.
- [86] M. I. Miller, “Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms,” *NeuroImage*, vol. 23, pp. S19–S33, 2004.
- [87] K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [88] G. D. Hager and P. N. Belhumeur, “Efficient region tracking with parametric models of geometry and illumination,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 10, pp. 1025–1039, 1998.
- [89] R. B. Fisher and P. Oliver, *Multi-variate Cross-Correlation and Image Matching*. University of Edinburgh, Department of Artificial Intelligence, 1995.
- [90] C. Nastar, B. Moghaddam, and A. Pentland, “Generalized image matching: Statistical learning of physically-based deformations,” in *Computer Vision ECCV’96*. New York : Springer, 1996, pp. 589–598.
- [91] M. J. Jones and T. Poggio, “Multidimensional morphable models,” in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 683–688.
- [92] T. Vetter, “Learning novel views to a single face image,” in *Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on*. IEEE, 1996, pp. 22–27.
- [93] Y. Wang and L. H. Staib, “Elastic model based non-rigid registration incorporating statistical shape information,” in *Medical Image Computing and Computer-Assisted Intervention MICCAI98*. New York : Springer, 1998, pp. 1162–1173.
- [94] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, “Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain,” *MEDIA*, vol. 12, no. 1, pp. 26–41, 2008.
- [95] J. A. Maintz and M. A. Viergever, “A survey of medical image registration,” *MEDIA*, vol. 2, no. 1, pp. 1–36, 1998.

- [96] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas, "Deformation-based mapping of volume change from serial brain mri in the presence of local tissue contrast change," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 5, pp. 626–639, 2006.
- [97] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. Hawkes, "Non-rigid registration of breast mr images using mutual information," *Medical Image Computing and Computer-Assisted InterventionMICCAI98*, pp. 1144–1152, 1998.
- [98] C. Studholme, D. L. Hill, D. J. Hawkes *et al.*, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.
- [99] W. G. Negendank, A. Al-Katib, C. Karanes, and M. Smith, "Lymphomas: Mr imaging contrast characteristics with clinical-pathologic correlations," *Radiology*, vol. 177, no. 1, pp. 209–216, 1990.
- [100] T. Watanabe, H. Ito, A. Sekine, Y. Katano, T. Nishimura, Y. Kato, J. Takeda, M. Seishima, and T. Matsuoka, "Sonographic evaluation of the peripheral nerve in diabetic patients the relationship between nerve conduction studies, echo intensity, and cross-sectional area," *J Ultrasound Med*, vol. 29, no. 5, pp. 697–708, 2010.
- [101] Y. Kita, "Change detection using joint intensity histogram," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 2. IEEE, 2006, pp. 351–356.
- [102] D. Tell and S. Carlsson, "Combining appearance and topology for wide baseline matching," in *Computer VisionECCV 2002*. New York : Springer, 2002, pp. 68–81.
- [103] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4itk: improved n3 bias correction," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 6, pp. 1310–1320, 2010.
- [104] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier *et al.*, "Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration," *Neuroimage*, vol. 46, no. 3, pp. 786–802, 2009.
- [105] S. Joshi, B. Davis, M. Jomier, G. Gerig *et al.*, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, no. 1, p. 151, 2004.
- [106] F. Shi, Y. Fan, S. Tang, J. H. Gilmore, W. Lin, and D. Shen, "Neonatal brain image segmentation in longitudinal mri studies," *Neuroimage*, vol. 49, no. 1, pp. 391–400, 2010.
- [107] E. Armstrong, A. Schleicher, H. Omran, M. Curtis, and K. Zilles, "The ontogeny of human gyrification," *Cerebr. Cortex*, vol. 5, no. 1, pp. 56–63, 1995.
- [108] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of mr images of the brain," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 10, pp. 897–908, 1999.
- [109] T. Kailath, "The divergence and bhattacharyya distance measures in signal selection," *Communication Technology, IEEE Transactions on*, vol. 15, no. 1, pp. 52–60, 1967.
- [110] D. Pollard, "Asymptopia," *Manuscript in progress*). Available at <http://www.stat.yale.edu/~pollard/>, 2000.

- [111] C. R. Rao, "A review of canonical coordinates and an alternative to correspondence analysis using hellinger distance," *Questiões: Quaderns d'Estadística, Sistemes, Informàtica i Investigació Operativa*, vol. 19, no. 1, pp. 23–63, 1995.
- [112] R. Arandjelović and A. Zisserman, "Three things everyone should know to improve object retrieval," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2911–2918.
- [113] A. L. Gibbs and F. E. Su, "On choosing and bounding probability metrics," *Int. Stat. Rev.*, vol. 70, no. 3, pp. 419–435, Jan 2002.
- [114] A. Barmpoutis, B. C. Vemuri, and J. R. Forder, "Registration of high angular resolution diffusion mri images using 4 th order tensors," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*. New York : Springer, 2007, pp. 908–915.
- [115] Y. Qiao and N. Minematsu, "f-divergence is a generalized invariant measure between distributions." in *INTERSPEECH*, 2008, pp. 1349–1352.
- [116] A. Vardhan, M. Prastawa, N. Sadeghi, C. Vachet, J. Piven, and G. Gerig, "Joint longitudinal modeling of brain appearance in multimodal mri for the characterization of early brain developmental processes," in *Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data*. New York : Springer, 2015, pp. 49–63.
- [117] M. A. Styner, H. C. Charles, J. Park, and G. Gerig, "Multisite validation of image analysis methods: assessing intra- and intersite variability," pp. 278–286, 2002. [Online]. Available: <http://dx.doi.org/10.1117/12.467167>
- [118] J. Klein, S. Barbieri, H. Stuke, H. K. Hahn, M. H. Bauer, J. Egger, and C. Nimsky, "On the reliability of diffusion neuroimaging," 2010.
- [119] S. Y. Bookheimer, "Methodological issues in pediatric neuroimaging," *Ment Retard Dev D R*, vol. 6, no. 3, pp. 161–165, 2000.
- [120] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany *et al.*, "Reliability of mri-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer," *Neuroimage*, vol. 32, no. 1, pp. 180–194, 2006.
- [121] J. Leo and D. Cohen, "Broken brains or flawed studies? a critical review of adhd neuroimaging research." *Journal of Mind and Behavior*, 2003.
- [122] J. Jovicich, M. F. Beg, S. Pieper, C. Priebe, M. M. Miller, R. Buckner, and B. Rosen, "Biomedical informatics research network: Integrating multi-site neuroimaging data acquisition, data sharing and brain morphometric processing," in *Computer-Based Medical Systems, 2005. Proceedings. 18th IEEE Symposium on*. IEEE, 2005, pp. 288–293.
- [123] S. Gouttard, M. Styner, M. Prastawa, J. Piven, and G. Gerig, "Assessment of reliability of multi-site neuroimaging via traveling phantom study," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*. New York : Springer, 2008, pp. 263–270.

- [124] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. Hawkes, "Non-rigid registration of breast mr images using mutual information," *Medical Image Computing and Computer-Assisted Intervention MICCAI98*, pp. 1144–1152, 1998. [Online]. Available: <http://www.springerlink.com/index/F272V611L25G27Q7.pdf>
- [125] J. H. Gilmore, C. Kang, D. D. Evans, H. M. Wolfe, J. K. Smith, J. A. Lieberman, W. Lin, R. M. Hamer, M. Styner, and G. Gerig, "Prenatal and neonatal brain structure and white matter maturation in children at high risk for schizophrenia," *AJP*, 2010.
- [126] B. Casey, N. Tottenham, C. Liston, and S. Durston, "Imaging the developing brain: what have we learned about cognitive development?" *Trends Cogn. Sci.*, vol. 9, no. 3, pp. 104–110, 2005.
- [127] A. Karmiloff-Smith, "Neuroimaging of the developing brain: Taking developing seriously," *Human Brain Mapping*, vol. 31, no. 6, pp. 934–941, 2010.
- [128] J. Pinheiro and D. Bates, *Mixed-effects models in S and S-PLUS*. Springer Science & Business Media, 2006.
- [129] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, pp. 963–974, 1982.
- [130] M. J. Lindstrom and D. M. Bates, "Nonlinear mixed effects models for repeated measures data," *Biometrics*, pp. 673–687, 1990.
- [131] A. Vardhan, M. Prastawa, A. Sharma, J. Piven, and G. Gerig, "Modeling longitudinal MRI changes in populations using a localized, information-theoretic measure of contrast," in *Proc. 10th IEEE Int Biomedical Imaging (ISBI) Symp*, 2013, pp. 1396–1399.
- [132] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar *et al.*, "Linear and nonlinear mixed effects models," *R Package Version*, vol. 3, p. 57, 2007.
- [133] H. C. Hazlett, H. Gu, R. C. McKinstry, D. W. Shaw, K. N. Botteron, S. R. Dager, M. Styner, C. Vachet, G. Gerig, S. J. Paterson *et al.*, "Brain volume findings in 6-month-old infants at high familial risk for autism," *AJP*, vol. 169, no. 6, pp. 601–608, 2012.
- [134] L. Wang, F. Shi, P.-T. Yap, J. H. Gilmore, W. Lin, and D. Shen, "4d multi-modality tissue segmentation of serial infant images," *AJP*, 2012.
- [135] M. Prastawa, J. H. Gilmore, W. Lin, and G. Gerig, "Automatic segmentation of neonatal brain mri," in *MICCAI (1)*, 2004, pp. 10–17.
- [136] J. M. Tyszka, D. P. Kennedy, L. K. Paul, and R. Adolphs, "Largely typical patterns of resting-state functional connectivity in high-functioning adults with autism," *Cerebr. Cortex*, vol. 24, no. 7, pp. 1894–1905, 2014.
- [137] M. Zilbovicius, B. Garreau, Y. Samson, P. Remy, C. Barthelemy, A. Syrota, and G. Lelord, "Delayed maturation of the frontal cortex in childhood autism," *AJP*, vol. 152, no. 2, pp. 248–252, 1995.

- [138] S. C. Deoni, E. Mercure, A. Blasi, D. Gasston, A. Thomson, M. Johnson, S. C. Williams, and D. G. Murphy, "Mapping infant brain myelination with magnetic resonance imaging," *J. Neurosci.*, vol. 31, pp. 784–791, Jan 2011.
- [139] A. MacKay, I. Vavasour, A. Rauscher, S. Kolind, B. Mädler, G. Moore, A. Traboulsee, D. Li, and C. Laule, "Mr relaxation in multiple sclerosis," *Neuroimaging Clin N Am*, vol. 19, no. 1, pp. 1–26, 2009.
- [140] J. Pujol, A. López-Sala, N. Sebastián-Gallés, J. Deus, N. s. Cardoner, C. Soriano-Mas, A. Moreno, and A. Sans, "Delayed myelination in children with developmental delay detected by volumetric mri," *Neuroimage*, vol. 22, no. 2, pp. 897–903, 2004.
- [141] A. Vardhan, M. Prastawa, S. Gouttard, J. Piven, and G. Gerig, "Quantifying regional growth patterns through longitudinal analysis of distances between multimodal mr intensity distributions," in *Proc. 9th IEEE Int Biomedical Imaging (ISBI) Symp*, 2012, pp. 1156–1159.
- [142] T. Kanda, K. Ishii, H. Kawaguchi, K. Kitajima, and D. Takenaka, "High signal intensity in the dentate nucleus and globus pallidus on unenhanced t1-weighted mr images: relationship with increasing cumulative dose of a gadolinium-based contrast material," *Radiology*, vol. 270, no. 3, pp. 834–841, 2013.
- [143] B. Drayer, P. Burger, B. Hurwitz, D. Dawson, and J. Cain, "Reduced signal intensity on mr images of thalamus and putamen in multiple sclerosis: increased iron content?" *AJNR*, vol. 8, no. 3, pp. 413–419, 1987.
- [144] J. Grimaud, J. Millar, J. Thorpe, I. Moseley, W. McDonald, and D. Miller, "Signal intensity on mri of basal ganglia in multiple sclerosis." *Journal of Neurol. Neurosurg. Psychiatry*, vol. 59, no. 3, pp. 306–308, 1995.
- [145] S. Venkataraman, R. C. Semelka, L. Braga, I.-M. Danet, and J. T. Woosley, "Inflammatory myofibroblastic tumor of the hepatobiliary system: Report of mr imaging appearance in four patients 1," *Radiology*, vol. 227, no. 3, pp. 758–763, 2003.
- [146] S. K. Iyer, T. Tasdizen, and E. V. DiBella, "Edge-enhanced spatiotemporal constrained reconstruction of undersampled dynamic contrast-enhanced radial mri," *MRI*, vol. 30, no. 5, pp. 610–619, 2012.
- [147] D. R. Haynor, L. A. Mack, M. R. Soules, W. Shuman, M. Montana, and A. Moss, "Changing appearance of the normal uterus during the menstrual cycle: Mr studies," *Radiology*, vol. 161, no. 2, pp. 459–462, 1986.
- [148] L. G. Nyu and J. K. Udupa, "On standardizing the mr image intensity scale," *IEEE TMI*, vol. 19, no. 2, pp. 143–150, Feb 2000.