FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Lúcia Teresa Neto Fonseca

Automatic atlas-based segmentation of brain white-matter in neonates at risk for neurodevelopmental disorders

University of Coimbra, July 2014

• U C •

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Lúcia Teresa Neto Fonseca

Automatic atlas-based segmentation of brain white-matter in neonates at risk for neurodevelopmental disorders

BMIA Master Thesis Number 14/06

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Biomédica - Thesis presented to Coimbra University for Master of Biomedical Engineering.

Supervisors:
dr. Nicolás Lori[1]
dr. ir. Miguel Morgado[2,3]
dr. Anna Vilanova[4]
dr. ir. Carola van Pul[5]

University of Coimbra, July 2014

---

[1] *Instituto Superior de Engenharia de Lisboa*, Lisbon, Portugal – previously at Coimbra University
[2] Physics Department, Faculty of Science and Technology, Coimbra University, Portugal
[3] Formal supervisor at Coimbra University after transition of dr. Lori from Coimbra University
[4] Biomedical Image Analysis, Technical University of Eindhoven, Eindhoven, The Netherlands
[5] Clinical Physics, Máxima Medical Center, Veldhoven, The Netherlands

This work was developed in collaboration with:

Biomedical Image Analysis (BMIA) Group, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands.

Clinical Physics Department and Neonatology Department, Máxima Medical Center, Veldhoven, The Netherlands.

Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands

Acknowledgement

I would like to thank my supervisors, dr. Anna Vilanova, dr. ir. Carola van Pul and dr. Nicolás Lori, for all their interest and high quality supervision. In addition I also thank: dr. Lauren O'Donnell for providing part of the library for the developed pipeline, ir. Mehmet Yusufoglu for his time in helping me making the pipeline automatic, the neonatologist doctors Pieter Andriessen and Jan Buijs for the very valuable clinical insight, ir. Rieneke van den Boom for helping me better understand the neonatal atlas, dr. Floris Groenendaal and the group in UMCU for providing part of the dataset, and dr. Rainer Goebel and dr. Alard Roebroeck for the support given during my research at Maastricht University. Finally, I also thank prof. dr. ir. Pieter Wijn, prof. dr. ir. Bart ter Haar Romenij from TU/e and dr. ir. Sónia Gonçalves, dr. ir. Luís Cruz and dr. ir. Miguel Morgado from UC for their interest and time as committee members.

Abstract

Each year over half a million babies are born prematurely in Europe. The preterm neonates born before 32 weeks of gestation are especially at high risk for neurodevelopmental disorders. For these very preterm babies, cognitive deficits without major motor deficits are the dominant outcome (40%). Rehabilitative interventions at such an early age can prevent these cognitive deficits, due to the high plasticity of the brain at this period. However, there is currently no predictive biomarker for cognitive deficits at such early stage. Our aim is to label and quantify white matter (WM) maturation of preterms imaged at term equivalent age (TEA) and investigate its usefulness as a possible predictive marker for neurodevelopmental disorders.

We implemented an automatic pipeline for atlas-based segmentation of WM diffusion tensor imaging (DTI) tractography. We used a previously constructed tractography atlas for very prematurely born babies scanned at TEA. Main contributions of the current study involved a tract-wise tractography registration, a skull-stripping method tuned for neonatal data, a tract sampling method and automation of the entire pipeline. The tract-wise registration was inspired by the work of O'Donnell (2012). This new registration approach allows using the global directional and connectivity information entailed in the tractography pattern. Compared with manual segmentation methods, our method is less time consuming and less user dependent.

The result analysis is promising, as only 12% of the segmentations contained more than 30% mislabeled tracts. Segmentation performance showed not to be influenced by presence of WM pathology among subjects. The automatically segmented corpus callosum (CC) structure was further analyzed by studying the respective volume and anisotropy measurements per subject. For the studied measurements, volume and mean diffusivity (MD) showed a significant trend with degree of WM injury. These trends are in accordance with previous findings about how WM injury influences DTI derived anisotropies. This work shows that tractography can be segmented into the main WM anatomical structures that are of interest in neonates at risk for neurodevelopmental disorders and can be of added value for clinical evaluation.

Resumo

Na Europa, mais de meio milhão de bebés nasce prematuramente por ano. Os recém-nascidos com menos de 32 semanas de gestação estão especialmente em risco para desordens de desenvolvimento neuronal. Para estes bebés, os principais problemas de desenvolvimento surgem a nível cognitivo (40%). Reabilitação é possível, principalmente se for feita nos primeiros tempos de vida quando o cérebro é caracterizado pela sua enorme plasticidade. No entanto, não existem bio-marcadores que possibilitem prever quais os bebés prematuros que estão em risco. Este trabalho tem como objetivo analisar a maturação da matéria branca do cérebro em bebés prematuros e investigar a sua usabilidade como possível marcador para desordens de desenvolvimento neuronal.

Um *pipeline* automático para segmentação atlas-based de matéria branca visualizada com tratografia de Diffusion Tensor Imaging (DTI) foi implementado. O atlas usado foi construído previamente com tratografias de bebés prematuros em *term equivalente age* (TEA). Principais contribuições correspondem à automatização do *pipeline* e desenvolvimento de algoritmos específicos para tratografias neonatais para: registo entre tratografias, *skull-stripping* e *sampling*. O algoritmo para registo entre tratografias foi inspirado no trabalho de O'Donnell (2012). Este tipo de registo utiliza a informação relativa à conectividade global de regiões de matéria branca no cérebro, característica dos dados de tratografia. Em comparação com métodos de segmentação manual, este método consome menos tempo e é menos dependente do utilizador.

Os resultados são promissores, apenas 12% das segmentações contêm mais de 30% de fibras erroneamente segmentadas por estrutura anatómica. A performance da segmentação não foi influenciada pela presença de patologias da matéria branca nos pacientes. As estruturas anatómicas automaticamente segmentadas do corpus callosum foram também analisadas relativamente aos seus volumes e valores de anisotropia. Volume e difusão média são significamente correlacionados com a intensidade de patologia da matéria branca. Estes resultados estão de acordo com descobertas prévias sobre como patologia na matéria branca influencia os valores de anisotropia. Em conclusão, tratografia neonatal pode ser segmentada nas principais estruturas anatómicas de interesse para estudo de desordens do desenvolvimento neuronal.

Contents

List of Abbreviations

ADC	Apparent Diffusion Coefficient
ADHD	Attention Deficit Hyperactive Disorder
AP	Affinity Propagation
CC	Corpus Callosum
CG	Cingulum
Cl	Case Linear or Linear Diffusion
Cp	Case Planar or Planar Diffusion
CR	Corona Radiata
CT	Computed Tomography
DICOM	Digital Imaging and Communication in Medicine
DOF	Degrees of Freedom
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Imaging
FA	Fractional Anisotropy
FX	Fornix
GM	Gray Matter
LLS	Linear Least Square
MC	Mean of Closest Points
MCP	Middle Cerebellar Peduncles
MD	Mean Diffusion Coefficient
MRI	Magnetic Resonance Imaging
NifTi	Neuroimaging Informatics Technology Initiative
PFG	Pulse Field Gradients
Ol	Oligodendrocyte
ROI	Region of Interest
SENSE	Sensitivity Encoding
SNR	Signal to Noise Ratio
SS	Sagittal Stratum
TEA	Term Equivalent Age
US	Ultrasound
WM	White Matter
WVS	Whole Volume Seeding

Chapter 1

Introduction

Each year over half a million babies are born prematurely in Europe [EFCNI, 2010]. About 8-9% of all infants that were born prematurely suffer from severe disabilities, and 19-22% needs special education [Kooij, 2011]. Especially preterm neonates born before 32 weeks of gestation are at high risk for neurodevelopmental disorders. This condition is thought to be related with abnormal White Matter (WM) maturation of the brain. Therefore, early detection of WM abnormalities is important; not only for giving an adequate prognosis to the patient and parents, but also in the design of preventive, protective and rehabilitative strategies for the management of the development of the preterm infant. [Jong, 2012; Glass, 2011; Latal 2009; Volpe, 2009]

Contemporary techniques are unable to accurately distinguish which neonates are at risk for neurodevelopmental deficits. Collaboration has been started between the University Medical Center Utrecht, the Maxima Medical Center Veldhoven and the Biomedical Engineering department of Eindhoven University of Technology, in order to explore the usefulness and possible applications of Diffusion Tensor Imaging (DTI) for preterm neonatal patients. The main aim of this collaboration project is to investigate whether MRI-observed abnormalities in the brain structures after premature birth can help predict abnormalities in cognitive development.

Diffusion Tensor Imaging (DTI) is a non-invasive technique, which allows in-vivo reconstruction and visualization of the brain WM structures, which is referred to as fiber-tracking or tractography [Vilanova, 2004]. Analysis of these WM structures can give a powerful insight about the global arrangement of brain connectivity [Pul, 2012; Dubois, 2008; Huppi, 2006]. However, when performed automatically in the entire brain, tractography can generate hundreds of tracts per subject. Such output is not immediately useful to clinicians or researchers when it comes to assess risk of neurodevelopmental disorders. It makes it hard to recognize all anatomical structures, especially due to information cluttering.

Specifically for the neonatal data, it is wished to analyze particular anatomic WM structures, which are still in development in these patients, and carry meaningful physiological information. For this, the different tract structures need to be anatomically categorized, a process referred as segmentation. Although segmentation can be done manually [Pul 2012, Kooij 2011, Wakana, 2004], that requires extensive knowledge about complex WM tract anatomy, introduces user bias, and can also become very time consuming. New segmentation approaches have emerged, where the tracts are successfully segmented by automatic clustering methods [Visser, 2011; Leemans, 2009; Moberts, 2005]. However they do not automatically attach anatomical labels to the clusters, or find corresponding clusters across subjects. For this to happen, a priori information about the anatomical structures location, shape and other properties must be taken into account.

One way to incorporate a priori anatomical information into the segmentation procedure is via a WM brain atlas; where the unlabeled data is compared with the labeled atlas tracts [Hua, 2008; Wakana, 2007]. Because DTI data in neonates suffers from extra constraints such as noise and low anisotropy, tractography results from neonates will highly differ from adults. Consequently, tractography segmentation methods need to be tuned for handling neonatal data. Although some software packages allow for adult tractography segmentation [http://www.trackvis.org/, O'Donnell, 2012], none of them allow for neonatal tractography segmentation.

Previous work developed in the collaboration project includes: study of segmentation clustering techniques of WM tracts specific for preterm neonates [Boom, 2011; Hoskam, 2009]; and the creation of an automatic atlas-based tractography segmentation algorithm, in which a tractography atlas from non-pathologic preterm neonates at term equivalent age (TEA) was created [Boom, 2011]. Although the initial results are promising, several improvements are necessary for applying it directly to a large group of data sets. The current master thesis project aimed to continue that development and solve main limitations of a clinical pipeline for automatic segmentation of anatomical brain structures in preterm neonatal patients. In concrete, the main contributions of this thesis are the following:

1. Tract-wise registration algorithm for alignment of patient with the atlas;
2. Improvement of skull-stripping algorithm specifically for neonatal data;
3. Study of atlas-based segmentation performance for patients with different degrees of WM injury;
4. Actualization, improvement and automation of the clinical pipeline.

In this thesis report, Chapter 2 provides an overview of premature birth and its consequences on brain development, especially for WM structures. Chapter 3 clarifies how the tractography results are produced. Chapter 4 discusses possible methods for segmentation of tractography results. Here the atlas-based approach is introduced, and more details are given about the employed atlas. Chapter 5 goes through the constructed pipeline, and the implemented algorithms. The results and their analysis are presented in Chapter 6, and discussion and conclusion in Chapter 7. The appendixes constitute a schematic representation of the tractography data flow used for the proposed method, a compilation of the study plots for the result analysis, and the first page (including the abstract) of two soon to be submitted papers. The first of these papers corresponds to the work developed under this master thesis project; the second paper corresponds to work of a previous externship in DTI.

Chapter 2

Brain Development for Premature Infants

This chapter focuses on the clinical and physiological aspects of brain maturation. Its goal is to provide the reader with an elementary understanding about premature birth and its implications, the anatomy and physiology of the brain, and near term brain maturation. The final section considers the possible triggers of development disruption and its consequences, especially for WM formation.

2.1 Premature Infants

2.1.1 Social Impact

Approximately one in ten neonates in Europe is born prematurely, i.e., before 37 weeks of gestation; this corresponds to about half a million babies every year. Babies born with less than 32 weeks of gestation are considered to be very preterm. In the Netherlands these very preterm constitute 1.5% of all births [EFCNI, 2010]. During the last two decades, with the growing use of assisted reproductive technologies plus an older maternal age, the number of preterm births increased [Latal, 2009] - with a 4% increase in very preterm births [Jong, 2012]. These children, together with their parents, experience enormous physical, emotional and financial challenges. They also represent a significant burden on our healthcare systems [EFCNI, 2010].

2.1.2 Health Complications

While the mortality rate of preterm neonates is decreasing, they remain at risk for health complications during their life, especially the very preterm infants. In principle, complications can be categorized into severe and moderate-to-mild. Severe deficits are summarized as functional disability, usually defined as an aggregate group comprising cerebral palsy, mental retardation i.e., developmental or intellectual quotient of less than 70, and severe visual or hearing impairment. Cognitive deficits without major motor deficits are by far the dominant neurodevelopmental outcome [Kooij, 2011; Volpe, 2009; Latal, 2009]. In a cohort of Dutch adolescents born before 32 weeks, 38.4% had mild neuromotor problems and 3.2% severe neuromotor problems. Mild cognitive impairments were seen in 14.8% and only 4.3% suffered from moderate cognitive deficits [The Dutch Project on Preterm and Small for Gestational Age Infants at 19 years of age, 2007].

The combination of these mild degree deficits in cognitive, motor, behavioral performance and in interrelated functions, often lead to a lower school performance, and ultimately to a lower academic achievement and social integration [Latal, 2009]. Especially attention deficit hyperactivity disorder (ADHD) and other attention problems are more frequently reported for the very preterm children [Jong 2012; Volpe, 2009].

2.1.3 Intervention and Assistance

Currently, drug interventions are not able to reduce and rehabilitate brain injury. However, cohort studies found that individual differences in cognitive development at infant age were related with differences in parent-child interaction, quality of caregiver stimulation, differences in socioeconomic background and different society attitudes toward disabilities and schooling systems [Jong 2012, Latal 2009]. Furthermore, a delayed and alternative development of these children was observed, like the engagement of different pathways for language and the decrease of ADHD prevalence as entering adolescence [Mullen 2011; Latal 2009]. These findings suggest that environmental enrichment enhances cerebral plasticity and reorganization of cortical maps, and can also improve functional outcomes. These studies provided a good theoretical basis for early-intervention programs, at an age where the brain is characterized by a high degree of plasticity. Moreover, recent analysis concluded that early-intervention programs, extra parental guidance and physiotherapy, indeed positively influence cognitive outcome in the preterm population [Glass, 2011].

Therefore, early knowledge of the spectrum and severity of neurodevelopmental injury is necessary for counseling care-givers and for tailoring therapeutic interventions. Although major functional deficits in early childhood can be detected by the age of 2, this is not the case for moderate and mild long-term functional disabilities: moderate-to-mild deficits may stay undiagnosed until school age [Jong, 2012; EFCNI, 2010; Latal, 2009].

2.2 Brain in a Nutshell

The brain is an organ which exerts centralized control over the other organs of the human body. It collects high quantities of information from an always changing environment, integrates it, processes it, and accordingly engages patterns of muscle activity and secretion of chemicals, named hormones, on the rest of the body. It is therefore responsible for a wide range of processes, from simple unconscious reflexes until sophisticated purposeful behaviors. All these myriad of responses are the key for survival and prosperity.

2.2.1 Cell-scale

Neurons and glial are the two types of cells present in the brain, all together they make more than 170 billion cells [Herculano-Houzel, 2009]. Neurons have as function to process and transmit information. A typical neuron will consist of a cell body and some fiber projections: usually short dendrites and a long axon (see figure 1.1). Axons propagate information along its length by a flow of chemical-electrical exchanges with the extra-cellular surrounding. The communication between neurons occurs at the synapses, between the axon of the pre-synaptic neuron with the dendrites of the post-synaptic neuron(s). Connections can be made with up to hundreds of other neurons.

Glial cells are responsible for the neurons' well-being: supplying nutrients and oxygen, providing structure and protection, and assisting neuronal connections. White lipid layers, called myelin sheets, are wrapped around axons by a specific type of glial cell, the oligodendrocytes (Ols). This wrapping occurs at localized portions of the axon, blocking any possible exchange of chemical-electrical signals with the surroundings. Between these wrapped portions, unwrapped pieces of axon are maintained, the Ranvier nodes (see again figure 2.1). These nodes are the only locations at the axon which can make extra-cellular exchanges. Therefore, for myelinated axons, the propagation of signal corresponds to a saltatory exchange flow between the Ranvier nodes. This behavior increases the speed of signal propagation along the axon.

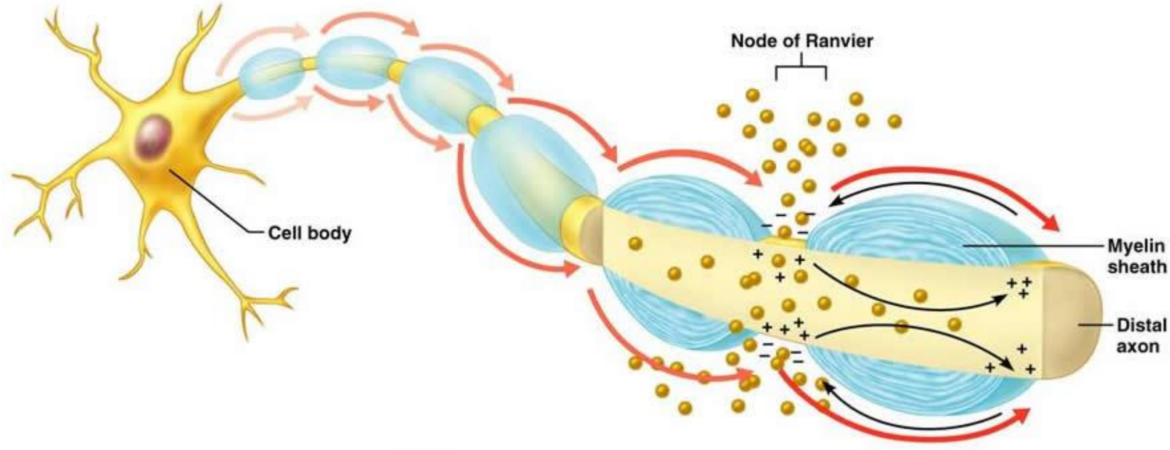


Figure 2-1 - A neuron cell body with small perturbations, the dendrites, and a long perturbation, the axon. Through the axon length, a flow of electrical-chemical exchanges with the surrounding occurs at consecutive Ranvier Nodes. In blue the myelin sheaths and in red the flow direction along the axon. In the human brain, the axon diameter has a length on the micrometer scale and the axon length can reach the millimeters range. (image from http:// http://classes.midlandstech.edu/)

2.2.2 Tissue-scale

At the tissue scale another division is observed: the neuron cell bodies, dendrites and glial cells make the Gray Matter (GM), axons bundles the White Matter (WM) (see figure 2.2). GM is primarily located on the outer surface of the brain, called the cerebral cortex. WM forms the pathways within the various locations of GM and for the rest of the nervous system.

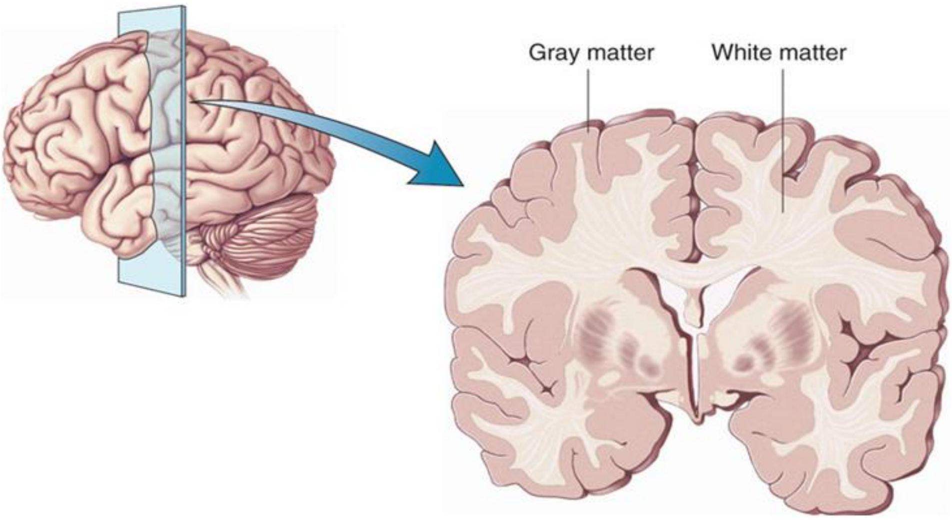


Figure 2-2 - White Matter and Gray Matter in the Human Brain. (image from http://hms.harvard.edu/)

2.2.3 White Matter Tracts

Some WM pathways can be anatomically distinguished and grouped in classes based on their structure, location and interrelated function. Main tract groups identified are the commissural tracts, running between the two brain hemispheres, the left and right partitions of the brain; the projection tracts - which can be afferent, bringing sensorial information to the brain, or efferent, sending information to the motor centers - uniting the cortex with lower parts of the brains and with the spinal cord; the association tracts, uniting different parts of the same cerebral hemisphere; and the brain stem tracts, at the posterior part of the brain, which join and are continuous with the spinal cord.

Characteristic anatomic WM structures are the Corpus Callosum (CC), the Corona Radiata (CR) left and right, the Fornix (FX), the Cingulum (CG), and the Cerebellar Peduncles (CP) (see figure 2.3). The CC corresponds to a commissural WM tract, and partial or complete absence of CC can result in motor control deficits, development impairments and visual and auditory memory losses. The CRs are efferent projection tracts, which play major roles in the transmission of motor commands to lower motor neurons. The FX and the CG are both association tracts. The first connects the hippocampus to the hypothalamus, which are both important structures for emotional and motor functions. The CG connects with the limbic system, thereby being involved with emotion, motivation pleasure and the emotional memory. Finally, cerebellar peduncles, brain stem type tracts, are responsible for the conveyance of sensations of touch, vibration and proprioception of human beings. [Mori, 2005]

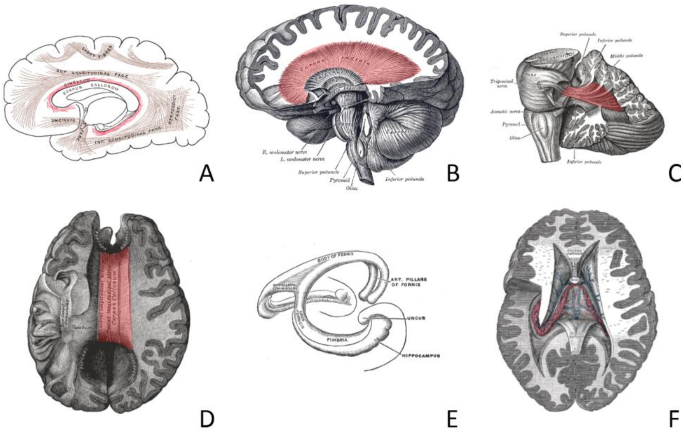


Figure 2-3 – Main WM Tracts at the Brain. The structures are depicted in a light shade of red: A – Cingulum, B – Corona Radiata, C – Middle Cerebellar Peduncles, D – Corpus Callosum, F – Fornix. At E, a scheme of the complex Fornix shape. (images from Gray's Anatomy http://en.wikipedia.org/)

2.3 Near Term Brain Development

Normal cognitive development is assumed to be correlated with structural maturation of congruous functional networks at the brain. Between 32 and 37 weeks' gestation, the brain undergoes rapid growth and microstructural evolution [Volpe, 2009]. This cascade of events strongly benefits from the womb unique environment: ranging from the protection it offers from the outside world, to the influence of the mother's natural body processes [Jong, 2012].

Projection, commissural, and association tracts are in a phase of rapid growth during the near term period. Over the last trimester of gestation and in the early postnatal period, axonal development is remarkably exuberant in the cerebrum. Pre-oligodendrocytes (pre-Ols), a still differentiating form of Ols, are in a phase of active development during 24 to 40 weeks' gestation. These cells already ensheath axons in preparation for full differentiation to myelin-producing Ols (see figure 2.4). This process is usually named pre-myelination. Mature myelin-producing oligodendrocytes do not become abundant in cerebral WM until after term. [Volpe, 2009]

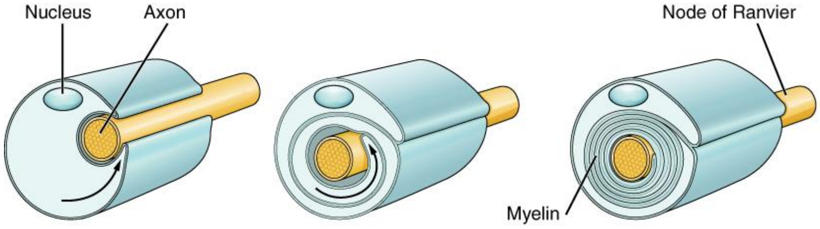


Figure 2-4 - Myelination process. In yellow the axon and in blue the glial cell and correspondent myelin sheet. (image from http:// http://classes.midlandstech.edu/)

Myelination is a long sequential nonlinear process that runs from the last trimester of gestation to at least 20 years of age, with a peak in the first postnatal year until the age of approximately 2 years. It progresses in an inferior-to-superior and posterior-to-anterior direction and from central to peripheral regions. As a result some WM structures show myelination earlier than others: motor pathways become myelinated before sensor pathways, projection tracts before association tracts, and occipital-parietal regions before temporal-frontal regions (see figure 2.5). The posterior and anterior part of the CC are myelinated in the first few months after birth and the middle part is only later myelinated [Volpe, 2008].

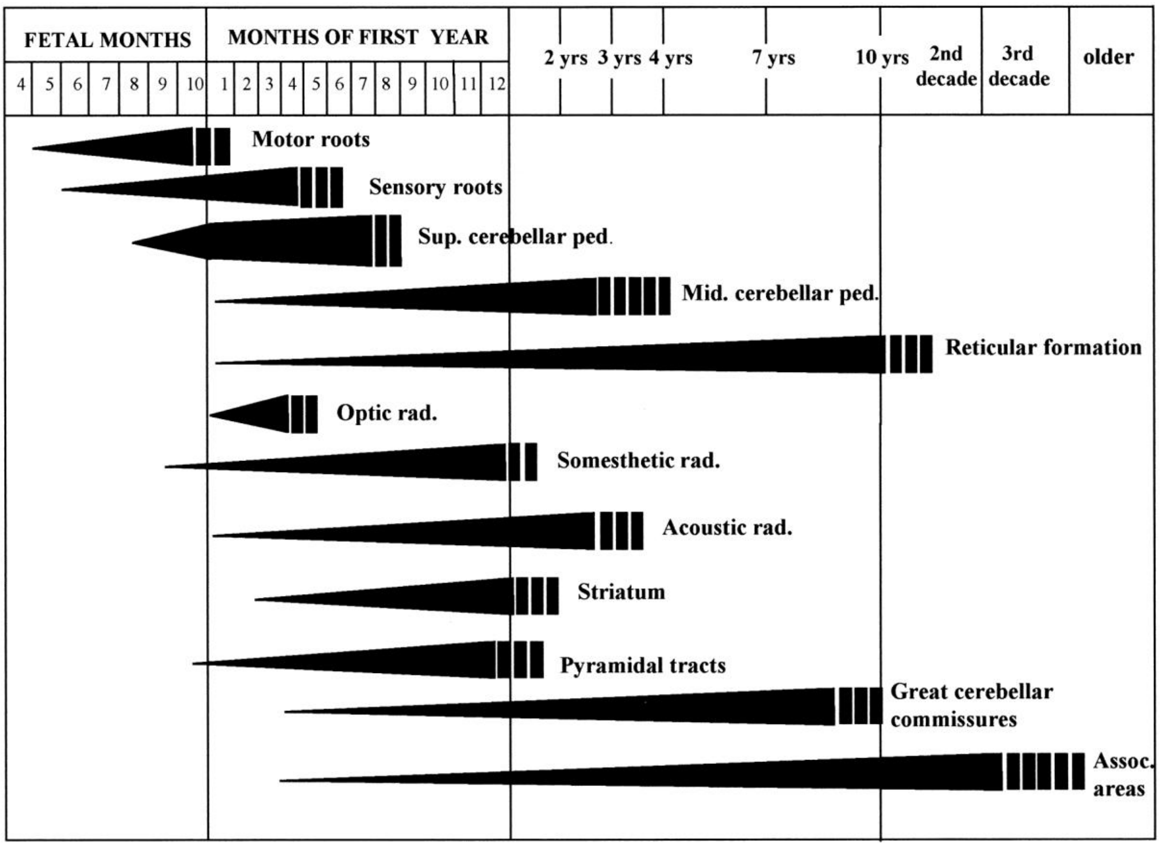


Figure 2-5 - Sequential process of myelination over time for main WM brain structures. (image from http://physrev.physiology.org/)

2.4 Encephalopathy of Prematurity

Premature birth is a risk for the baby, since part of the delicate process of brain maturation has to occur without benefitting from the womb's special environment. Moreover, other organs such as the lungs and the heart also have to adapt to the extrauterine life at an earlier stage of development, adding stress to the undeveloped brain. [Jong, 2012]

Neonatal neurodevelopment presents extreme vulnerability to exogenous and endogenous insults that often co-exist and can potentiate each other. The more common insults are hypoxia-ischaemia, inflammation and subsequent microglia harmful response, excitotoxicity, and free-radical attack [Volpe, 2009]. Most of the insults are potentially initiated during or shortly after preterm delivery, by chronic exposure to infection, and also by reperfusion as the infant is resuscitated after hypoperfusion and/or hypoxemia events [Huppi, 2006].

WM tissue in preterm neonates is thought to be especially vulnerable to these insults due to its elevated blood supply [Huppi, 2006] and still immature cerebral blood flow regulation [Kooij, 2011]. The primary consequences for this tissue are hypothesized to correspond to destructive processes, i.e., injury, with the subsequent maturational/ developmental disturbances being secondary. Injuries correspond to microscopic areas of necrosis involving all cellular elements; thus loss of pre-OLs, axons, and late-developing neurons, are to be expected. Furthermore, failure of pre-OLs to mature into OLs has been well documented in a neonatal animal model. [Volpe, 2009]

Many clinical findings support these processes of WM injury, as term equivalent age (TEA) preterms have been associated with reduced cerebral volume and WM immaturity, such as thinning of the CC, widening of the ventricles, and CC microstructural changes [Huppi, 1998; Pul, 2011]. Still, prediction of outcome is hard to assess, as there is a multitude of possible outcome patterns varying with onset times, severity, and duration of insults.

Chapter 3

White Matter Tractography

In the previous chapter, the anatomical and clinical backgrounds were introduced. The current chapter starts by discussing the relevance of neurological images for predicting developmental deficits. After this, the imaging modality specifically used for this thesis is presented in more detail. All the main concepts and techniques behind tracking WM based on water diffusion in the brain are clarified.

3.1 Neuroimaging as a Prediction Tool

It is known that intervention among preterms at risk for neurodevelopmental disorders is more successful at an early post-natal age. This is a stage characterized by an optimal neuroplasticity [Latal, 2009]. Such interventions need to be guided by an early diagnosis and prognosis predictions. Due to the complex and varied neurodevelopmental outcome after preterm birth, this prediction can become very challenging.

Biochemical markers have been shown to play a secondary role for such early assessment. Thus, the focus has shifted to near-term neuroimaging. A possible biomarker can consist of an imaging parameter displaying the WM development. Myelination, only one of the numerous cerebral maturational processes, has been shown to be closely related to cognitive development during the human life span [Dubois, 2008]. Other possible biomarkers can also arise from the study of gray matter development and brain gyration degree [Pul, 2012].

As for the neuroimaging modalities, cranial ultrasound (US), computed tomography (CT) and magnetic resonance imaging (MRI) are typically used for investigating the neonatal brain. US is a bedside tool used for initial assessment. US can detect important and common intracranial pathologies, after which the decision for further imaging using CT or MRI can be made. CT scans require a high dose of radiation in neonates and is therefore not frequently used, since most clinical questions can be answered using MRI.

MRI is the most sensitive imaging modality to investigate white and gray matter, being able to detect pathologies associated with neuromotor and cognitive impairment. Pitfalls of using MR with neonates relates with a necessary imaging sequence adaption to account for the higher water content in the neonatal brain. Furthermore, more movement artifacts and lower signal to noise ratio (SNR) occur in comparison with adult patients. Movement artifacts can be reduced using sedation protocols; however, these protocols comprise a risk to the patient.

Recently, advanced magnetic resonance imaging techniques such as volumetry and morphometry, diffusion tensor imaging (DTI) and tractography have been used to better determine the full spectrum of brain injury in these neonates. However, these do not yet belong to the common clinical practice. For this situation to change, new software specifically dedicated to clinical users, with fast learning curves and easy usability, must be created and made available. [Latal, 2009; Glass, 2011]

Specifically at the MMC in Veldhoven and the UMC in Utrecht, a MRI scan of the premature infants is obtained at term equivalent age (TEA), i.e., the date in which they would have been born considering a standard gestational period of 37 weeks. The usual sequences acquired are T2 and T1 weighted, and Diffusion Weighted Imaging (DWI). Study of anisotropy maps is part of routine clinical practice at both centers (anisotropy maps concept will be further explained in this chapter). The complete clinical routine also includes a complex continuous monitoring of respiratory system, cardiovascular system and brain physiological activity, and diagnostic tests comprising other neuroimaging techniques and clinical observations. On top of this, at the UMC in Utrecht, at least once a year, neurodevelopmental assessments are conducted to test motor skills and cognitive development. If necessary, therapy can be started.

3.2 Water Diffusion in the Brain

Diffusion is an essential transport mechanism in living organs. If molecules can move freely through a volume, the path each molecule describes will not have a preferred direction; it will be an isotropic displacement. The displacement ($\|\boldsymbol{r}\|$) will only depend on the observation time (t), temperature (T), and viscosity of the fluid on the particle (v), where k_B is the Boltzmanns's constant [J/K] (see equation 3.1).

$$\|\boldsymbol{r}\| = 6\, t\, k_B\, T\, \frac{1}{v}$$

Equation 3.1 – Brownian Movement Equation.

These last two variables represent the diffusion properties of the environment and can be represented instead by a diffusion coefficient (D) (see equation 3.2). In pure water, at body temperature, 37^{o} Celsius, the diffusion coefficient is 3×10^{-3} mm^2/s.

$$D = \frac{k_B T}{v}$$

Equation 3.2 – Diffusion Coefficient Definition.

As the brain develops, the water content in the brain decreases, extracellular spaces diminish in size, and intra- and intercellular microstructures become more complex and organized, constraining the water diffusion movement [Dubois, 2008]. Molecules will then encounter physical constraints, i.e., obstacles, to their random diffusion. These obstacles are not distributed homogeneously in all directions. Water diffusion will be more restricted for some directions than others. For the same amount of time, diffusion displacement will be larger in non-restricted directions than in restricted directions. This type of diffusion behavior is called anisotropic. In the brain, diffusion of water along the WM tracts will be less restricted than diffusion in the direction perpendicular to the tracts. Even for non-myelinated axonal bundles of premature infants, diffusion is already anisotropic due to higher restriction transverse to axons [Berman, 2005]. Figure 3.1 illustrates different diffusion behaviors: a non-restricted isotropic diffusion; a restricted but still isotropic, i.e., constraints are the same for all directions; and a restricted and anisotropic diffusion, like the one in the WM parts of the brain.

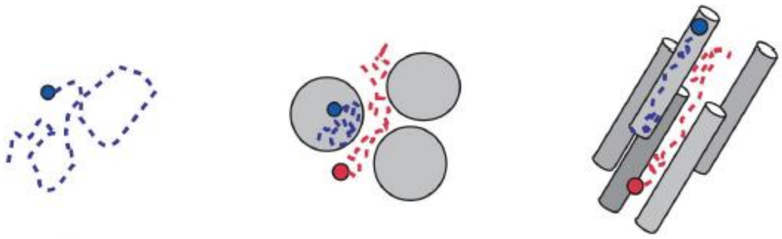


Figure 3.1 – Diffusion of molecules. Diffusive molecules are depicted in red and blues circles, the correspondent diffusion paths are illustrated with traced trajectories. From left to right: non-constrained isotropic diffusion, constrained isotropic diffusion, anisotropic restricted diffusion. [Pul, 2004]

3.3 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a medical imaging modality that uses the magnetic characteristics of the hydrogen proton. The imaged volume corresponds to stacks of slices, with a given thickness. The volume is said to be composed of several volume elements, so called voxels.

The hydrogen protons are highly abundant in the human body, since they are an elementary part of water molecules. The hydrogen proton possesses a property called spin. This is a purely quantum mechanical characteristic, which can be imagined as a rotation of the proton around an axis of itself, just like the spinning of a top. The important physics observation is that the spin possesses a magnetic field oriented along its axis. The spin will always rotate around itself with the same frequency, and can only vary the orientation of its axis. This variation is called precession and can have its own changeable frequency and phase (see figure 3.2).

Inside the MR scanner, a strong magnet produces an external magnetic field, B_0. When the water proton is placed in this external magnetic field, the spin axis of the proton aligns itself with this external field, just like a small magnet meeting a bigger magnet. The MR signal arises from disturbing this initial alignment. The disturbance is made possible by the existence of a coil in the scanner. This coil can create an alternating current, RF. This alternated current creates then a pulsed magnetic field. However, there is one condition that needs to be met in order for the perturbation to occur: the current must have the same frequency as the spin, which is called the resonance condition. When the RF stops, the spins return to their initial alignment with the B_0 magnetic field. The returning is characterized by a precession of the spin axis around B_0 (see figure 3.2). This behavior of the spin induces an alternate magnetic field. According to Maxwel's equations this induces an alternate electric field in the coil. In accordance with Ohm's law this creates an alternate electric current. This current corresponds to the magnetic resonance signal.

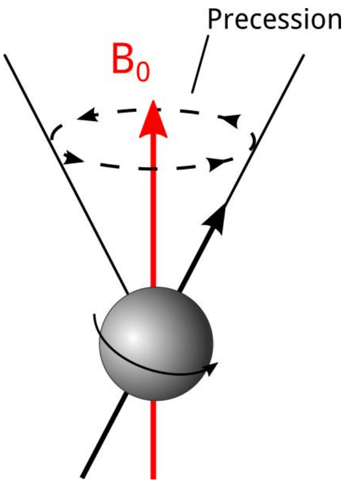


Figure 3.2 – A schematic representation of the precession of a spin. The proton is depicted as a sphere, which spins around an axis of itself. This axis varies its orientation in time, in a behavior called precession. After magnetic disturbance by the RF current, the spin gets to equilibrium by precessing around and finally aligning with the initial applied magnetic field (B0).

It is possible to encode spatial information in the signal, if each voxel experiences a different magnetic field. This is created by applying magnetic gradients in three different directions. The first gradient allows for slice-selection by generating a defined spatial region where nuclear spins will resonate; outside the slice, the nuclear spins are not affected by the RF pulse, i.e., the resonance condition is not met. The other gradients allow for a phase and frequency encoding of the spin axis precession; for the previously selected slice, spin axis will have different phase and frequency according to their position. Finally, a Fourier Transform is capable of depicting all these different frequencies and phases from the MR acquired signal, and translating them to spatial information. [Hornak, 1996-2014]

3.4 Diffusion Weighted Imaging

Diffusion Weighted Imaging (DWI) is a MR technique that is made sensitive for diffusion in living tissues. It does so by using a MR sequence that, apart from the standard magnetic gradients used for spatial encoding of the image, has two additional magnetic gradient pulses to measure diffusion along the gradient direction i: the Pulsed Field Gradients (PFGs) (see figure 3.3). These two PFG are identical in amplitude and width (δ), separated by a time Δ, and placed symmetrically about the standard 180 degree pulse of a MR sequence. The function of the first PFG pulse is to dephase the magnetization vector generated by the addition of the spins (mostly the spin of the Hydrogen proton in free water). The second PFG refocuses the spins which did not diffuse. Spins which have diffused to a new location in the period Δ will acquire a phase difference. Indeed, this period of time cannot be too short, or molecules will not have the time to encounter physical barriers, resulting in an isotropic diffusion profile. By increasing the time between the gradients, more barriers to the diffusion are encountered and then the diffusion coefficient will reflect the underlying tissue structure. Hence the diffusion coefficient D_i is renamed *apparent* diffusion coefficient (ADC_i). It is the fine tuning of the DWI parameters which allows for looking into water diffusion long enough, 10^{-2} s, for it to spread a distance of approximately an axon average diameter, 5-10 µm [Alexander, 2006].

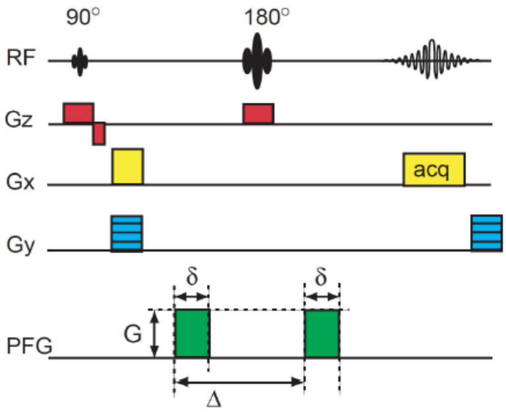


Figure 3.3 – Diffusion Weighted Spin Echo Sequence. The Pulsed Field Gradients have identical strength G, and width δ, and a time duration Δ between the gradients. [Pul, 2004]

Figure 3.4 illustrates the effect of the PFG for static and diffusing spins in a voxel. The PFG pulses have no effect on stationary spins. Both stationary and diffusing spins will acquire a phase when exposed to the first PFG. With the second pulse, the stationary spin will acquire an equal but opposite phase, since the pulses are on different sides of the 180 degree RF pulse, cancelling each other. The stationary spin comes back into phase, indicating a positive contribution to the echo. However, when the molecules in the voxel have moved due to diffusion, the rephasing effect of the second PFG signal will not correct the first PFG dephasing. The diffusing spin does not come back into phase, diminishing the total net magnetization signal M, the net sum of all spins. The net phase effect in the voxel will be zero, since the phase of all spins in the voxel together will be randomly distributed.

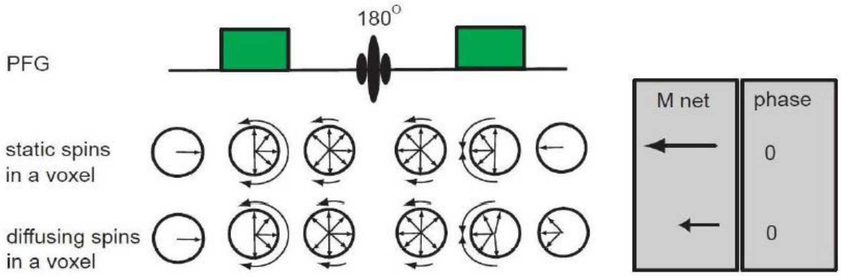


Figure 3.4 – Spin phase evolution during a DWI sequence and correspondent final outcome for phase and magnetization signal strength. The first line depicts the behavior of static spins and the second line, the behavior of diffusing spins. [Pul, 2004]

The common mathematical model for calculating the ADC from the magnetization signal assumes exponential signal decay and Gaussian-distributed diffusion for each voxel (see equation 3.3). S is the signal obtained in the presence of a PFG in the i direction (G_i), ADC_i the diffusion coefficient in the same direction, and γ the gyromagnetic ratio of the hydrogen proton. S_0, also referred as the b0 measurement, corresponds to a measurement not made sensitive to diffusion, i.e., with a sequence without PFGs.

$$S = S_0 \, e^{-ADC_i \, (G_i \gamma \delta)^2 \, (\Delta - \delta/3)}$$

Equation 3.3 – Stejskal and Tanner Equation.

Equation 3.3 gives a simplified version of all the factors which can in reality contribute for decreasing the signal. It is known that also a complex interaction between all the applied gradients can have an impact on the measured signal. In order to describe in a simpler way all these possible factors, a unique value can be used instead, the b-value (b) (see equation 3.4). The b-value is a factor which describes how sensitive is the sequence for diffusion in the probed direction i. For adults, the b-value for DTI studies is typically in the order of 1000 mm^2/s. For the infant brain, which has higher values for the ADC, b values are usually used in the order of 700-800 mm^2/s [Huppi, 2006].

$$S = S_0 \, e^{-ADC_i \, b}$$

Equation 3.4 – Le Bihan Equation.

3.5 Diffusion Tensor Imaging

Diffusion Tensor Imaging (DTI) is a specific case of DWI, more specifically a post-processing model of DWI. It measures the diffusion coefficient in at least six directions, allowing reconstructing a diffusion tensor per voxel (see equation 3.5 and 3.6).

$$\boldsymbol{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$$

Equation 3.5 – Diffusion Tensor.

$$D_i = \boldsymbol{g}_i{}^T \boldsymbol{D} \boldsymbol{g}_i$$

Equation 3.6 – Relation between the diffusion tensor D, and diffusion coefficient D_i, for specific gradient direction g_i.

To reduce noise effects, usually at least 32 directions are probed. This corresponds to acquire 32 DWI volumes. One extra volume is acquired for the b0-measurement. The two main artifacts intrinsic to DWI acquisitions that may destroy the voxel-wise correspondence across the volumes are eddy current distortions and motion. Scanner software is usually designed for tackling these artifacts by performing an automatic alignment between the volumes.

For each gradient direction, an equation can be written in matrix form (see first part of equation 3.7); where $\boldsymbol{S}$ is a vector containing the signal values for each gradient direction, $\boldsymbol{B}$ is the B-matrix containing direction-dependent b-value information, and $\boldsymbol{d} = [\mathrm{D}_{xx}, \mathrm{D}_{yy}, \mathrm{D}_{zz}, \mathrm{D}_{xy}, \mathrm{D}_{xz}, \mathrm{D}_{yz}]^{\mathrm{T}}$ is a vector containing the six unique values of the diffusion tensor. This system of linear equations can be solved by a Least Square solution (see second part of equation 3.7).

$$\boldsymbol{S} = \boldsymbol{B} \, \boldsymbol{d} \Rightarrow \boldsymbol{d}' = (\boldsymbol{B}^T \boldsymbol{B})^{-1} \boldsymbol{B}^T \boldsymbol{S}$$

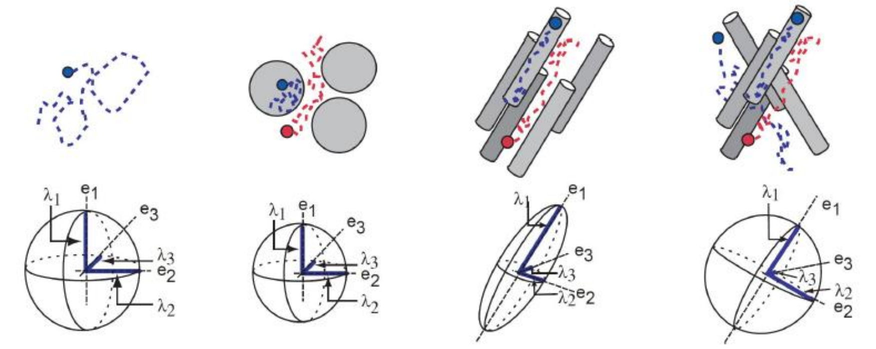
Equation 3.7 – Estimation of the six unique values of the diffusion tensor, d’, by using Least Square Regression.

The information given by the diffusion tensor can be better understood after eigenanalysis (see equation 3.8). The first eigenvector will give the principle direction of displacement and it is typically parallel to the WM fascicles. On the contrary, the second and third eigenvectors describe the diffusivity transverse to the axonal bundles [Dubois, 2006].

$$\boldsymbol{D}\,\boldsymbol{e} = \lambda\,\boldsymbol{e}$$

Equation 3.8 – An eigenvector of the square matrix D is a non-zero vector e that, when the matrix is multiplied by e, yields a constant multiple of e, the eigenvalue λ.

Commonly, tensors are visualized using eigenvectors scaled to their eigenvalues, which is referred as tensor shapes or glyph visualization. The most used shape is the ellipsoid (see figure 3.5). 3D ellipsoidal shape elongates along the preferred diffusion direction and squashes along restricted diffusion directions. However, when many data points are being studied, the density of the glyphs will lead to cluttered visualizations and make it difficult to distinguish between the different information per glyph. On top of this, glyphs can only represent local information at discrete data points. [Brecheisen, 2012]



(a) Free diffusion (b) Isotropic restricted diffusion (c) Anisotropic restricted diffusion (d) Crossing fibers

Figure 3.5 – Different diffusion behaviors and correspondent elliptical glyphs. λ corresponds to the eigenvalues and e to the eigenvectors. From left to right: a) non-restricted isotropic diffusion, b) restricted isotropic diffusion, c) anisotropic restricted diffusion for voxel with one orientation population, d) anisotropic restricted diffusion for voxel with more than one tract orientation population. [Pul, 2004]

Other common approach to analyze the tensor data is the construction of scalar metrics from the eigenvalues per voxel, see table 3.1. In figure 3.6, different visualizations of DTI are represented. The most popular anisotropy metrics are fractional anisotropy (FA) that measures the degree of diffusion directionality and mean diffusivity (MD) that corresponds to the directionally averaged magnitude of water diffusion. Many studies focus on the relation between these values and anatomy-physiology of the human brain. Anisotropy scalars trends are hypothesized to be related with normal and abnormal brain development. With increase of maturation, an increase of FA and decrease of MD is usually observed. Also linear (Cl) and planar (Cp) diffusion are often used for discrimination between the water diffusivities parallel and perpendicular to the WM tracts, with implications for axonal and myelin integrity, respectively. [Rose, 2014; Liu, 2012; Pul, 2012; Chen, 2011; Hasegawa, 2011; Lindqvist, 2011; Mullen, 2011; Liu, 2011; Dubois, 2008; Dubois, 2006; Huppi, 2006]

Name	**Equation**
Fractional Anisotropy (FA)	$FA = \frac{1}{2}\sqrt{2}\sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{(\lambda_1{}^2 + \lambda_2{}^2 + \lambda_3{}^2)}}$
Relative Anisotropy (RA)	$RA = \frac{1}{2}\sqrt{2}\sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{(\lambda_1 + \lambda_2 + \lambda_3)^2}}$
Volume Ratio (VR)	$VR = 1 - 27\frac{\lambda_1 \lambda_2 \lambda_3}{(\lambda_1 + \lambda_2 + \lambda_3)^3}$
Case Linear (Cl)	$Cl = \frac{(\lambda_1 - \lambda_2)}{\lambda_1 + \lambda_2 + \lambda_3}$
Case Planar (Cp)	$Cp = \frac{2(\lambda_2 - \lambda_3)}{\lambda_1 + \lambda_2 + \lambda_3}$
Mean Diffusivity (MD)	$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$

Table 3.1 – Commonly used anisotropy values.

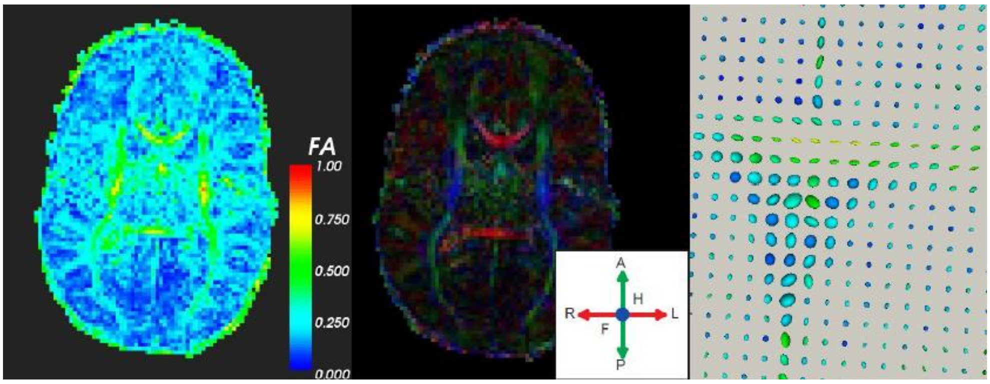


Figure 3.6 – Different DTI information visualizations. From left to right: Scalar anisotropic map for FA for a brain slice; main eigenvector direction for the same brain slice; glyph visualization for some voxels. [Boom, 2011]

3.6 Tractography Model

Tractography corresponds to tracing 3D paths through diffusion vector fields. These vector fields are often created by numerically integrating the main eigenvector of the diffusion tensor (see left side of figure 3.7). The tracing models can be divided into two general classes: deterministic and probabilistic. The first gives always the same output for same input. The second does not; as it introduces an element of randomness into the tracking process in order to simulate the effects of data disturbances. So far, the simpler deterministic method has been used for when dealing with neonatal brains.

For the deterministic methods, the most common approach is the streamline tracing [Vilanova, 2004]. This implementation requires: seeding points, which define the initial conditions for solving the integration; numerical integration along the main eigenvector direction; interpolation for calculating the next integration step vector - usually for a pre-defined step size of around 0.2 mm; and stopping criteria. The stopping criteria prevent the algorithm from tracing into regions where the vector field is not reliably defined, such as gray matter. Often used stopping criteria are minimum anisotropy index like Cl, and maximum angle to avoid high curvature of the reconstructed tracts.

Allocation of seed points can be done manually or automatically. Manual placing requires specification of one or more regions of interest (ROI) by the user. The interior of the ROI is discretely sampled and used as seed points. The automatic approach is called whole volume seeding (WVS) [Vilanova, 2004] and creates tracts throughout the entire volume using a user-defined density (see right side of figure 3.7). Manually defining seed points is sensitive to user bias, whereas WVS gives reproducible results. On the other hand, tractography using WVS produces cluttered data which is difficult to interpret and is more susceptible to artifacts. [Mori, 2002]

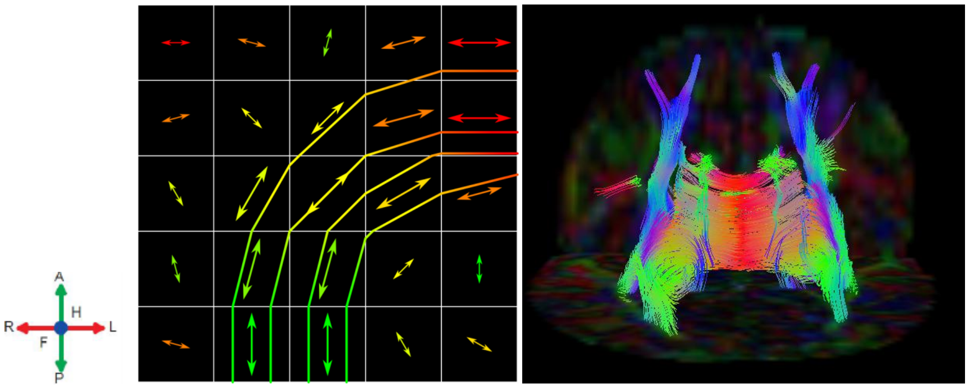


Figure 3.7 – Tractography result. From left to right: Tracing of principal eigenvectors through a diffusion tensor field of a transverse plane; WVS tracking of a preterm neonatal patient scanned at TEA, with frontal and transverse plane depicting an anisotropy map.

Tractography results are only expected in the WM regions of the brain. Sometimes due to scanning artifacts, tracts are reconstructed outside the WM region – in GM regions and especially on the outer side of the head. Correction is usually performed by applying a skull-stripping algorithm, removing the spurious areas from analysis. Automatic skull-stripping for neonatal tractography can be performed by morphological inspection of anisotropy maps - i.e., by probing the connectivity between regions with similar intensity values [Boom, 2011; Hoskam, 2009]. However, Boom (2011) found this method to be insufficient for some of her tested subjects. Notestine et al (2006) presented a detailed comparison of common skull stripping methods and tested their efficiency for different datasets. They concluded that especially for young subjects, best results were obtained from a hybrid approach between two methods. This hybrid approach combined morphological operators with a threshold method which defined minimum and maximum values along an intensity histogram.

3.7 White Matter Reconstructed Tracts

Although the 3D curves originating from tractography are commonly referred to as tracts or fibers, they do not represent individual axons. Instead, these curves are an estimation of how the diffusion of water is constrained along large WM fascicles. While neurons are microscopic in size, the DTI voxels have a size in the millimeters range. Usually for neonatal MRI in clinical practice, the voxels have a range between 2 and 8 mm^3 [Rose, 2014; Geng, 2012; Liu, 2012; Pul, 2012; Chen, 2011; Hasegawa, 2011; Lindquist, 2011; Liu, 2011; Dubois, 2008 and 2006; Huppi, 2006; Berman, 2005]. Therefore, many thousands of axons can fit into one DTI voxel. The DTI measurements correspond to an averaged effect of the collisions of water molecules against all the cellular components present per voxel. When noise, acquisition artifacts, or when different tract orientation populations are present at one single voxel, like in figure 3.8, this averaging can lead to an incorrect estimation of tract direction. Consequently, the tracking algorithm can continue tracking in a voxel that actually contains no real underlying structure, but is anisotropic, or the other way around, which will result in a misrepresentation of brain structures.

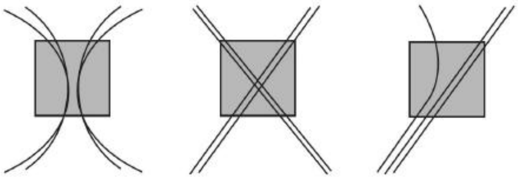


Figure 3.8 – More than one tract orientation population can be present by voxel. From left to right different combinations of two different tract orientations: kissing tracts, crossing tracts, and converging/diverging tracts. [Pul, 2004]

Reconstructed tracts from voxels contain multiple tissue structures, like the ones depicted at figure 3.8, which can lead to an effect called the partial volume effect. It can occur also when different kinds of tissue are present in one voxel, like at the boundary of gray and white matter. The bigger the voxel size, the more pronounced the partial volume effect will be. In DTI, it refers usually to multiple tract directions within one voxel. For these voxels, errors are introduced since DTI assumes a simple 3D Gaussian distribution for modeling diffusion per voxel. To accurately capture the diffusion profile arising from such complex tract configurations, many more gradient directions are required as well as complex post-processing mathematical models. However, this requires acquisition in many directions making clinical scan times large, which is not easily feasible in neonates. Also, these high-diffusion directions are not commonly implemented in clinical scanners. Still, the visualization of tractography allows for better depiction of global information from DTI than all the other referred approaches, like anisotropy scalars or glyph visualization.

Chapter 4

Atlas-based Tractography Segmentation

This chapter introduces the research domain of atlas-based tractography segmentation and discusses relevant previous work in this field. Creation of tractography atlases is addressed, and a more detailed explanation is given about the previously constructed neonatal tractography atlas used for this thesis work. Based on previous works, a structure of processing steps is delineated. Steps with a higher complexity are discussed in their own section.

4.1 Why an atlas-based segmentation?

When tractography is performed for the entire brain, the output will consist of hundreds of tracts. Such output is not immediately useful to clinicians or researchers when it comes to assess risk of neurodevelopmental disorders. It is hard to recognize and analyze all anatomical structures, especially due to information cluttering. For the neonatal data, it is helpful to analyze particular WM anatomic structures, which are still in development, and carry meaningful physiological information. Therefore, after tractography, the different reconstructed tracts need first to be categorized into the correspondent WM anatomical structures. This is a process referred to as segmentation.

Segmentation can be done interactively using expert knowledge. This method is based on manually defining regions of interest (ROIs). Tracts are labeled as the same anatomical structure when passing through the same ROIs. One of the strongest examples of this method is the work presented by Wakana (2004). Also manual segmentation has already been successfully performed in neonates from the collaboration project between UMCU-MMC-TU/e [Pul 2012, Kooij 2011]. Although this is the most common approach for tractography segmentation in clinical studies, it poses some drawbacks. It requires extensive knowledge about complex WM tract anatomy, it is user biased and it can become very time consuming. On top of this, manual selection might be difficult to perform for some anatomical structures with complex shapes, like the fornix (FX) which requires multiple ROI's to be selected. Figure 4.1 depicts ROI drawing for FX segmentation in the coronal plane (vertical from head to feet orientation plane) and subsequent result. This image represents the best manual segmentation result from Pul (2012). In figure 4.2, the more common result obtained for FX segmentation is shown. For these cases, only a small part of the FX can be segmented.

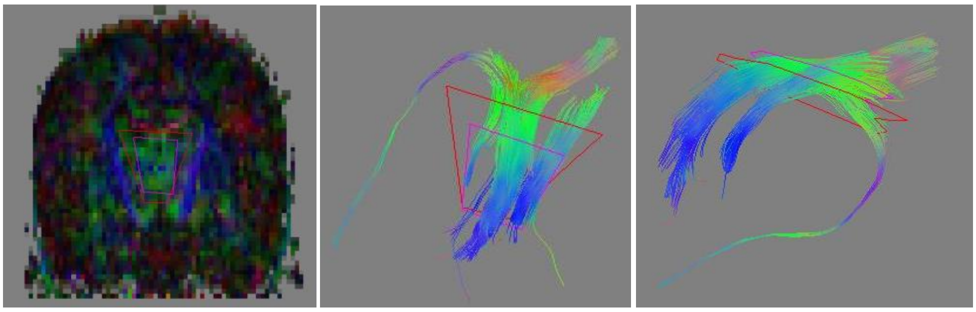


Figure 4.1 – Manual Segmentation of ROI in red – in the left, the ROI drawing; in the center, ROI and segmentation results in frontal view; in the right same results in traversal view.

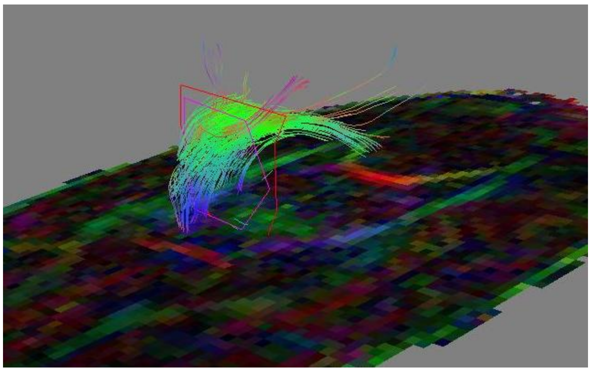


Figure 4.2 – ROI and typical segmentation result of FX.

Another common approach for tract segmentation is to automatically organize similar tracts in groups, i.e., to cluster them [Boom, 2011; Visser, 2011; Hoskam, 2009; Leemans, 2009; Moberts, 2005]. Similarity between tracts is decided usually based on the distance between tracts (e.g., [Lori, 2002]). Tracts that are similar to each other, based on this metric, are grouped together. The grouping is very dependent on the type of similarity metric, clustering method and the clustering algorithm parameters [Boom, 2011; Moberts, 2005]. This type of approach has two inter-related drawbacks. First, although similar tracts are grouped together, automatic anatomical labeling is not included in the algorithm. Second, in order to avoid having clusters containing tracts from more than one anatomical structure, tracts from the same anatomical structure will be also separated into different clusters. For neonatal data, the optimal number of clusters to ensure that one cluster contains only one anatomical structure is in the order of 200, which is clearly superior to the number, less than 10, of anatomical structures wished to segment (see figure 4.3; and see chapter 2, section 2.2 for the anatomical structures of interest). In conclusion, clustering alone shows to be not enough for producing an anatomically meaningful automatic segmentation.

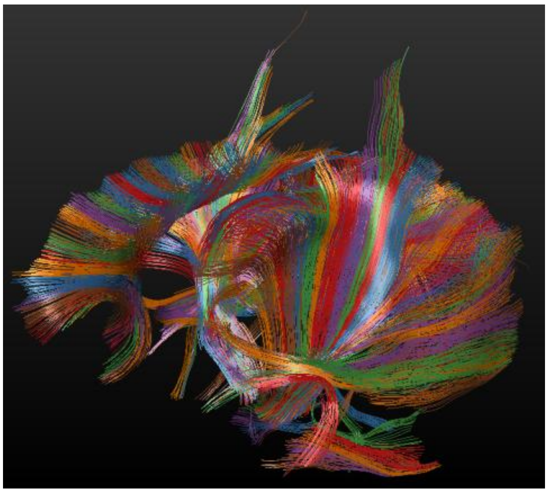


Figure 4.3 – Clustered tractography of a preterm scanned at TEA using Boom clustering method (2011).

To overcome this, another segmentation approach has emerged that uses a priori information about the WM anatomy. The information of interest comprises anatomical structures location, shape and other differentiating properties of the WM tracts. Due to the complexity of this information, the creation of such model is usually accomplished by manually labeling tractography results from subject's data. These models are called tractography atlases. Atlas and subject are compared, and the tract from the new subject will then inherit the label of the atlas tract to which it is most similar. Atlas creation is a step performed once. This segmentation approach is called atlas-based segmentation.

4.2 Atlas Creation

A brain atlas is a map of the brain that consists of pictures and/or tables and charts that label the anatomical structures, i.e., the atlas labels. The Talairach-Tournoux atlas is probably the first successful and most widely used atlas space for mapping of the human brain [Talairach and Tornoux, 1988] (see left image in figure 4.4). This atlas is based on a single post-mortem dissection of a human adult brain and labeled according to the Brodmann's brain regions. Its coordinates are defined by making two anchors, the anterior commissure and the posterior commissure, lying on a straight horizontal line: the AC-PC line.

Currently, many more alternative atlases exist, including atlases from in-vivo data. Mori et al. (2005) created the first WM brain atlas, by hand segmenting volumes of a standard-space average of diffusion MRI tensor maps from 81 healthy subjects (see right image in figure 4.4). Wakana et al. (2007) and Hua et al. (2008) created a tractography atlas, by identification of 20 structures from an average of 28 deterministic tractography results of normal subjects.

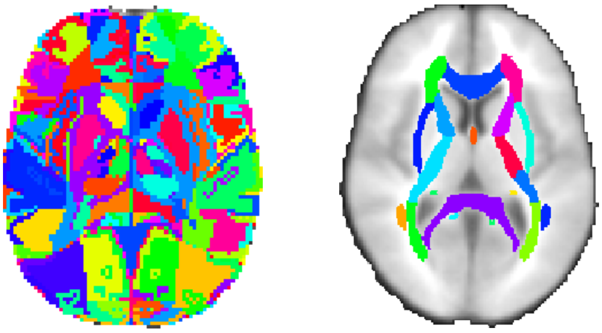


Figure 4.4 – Representation of two brain atlas in tranversal view. From left to right: Talairach-Tournoux atlas; Mori WM atlas. (images from http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases)

Preterm neonatal brain tractography results are quite different from tractography of healthy normal adult subjects (see figure 4.5). The difference comes from the low brain development degree. First, lower tract organization and myelination degree translate in different tract patterns. Tractography from neonates is characterized by a lower number of tracts, presence of smaller and broken tracts, and even missing anatomical structures yet to develop. Second, neonates have a smaller brain size, making tractography more sensitive to partial volume effects. Third, the neonatal patient group presents a large variation in gestational age, birth weight, and severity of illness which translates into a high degree of variability across subject brains. Therefore, adult atlases are of little use for segmenting neonatal data [Boom, 2011].

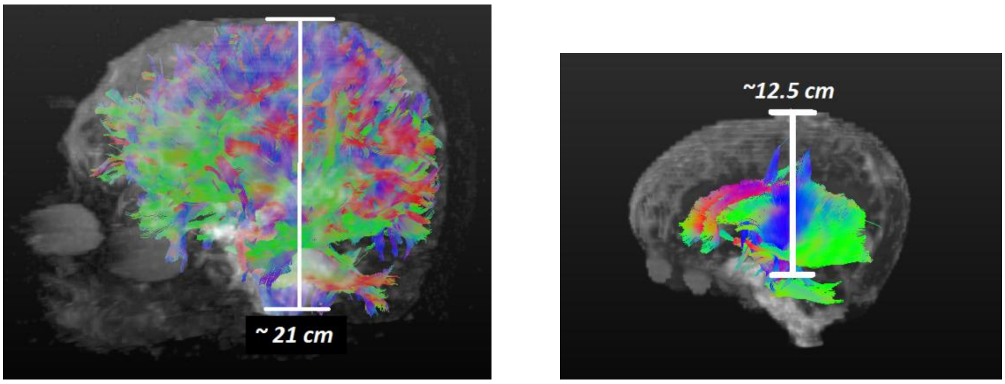


Figure 4.5 – Tractography results of non-pathologic patients with head shape in transparency. Tracts are color labeled: blue for feet to head orientation, green for back to front and red for left to right. From left to right: adult patient tractography, with length from bottom of ear to top of head of approximately 21 cm; premature at TEA tractography, with length from bottom of ear to top of head of approximately 12.5 cm.

During literature search, the only tractography atlases for neonatal data found corresponded to the work of Boom (2011). This is the atlas used for the current study. This atlas was built from three tractography datasets of preterm neonates imaged with DTI at term equivalent age (TEA) in UMCU. The tracts were constructed using software developed at the Biomedical Image Analysis Group from the University of Eindhoven, vIST/e [http://bmia.bmt.tue.nl/software/viste/]. The tracking parameters used were specific for the patient type of neonates at TEA, as suggested by Pul (2004). After tractography, the results were evaluated by experts, to make sure they provided an accurate representation of the brain WM anatomy structure as expected for this degree of development. The tractography results from two of the patients were aligned, i.e., registered, with the third patient. After registration, all tracts from the three patients were clustered together in small groups by applying an Affinity Propagation Method [Frey, 2007]. All clusters were visually inspected by three experts and labeled according to anatomical structure based on the WM atlases by Mori (2005) and Wakana (2004). In this way all the tracts were divided into the following labels: CC (corpus callosum), CR (corona radiata, left and right), SS (sagittal stratum, left and right, CG (cingulum), FX (fornix), MCP (middle cerebellar peduncle), Cheeks, Artifacts and Other. For more details the reader is referred to the master thesis of Boom (2011), “Automatic Atlas based White Matter Bundle Labelling for Neonates”. The atlas is presented in figure 4.6 and 4.7.

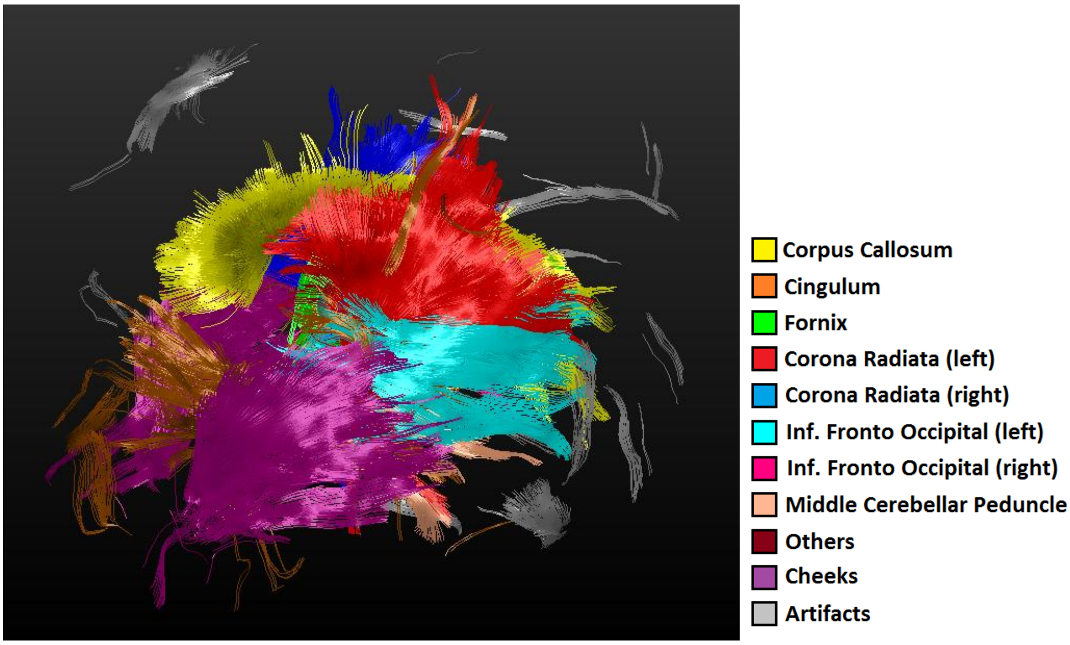


Figure 4.6 – Atlas of preterm neonates imaged at TEA – lateral view. Label of segmented structure is depicted by a color-key, at right side of the image.

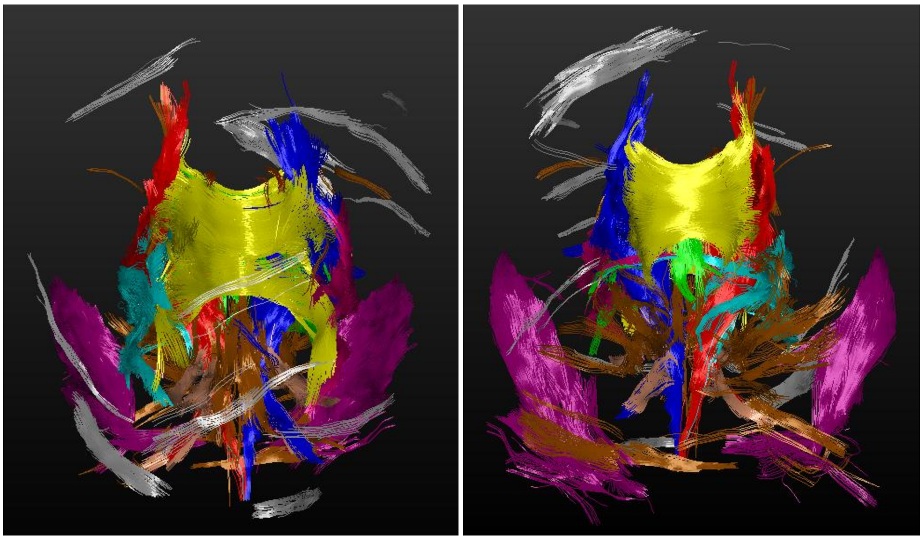


Figure 4.7 – Atlas of preterm neonates imaged at TEA – frontal and anterior views. Same color key as in figure 4.6.

4.3 Segmentation Procedure

After creation of the tractography atlas, tractography data of new subjects can be segmented by comparing them with the atlas. The new subject tracts will inherit the label of the atlas tract with which they have the highest similarity, the *Label Attribution* step. This comprises previous computation of a similarity metric between the atlas tract set and the new subject atlas tract set, the *Similarity Metrics* step. By definition, similarity metric is distance-based. Therefore, it is critical to have the subject and atlas geometrically aligned previously to similarity computations between them. The process of finding the best alignment between the atlas tract set and the new subject tract set is the *Registration to Atlas* step. Previous to registration, some specific algorithms can be employed for improving performance, like *Down-Sampling* of the tracts. *Down-Sampling* requires also an extra final step at the end of the segmentation flow, *Label Propagation;* such that the labeled sampled tracts information is propagated to the entire tractography set. The correspondent processing procedure, i.e. pipeline, of these steps is depicted in figure 4.8.

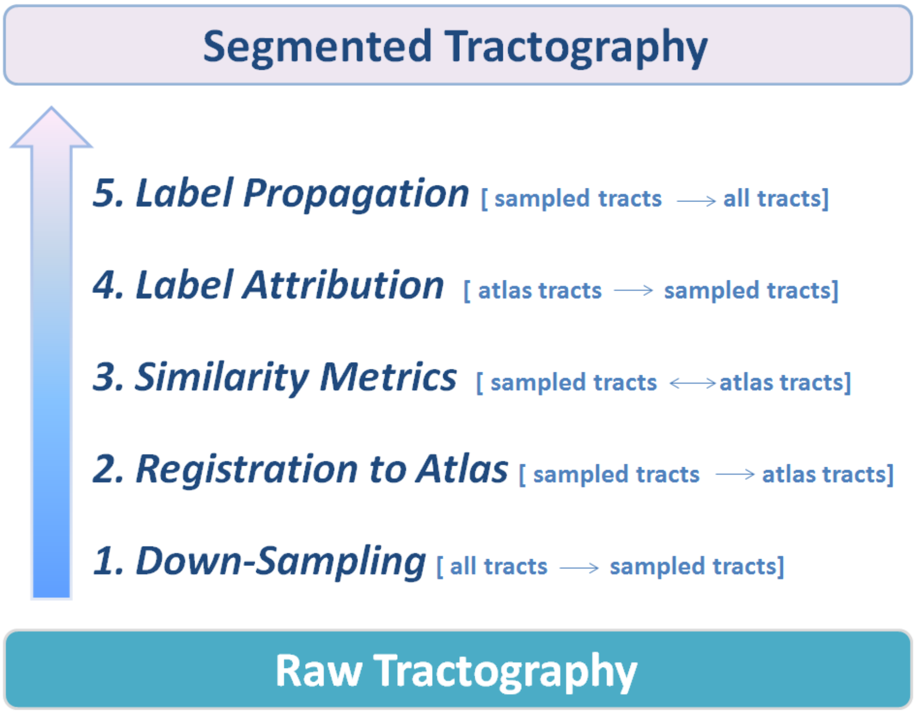


Figure 4.8 – Procedure for segmentation of new subject using a previously constructed atlas: down-sampling of new subject tractography, registration to atlas of the sampled tracts, similarity metrics computation between atlas and sampled tracts, sample tracts label attribution and label propagation from sample tracts to all the tracts.

4.4 Similarity Metrics

Similarity is usually based on the point-wise spatial distance between the points p that represent the tracts t. There are several choices on how to define the similarity between curves. Moberts (2005) showed that the *Mean of Closest Points* (MC) distance gave the best results for computing accurate similarity for adult brains between two tracts t_i and t_k represented by r and l points correspondently (see equation 4.1).

$$d_{MC}(t_i, t_k) = \mathrm{mean}_{\mathrm{p_r} \in \mathrm{t_i}}(\, \mathrm{min}_{\mathrm{p_l} \in \mathrm{t_k}}(\, \|p_r - p_l\| \,)\,)$$

Equation 4.1 – Mean of closest points (MC) distance.

However, for representing similarity between neonatal tractography, this distance measure is not optimal. As already mentioned, neonatal tractography can present broken tracts. Normally, when two tracts cross, the MC will be large, because the mean of all distances of the two tracts will be taken into account. When one tract is shorter than the other, the MC distance between them will be small, and can be equivalent to two parallel tracts with a reasonable distance. This is an undesired situation for neonatal brain imaging since tracts that cross have different orientation and therefore should be classified as non-similar.

Another distance that can be used is the *Classic Hausdorff (*see equation 4.2). The *Classic Hausdorff* distance similar to the MC, however it computes the maximum distance between the closest point pairs instead of the mean.

$$d_{Classic\ Hausdorff}(t_i, t_k) = \mathrm{max}_{\mathrm{p_r} \in \mathrm{t_i}}(\, \mathrm{min}_{\mathrm{p_l} \in \mathrm{t_k}}(\, \|p_r - p_l\| \,)\,)$$

Equation 4.2 – The Classic Hausdorff distance.

The *Classic Hausdorff* is a directed distance, i.e., the distance from tract *i* to tract *k* can differ from tract *k* to *I (*see figure 4.9). For transforming it in a symmetric distance the maximum of these two distances is used (equation 4.3). This is a conservative distance; tracts will only be similar when all points have small distances.

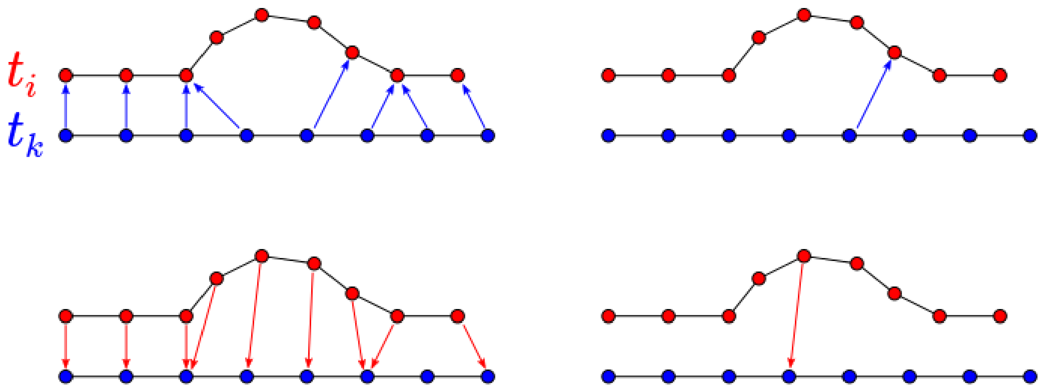


Figure 4.9 – The Classic Hausdorff distance as a direct type distance. The Classic Hausdorff distance is different between tract k to tract i, first line of image, with tract i to tract k, second line. The first row depicts the first step of the Classic Hausdorff distance, with the distance between the points of the study tract with its closest tract points. The second row depicts the final step of taking the maximum of the previous computed distances.

$$d_{conservative\ Hausdorff}(t_i, t_k) = \max\ (\ d_{Classic\ Hausdorff}(t_i, t_k)\ , d_{Classic\ Hausdorff}(t_k, t_i)\ )$$

Equation 4.3 – Classic Hausdorff distance made conservative by taking the maximum of the two possible directed Classic Hausdorff distances for a pair of tracts.

In the neonatal brain, also a broken tract can run parallel with a tract with a larger length. The *Symmetric Hausdorff* overestimates the distance for two parallel tracts with very different lengths. The *Adapted Hausdorff* [Boom, 2011] is still a symmetric distance, but instead of taking the maximum of the two possible direct distances, it takes the minimum of these two. In conclusion, the *Adapted Hausdorff* solves the problem of broken tracts, as it does not overestimate the distance between two parallel tracts of different length, or underestimates the distance between two crossing tracts of different lengths. In figure 4.10 these two situations are depicted.

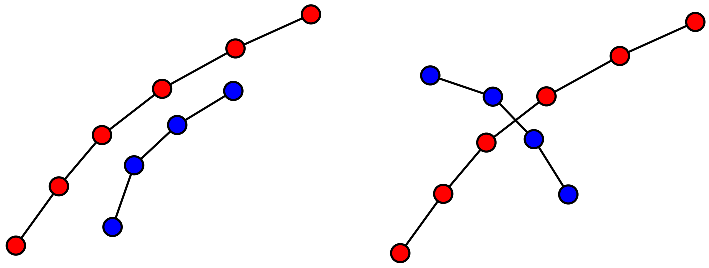


Figure 4.10 – Pair of tracts with very different lengths due to presence of broken tracts for neonatal data. From left to right: parallel tracts; crossing tracts.

$$d_{Adapted\ Hausdorff}(t_i, t_k) = \min\ (\ d_{Classic\ Hausdorff}(t_i, t_k)\ , d_{Classic\ Hausdorff}(t_k, t_i)\ )$$

Equation 4.4 – The Adapted Hausdorff distance between tract t_i and t_k.

4.5 Registration to Atlas

In order to maximize the accuracy of the distance-based similarity computation, the tractography results from the new subject need first to be geometrically aligned with the atlas tracts. This correction is necessary due to two factors. The first arises from the fact that MRI scans of different patients are not always acquired in exactly the same position, so position and orientation of the datasets usually differ across patients. A second factor comes from variation in head size and shape among subjects. Consequently, a good registration will normalize the subject variability, having the atlas as reference. However, the registration process should not alter the anatomical features specific for each structure, such as the structures relative location to each other and the different shapes of tract bundles. These are the features necessary to maintain for an accurate computation of similarity between atlas tract set and the new subject tract set, necessary for the anatomical labeling of the subject tracts set.

The process of finding the best alignment of the new subject tract set with the atlas tract set corresponds to an optimization problem. Alignment of the subject tracts with the atlas tracts involves applying a coordinate transformation model to the subject tracts points. Finding the best alignment comprises a study of similarity between subject and atlas tracts. Search for the highest similarity, considering the deformation model applied, can be performed in different domains: atlas and subject can be compared at scalar-domain, tract-domain or even feature-domain.

4.5.1 Transformation Models

Transformation models for registration purposes cover a high range of possibilities: from global rigid deformation [Leemans, 2006], where only translation and rotation are allowed, until fluid models based on physical processes which allow different local deformations [Zitova, 2003]. Considering tractography registration, the most used transformation corresponds to the Affine type, where not only translation and rotation are allowed but also scaling and shearing [O’Donnell, 2012; Mayer, 2011; Zvitia, 2011]. Deformations more complex than affine, like elastic or fluid deformations, were not found in literature for registering tractography. Indeed even for voxel-wise registrations these more complex deformations need to be used with care. Although they allow for a more detailed alignment, they can corrupt the local spatial variability characteristic of each tract population which needs to be preserved [Murgasova 2011].

4.5.2 Similarity Metric Search Domain

Although the goal of the registration is to align the subject tractography with the atlas tractography, search for the best alignment can be performed by using other DTI-derived data. The most commonly used data corresponds to scalar information extracted from the DT-MRI images, like anisotropy maps [Anjari, 2007]. One advantage of performing tractography registration based on scalar images is that classical intensity based registration methods can be used. Boom (2011) used this type of registration for constructing the atlas. The atlas tractography registration was conducted by employing a transformation to the tracts that was found by comparison of Cl anisotropy maps between subjects and reference. Boom also applied the same registration procedures to register new subjects to her atlas. This approach showed to be problematic, as for more than 10% of subjects the registration failed without any specific reason being elucidated. When working with anisotropy maps, not all the directional information contained in the tensors is being used. In addition, since every voxel of the data set is included, this approach can be inefficient, as many voxels contain no information about the WM structure and can be seen as not contributing to the real optimization process. [Hutton, 2003; Zitova, 2003]

Another option is to perform the registration directly between the tracts, i.e., by searching at tract-domain [O'Donnell, 2012; Mayer, 2011; Wasserman, 2010; Zvitia, 2010; Mayer, 2007; O'Donnell, 2007; Leemans, 2006]. For this approach, the directional information contained in the diffusion tensors is naturally inherited by the reconstructed WM tracts and further enhanced by the connectivity information described by each tract. This approach then uses global information as opposed to the scalar-wise approach, which uses only local information. In addition, with the eventual goal of modeling and analysis of WM tracts, it may be advantageous to register the tracts themselves, as the quantity being optimized during registration will be closely related to the final goal.

Drawbacks from directly registering tractography data can arise from three sources. First, tractography registration focuses only on the tensor information contained in the first eigenvector, in regions where anisotropy values are high enough for reconstructing WM tracts. By not considering information from outside these regions, i.e., information from GM regions, it might be argued that information of added value is being discarded. Secondly, tractography registration requires extra computation, a quadratic increase, as each tract of the new subject is compared with each of the tracts from the reference tractography. Finally, any possible errors on the tractography procedure will be propagated to the registration.

As for the feature domain, it corresponds to reducing the tract to some geometrical features, like curvature and length. Each tract is then mapped into a feature space, and similarities are computed there [Leemans, 2006; O'Donnell, 2007]. Problems arise with the fact that it is not clear which are the best geometrical features for describing similar WM tracts.

4.6 Down-Sampling

Registration time and accuracy can be improved by reducing the number of tracts fed to its algorithm. In this way the complexity of registration is reduced. This can be performed automatically by sampling methods. All important anatomical structures should continue to be tract-represented after the sampling. However, for neonatal data, some anatomical structures, e.g. like the cingulum (CG), contain no more than 5-8 reconstructed tracts. For including in the sample a representation of anatomical bundles with a reduced number of tracts, a new clustering approach is suggested in this thesis. The sampled data-set is constituted by an existent tract from each cluster that is the cluster center tract. The specific algorithm used for the current work will be presented in the next chapter.

Chapter 5

Pipeline Implementation

The previous chapters laid the foundation for the proposed method: the importance of this type of research direction was clarified, tractography type data and its usability were elucidated, and plus its possible drawbacks and the methods probed so far for segmenting neonatal tractography were discussed. In this chapter, the contribution of this thesis to the field of neonatal atlas-based tractography segmentation is explained.

First, a schematic overview of the proposed pipeline flow is given. For some of the steps previous work was mostly not altered, while for other steps major alterations were implemented. Input data, output data, parameter settings, and implementation details are also presented for each processing step section.

5.1 Pipeline Flow

How WM tractography arises from DWIs was clarified in chapter 3. Then in chapter 4, the processing steps from raw tractography to segmented tractography were also explained. The pipeline in this thesis consists of both parts, as it starts from the DWIs acquired, to attain in the end a segmented tractography. Consequently, eight main processing steps can be identified (see also figure 5.1): skull stripping, tensor fitting, tractography, down-sampling, registration to atlas, computation of similarities between subject and atlas, label attribution and label propagation. In appendix I, another scheme is depicted, this one focusing on the data flow from raw tractography to segmented tractography.

Implementation

Previous to skull stripping, the scanned data is converted from the standard scanner data type, Digital Imaging and Communication in Medicine (DICOM), to the standard neuroscience data type, NifTi (Neuroimaging Informatics Technology Initiative). This data conversion is performed with dcm2nii software [http://www.mccauslandcenter.sc.edu/].

The pipeline runs on Matlab®. Some steps were written in a different language; therefore, executables for these steps were created and then called via Matlab®. Three libraries are necessary for the pipeline, a Matlab® function library called VISTASoft [http://vistalab.stanford.edu/] for processing of medical images, the VTK library [http://www.vtk.org/] for tract representation and visualization, and the O'Donnell library for registration of WM reconstructed tracts [O'Donnell, 2012]. The entire pipeline runs automatically, and the user does not need to tune any parameter. There is only one exception: the user needs to manually initialize the tractography algorithm in the pipeline. This is done in vIST/e, a software tool developed at the Biomedical Image Analysis Group from the University of Eindhoven [http://bmia.bmt.tue.nl/software/viste/]. After tractography, the pipeline continues automatically.

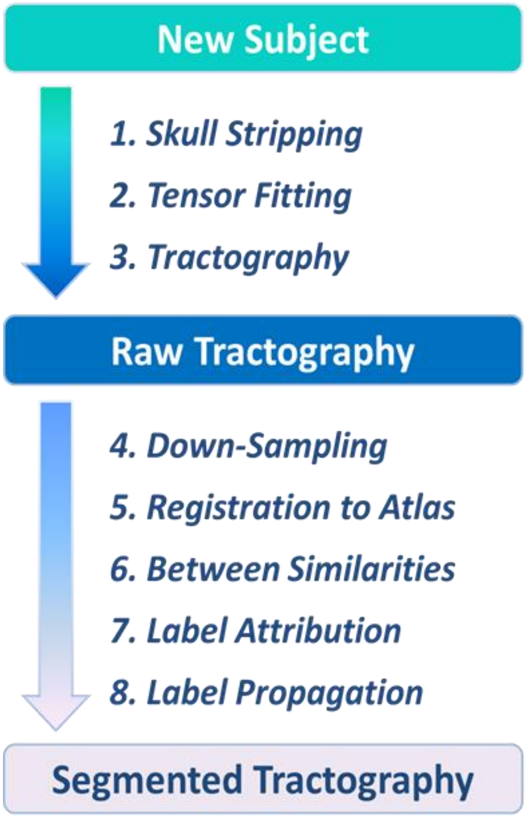


***Figure 5.1** – Pipeline flow. The flow from raw tractography to segmented tractography was already presented on the previous chapter. Now, also the procedures from the scanned image data to its raw tractography are depicted.*

5.2 Skull Stripping

The skull stripping is performed as a first step, before fitting the tensors. Its constitutive steps are depicted in figure 5.2, and its effects on the tractography dataset in figure 5.3.

It takes as input the b0 image (see section 3.5). The used algorithm is similar to the one proposed by Hoskam (2009) and Boom (2011), as it also applies image morphological operators; more specifically one dilation step and two erosions. For the current pipeline, the morphological operators are extended from simple dilations and erosions to more complex operations, referred as smart morphological operations, and a threshold step is performed previous to these. This is done with the aim of reducing spurious tracts pervious to segmenting the tractography.

The threshold step analyses the intensity histogram of the b0 image. The b0 image is clipped according to a lower and upper intensity threshold. The clipping is special, as the voxels with intensities bigger than the upper threshold are clipped to 0, instead of the normal approach of clipping to the upper threshold value. The upper and lower thresholds are defined in percentile values, i.e., percentages of the intensity value distribution, see table 5.1, percentile parameters. After clipping, the intensity values are transformed in a logical mask, where all intensity values bigger than 0 are assigned to the logical value true.

This mask is then further processed by application of morphological operators. First, satellites are removed. For this, voxels are grouped by a 3-dimensional 26-connectivity, i.e., voxels are considered neighbors if they touch each other faces, or edges, or corners. Groups with less than a predefined number of voxels, see table 5.1 satellite size parameter, are seen as satellites, and their logical value changes to 0. Secondly, holes are filled by using a smart closing operator. A smart closing operator is defined similar to a normal closing operator, i.e., dilation followed with erosion, having in addition the property of conserving the initial mask outside contour. It does so by comparing the after-closing mask with a complementary image of the before-closing mask [Vicent, 1993]. The structural element comprises a 3-dimensional 6-connectivity, i.e., voxels are considered neighbors only if they touch each other faces.

Thirdly, the image is smoothed by a 3-dimensional Gaussian kernel, see table 5.1 sigma size parameter. The smoothed result is again converted into a mask by assigning the value true to voxels higher then 0.5; this can be seen as a second erosion procedure. After smoothing, the mask gets its satellites removed and holes filled one more time. The mask is then applied to the multiple DW images. The DW images are already aligned between themselves by the scanner software for correction of eddy current and motion artifacts.

Figure 5.2 – Skull stripping consecutive steps. First, b0 image is clipped and transformed into a mask. Then this mask passes through morphological operations which allow for satellite removal, then hole filling and after smoothing. After smoothing, step 2 and 3 are repeated.

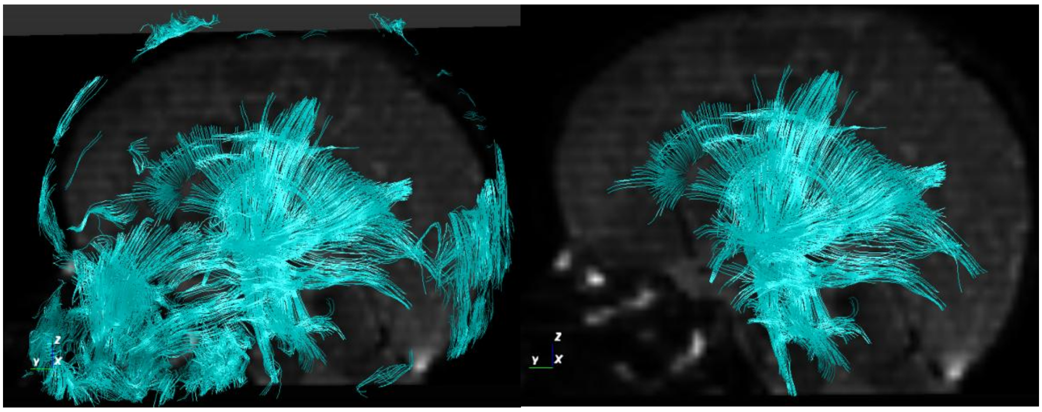


Figure 5.3 –Effects of skull stripping in tractography results with b0 image in side plane. Left: No Stripping applied; Right: Tractography results from a tensor image for which skull stripping was performed.

Implementation

VISTASoft was the basis for implementing the skull stripping step. Before the implementation at Matlab®, some threshold methods were probed using the software ImageJ [http://imagej.nih.gov/]. For the tested subjects, the percentile threshold performed better in comparison with other threshold methods available at ImageJ. In table 5.1 the main functions and corresponding parameters can be found. Taking into account the differences between neonatal and adult brain (see section 4.2, paragraph 3), an empirical investigation of the optimal parameters for neonatal datasets was conducted. Parameter values which generated a skull stripping mask containing holes in the brain region were disregarded. Optimal parameters were seen as the ones which did not result in these holes and contained less amount of non-brain region, like facial and neck area and above the skull regions.

Step	Functions	Parameters	
Clipping	mrAnatHistogramClip.m, modified for clipping to 0	Percentiles (%)	85% 99.9%
Remove Satellites	dtiCleanImageMask.m	Satellite size (voxels)	100
Fill Holes	imfill.m		-
Smoothing	dtiSmooth3.m	Sigma size (voxels)	10

Table 5.1 – Skull stripping functions and correspondent parameters.

5.3 Tensor Fitting

Tensor fitting is performed with the simplest and more common algorithm, the Linear Least Square (LLS) (section 3.5).

Implementation

VISTASoft library functions were again chosen for this implementation. dtiRawFitTensor.m was slightly modified for allowing fitting with standard LLS and posterior tractography processing.

5.4 Tractography

Tractography is performed with a deterministic full brain seeding algorithm (section 3.6).

Implementation

Tractography was performed by vIST/e plug-in *Fiber-Tracking*. Pul (2004, 2012) specific tractography parameters for neonates were used (see table 5.2). These parameters have especially taken into account the lower anisotropy values, and smaller sizes of WM anatomic regions in neonatal DTI in comparison with adult DTI.

Parameter name	Value
Minimum tract length	20 mm
Minimum anisotropy index (0 - 1)	Case liner (Cl) of 0.12
Maximum tract angle	10°
Minimum seed distance	0.5 voxels

Table 5.2 – Tractography constrains applied from Pul (2012).

5.5 Down-Sampling

Down-sampling step starts by computing distance-based similarities for each pair of tracts within the subject tractography. These similarities are saved in a matrix format and constitute the input of the clustering algorithm. Next, clustering is performed and finally the sampled dataset is made of the cluster center tracts, which emerge from the clustering algorithm. This process is shown in figure 5.4. The subject tract set and the atlas tract set are both down-sampled prior to registration. The atlas tract down-sampled set is presented in figure 5.5

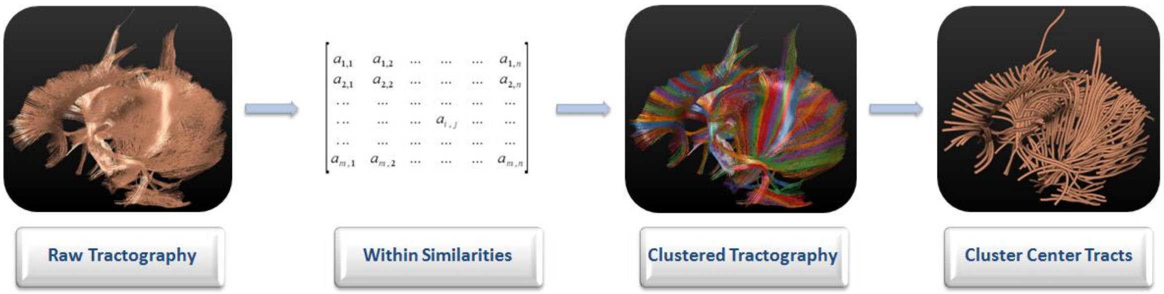


Figure 5.4 – Down-sampling data flow. The input data corresponds to the subject raw tractography. After, tract pair distance-based similarity is computed. The clustering algorithm uses then these similarities for clustering the tracts. The cluster centers become the down-sampled dataset.

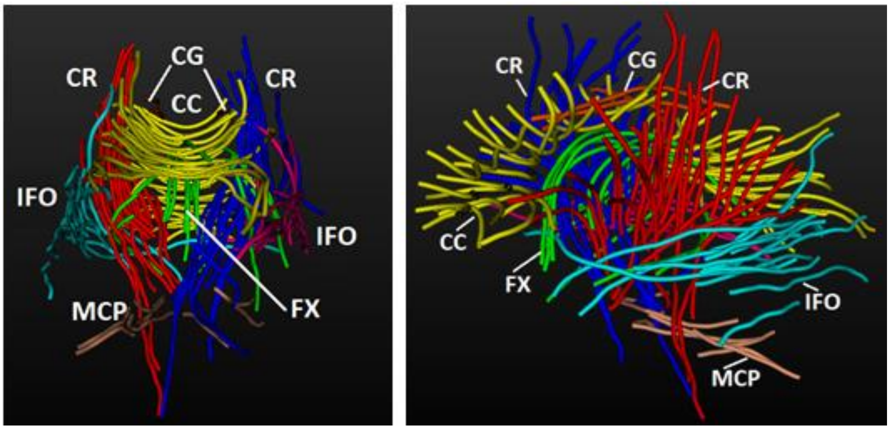


Figure 5.5 – The atlas down-sampled set by the proposed method of cluster down-sampling.

Affinity Propagation (AP) [Frey, 2007] is the clustering algorithm implemented. From the studied clustering methods, AP produced clusters containing reconstructed tracts from not more than one WM anatomic structure [Boom, 2011; Leemans, 2009]. In addition, AP can take as input general similarities like the Adapted Hausdorff distance, which for the neonatal preterm dataset is clearly an advantage (section 4.4).

AP algorithm clusters by studying distance-based similarity between the tract pair and their neighbor tracts. It initially considers every tract simultaneously as a potential cluster center and iteratively exchanges messages between the tracts until a set of cluster centers and corresponding clusters emerge. For the current implementation, distance, $d(t_i, t_k)$, is taken into account in the similarity metric by computing the negative squared product of the Adapted Hausdorff distance proposed by Boom (2011), $d_{Adapted\ Hausdorff}(t_i, t_k)$ (see equation 5.1).

$$d(t_i, t_k) = -\, d_{Adapted\ Hausdorff}(t_i, t_k)^2$$

Equation 5.1 – Negative squared Adapted Hausdorff distance as the difference metric desired to minimize between tracts t_i and t_k. See equation 4.4 for the expression of the Adapted Hausdorff.

There are two types of iteratively exchanged messages which measure how appropriate it is for t_k to be the cluster center of t_i, responsibilities and availabilities. Responsibility, $r(t_i, t_k)$, takes into account that t_i can have other more appropriate cluster centers. It does so by studying how far away t_i is from t_k, and the best competitor of tract t_k for being the cluster center of t_i (see equation 5.2 and left side of figure 5.6).

$$r(t_i, t_k) = d(t_i, t_k) - \ max_{k', k' \neq k} \{\, a(t_i, t_{k'}) + \ d(t_i, t_{k'}) \,\}$$

Equation 5.2 – Responsibility of t_k to be cluster center of t_i.

Availability, $\boldsymbol{a(t_i, t_k)}$, takes into account the other possible tracts for which $\boldsymbol{t_k}$ can be considered an appropriate cluster center. It does so by studying the responsibility of $\boldsymbol{t_k}$ to $\boldsymbol{t_i}$, and all the positive responsibilities of $\boldsymbol{t_k}$ for being the cluster center of other tracts (see equation 5.3). Self-availability, $\boldsymbol{a(t_k, t_k)}$, is computed differently as the sum of all the positive responsibilities of tract $\boldsymbol{t_k}$ (see equation 5.4 and right side of figure 5.6).

$$a(t_i, t_k) = \ \min \{\, 0, r(t_k, t_k) + \sum_{i' \notin \{i,k\}} \max \{\, 0, r(t_{i'}, t_k) \,\} \,\}$$

Equation 5.3 – Availability of t_k to be cluster center of t_i.

$$a(t_k, t_k) = \sum_{i', i' \neq k} \max\{\, 0, r(t_{i'}, t_k) \,\}$$

Equation 5.4 – Self-availability of tract t_k.

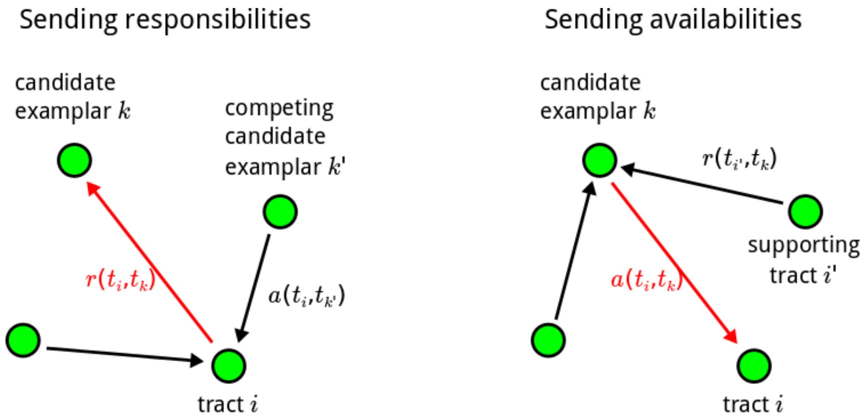


Figure 5.6 – Exchanged values between tracts of a tractography data-set.

Implementation

Computation of similarity metrics was performed in Matlab®. The computation of the correspondent distances is performed on C++ from adaptation of previous work from Boom (2011) and Moberts (2005). The clustering algorithm was made available by Frey in a Matlab function, apcluster.m [Frey, 2007]. There is only one parameter which needs to be tuned, the initial responsibility of each tract. Currently, for the approach followed in this thesis, there is no known a priori information about which tracts are more appropriated to become cluster centers and constitute the down-sampled dataset. Therefore, initial responsibility is made equal for al tracts. This parameter was empirically found by Boom (2001) to be optimal at the value of -200. Although Boom (2001) did not use the AP algorithm for sampling purposes, she empirically found that for typical preterm neonate at TEA tractography data, this value produced an optimum number of clusters (see figure 4.3, section 4.1).

5.6 Registration to Atlas

The deformation model used for registering the new subject to the atlas corresponds to an affine transformation of the subject tracts [O'Donnell, 2012; Boom, 2011], presented at sub-section *"5.6.1 Affine Transformation of Tracts"*. Registration is performed directly between the tracts, by studying the similarities between the down-sampled new subject tracts with the down-sampled atlas tracts. The specific similarity metric corresponds to a probability density function [O'Donnell, 2012; Wasserman, 2010; Zvitia, 2010] and its concept is clarified in sub-section *"5.6.2 Similarity Probability Density Function"*. The optimization algorithm, responsible for searching for the best transformation, is presented at sub-section "*5.6.3 Similarity Optimization Algorithm*". As a registration output example see figure 5.7.

The registration algorithm was inspired by the group-wise registration from O'Donnell (2012). The main difference between the used implementation and O'Donnell's method, it is that the implemented method aims to find the best alignment between a group of tractographies by giving a reference tractography for performing the alignment; the atlas is the reference for alignment. Thus, in O'Donnell's work, all the input tractographies can be deformed, while searching for the best alignment between all of them. While in the current pipeline, only one tractography is deformed, the subject's tractography.

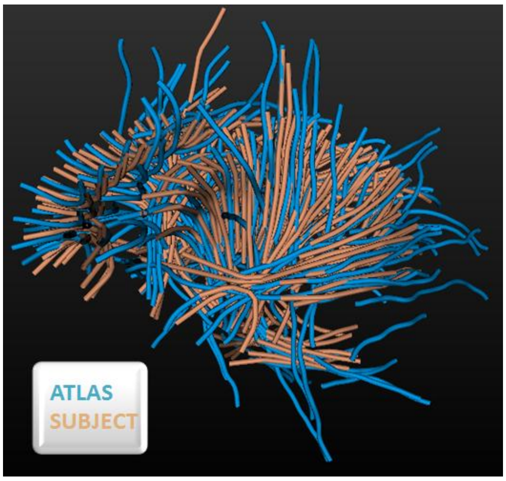


Figure 5.7 – Registration of subject cluster center tracts to atlas cluster center tracts. The subject is deformed to the atlas as reference. Subject depicted in orange and atlas in blue.

5.6.1 Affine Transformation of Tracts

Each subject tract from the down-sampled subject set, t_s, is represented by a set of points, $p \in t_s$. An affine transformation τ can be expressed as a matrix which acts on the coordinates (x_p, y_p, z_p) of every of these points p (see equation 5.5). The different entries of the transformation matrix are depicted as $tm_{i,j}$.

$$\tau(p) = \begin{bmatrix} tm_{11} & tm_{12} & tm_{13} & tm_{14} \\ tm_{21} & tm_{22} & tm_{23} & tm_{24} \\ tm_{31} & tm_{32} & tm_{33} & tm_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_p \\ y_p \\ z_p \\ 1 \end{bmatrix}, p \in t_s$$

Equation 5.5 – 3-dimensional affine transformation.

In order to accelerate convergence, prior to registration and after down-sampling, the subject down-sampled tract set center of mass is aligned with the center of mass of the down-sampled tract set atlas.

5.6.2 Similarity Probability Density Function

For each iteration of the registration algorithm, a different affine transformation is probed and applied to the subject tracts. For each probed transformation, corresponding similarity metrics between all the subject tracts with all the atlas tracts are computed. This means, the similarity metrics between the atlas and the new subject are computed as many times as the algorithm iterates. Similarity metrics are distance-based by definition. Computing distances between two tracts corresponds to studying all the possible pair of points between the pair of tracts. Consequently, computation of distance-based similarity metrics for registration purposes is very computationally expensive.

O'Donnell proposes a distance which does not study all possible pair of points between the subject tract and the atlas tract. Instead, the pairs of points between the two tracts are fixed: the first point of *tract* $\boldsymbol{t_s}$ always pairs with the first point of tract $\boldsymbol{t_a}$, the second point of tract $\boldsymbol{t_s}$ pairs always with the second point of tract $\boldsymbol{t_a}$, and so forth until the last point.

However, the number of points differs between tracts, as it is dictated by the tract length and the tractography parameter of the space between tract points (see section 3.6). Making the number of points per tract uniform is a compromise between fidelity to the original tract shape and computational complexity. It was earlier empirically found by Mayer et al. (2011) and O'Donnell et al. (2011) that a good balance between these two factors can be attained with a re-description by five equally spaced points: endpoint, midpoint, and two intermediate points. If necessary, the new points are found by applying a nearest neighbor interpolation. This re-description of the tract by five equally distant points is depicted in figure 5.8.

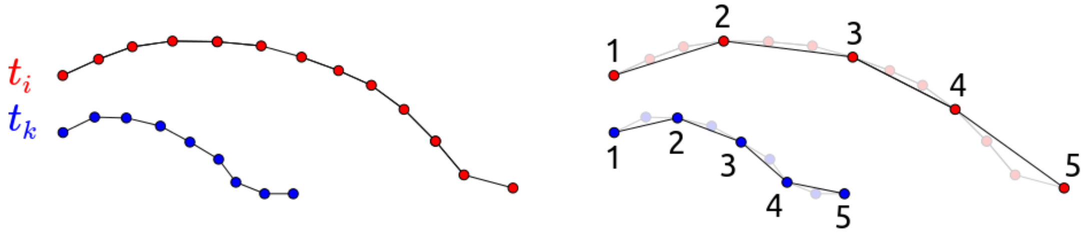


Figure 5.8 – Tract re-description of tracts t_i and t_k using five equally distant points.

The distance proposed by O'Donnell corresponds then to using the maximum distance of the five distances that arise from the five fixed pairs between points. Because point ordering along the tract is not known a priori, the maximum distance of correspondent pairs is computed twice, the second time with an inverse ordering for one of the tracts. The final pair-wise tract distance corresponds to the minimum of these two maximums (see equation 5.6).

$$d_{O'Donnell}(t_s,t_a) = \min\left(\max_{\mathrm{p_r}\in \mathrm{t_s}} (\|p_r - p_l\|_{\mathrm{p_l}\in \mathrm{t_a}, \mathrm{r=l}}) \ , \max_{\mathrm{p_r}\in \mathrm{t_s}} (\|p_r - p_l\|_{\mathrm{p_l}\in \mathrm{t_a}, \mathrm{r=6-l}}) \right)$$

Equation 5.6 – O'Donnell distance between tract t_s and t_a. Points of t_s are referred as p_r, points of t_a as p_l.

The distance between points is transformed to a probability density function $\delta(t_s,t_a)$ (see equation 5.7). This pdf describes how probable it is for that subject tract to have close-by neighbor atlas tracts. However, it is important to understand that this function is only defined for a positive domain, as distance is always a positive value. In addition, as a pdf it will take only values between 0 and 1. σ allows the definition of a radius of interest; where distances outside this radius will have similarities close to 0. Z is normalization constant, and therefore will not influence the optimization procedure. The distance metric used corresponds to the previously defined in equation 5.6.

$$\delta(t_s,t_a) = \frac{e^{-[\frac{d(t_s,t_a)}{\sigma}]^2}}{Z}$$

Equation 5.7 - Similarity metric as an affinity probability density function. $d(t_s,t_a)$ corresponds to the O'Donnel distance between tract t_s from the subject and tract t_a from the atlas defined in equation 5.6.

Taking into account all the pdf's of a subject tract to all the atlas tracts $\delta(t_s, \mathrm{A})$, allows to describe how probable it is for that subject tract to have close-by neighbor atlas tracts (see equation 5.8).

$$\delta(t_s, \mathrm{A}) = mean_{t_a \in \mathrm{A}}\ (\ \delta(t_s, t_a)\ )$$

Equation 5.8 – Similarity probability distribution function of tract t_s to all the atlas tracts, t_a ∈ A.

5.6.3 Similarity Optimization Algorithm

The best alignment between subject tracts with atlas tracts can be found by maximizing the probability of all the subject tracts having close-by atlas tracts. Information theory attains this by minimizing the entropy of this distribution [Maes, 1997; Collignon, 1995; Shannon, 1948]. The Shannon Entropy (H) corresponds to the expected value (E) of the negative log-probability of the subject tracts. By the weak law of large numbers [Poisson, 1837], the expected value can be replaced by the sample average value. This concept is the foundation for the construction of the cost function used by the optimization problem (see equation 5.9). The average was obtained by dividing per the number of tracts in the atlas tract set. It might be pointed that dividing by the number of tracts in the subject tract set would be more accurate for computing the average. However, this value n_A is a constant through the entire algorithm, and therefore will not have a influence in the optimization search.

$$H(\tau(t_s), \mathrm{A})\ =\ E\ (\ \log\frac{1}{\delta(\tau(t_s), \mathrm{A})}\ )\ = [\sum_{t_s \in \mathrm{S}} \log\frac{1}{\delta(\tau(t_s), \mathrm{A})}\ ] \ / \ n_A$$

Equation 5.9 – Cost function as the subject tracts similarity probability distribution functions entropy. H is the Shannon entropy, τ the affine transformation performed at the subject tracts, A the set of all atlas cluster center tracts, S the set of all subject cluster center tracts, and δ the similarity probability distribution function of tract t_s to all the atlas tracts, t_a ∈ A. n_A is the number of tracts in set A.

The optimization algorithm searches for the optimal transformation, $\hat{\tau}$, in the space of all affine transformations, $\tau \in \mathcal{T}$ (see equation 5.10). The algorithm sees the optimal affine transformation, as the one that, once applied to the cluster center subject tracts, brings the minimal value for the cost function defined in equation 5.9.

$$\hat{\tau} = argmin\, H(\tau(t_s), \mathrm{A}), \tau \in \mathcal{T}, t_s \in S$$

Equation 5.10 – Optimal affine transformation, $\hat{\tau}$, as the argument which brings the cost function, H, to its minimum. Whereas τ is the affine transformation applied to the subject tract points coordinates, $\mathcal{T}$ is the space of all possible affine transformations, A the set of all atlas cluster center tracts, and S the subject cluster center tract set.

The space of all affine transformations comprises twelve degrees of freedom (dof). Therefore, the optimization problem is a multi-variable problem. In addition, there is no analytical solution for the derivative of this cost function. Consequently, the algorithm applied is a direction set method, where the optimal function is reached by linear approximations. The multiple variables are represented as vertices of a simplex geometrical (see figure 5.9). Both size of the simplex and size of neighborhood for next vertex interpolations are constrained by a spherical region of trust. Constraint parameters correspond then to the expected initial step size, *ρbeg*, and final step size, *ρend*, for each variable. The *ρend* is responsible for determining convergence. The algorithm was first proposed by Powell on 1998, and is usually referred by the name of COBYLA, which stands for Constrained Optimization by Linear Approximation (Powell, 1998).

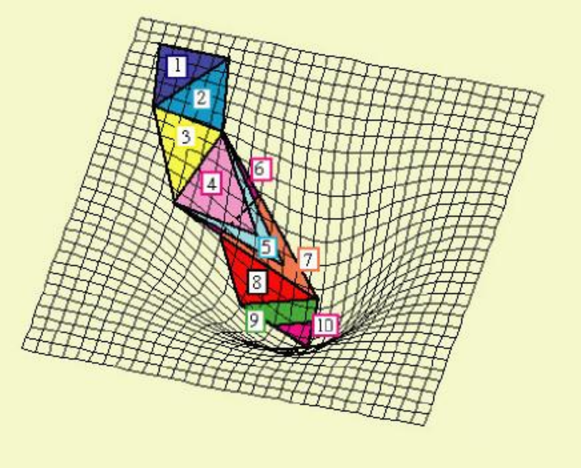


Figure 5.9 – Representation of a simplex direction set method for a 3-variable optimization problem. The simplex is seen as a triangle in a 3D space, in which each of its vertices corresponds to one of the 3 variables. 10 iterations are represented in direction of the (local) minimum of the 3-D cost-function, depicted as a grid surface. (image from http://www.princeton.edu/)

Implementation

The registration step implementation was based on a python library developed by Lauren O'Donnell's research group [O'Donnell, 2012]. Classes for description of the reconstructed WM tracts, similarity metric, transformation model, and optimization algorithm were used without any modification besides parameter tuning. The optimization algorithm was performed with the python toolkit scipy.optimize [http://www.scipy.org/] which already includes a COBYLA package. Main modifications were performed in the cost function definition, as for the current pipeline a pair-wise registration was of interest, instead of the group-wise method implemented by O'Donnell.

One of the more important parameters to empirically tune was the similarity pdf standard deviation value, σ. For adult tractography, optimal results are found by running several iterations for different σ, from 30 mm to 5 mm. For the neonatal tractography data, the results were optimal for the value of 3 mm, and iterations were not found to produce any observable improvements.

As for the constraint parameters, *ρbeg* and *ρend* were tuned in order to avoid an unnecessary overall rotation or translation of the tracts, and to avoid the shrink to a point solution that artifactually reduces entropy, table 5.3. This was done by taking into account the expected transformation variable magnitudes. Different parameter combinations were empirically probed, with O'Donnell's empirical parameters for a typical adult dataset performing the best.

	ρbeg	ρend
Rotation (degrees)	5	0.1
Translation (mm)	5	0.1
Scale (factor)	0.01	0.001
Shear (degrees skewed)	2	0.5

Table 5.3 – COBYLA constraint parameters [O'Donnell, 2012].

It is also important to remark that for the type of data tested, the best results were not found while optimizing simultaneously in a domain of twelve dofs. Instead, they were found while optimizing iteratively between four distinct deformation modes of translation, rotation, scaling and shearing. For each iteration, the optimization probes then in a domain of only three dofs, one for each spatial coordinate. O'Donnell also proposes an iterative optimization in her implementation. However, for the tested dataset, a simpler iteration sequence than the one proposed by O'Donnell was found to give the best results. This iteration corresponds to first optimizing using only translation, then rotation, scaling, and then shearing. Due to the difficulty of registering the cingulum

(CG) cluster center tracts for some subjects, the modes of translation and rotation are performed one more time.

5.7 Between Similarities

After registration, another computation of similarity is performed, between every registered cluster center tract and every cluster center atlas tract. The metric used corresponds again to the distance-based metric suggested by Boom (2001), the *Adapted Hausdorff* (4.4 Similarity Metrics). This is a computationally expensive metric, since it studies all the pair of points between tracts. For this step, differently from the registration step, the metric only needs to be computed once; and differently from the down-sampling method, it is computed between two down-sampled tract sets, consequently between a reduced number of tracts.

Implementation

The computation of the correspondent distances is performed on C++ from an adaptation of previous work from Boom (2011) and Moberts (2005). Distances are loaded in Matlab® as a matrix. As oppose to the distance matrix from the down-sampling metric, it is important to have in mind that this matrix will not be square or symmetric.

5.8 Label Attribution

Each cluster center registered tract receives the label of the atlas cluster center tract with which it has the highest similarity, i.e., the lowest *Adapted Hausdorff* distance.

Implementation

This attribution is performed in Matlab®. Besides the "between-similarities", also the key of atlas cluster center tract identification number with anatomical label need to be assessed. A label key is created for the cluster center tracts. The label key corresponds to a file which connects tract identification number with anatomical label. Because identification number of the cluster center tracts does not change with the registration procedure, this key is valid both for un-registered and registered cluster center tract sets. Visualization of such intermediate results, like the one represented at the left side of figure 5.10, can be performed in vIST/e, using the *Clustering* plug-in.

5.9 Label Propagation

The label of the cluster center tract is finally propagated to all the tracts which belong to the cluster they are representing (see right part of figure 5.10). It is therefore possible to return to the full tractography dataset, without any deformation performed to the tracts.

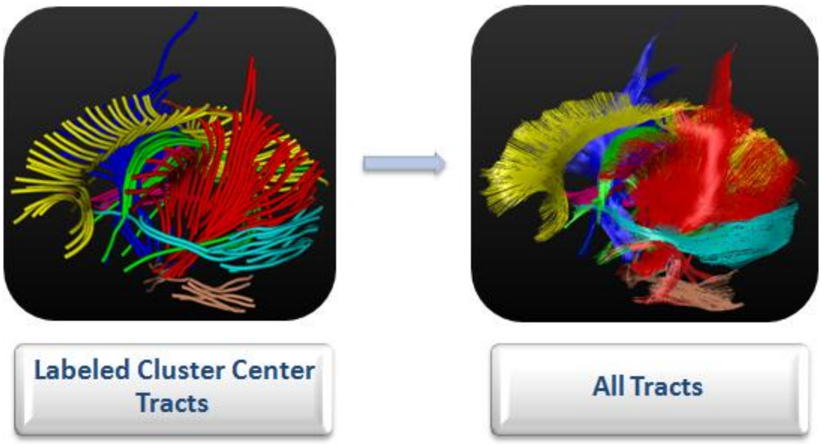


Figure 5.10 – Label Propagation from cluster center tracts to all subject tracts.

Implementation

Two keys are now given as input, the key with the correspondence between cluster center tract and the tracts which belong to that cluster and the key with the anatomical label for each cluster center tract. The produced output corresponds to a key of all tracts identification number with correspondent anatomical label. The final segmentation results can be visualized in vIST/e using the *Clustering* plug-in. Besides qualitative evaluation, vIST/e also allows for quantitative evaluation as anisotropy values can be generated for each anatomical bundle of tracts. This evaluation will be further explored in the next chapter.

Chapter 6

Experimentation and Evaluation

The present chapter presents the evaluation of the segmentation results for clinical data. First the dataset used to test the pipeline is introduced. The evaluation/analysis of results is divided in two sections: the first section centers on a qualitative score regarding performance of segmentation; the second section focuses on the volume and anisotropic values of the segmented anatomical structures.

6.1 Dataset

Both the University Medical Center Utrecht (UMCU) and the Maxima Medical Center (MMC) provided clinical data for this study. The patient data from UMCU was part of a larger cohort that has been published previously (Kooij 2011, Pul 2012). The use of patient data from the UMCU was approved by the Institute's Medical Ethical Committee and parental informed consent was obtained. For the MMC, the scans are part of routine clinical practice and as the method is used for quality control to improve the image processing chain, no ethical approval was necessary. 54 scans were included in this study (23 from UMCU and 31 from MMC). The neonates were all scanned at term equivalent age (TEA). For a full-term neonate this corresponds to its birth date +10 days later. As for a preterm neonate, this corresponds to approximately the date on which she/he would be born if they would have been inside their mother's womb for the standard gestation period of 40 weeks.

The data was acquired on a Philips Achieva 3.0T MRI-scanner, in both hospitals and the DTI sequence was almost similar, except for the SENSE factor. The DTI-sequence is performed with Sensitivity Encoding (SENSE) factor 3 (UMCU) and factor 2 (MMC - in order to reduce artifacts), b-values 0 and 800 s/mm2 in 32 directions using a single-shot-EPI sequence. The use of this sequence allows for reducing scan time, and therefore reducing the influence of motion artifacts; each sequence took less than 5 minutes. The dataset consists of 50 adjacent slices, each slice with 128x128 voxels, each voxel corresponds to a size of 1.44 by 1.41 by 2 mm. For correction of movement artifacts and eddy current distortions, registration between the DWI images is performed with software from the scanner workstation.

Quality control was then performed by visual inspection using research image analysis tools, such as MRIcron [http://www.mccauslandcenter.sc.edu/], ImageJ [http://imagej.nih.gov/] and vIST/e, for checking if all the images were correctly imported, sorted, and free of potential artifacts. Of the 54 scans, 2 were excluded due to corruption of their DICOM file (both from UMCU) - as their header had not been appropriately saved, 7 were excluded due to presence of scanner artifacts which highly disrupted the DTI anatomical structure information (1 from UMCU and 6 from MMC), and 2 due to unavailable clinical information (both from MMC).

Therefore, 43 scans were eligible for automatic segmentation (20 from UMCU and 23 from MMC) and 11 were dismissed. Of the 43 included scans, clinical information was available from the scanned patient, such as gestational age, and a MRI based injury scoring. This injury scoring is an internationally accepted score [Woodward, 2006], and takes into account WM deviations/abnormalities, sub-arachnoids' space, cysts presence and ventricular dilation; scores from 5 to 6 represent patients for which no abnormality was observed; values from 7 to 9 for patients with mild abnormalities, 10 to 12 for moderate abnormalities and 13 to 15 for severe abnormalities [Pul 2012; Kooij 2011, Woodward, 2006]. The 14 scanned patients were full term neonates and 29 were very preterm neonates, i.e., with less than 32 weeks of gestation age. In the studied dataset, there were no patients between 32 and 37 weeks of gestation. Pathology present in the dataset included damage in the basal ganglia, widened ventricles, hemorrhage and large WM deviations. Using this scoring system, abnormalities in the basal ganglia are not scored. Moreover, 11 of the included patients had no abnormalities, 28 had mild abnormalities, and 5 had moderate abnormalities. Patients with MRI injury scoring of 13 to 15, were not present in the studied dataset.

6.2 Segmentation Evaluation

For the current pipeline, visualization of the segmentation results is done by loading the pipeline output in the visualization software vIST/e. This allows for visual inspection of the segmentation result. A global evaluation of the segmented tractography was performed by a MRI expert. For each segmented structure, the number of incorrectly labeled tracts is evaluated. The ratio of incorrectly labeled tracts with the total number of tracts per segmented structure is further referred to as percentage error. For minor structures, FX, CG, MCP, IFOs, segmentation errors are seen as less severe than segmentation errors at major structures, CRs and CC. Therefore, segmentation performance is divided in 4 performance classes. They are:

4 - Good: When all segmented structures, major and minor, have less than 10% of error;

3 - Well: When major structures continue to have less than 10% of error but minor structures have between 10% and 50% of error;

2 - Moderate: When major structures have now more than 10% of error but still less than 30%;

1 - Bad: When major structures have more than 30% of error; the segmentation is said to have failed for these patients.

The results distribution is presented with a pie chart in figure 6.1. From the eligible patients, 19 - 44%, were classified as Good. 12 patients - 23%, and 7 patients - 16%, were classified as Well and Moderate correspondently. 5 patients - 12%, were classified as Bad. In figure 6.2, one exemplar segmented tractography is depicted for each of the four classes.

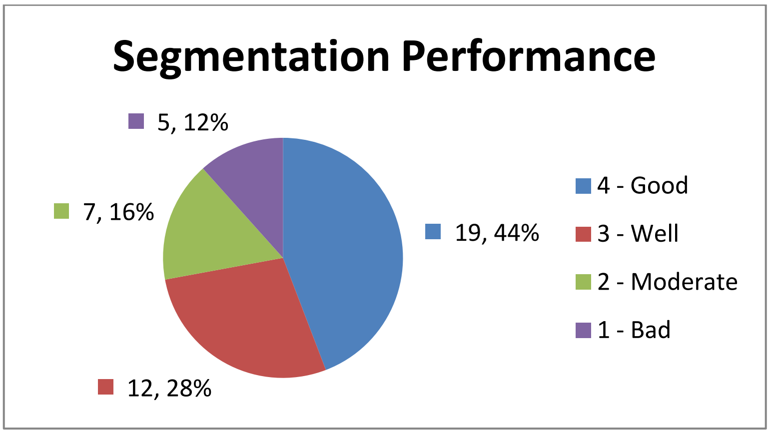
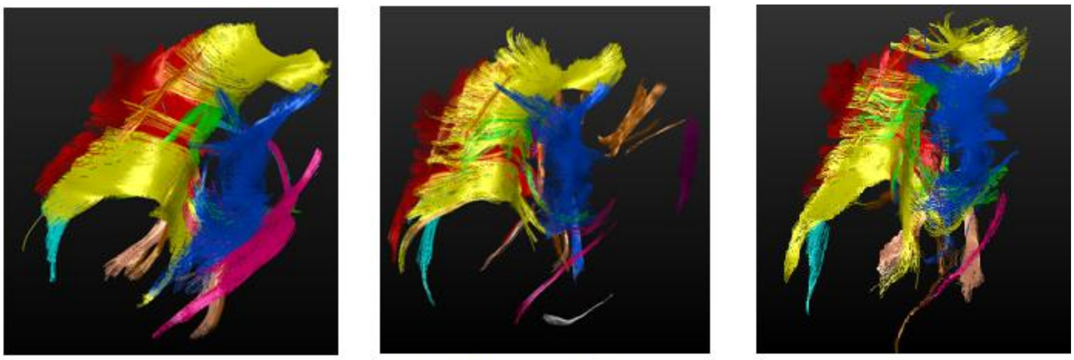


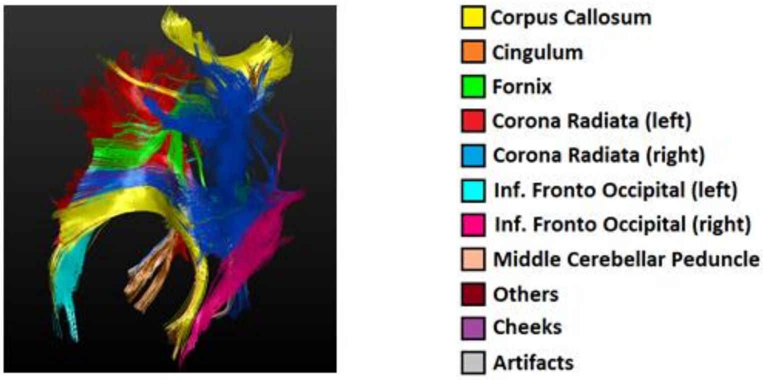
Figure 6.1 – Segmentation performance among the eligible scans.



4 - Good: < 10% error

3 - Well: Major < 10% error
Minor < 50% error

2 - Moderate: Major < 30% error
Minor < 50% error



1 - Bad: > 30% error

Figure 6.2 – Representative tractography segmentation example for each quality class, posterior view. From left to right and top to bottom: Good segmentation performance example; Well segmentation performance example, Moderate segmentation performance example; Bad segmentation performance example; key of structure color with anatomical label.

There might be two sources of segmentation error:

1. Errors because the atlas variability is not enough, i.e., not all necessary anatomical information for labeling the tested dataset is present in the atlas.
2. Errors because pipeline algorithms are failing, i.e., they are not propagating correctly the anatomical information from the atlas to the new subject.

For understanding better the contributions of these two possible sources of errors, some variables that might be influencing the segmentation performance are studied. First, it is probed if segmentation performance is affected by the presence of pathology. Second, the relation between registration performance and segmentation performance is assessed. Third, influence of spurious tracts in segmentation performance, due to an ineffective skull-stripping, is also addressed.

6.2.1 Presence of Pathology

Presence of pathology is here assessed by the MRI based injury scoring [Pul 2012; Kooij 2011]. Segmentation performance with injury scoring is presented at plotted at figure 6.3. There is no significant linear trend, as the R-squared is in the range of 10^{-4}.

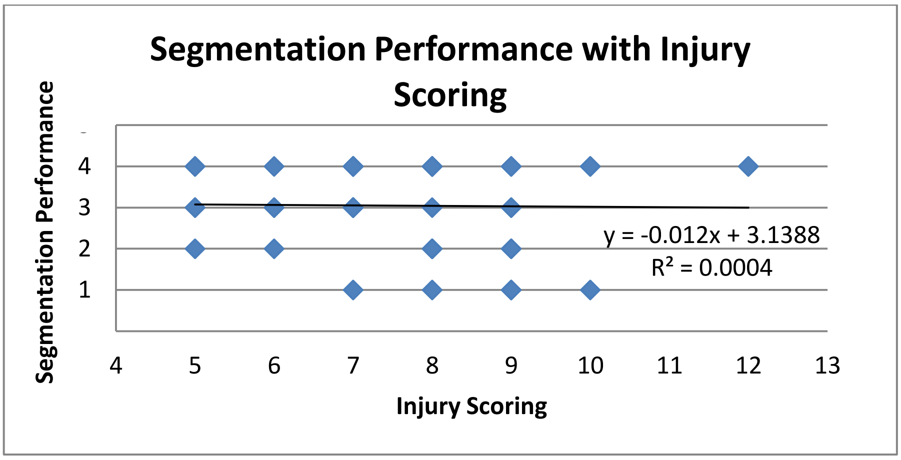


Figure 6.3 – Plot of segmentation performance per injury scoring. A linear regression was probed with result of non-significant R-squared of 0.004.

Average, minimum value and maximum value of segmentation performance per each different injury scoring values is presented in a graph in figure 6.4. Also here, the amount of patients per different injury scoring is depicted in bars. Average of performance for all different injury scoring values is around 3, the well segmentation performance class.

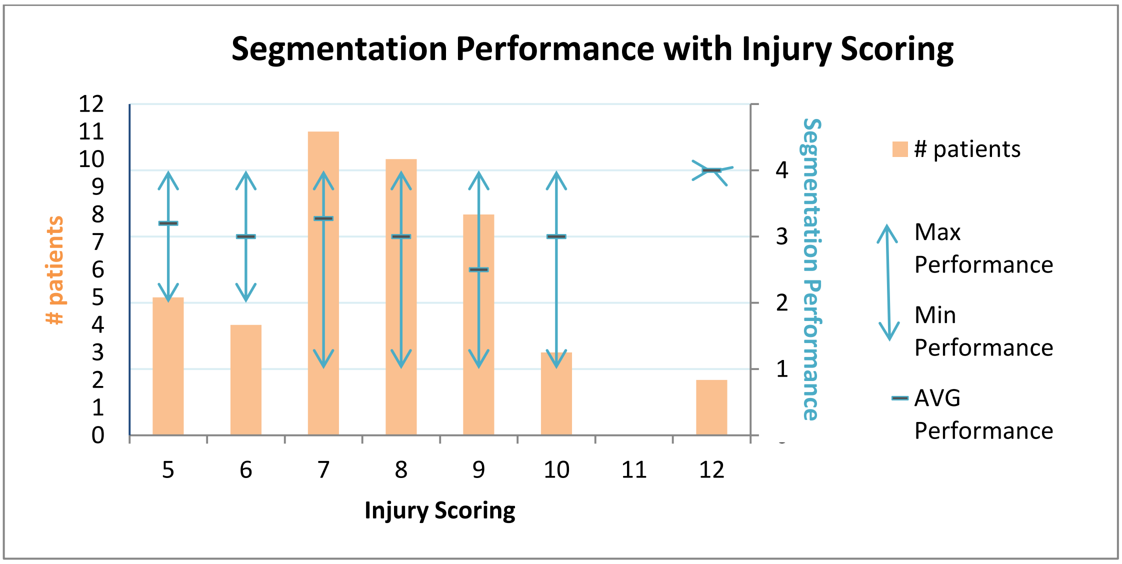


Figure 6.4 – Average, minimum value and maximum value of segmentation performance per different injury scoring.

Distribution of different abnormality degrees per class of segmentation performance is summarized in table 6.1. The patients with no abnormalities were never classified as having a bad segmentation performance. And of the 5 patients with moderate abnormalities, 4 were classified as presenting good segmentations and 1 presented a bad segmentation. Patients with mild abnormalities are distributed among all the four classes of segmentation performance, but are also in greater number in classes of well and good segmentation performance.

Segmentation Performance	No Abnormality	Mild Abnormality	Moderate Abnormality	Severe Abnormality
1 - Bad	-	4	1	-
2 - Moderate	2	5	-	-
3 - Well	4	8	-	-
4 - Good	3	12	4	-

Table 6.1 – Distribution of abnormalities degree per class of segmentation performance.

Figure 6.5 presents the segmentation results for 3 patients with severe WM deviations. The first has CC agenesis, i.e., inexistence of corpus callosum (CC). No CC was segmented for this patient, i.e., no tracts are colored in yellow in the picture. A yellow arrow was inserted in the typical CC region only for ease of interpretation. This patient was classified as having a good segmentation performance. The second patient was also classified as having a good segmentation. The third shows poor segmentation performance and it was inserted in performance class 1, since the CC (a main structure) is clearly showing mislabeled tracts. For this last patient the frontal part of the CC, in yellow, is labeled as FX, in green.

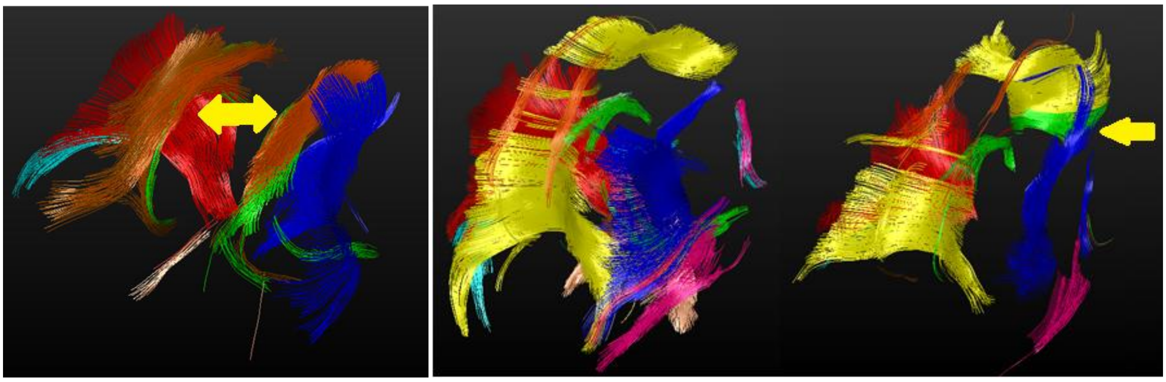


Figure 6.5 – Segmentation results for patients with severe WM deviation, posterior views. From left to right: patient with CC agenesis and with good segmentation performance; another patient with severe WM deviation and good segmentation performance; patient with severe WM deviation and poor segmentation performance, anterior part of the CC is segmented as FX.

6.2.2 Registration Performance

Registration performance is here assessed by studying optimization cost function (CF); the entropy from the similarity probability density functions that should be as small as possible. A plot of the distribution of last CF value per segmentation performance class is presented at figure 6.6. The relation between the last CF as a function of injury shows that the last CF value decreases with increase of segmentation performance. R-squared correlation measure is of 0.35 for 42 data-points with ANOVA significance of less than 0.001.

The sampled tracts of some subjects of interest were visualized after and before registration. Patients with last cost value bigger than 35 were individually assessed; this corresponded to 7 patients. Only one of these patients presented no abnormalities, the other 6 had mild abnormalities. It was also found that one of two factors was always present for these patients: presence of spurious tracts and highly curved anterior part of the CC. The same type of visual assessment was also conducted for the patients with last cost value under 35 which were classified as having a bad segmentation performance; this corresponded only to 2 patients. These patients did not have spurious tracts. However, they also had a highly curved anterior part of the CC, elucidated in figure 6.7. They both presented abnormalities, one mild and the other moderate. For most of these 9 patients, registration always seemed to result in excessive shrinking along superior-inferior and anterior-posterior orientations. An example is presented in figure 6.8.

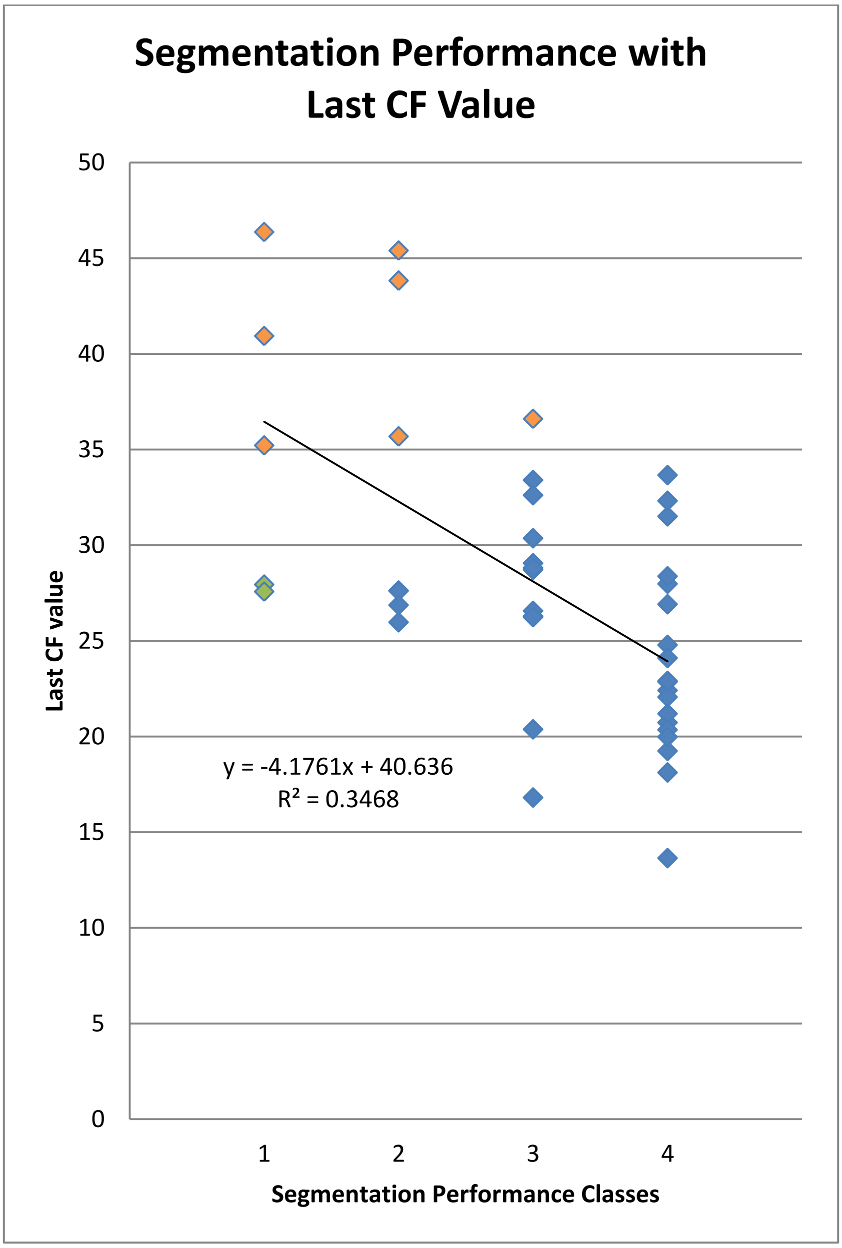


Figure 6.6 – Distribution of last CF value per segmentation performance class. In orange the points with last CF value of more than 35. In green the points from class 1 with less than 35 for last CF value. A linear regression was probed with result of R-squared of 0.35.

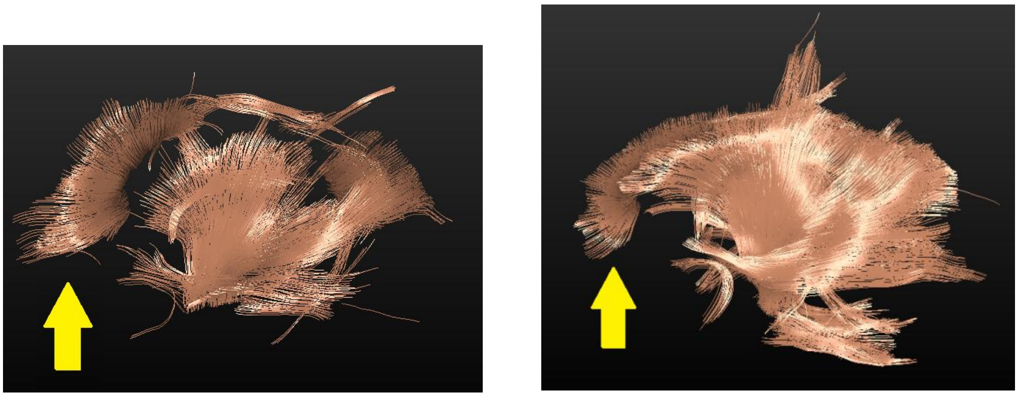


Figure 6.7 – Tractography with anterior part of CC highly curved in the left. Tractography with standard anterior part of CC in the right. Lateral view.

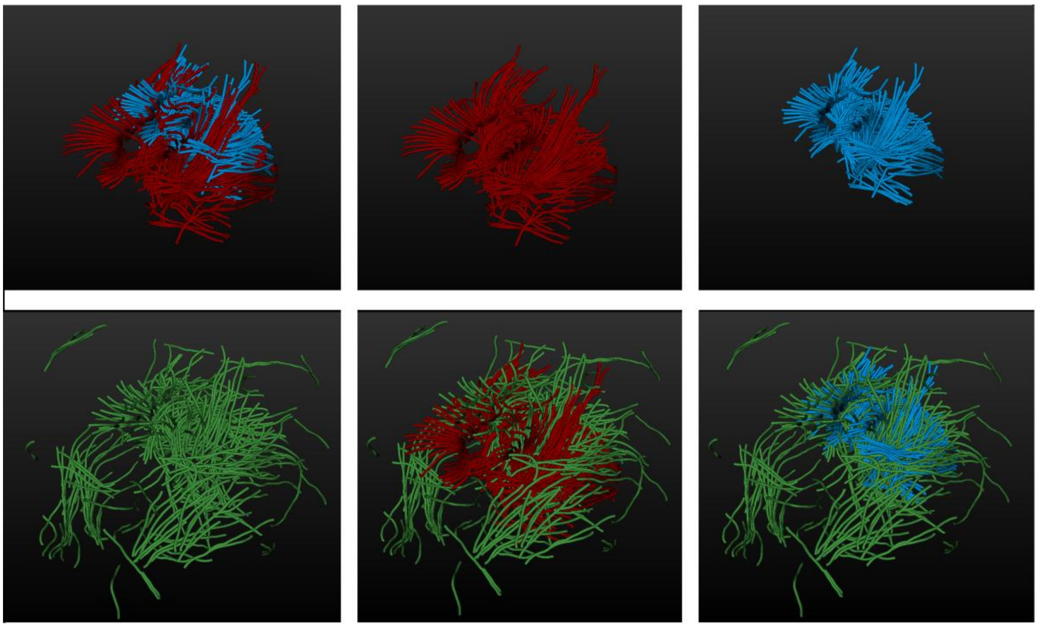


Figure 6.8 –Example of registration with excessive shrinking. In green the down-sampled tracts of the atlas, in red the down-sampled tracts previous to registration, in blue the down-sampled tracts after registration. Anterior views.

6.2.3 Skull Stripping Performance

Skull stripping performance is assessed in terms of existence of spurious tracts, i.e., tracts outside the WM region. Existence of these spurious tracts is seen as an error from the skull stripping method, and therefore a lower performance of this algorithm. These tracts correspond usually to spurious tracts above the skull. For the neonate dataset it is also usual to have spurious tracts coming from facial muscular structures, usually from the cheeks.

For the complete analyzed dataset, only 30% of the patient tractographies contained spurious tracts. Existence of spurious tracts is almost evenly present for all the four performance classes, as it can be observed in the bar plot on image 6.9. Boom also used part of this dataset in her work [Boom, 2011]. For this specific part of the dataset, Boom results contained spurious tracts from all patient tractographies, i.e., 100% of spurious tracts presence. When applying the skull-stripping method proposed in the current work, this number is reduced to 55%, a decrease of almost half.

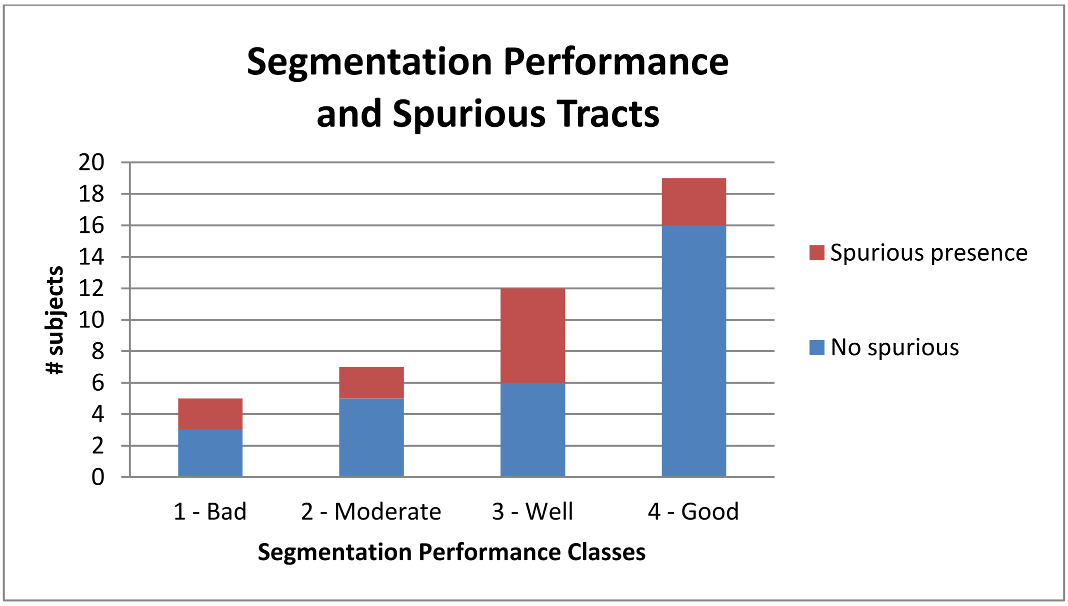


Figure 6.9 – Spurious tracts presence distribution among segmentation performance classes.

6.3 Volume and Anisotropy Trends

The vIST/e software allows to output quantitative information about each segmented structure; such as volume and anisotropy values. Volume value is calculated by knowing the number of voxels occupied by the structure tracts and the size of the voxel. Anisotropy value per structure corresponds to a weighted average of voxel anisotropy values: the anisotropy value of the voxels are included as many times as the number of tracts passing through them.

6.3.1 Difference between manual and automatic segmentation

For 20 Utrecht patients a manual segmentation of the corpus callosum (CC) was available [Pul, 2012]. The manual segmented tracts belong to a tractography result produced with the same software and parameters used for the current automatic segmentation pipeline. In addition, the same tensor fitting algorithm was used. Volume and some anisotropy values (FA, Cl and MD) were available for the manually segmented CC. Therefore, these values were also generated for the current automatic segmentation, and compared.

The difference between segmentation methods was analyzed by calculation of p-value from a two tail paired t-test and via the difference plot Bland-Altman. Only segmented tractography from performance class 4, 3 and 2 are included, i.e., tractography segmentations with more than 30% of error for the CC are not included. This resulted in analyzing 14 outcome pairs.

In table 6.2 the correspondent p-values are presented. All tested parameters have a p-value < 0,05 and therefore parameter measurement with manual and automatic method are considered significantly different.

Measurement	p-value from 2 tails paired t-test
Volume	0.001
FA	0.024
Cl	0.028
MD	0.002

Table 6.2 – P-values from a two tail paired t-test between manual and automatic measurements for the segmented CC.

Bland-Altman is a type of plot that has the horizontal axis depicting the average of the two measurements from the different methods for each patient. The vertical axis depicts the difference between the pair of measurements, normalized with the expected value for that measurement. Measurement from the manual method is subtracted from the automatic method. Consequently, positive values in the vertical axis will indicate the automatic method measurement is bigger than the measurement from the manual method. An example of a Bland-Altman plot is shown in figure 6.10 and 6.11, for volume and FA comparison respectively.

All other plots can be found in appendix II. Usually, the automatic method produces bigger volume and MD values than the manual method. And the manual method produces bigger FA and Cl values than the automatic. Differences between the methods seem to occur more often for lower volumes, lower FA and lower Cl values. For higher MD values, higher difference between methods seems to occur more often.

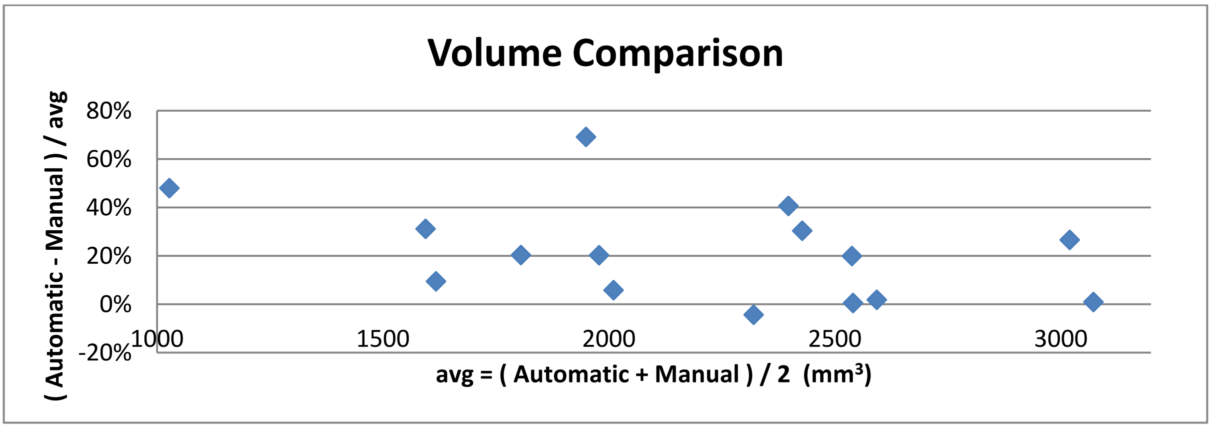


Figure 6.10 – Bland Altman CC volume comparison between the automatic and the manual method.

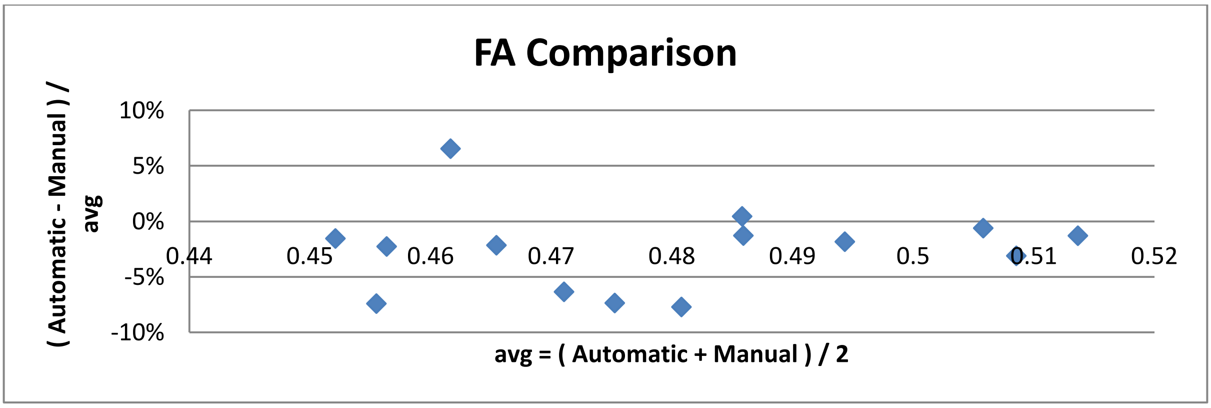


Figure 6.11 – Bland Altman CC FA comparison between the automatic and the manual method.

6.3.2 Volume and anisotropy as functions of injury scoring

It was also assessed if anisotropy and volume correlated with injury scoring. Two examples are shown in figure 6.12 and 6.13 of MD and volume respectively. This study was conducted only for one of the automatically segmented anatomical structures, the corpus callosum (CC). This structure was chosen due to its significance in prediction of neurodevelopment disorders at term equivalent age (TEA). Only CC's segmented with less than 30% of error are included for this study, i.e., quality segmentation classes 4, 3 and 2, excluding 1. Also the patient with CC agenesis was excluded. This resulted in the analysis of 37 patients. The plots of volume and anisotropy values with WM injury degree can be found in the appendix II. A summary of the statistic results can be found on table 6.3. Significant trends were found for MD and volume, where the p-value was < 0.05.

Qualitative Output	Type of correlation	R-squared	ANOVA significance
Volume	Negative	0.0951	0.043
FA	Negative	0.0542	0.165
Cl	Negative	0.0389	0.242
Cp	Positive	0.0010	N/A
MD	Positive	0.2401	0.02

Table 6.3 – Summary of volume and anisotropy correlation and significance with injury scoring for segmentations of quality class 1, 2 and 3.

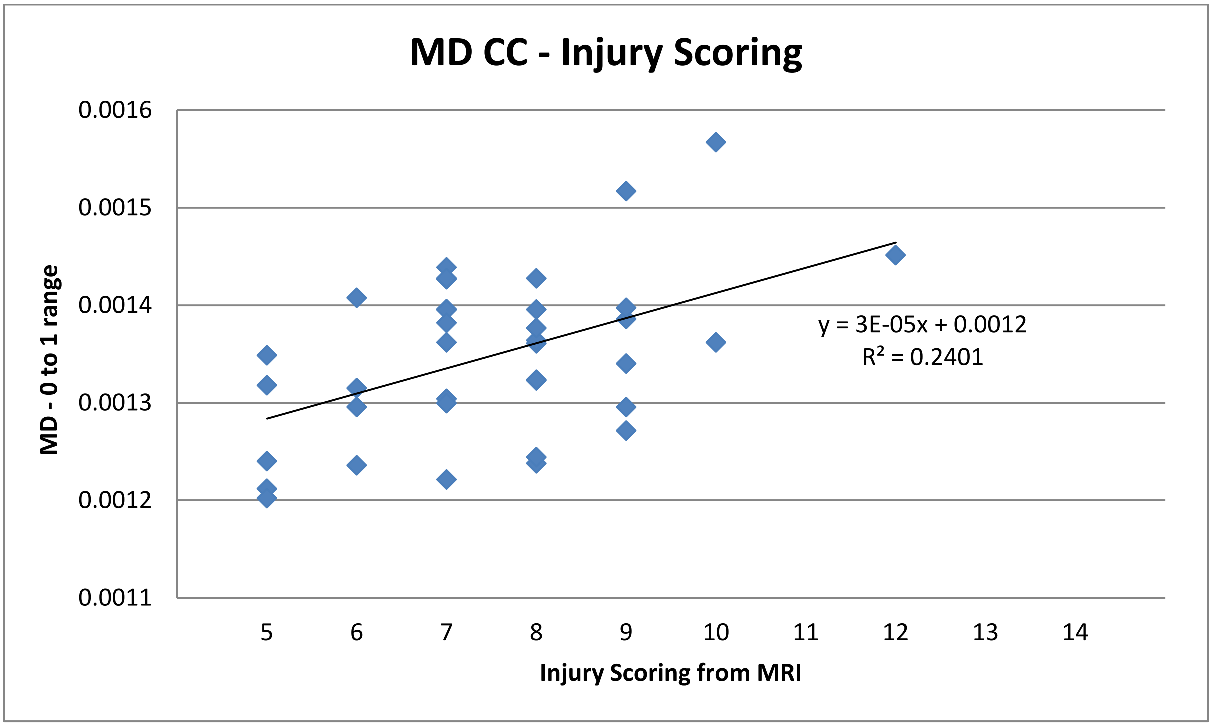


Figure 6.12 – Plot of CC anisotropic value MD per injury scoring. A linear regression was probed with the result having a R-squared of 0.24.

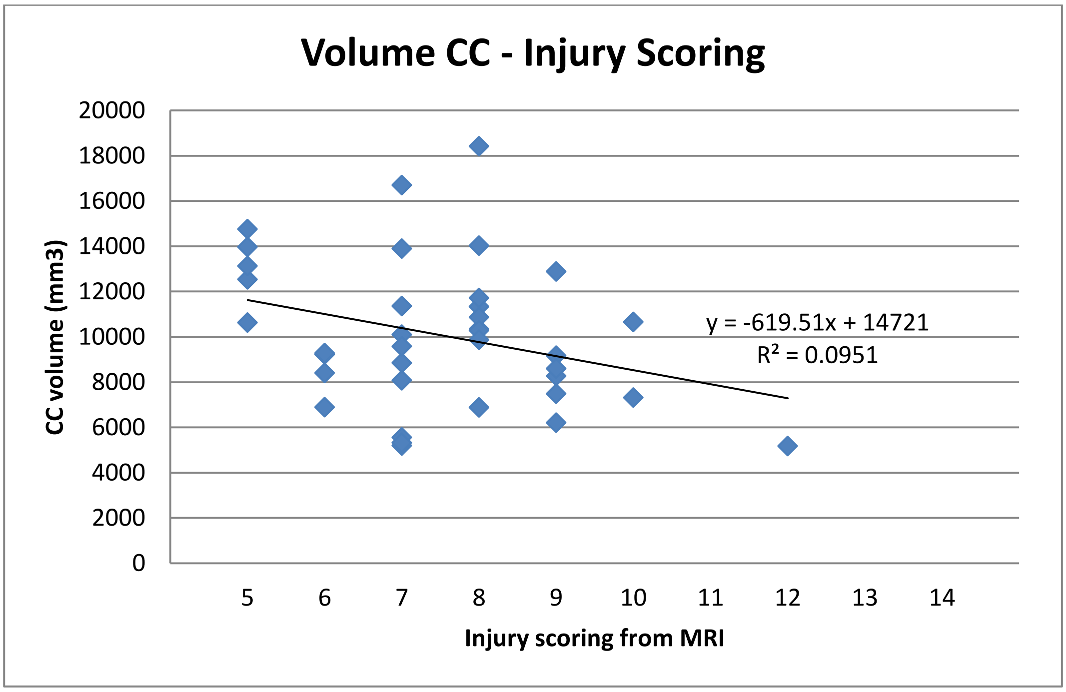


Figure 6.13 – Plot of CC volume per injury scoring. A linear regression was probed with the result having a R-squared of 0.09.

Chapter 7

Discussion and Conclusion

7.1 Discussion

This thesis had as starting point that tract-based assessment of the neonatal brain would be of added value in comparison with voxel-based assessments, such as the ones using MRI and DTI [Geng, 2012]. Tractography segmentation is seen as promising due to representation of global information as opposed to the voxel local information. It is hypothesized that this global information allows not only for a more complete study of the present WM anatomical structures, but can also improve atlas-based segmentation by allowing a better (global) comparison between atlas and patient.

Tractography registration performs an alignment of the subject tractography set with the atlas tractography set. The best alignment is the one for which the subject tract set is more similar to the atlas tract set. In this thesis, this similarity is computed directly in the tract domain [O'Donnell, 2012]. This method is referred to as tract-wise tractography registration, as opposed to voxel-wise tractography registrations

Tract-wise tractography registration has three potential drawbacks. First, it aligns patient with the atlas by only considering the tensor information that is contained in the first eigenvector, in regions where anisotropy values are high enough for reconstructing WM tracts, and therefore not even considering information from the GM areas. Second, it is computationally expensive, due to the similarity calculations between subject and atlas tracts. Third, inaccurate reconstruction of the WM bundles by the tractography model will result in erroneous tracts that can decrease registration accuracy.

It is important to understand that these three drawbacks are also the drawbacks of segmenting WM using tractography information, as opposed to using DTI or other scalar information. Indeed, for the factors hypothesized to be affecting segmentation performance, a correlation was only found with registration accuracy.

The first possible drawback relates specifically with the type of information contained in a tract, as opposed to the information contained per voxel. Because segmentation performance showed to not be affected by pathology presence, it was possible in the current study to assess the trends of volume and anisotropies with WM abnormality degree. Therefore, assessing the clinical relevance of tractography derived information. Analysis was only conducted for the CC anatomical structure. The trends for volume and MD with WM abnormality degree were found to be significant. MD was found to increase with WM abnormality, and volume to decrease with WM abnormality for the CC anatomical structure. This is coherent with the results based on a manual segmentation method by Pul et al (2012). Her work suggests that infants with WM injury will have reduced CC volume. This is also observed visually in our dataset, especially due to the fact that the middle part of the CC is often not tracked for patients with low degree of development, as the central part of the CC is known to normally mature later than the peripheral parts. Increase of MD is seen as a consequence of a less hindered diffusion for these patients. This model from Pul et al (2012) also suggests that WM injury will result in a decrease of FA and Cl. This last behavior was not statically significant for our dataset. The tested dataset comprised only 43 patients. This was quite a small group and for the MMC patients the pathology was quite heterogeneous. Testing the pipeline for bigger datasets with other forms of WM injury can lead to more significant results.

Still related with this topic, it is also of interest to comment on the differences in results between the proposed automatic segmentation method and the common manual segmentation method for the CC structure. The measurements between these two techniques showed some differences. Bigger volumes and MD values, and smaller FA and Cl values are observed for the automatic method. These differences are more pronounced for the range of lower volumes, FA and Cl values and higher MD values. A possible explanation might be that the automatic method includes more peripheral tracts from the anatomical structure than the manual method; with this difference being more pronounced for less mature CC. The comparison of manual versus automatic segmentation needs still to be investigated for the anatomical structure corona radiata (CR). Furthermore, an investigation of other structures will be of much interest. With the current method, fornix (FX) segmentation results still did not allow for an accurate investigation of this structure anisotropy. There are still too many erroneous tracts present. Therefore, future improvements should also focus on increasing FX segmentation accuracy. Guevara (2012) and O'Donnell (2007) also found for their methods that the association type tracts, FX and CG for the used atlas labels, were the ones more difficult to segment.

As for the computation complexity arising especially with calculation of similarities between tracts, two solutions are proposed. First, the skull-stripping method from Hoskam (2009) was improved successfully by a reduction of almost half of the spurious tracts presence, therefore reducing calculation of similarities for tracts without anatomical interest. Second, for most part of the segmentation pipeline a down-sampled tract set is used, both for subject and atlas. The down-sampled set proposed is created by the cluster-centers of the complete tractography set. This method allows to have represented in the down-sampled set all the anatomical structures of interest, even anatomical structures which have in total not more than 5-8 tracts. A pitfall from performing a cluster-based down-sampling arises with the fact that clustering in itself makes use of similarity metrics. Still, performing registration and label attribution using as input a down-sampled set was found to improve the accuracy of these pipeline steps, in contrast with using the complete tractography set for representing the atlas and the patient tracts. Down-sampling computation complexity can be improved by using less expensive similarity metrics, such as the one used proposed by O'Donnell (2012). In addition, future versions of the atlas should consider the removal of spurious tracts/labels representation from the atlas. Besides adding extra complexity to the similarity calculations, their existence is also source of mislabeling for some of the tested subjects. After all, the spurious tracts contained in the atlas are only examples of all the possible spurious tracts patterns.

The tractography model implemented corresponds to a deterministic whole volume seeding based in a DTI with 32 directions with voxel size of almost $4mm^3$ tuned for neonatal data [Pul, 2012]. The relation between tractography accuracy and segmentation performance was not evaluated. However, subjects with poor segmentation performance often showed partial volume effects. These occurred usually between the corpus callosum (CC) and fornix (FX), or between the CC and cingulum (CG). For some patients a partial volume effect was also observed between CC and the corona radiatas (CR). Improvement of tractography accuracy, i.e., reduction of partial volume effects, is directly related with DTI resolution and the accuracy of the implemented tractography model - which of course, also depends partially on DTI's voxels resolution. Currently, new scanner sequences and software are becoming available for improving both variables. However, DTI resolution corresponds to a delicate balance between acquisition time and signal to noise ratio, which for neonatal clinical practice is very much constrained. Within maximum of 5 minutes the scan needs to be performed, because the baby may get restless. Therefore, for the current clinical restrictions, reduction of partial volume effect should be mostly tackled by the improvement of tractography models.

Besides this drawback consideration, it is also of interest to report a special case of inaccurate registration/segmentation. For some subjects, a special shape of the anterior part of the corpus callosum (CC) was identified by visual inspection as a probable trigger for an inaccurate segmentation. For these patients, registration often lead to an excessive shrinking of the tracts. Excessive shrinking might be caused due to an ineffective tuning of the registration constraint parameters, and/or due to the fact this kind of CC shape is not represented in the atlas, and/or even due to a misleading similarity metric definition between atlas and subject tracts. Possible ways to tackle this can include: increasing the atlas variability by adding more anatomical labeled subjects to the atlas; applying a different deformation model which allows for local transformation, like elastic and fluid deformations - with the need of then defining optimal constraints for these models; finally, might be also of interest to tune the similarity metric being used for registration [O'Donnell, 2012] specifically for dealing with neonatal data [Boom, 2011]. Further research about the effect of these factors on the registration/segmentation performance should be conducted.

In addition, two other improvements should be also considered. First, segmentation results were qualitatively evaluated by only one DTI expert (due to time constraints). Further analysis from more than one expert, and in particular including at least a neonatologist and a radiologist experienced with DTI is a necessary improvement. Second and last, in this phase only a global evaluation of segmentation performance was performed. A more detailed performance analysis for each segmented anatomical structure is needed.

7.2 Conclusion

The new proposed pipeline performs automatic atlas-based segmentation of WM anatomical structures from DTI tractography in preterms imaged at term equivalent age (TEA). This method allows the study of WM maturation as a possible predictive marker for neurodevelopmental disorders of preterms at TEA, making the procedure less time consuming than a manual method and, more importantly, less user dependent. Main contributions in the current study involved a tract-wise tractography registration, a skull-stripping method, a tract sampling method specific for neonatal data and automation of the entire pipeline. The tract-wise registration was inspired by the work of O'Donnell (2012). This new registration approach allows using the global directional and connectivity information entailed in the tractography.

The resulting analysis is promising, as only 12% of the segmentations contained more than 30% mislabeled tracts. Segmentation performance showed not to be influenced by presence of WM pathology among subjects, even when anatomical structures were missing due to severe WM deviations. The CC structure automatically segmented by the pipeline was further analyzed, by studying its respective volume and anisotropy measurements per subject. For the studied measurements, volume and MD showed a significant trend with degree of WM injury. These trends are in accordance with previously findings about how WM injury influences DTI derived anisotropies [Rose, 2014; Liu, 2012; Pul, 2012; Chen, 2011; Hasegawa, 2011; Lindqvist, 2011; Mullen, 2011; Liu, 2011; Dubois, 2008; Dubois, 2006; Huppi, 2006].

These findings can be of added value when it comes to understanding WM global deviations from the standard WM pattern. Such understanding is critical for clinical evaluation, as many pathologies related with neurodevelopmental deficits, like asphyxia, are characterized by a WM disturbed pattern [Pul, 2004].

Future work should aim to increase atlas variability and also to probe similarity metrics between tracts that are less computational expensive but still adequated for neonatal tractography. In addition, it might be of interest to extend the atlas for representing all gestational ages, for allowing study of full-term neonates at risk of neurodevelopmental disorders.

References

Alexander DC. An introduction to computational diffusion MRI: the diffusion tensor and beyond. Chapter in "Visualization and image processing of tensor fields" Springer, 2006.

Anjari M, Srinivasan L, Allsop JM, Hajnal JV, Rutherford MA, Edwards AD, Counsell SJ. Diffusion tensor imaging with tract-based spatial statistics reveals local white matter abnormalities in preterm infants. Neuroimage, 35(3):1021-1027, 2007.

Berman JI, Mukherjee P, Partridge SC, Miller SP, Ferriero DM, Barkovich AJ, Vigneron DB, Henrya RG. Quantitative diffusion tensor MRI fiber tractography of sensorimotor white matter development in premature infants. NeuroImage 27:862 – 871, 2005.

Brecheisen R. Visualization of uncertainty in fiber tracking based on diffusion tensor imaging. PhD Thesis, Department of Biomedical Engineering, Technology University of Eindhoven, 2012.

Cabezas M, Olivera A, Lladóa X, Freixeneta J, Cuadrab MB. A review of atlas-based segmentation for magnetic resonance brain images. Computer Methods and Programs in Biomedicine 104:158-177, 2011.

Chen Y, An H, Zhu H, Jewells V, Armao D, Shen D, Gilmore JH, Lin W. Longitudinal regression analysis of spatial–temporal growth patterns of geometrical diffusion measures in early postnatal brain development with diffusion tensor imaging. NeuroImage 58:993–1005, 2011.

Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marchal G. Automated multimodality medical image registration using information theory. Computational Imaging and Vision 3:263–274, 1995.

Dubois J, Dehaene-Lambertz G, Perrin M, Mangin JF, Cointepas Y, Duchesnay E, Le Bihan D, Hertz-Pannier L. Asynchrony of the Early Maturation of White Matter Bundles in Healthy Infants: Quantitative Landmarks Revealed Noninvasively by Diffusion Tensor Imaging. Human Brain Mapping 29:14–27, 2008.

Dubois J, Hertz-Pannier L, Dehaene-Lambertz G, Cointepas Y, Le Bihan D. Assessment of the early organization and maturation of infants' cerebral white matter fiber bundles: A feasibility study using quantitative diffusion tensor imaging and tractography. NeuroImage 30:1121–1132, 2006.

European Foundation for the care of newborn infants (EFCNI). Too Little, Too Late? Why Europe should do more for preterm infants. EU Benchmarking Report 2009/2010, 2010.

Frey BJ, Dueck D. Clustering by Passing Messages Between Data Points. *Science* 315(972), 2007.

Glass HC, Bonifacio SL, Shimotake T, Ferriero DM, Neurocritical Care for Neonates. Pediatric Neurology - Current Treatment Options in Neurology 13:574–589, 2011.

Geng X, Gouttard S, Sharma A, Gu H, Styner M, Lin W, Gerig G, Gilmore J. Quantitative Tract-based white matter development from birth to age of two years. Neuroimage 61(3):524-557, 2012.

Guevara P, Duclap D, Poupon C, Marrakchi-Kacem L, Fillard P, Le Bihan D, Mangin JF. Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas. NeuroImage 61:1083–1099, 2012.

Hasegawa T, Yamada K, Morimoto M, Morioka S, Tozawa T, Isoda K, Murakami A, Chiyonobu T, Tokuda S, Nishimura A, Nishimura T, Hosoi H. Development of corpus callosum in preterm infants is affected by the prematurity: *in vivo* assessment of diffusion tensor imaging at term-equivalent age. Pediatric Research 69(3), 2011.

Herculano-Houzel S. The Human Brain in Numbers. Front Hum Neurosci. 3(31), 2009.

Hornak JP, The Basis of MRI – online book at http://www.cis.rit.edu/htbooks/mri/. 1996-2014

Hoskam G. Quantification of DTI in preterm children. Master Thesis, Department of Biomedical Engineering, Eindhoven University of Technology, 2009.

Huppi PS, Dubois J. Diffusion Tensor Imaging of Brain Development. Seminars in Fetal & Neonatal Medicine 11:489-497, 2006

Hutton BF, Braun M. Software for Image Registration: Algorithms, Accuracy, Efficacy Seminars in Nuclear Medicine. 33(3):180-192, 2003.

Huppi PS, Dubois J. Diffusion tensor imaging of brain development. Seminars in Fetal & Neonatal Medicine 11:489-497, 2006.

Jong M, Verhoeven M, Baar AL. School outcome, cognitive functioning, and behavior problems in moderate and late preterm children and adults: A review. Seminars in Fetal & Neonatal Medicine 17:163-169, 2012.

Kooij B. MRI analysis and neurodevelopmental outcome in preterm infants. PhD Thesis, Utrecht University, 2011.

Latal B. Prediction of Neurodevelopmental Outcome After Preterm Birth. Ped. Neuro. 40:413-419, 2009.

Leemans A, Jones DK. A new approach to fully automated fiber tract clustering using affinity propagation. Proc. Intl. Soc. Mag. Reson. Med. 17, 2009.

Leemans A, Sijbers J, De Backer S, Vandervliet E, Parizel P. Multiscale White Matter Fiber Tract Coregistration: A New Feature-Based Approach to Align Diffusion Tensor Data Magnetic Resonance in Medicine 55:1414–1423, 2006.

Lindqvist S, Skranes J, Eikenes L, Haraldseth O, Vik T, Brubakk AM, Vangberg TR. Visual function and white matter microstructure in very-low-birth-weight (VLBW) adolescents – A DTI study Vision Research 51:2063–2070, 2011.

Liu Y, Aeby A, Balériaux D, David P, Absil J, Maertelaer V, Bogaert P, Avni F, Metens T. White Matter Abnormalities Are Related to Microstructural Changes in Preterm Neonates at Term-Equivalent Age: A Diffusion Tensor Imaging and Probabilistic Tractography Study. AJNR 33:839–45, 2012.

Liu Y, Metens T, Absil J, Maertelaer V, Balériaux D, David P, Denolin V, Overmeire B, Avni F, Bogaert P, Aeby A. Gender Differences in Language and Motor- Related Fibers in a Population of Healthy Preterm Neonates at Term-Equivalent Age: A Diffusion Tensor and Probabilistic Tractography Study. AJNR 32:2011–16, 2011.

Lori NF, Akbudak E, Shimony JS, Cull TS, Snyder AZ, Guillory RK, and Conturo TE. Diffusion tensor fiber tracking of brain connectivity: acquisition methods, reliability analysis and biological results. NMR in Biomedicine 15:494-515, 2002

Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality Image Registration by Maximization of Mutual Information. IEEE Transactions on medical imaging 16(2), 1997.

Mayer A, Zimmerman-Moreno G, Shadmi R, Batikoff A, Greenspan H. A Supervised Framework for the Registration and Segmentation of White Matter Fiber Tracts. IEEE Transactions on medical imaging 30(1), 2011.

Moberts B., Hierarchical Visualization using Fiber Clustering. Master Thesis, Department of Biomedical Engineering, Technological University of Eindhoven, 2005.

Mori S, Zijl PCM. Fiber tracking: principles and strategies – a technical review. NMR in Biomedicine 15:468-480, 2002.

Mori S, Wakana S, Nagae-Poetscher L, Zijl P. MRI Atlas of Human White Matter. Elsevier, 2005

Mullen KM, Vohr BR, Katz KH, Schneider KC, Lacadie C, Hampson M, Makuch RW, Reiss AL, Constable RT, Ment LR. Preterm birth results in alterations in neural connectivity at age 16 years NeuroImage 54:2563–2570, 2011.

Notestine CF, I. Ozyurt B, Clark CP, Morris S, Bischoff-Grethe A, Bondi MW, Jernigan TL, Fischl B, Segonne F, Shattuck DW, Leahy RM, Rex DE, Toga AW, Zou KH. Morphometry Quantitative Evaluation of Automated Skull-Stripping Methods Applied to Contemporary and Legacy Images: Effects of Diagnosis, Bias Correction, and Slice Location. Hum Brain Mapp. 27(2):99–113, 2006.

O’Donnell LJ, Wells W.M. 3rd, Golby AJ, Westin C.F. Unbiased Groupwise Registration of White Matter Tractography. MICCAI Part III:123–130, 2012.

O’Donnell LJ, Rigolo L., Norton I., Wells W.M. 3rd, Westin C.F., Golby AJ, fMRI-DTI modeling via landmark distance atlases for prediction and detection of fiber tracts. NeuroImage, 2011.

O’Donnell LJ, Westin C.F. Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas IEEE Transactions on Medical Imaging 26(11):1562–1575, 2007.

Poisson SD, Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilitiés, 1837.

Powell MJD. Direct Search Algorithm for optimization Calculations. Acta Numerica 7 Cambridge University Press, 1998.

Pul C, Kooij BJM., Vries LS, Benders MJNL, Vilanova A, Groenendaal F, Quantitative Fiber Tracking in the Corpus Callosum and Internal Capsule Reveals Microstructural Abnormalities in Preterm Infants at Term-Equivalent Age. AJNR 33, 2012.
Pul C, Buijs J, Vilanova A, Roos GF, Wijn PFF. Infants with perinatal hypoxic ischemia: feasibility of fiber tracking at birth and 3 months. Radiology 240:203-214, 2006.

Pul C. Diffusion Tensor Imaging for the detection of hypoxic-ischemic injury in newborns. PhD Thesis Technological University of Eindhoven, 2004.

Rose J, Vassar R, Cahill-Rowley K, Guzman XZ, Stevenson DK, Barnea-Goraly N. Brain microstructural development at near-term age in very-low-birth-weight preterm infants: An atlas-based diffusion imaging study. NeuroImage 86:244–256, 2011.

Shannon CE. A Mathematical theory of Communication. Bell System Technical Journal 27(3):379–423, 1948.

Soares JM, Marques P, Alves V, Sousa N. A hitchhiker's guide to diffusion tensor imaging. Frontiers in NeuroScience 7(31), 2013.

Visser E, Nijhuis EHJ, Buitelaar JK, Zwiers MP. Partition-based mass clustering of tractography streamlines. NeuroImage 54:303–312, 2011.

Volpe JJ, Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol. 8:110–124, 2009.

Vries LS, Groenendaal F. Patterns of neonatal hypoxic–ischaemic brain injury. Neuroradiology 52:555–566, 2010

Vilanova A, Berenschot G, Pul C. DTI visualization with streamsurfaces and evenly-spaced volume seeding. Proceedings of the Joint Eurographics - IEEE TCVG Symposium on Visualization 173-182, 2004.

Vincent L. Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms. IEEE Transactions on Image Processing, 2(2): 176-201, 1993.

Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, Hua K, Zhang J, Jiang H, Dubey P, Blitz A, Zijl P, Mori S. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36(3):630-44, 2007.

Wakana S, Jiang H, Nagae-Poetscher LM, Zijl PCM, Mori S. Fiber Tract–based Atlas of Human White Matter Anatomy. Radiology 230:77–87, 2004.

Wassermann D, Bloyb L, Kanterakisb E, Vermab R, Derichea R. Unsupervised White Matter Fiber Clustering and Tract Probability Map Generation: Applications of a Gaussian Process framework for white matter fibers. Neuroimage 15(51,1): 228–241, 2010.

Welker KM, Patton A. Assessment of Normal Myelination with Magnetic Resonance Imaging Semin Neurol 32:15–28, 2012.

Woodward LJ, Anderson PJ, Austin NC, Howard K, Inder TE. Neonatal MRI to predict neurodevelopmental outcomes in preterm infants. New England Journal of Medicine 355(7):685-694, 2006.

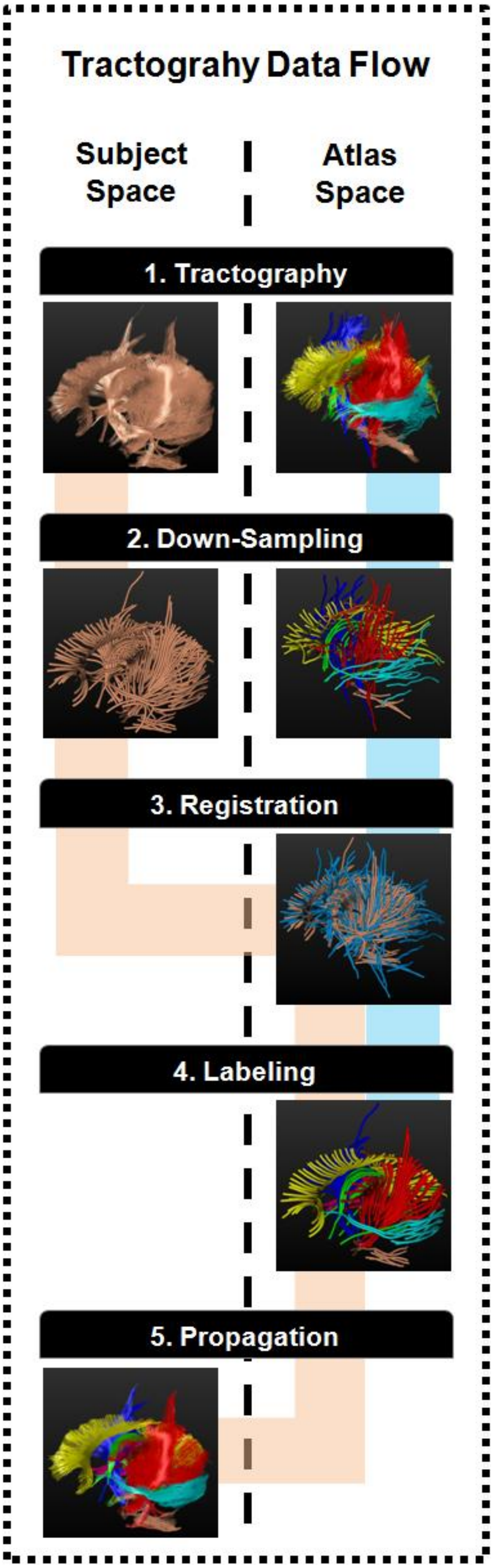
Zitová B, Flusser J. Image registration methods: a survey. Image and Vision Computing 21:977–1000, 2003.

Zvitia O, Mayer A, Shadmi R, Miron S, Greenspan H. Co-registration of White Matter Tractographies by Adaptive-Mean-Shift and Gaussian Mixture Modeling. IEEE Trans. Med. Imag. 29(1):132–145, 2010.

Appendix I – Tractography Data Flow

In this appendix, the tractography data flow is depicted. The main processing steps' titles appear in the black boxes. These are then followed with the correspondent tractography output images. The scheme has two different columns: the left column corresponds to the subject space and the right column to the atlas space. Data flow is depicted in two differently colored lines, orange for the patient tractography data and blue for the atlas tractography data.

Both atlas and subject full tractography (1.Tractography) are down-sampled (2.Down-Sampling). After, the subject down-sampled set is registered to the atlas down-sampled set, i.e., passing from subject space to atlas space (3.Registration). Then, still in the atlas space, labeling of the down-sampled subject is performed by measurement of similarities with the atlas down-sampled set (4.Labeling). The labeling is then propagated to the full tractography subject set, producing the complete segmentation result (5.Propagation). This last step makes the final passage from atlas space to subject space.



Appendix II – Result Analysis Plots

Here all the Bland-Altman plots for comparison of volume and anisotropies between manual and automatic method can be found, from figure A.1 to A.8. The difference between the measurements is plotted twice, the second plot takes the absolute value of the difference.

Also, the plots of volume and anisotropies with the MRI-based injury scoring are presented with the correspondent linear regression trend, from figure A.9 to A.13.

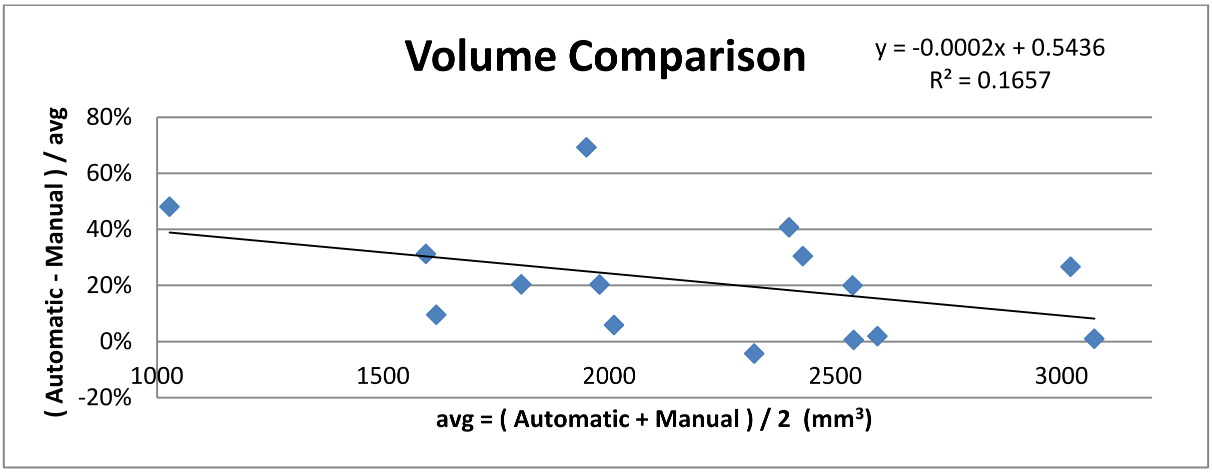


Figure A. 1 - Bland-Altman plot for difference of CC volume measurements between automatic and manual method.

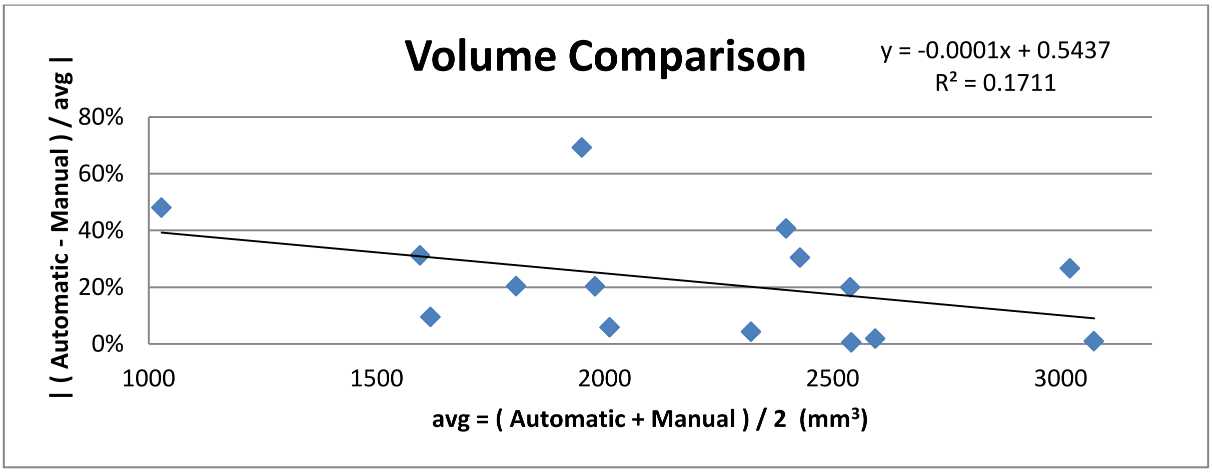


Figure A. 2 - Bland-Altman plot for absolute difference of CC volume measurements between automatic and manual method.

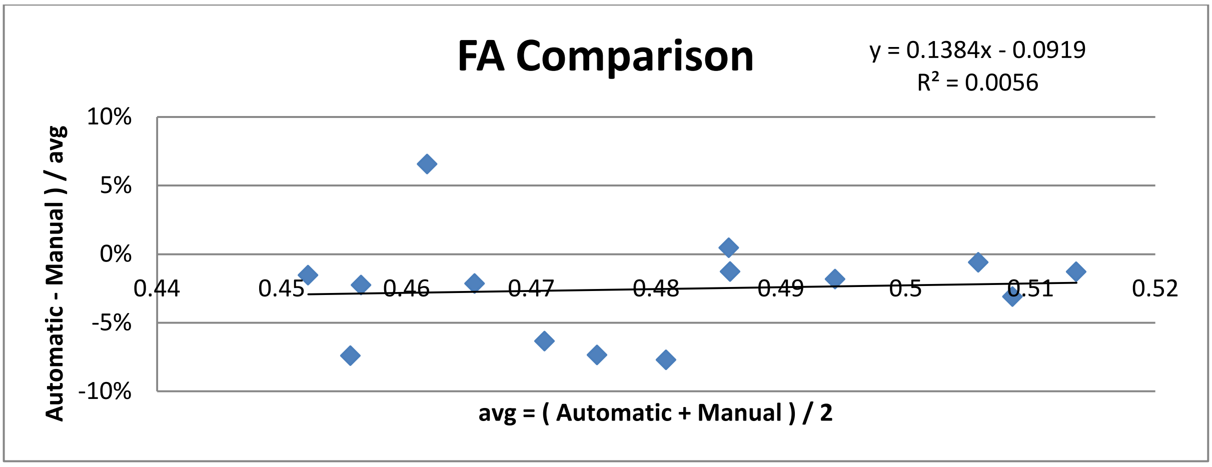


Figure A. 3 - Bland-Altman plot for difference of CC FA measurements between automatic and manual method.

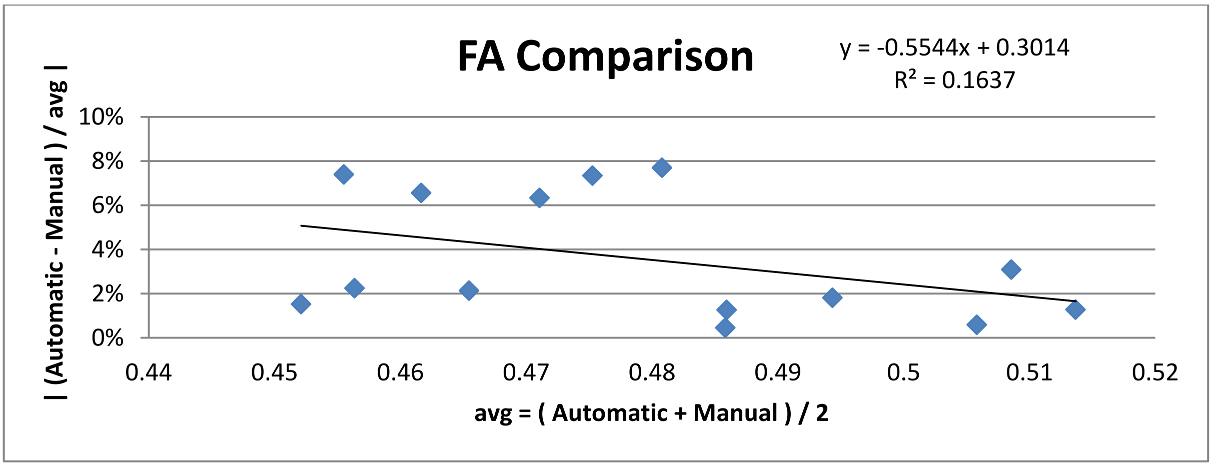


Figure A. 4 - Bland-Altman plot for difference of CC FA measurements between automatic and manual method.

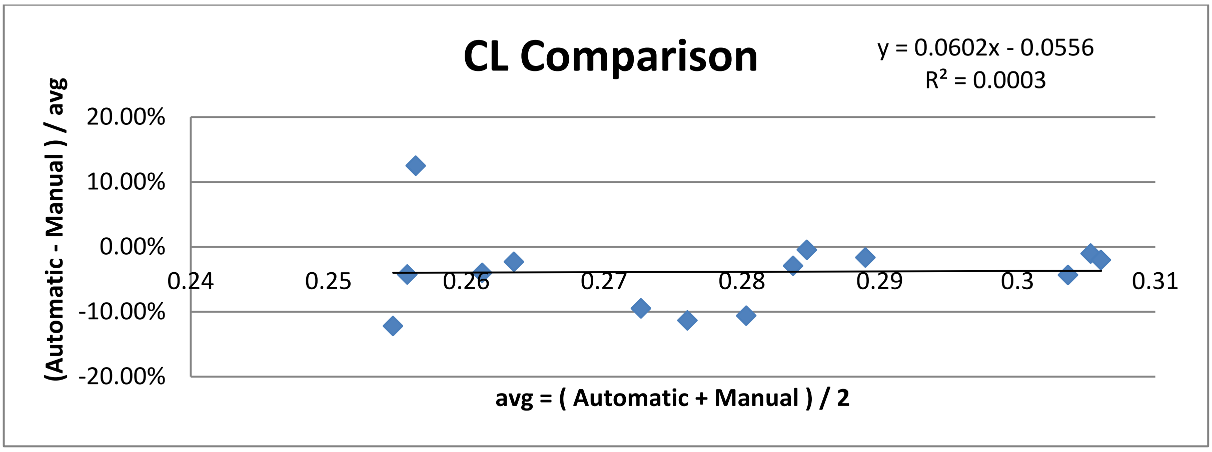


Figure A. 5 - Bland-Altman plot for absolute difference of CC Cl measurements between automatic and manual method.

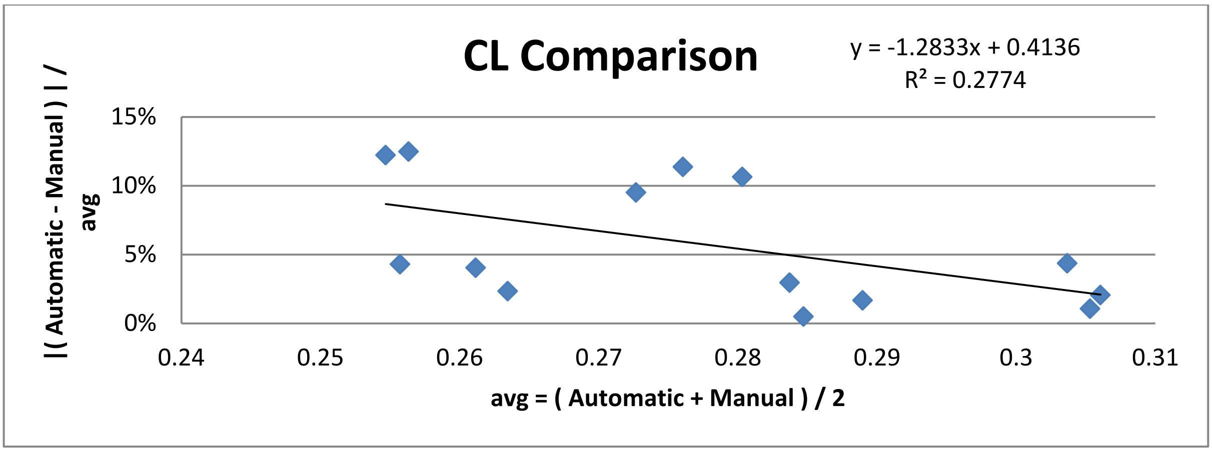


Figure A. 6 - Bland-Altman plot for absolute difference of CC Cl measurements between automatic and manual method.

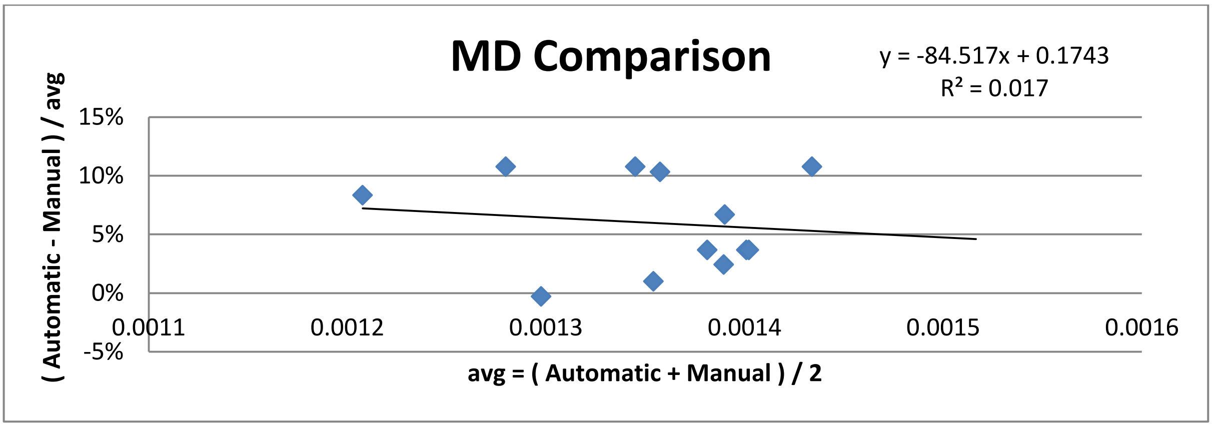


Figure A. 7 - Bland-Altman plot for difference of CC MD measurements between automatic and manual method.

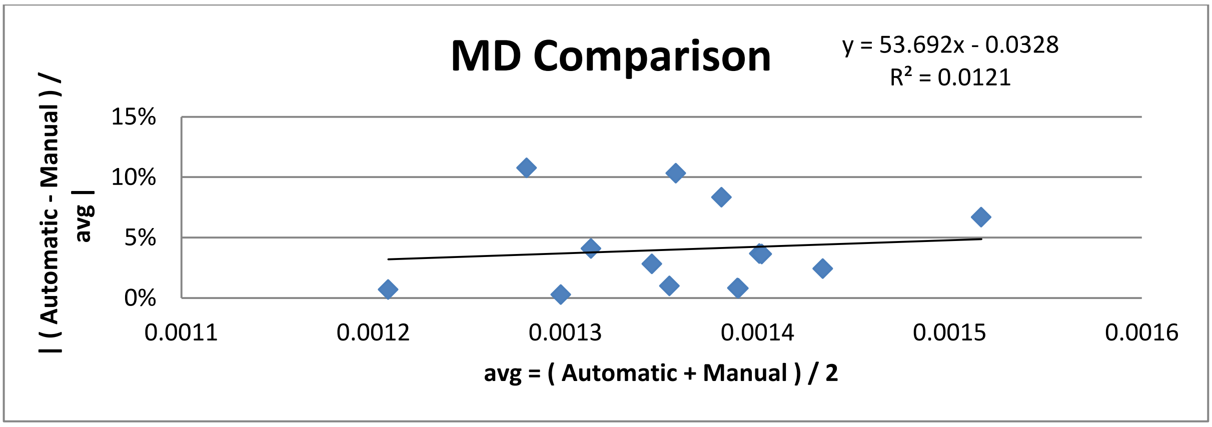


Figure A. 8 - Bland-Altman plot for absolute difference of CC MD measurements between automatic and manual method.

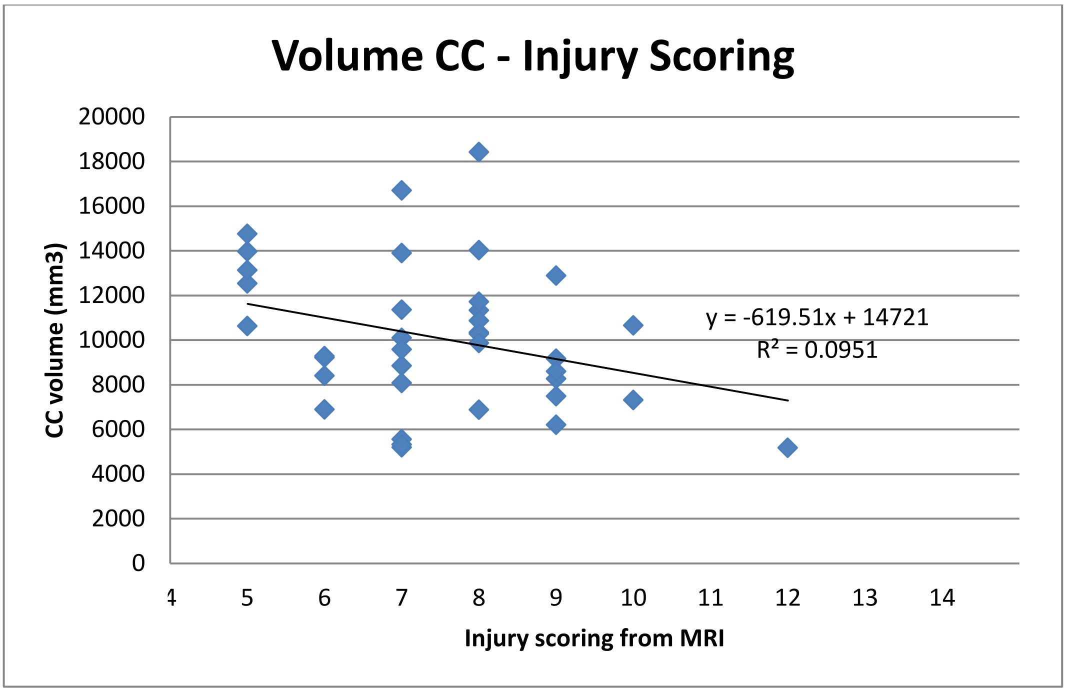


Figure A. 9 – Volume of CC with MRI-based injury scoring and correspondent linear regression.

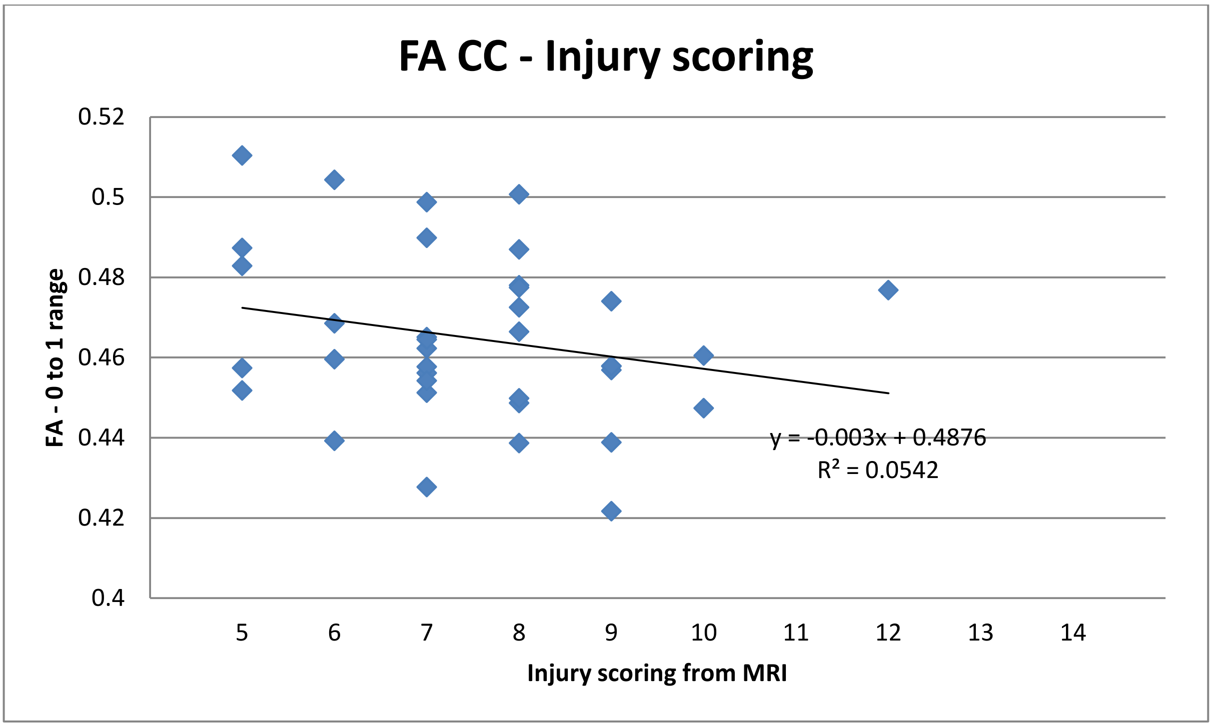


Figure A. 10 – FA of CC with MRI-based injury scoring and correspondent linear regression.

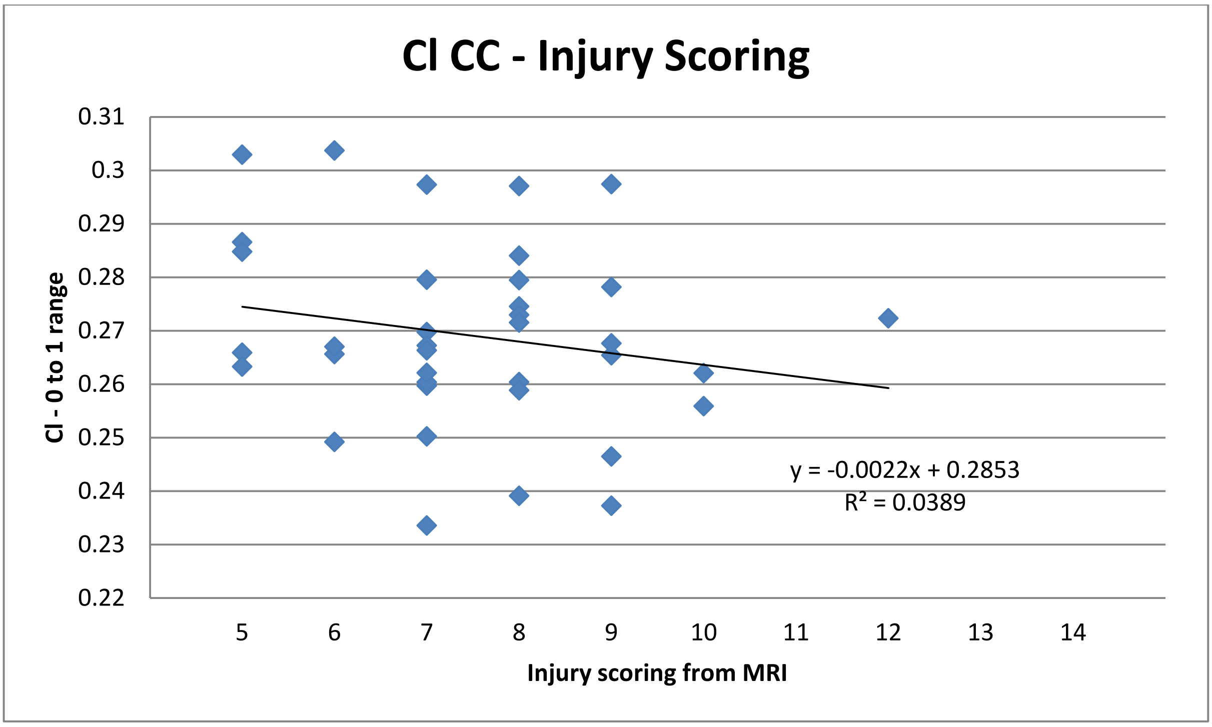


Figure A. 11 – Cl of CC with MRI-based injury scoring and correspondent linear regression.

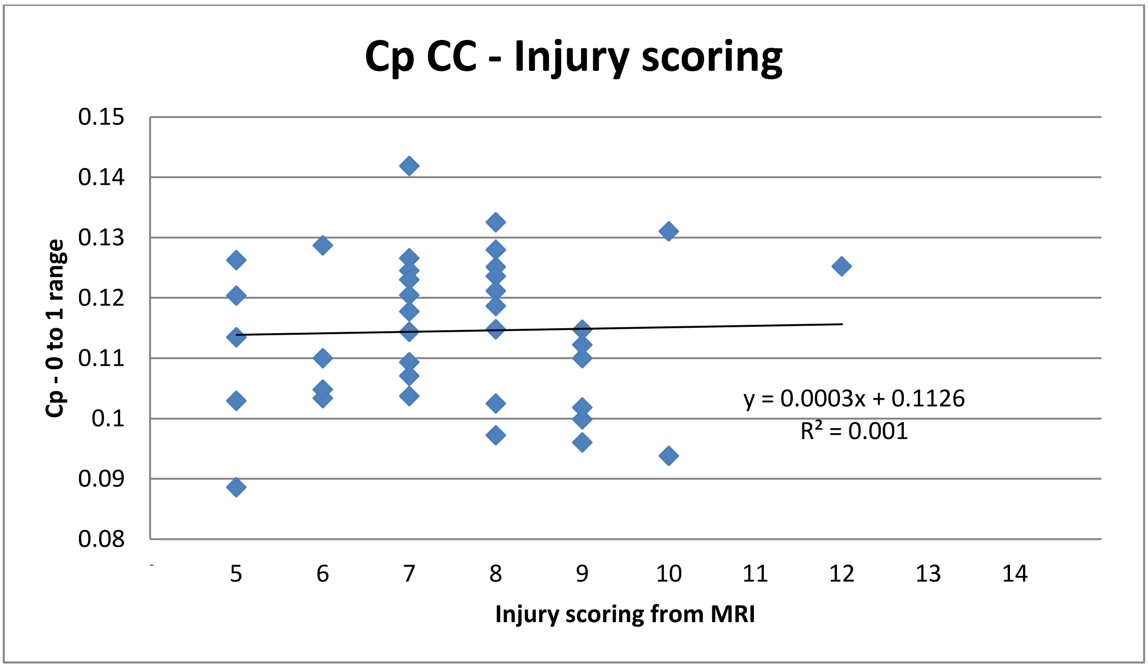


Figure A. 12 – Cp of CC with MRI-based injury scoring and correspondent linear regression.

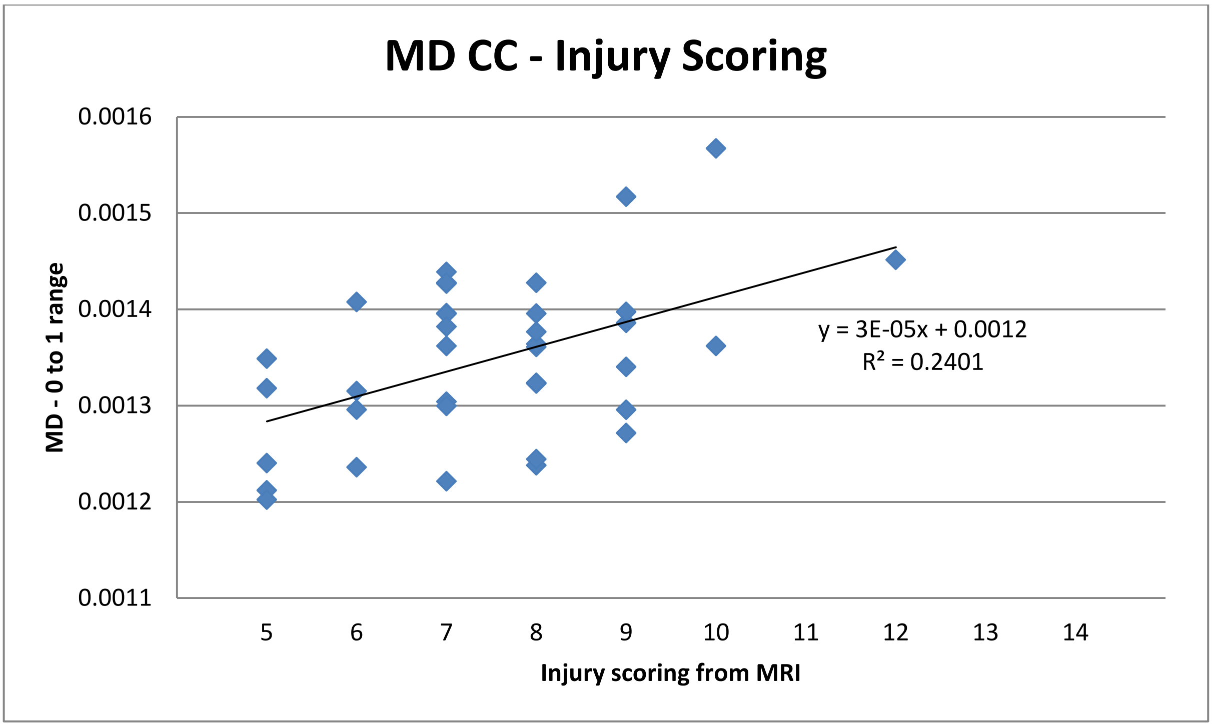


Figure A. 13 – MD of CC with MRI-based injury scoring and correspondent linear regression.

Appendix III – Papers for Submission

Title: Automatic atlas-based segmentation of brain white matter in neonates at risk for neurodevelopmental disorders.

Order of Authors: Lucia Fonseca[1,2,3]; Carola van Pul[2]; Nicolás Lori[4,5]; Rieneke van den Boom[1]; Linda S. de Vries[6]; Manon J.N.L. Benders[6]; P. Andriessen[2]; J. Buijs[2]; Floris Groenendaal[6]; Anna Vilanova[1]

1 Department of Biomedical Image Analysis, Eindhoven University of Technology, The Netherlands
2 Department of Clinical Physics, Maxima Medical Center, Veldhoven, The Netherlands
3 Physics Department, Coimbra University, Coimbra, Portugal
4 Medical School, University of Coimbra, Coimbra, Portugal
5 CIPROMEC, Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal
6 Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands

Suggested Reviewers: Carl-Fredrik Westin - Department of Radiology, Brigham and Women's Hospital - westin@bwh.harvard.edu; Lauren O'Donnell -Department of Radiology, Brigham and Women's Hospital - odonnel@bwh.harvard.edu

ABSTRACT

Maturation degree of anatomical white matter (WM) structures at the brain can constitute a biomarker for prediction of neurodevelopmental disorders already at early post-natal period. Therefore, we propose a new in-vivo and non-invasive global analysis of the WM patterns at the brain: we perform an automatic atlas-based segmentation of Diffusion Tensor Imaging (DTI) tractography; which then allows measuring the maturation degree per segmented anatomical structure. Due to the under-developed stage of the neonatal brain, segmentation approaches which are typically used for adult tractography datasets were rethought and redesigned having into account the neonatal tractography specific characteristics.

Forty-three neonates scanned at term equivalent age (TEA) were automatically segmented by our proposed method. Segmentation performance showed not to be influenced by the presence of WM pathology among subjects. The automatically segmented corpus callosum (CC) structure was further analyzed by studying the respective volume and anisotropy measurements per subject. For the studied measurements, volume and mean diffusivity (MD) showed a significant trend with degree of WM injury. These trends are in accordance with previous findings about how WM injury, and subsequent WM maturation degree, influences DTI derived anisotropies.

Title: Occurrence of 3-axis grids in white matter reduces the amount of isotropic distribution of axonal orientations.

Order of authors: R. Lavrador[1]; L. Fonseca[2,3]; A. Seehaus[4], C. Santos[1]; R. Travasso[1,2];R. Galuske[5], N. F. Lori[1,6]

1 Institute of Biomedical Imaging and Life Sciences (IBILI), Faculty of Medicine, University of Coimbra, Coimbra, Portugal.
2Physics Department, Faculty of Science and Technology, University of Coimbra, Coimbra, Portugal.
3Eindhoven University of Technology, Eindhoven, Netherlands.
4Maastricht University, Maastricht, Netherlands.
5Darmstad University of Technology, Darmstad, Netherlands.
6CIPROMEC, Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal.

ABSTRACT

We developed diffusion MRI data processing methodology that extracts the white matter (WM) axonal orientation distribution function (ODF), and estimates the isotropic and anisotropic fractions of the orientational distribution of WM fibers. The obtained axonal ODF was compared to the structural arrangement of fiber bundles obtained by an advanced tractography method. In our model, we defined the isotropic component of the ODF as a sphere of radius R, and the anisotropic component as Gaussian-like peaks above the isotropic sphere, this model agreeing with microscopy results. In addition, we determined parameters that build the ODF which best approximates the obtained experimental diffusion MRI data. We found that the regions that form a 3-axis grid with the neighboring voxels have a lower isotropic fraction than the regions that form a plane or a simple connection. We also obtained that in the corpus callosum the isotropic fraction is low enough to consider that there is no isotropic axonal ODF.