Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

Inger Sofie Vorren

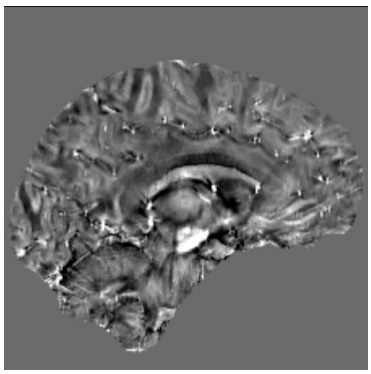
Automated Segmentation for Quantitative Susceptibility Mapping in Healthy Volunteers at 7T

Master's thesis in Applied Physics and Mathematics
Supervisor: Pål Erik Goa
July 2023

NTNU
Norwegian University of
Science and Technology

Inger Sofie Vorren

Automated Segmentation for Quantitative Susceptibility Mapping in Healthy Volunteers at 7T



Master's thesis in Applied Physics and Mathematics
Supervisor: Pål Erik Goa
July 2023

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

Abstract

Parkinson's disease (PD) is a prevalent neurodegenerative disorder and is associated with iron accumulation in subcortical brain structures. The MR imaging technique Quantitative Susceptibility Mapping (QSM) provides quantitative measures of the magnetic susceptibility values in tissue that have been shown to correlate with iron concentrations. QSM is for this reason promising for identification of new biomarkers in neurodegenerative diseases. It is of interest to establish a fully automated segmentation pipeline for extraction of QSM values at 7T to replace time consuming manual susceptibility extraction, and further explore the diagnostic potential of QSM.

The novel Deep Learning (DL) based segmentation tool SynthSeg 2.0 was applied to extract susceptibility values of the thalamus, caudate, putamen, pallidum and hippocampus from 29 healthy volunteers with mean age 29.97 ± 5.73. The QSM images were reconstructed with the integrative Total Generalized Variation method from 7T Multi-Gradient Echo acquisitions. Two different segmentation pipelines were evaluated and compared. The first generated the segmentation map with SynthSeg from QSM input data, and the second with T1w input data. The QSM image was co-registered to the T1w based segmentation, using the FSL FLIRT software. A CNN U-net previously developed in the MR physics group at NTNU was used to segment the Substantia Nigra (SN), the Red Nucleus (RN) and the OMEGA of the motor cortex. A correlation analysis was performed to investigate how well the extracted susceptibility values from the RN and OMEGA correlated to values manually extracted by a radiologist.

The analysis found the SynthSeg segmentation with 7T QSM input to be unreliable, failing to segment the hippocampus and putamen. No statistically significant difference between the segmentation pipelines was found in the susceptibility value extracted for the caudate and pallidum, while the susceptibility values of the thalamus, hippocampus and putamen were found to change significantly ($p < 0.005$). The quantitative comparison of the T1w and QSM SynthSeg masks found Dice Scores (DSs) in the range of 0.83-0.87 for the thalamus, caudate and pallidum, while the left and right putamen and hippocampus scored in the range 0.74-0.80 and 0.51-0.57, respectively.

As mislabeling of the QSM based segmentation were observed mainly in the lateral direction, the failure to segment the putamen and hippocampus is likely partly due to susceptibility artifacts near the air-filled cavities of the ears. The

T1w SynthSeg segmentation produced robust results, and is suggested to be used further for the automated susceptibility extraction pipeline. The raw susceptibility extracted in healthy volunteers from the T1w segmentation of the left thalamus, caudate, putamen, pallidum and hippocampus were found to be 0.45 ± 1.80, 17.05 ± 3.19, 5.23 ± 3.18, 53.86 ± 10.52 and 1.60 ± 2.41 ppb, respectively. The raw susceptibility were found to be 70.16 ± 12.10 ppb for the SN, 47.78 ± 11.75 ppb for the RN and 21.14 ± 4.50 for the OMEGA.

The correlation analysis of the mean SN susceptibility to manually extracted values found by linear regression a R^2-value of 0.58 ($p<0.0001$), suggesting that there is potential for the use of automated values as an alternative to manually extracted susceptibility values for diagnostic purposes. The mean susceptibility extracted from the OMEGA was not found to correlate significantly with the manual values, but further analysis with increased sample size and inclusion of data from patients is suggested.

Sammendrag

Parkinsons sykdom (PD) er en utbredt nevrodegenerativ lidelse og er assosiert med opphopning av jern i subkortikale hjernestrukturer. MR-avbildningsmetoden Quantitative Susceptibility Mapping (QSM) gir kvantitative mål for magnetiske susceptibilitetsverdier i vev som korrelerer med jernkonsentrasjoner. QSM er derfor lovende for å identifisere nye biomarkører for nevrodegenerative sykdommer. Det er av interesse å etablere en fullstendig automatisert segmenterings-pipeline for uthenting av QSM-verdier ved 7T feltstyrke for å erstatte tidkrevende manuell måling av susceptibilitet og videre utforske det diagnostiske potensialet til QSM.

Det nyskapende segmenteringsverktøyet SynthSeg 2.0 er basert på Dyp Læring (DL). SynthSeg ble brukt til å segmentere og hente ut susceptibilitetsverdier for hjernedelene thalamus, caudate, putamen, pallidum og hippokampus fra 29 friske frivillige med en gjennomsnittsalder på 29,97 ± 5,73. QSM-bildene ble rekonstruert med den integrative Total Generalized Variation-metoden fra 7T Multi-Gradient Echo-sekvens data. To ulike segmenterings-metoder ble evaluert og sammenlignet. Den første metoden genererte segmenteringskartet med SynthSeg basert på QSM data, og den andre basert på T1w data. QSM-bildet ble koregistrert til T1w segmentering ved hjelp av programvaren FSL FLIRT. Et CNN U-net tidligere utviklet ved MR-fysikkgruppen ved NTNU ble brukt til å segmentere Substantia Nigra (SN), Red Nucleus (RN) og OMEGA i motorisk cortex. En korrelasjonsanalyse ble utført for å undersøke hvor godt susceptibilitetsverdiene trukket fra den automatiske segmenteringen av RN og OMEGA korrelerte med verdier målt manuelt av en radiolog.

Analysen fant at segmenteringen utført med SynthSeg med 7T QSM input data var upålitelig, da segmenteringen hippokampus og putamen var av dårlig kvalitet. Det ble ikke funnet noen statistisk signifikant forskjell i segmenteringsresultatene mellom de ulike metodene for caudate og pallidum. Imidlertid viste susceptibilitetsverdiene for thalamus, hippokampus og putamen en betydelig endring ($p < 0{,}005$) avhengig av segmenteringsmetoden. Ved å sammenligne T1w og QSM segmenteringene kvantitativt, ble det observert Dice Scores (DSs) i området 0,83-0,87 for thalamus, caudate og pallidum, mens venstre og høyre putamen og hippokampus hadde verdier mellom 0,74-0,80 og 0,51-0,57, henholdsvis.

Feilmerking i den QSM-baserte segmenteringsmetoden ble hovedsakelig observert i den laterale retningen, og dette skyldes trolig susceptibilitetsartefakter fra luftrom i nærheten av øret. SynthSeg-segmenteringen med T1w input data

viste derimot robuste resultater og anbefales for videre bruk i automatisk analysering av susceptibilitetsverdier. Gjennomsnittlig susceptibilitetsverdiene for de 29 friske frivillige ble for T1 segmentene funnet til å være 0,45 ± 1,80, 17,05 ± 3,19, 5, 23 ± 3,18, 53,86 ± 10,52 og 1,60 ± 2,41 ppb for venstre thalamus, caudate, putamen, pallidum og hippokampus, hendholdsvis. For SN ble den susceptibiliteten funnet til å være 70,16 ± 12,10 ppb, for RN 47,78 ± 11,75 ppb, og for OMEGA 21,14 ± 4,50 ppb.

Korrelasjonsanalysen mellom gjennomsnittlig susceptibilitet i SN og manuelt målte verdier ved bruk av lineær regresjon fant en R^2-verdi på 0,58 ($p < 0{,}0001$), som antyder et potensial for å bruke automatiserte verdier som et alternativ til manuelle susceptibilitetsverdier for diagnostiske formål. Gjennomsnittlig susceptibilitet hentet fra automatisk segmentert OMEGA viste derimot ingen signifikant korrelasjon med de manuelle verdiene, men analyse av et større dataset og med inkludering av data fra pasienter vil være interessant for videre studier.

Acknowledgements

This thesis is a continuation of the specialization project initiated the autumn of 2022. First of all, I would like to thank my supervisor professor Pål Erik Goa for the initiation of the project and his guidance throughout the work on the thesis the spring of 2023. I am sincerely grateful for the insight I have gained from the help and feedback, and for the opportunity to work on such an engaging and interesting project. I am also thankful for all the support I have received from the MR physics group at NTNU. I would like to express my deepest appreciation for the friendliness and help extended to me by PhD Candidate Marc-Antoine Fortin and Dr. Anum Masood, as they were always more than happy to answer my questions and share their knowledge. I would like to give a special acknowledgement to Dr. Runa Unsgård. Thank you for sharing your knowledge with me in light of your expertise as a radiologist, and for providing the manually extracted susceptibility values that were crucial for the correlation analysis carried out in the thesis. I highly appreciate your generous feedback. Finally, I would like to express my appreciation to my family and friends. I truly value all your support and encouragement, and I am deeply grateful for all connections and memories I have made during my five years of studies in Trondheim.

Sofie Vorren, Trondheim, 1^{th} July, 2023

Acronyms

AI Artificial Intelligence

ALS Amyotrophic Lateral Sclerosis

ASPIRE A Simple Phase Image Reconstruction For Multi-Echodata

BG Basal Ganglia

BET Brain Extraction Tool

CS Centrum Semiovale

CNN Convolutional Neural Network

CSF Cerebrospinal Fluid

DL Deep Learning

FID Free Induction Decay

FSL FMRIB's Software Library

GRE Gradient Echo

MSE Mean Square Error

NMSE Normalized Mean Square Error

MEDI Morphology Enabled Dipole Inversion

MGRE Multi-echo Gradient Echo

MR Magnetic Resonance

MRI Magnetic Resonance Imaging

NMR Nuclear Magnetic Resonance

NTNU Norwegian University Of Science And Technology

PD Parkinson's Disease

PMC Primary Motor Cortex

PSSC Primary Somatosensory Cortex

QSM Quantitative Susceptibility Mapping

RF Radio Frequency

RN Red Nucleus

SD Standard Deviation

SHARP Sophisticated Harmonic Artifact Reduction For Phase

SN Substantia Nigra

STN Subthalamic Nucleus

SWI Susceptibility Weighted Imaging

Contents

Chapter 1

Introduction

1.1 Motivation

Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder and affects 1-2 % of the population above the age of 65 [1]. In the past two decades, a rapid increase in wold-wide prevalence is observed [2]. The disease is characterized by loss of neurons in the dopaminergic brain regions of the Substantia Nigra (SN) and interruptions of the motor-circuit of voluntary movement in the form of bradykinesia, rest tremor and rigidity [2]. A patient suffering fror PD have a prognosis of progressive disease development and premature death. Currently, there is no therapy to delay or prevent the progression of the disease [1], but several treatments have shown effective for suppressing the symptoms, including brain stimulation and dopaminergic therapy [2]. An early diagnosis is important to initiate a treatment of the symptoms and increase the life quality for patients suffering from PD.

PD is a complex degenerative disorder and occurs with a range of clinical presentations and causes, with some of the causes still unknown [2]. The symptoms arises from the lack of the neurotransmitter dopamine in pathways of the Basal Ganglia (BG) involved in movement. The dopamine depletion is caused by the death of neurons in the dopamine producing brain regions of the BG, particularly in the substantia nigra compacta (SNc). The causes of the degeneration are linked to the presence of α-synuclein aggregates, but are currently not fully understood [2]. Early stages of PD can be challenging to diagnose [3], and exploring new biomarkers are important for stratification of different subtypes of PD [4].

Magnetic resonance imaging (MRI), and particularly the MRI technique Quantitative Susceptibility Mapping (QSM) has shown potential for identifying new biomarkers for PD and other neurodegenerative diseases like Amyotrophic Lateral Sclerosis (ALS). QSM measures the magnetic response of tissue and shows good contrast for the iron rich brain regions of the BG

involved in PD. Voxel based extraction of susceptibility from QSM images have shown a statistically significant prediction of pathology [5]. PD is associated with increased iron accumulation in the SN, and for the later stage of the disease in the red nucleus (RN). Particularly, the SN is of interest as it is possible to detect an increase in susceptibility in the early stages of the disease, that can be helpful for a timely diagnosis. The susceptibility of several other brain regions have been investigated as possible biomarkers, such as the caudate, putamen, pallidum, thalamus and hippocampus [6]. The extraction of susceptibility values from QSM images are dependent on reliable identification of the brain regions of interest (ROIs). Traditionally, susceptibility values are extracted manually by a clinical professional. As this labor is time consuming, automating the process of extracting susceptibility values is advantageous, and will make large scale studies exploring biomarkers more feasible.

The freely available deep learning (DL) based automated segmentation tools from Freesurfer [7] and the Functional Magnetic Resonance of the Brain (FMRIB) Software Library (FSL) [8] does not currently segment all brain regions relevant for PD, and are sensitive to the contrast of the input data. The performance of DL-based segmentation tools are usually highly dependent on the characteristics of the training data, and have traditionally been optimized for MRI data at lower field strengths than 7 Tesla. Imaging at ultra-high field strength creates opportunities with increased spatial resolution, but also introduces increased susceptibility related artifacts and inhomogenities in the magnetic field. It is of interest to investigate whether automated segmentation can be performed directly on QSM images, removing the need for co-registering to an anatomical T1-weighted (T1w) image when extracting the susceptibility values from the segmentation masks. Additionally, automated segmentation based on less conventional image contrasts like QSM could potentially increase the segmentation quality of ROIs involved in PD, as they are of particularly high susceptibility contrast.

The novel segmentation tool SynthSeg [9] shows great promise for segmenting a number of regions of interest (ROIs), including the left and rigth caudate, putamen, pallidum, thalamus and hippocampus based on a large variety of input image contrasts. Sparse documentation is found on the performance of SynthSeg on T1w MRI data at 7 T, and the performance on QSM input data is currently not investigated in the literature. Recently, a Convectional Neural Network (CNN) was developed in the MR physics group at NTNU for segmentation of 7T QSM images, which showed promising results for accurate segmentation of 7T QSM images for PD relevant ROIs including the SN, RN and the OMEGA [10].

1.2 Research Goals

This project is a continuation of the specialization project initialized by the author the autumn of 2022. Through the course of the project and master thesis, QSM images from 29 healthy volunteers in the age range of 20-41 years have been reconstructed, co-registered to T1w images, segmented and analysed. The raw MRI data were acquired at the 7T MR Centre at the St. Olavs Hospital during previous projects [10] [11]. The study also use the QSM images acquired as a part of a currently ongoing study by radiologist Runa Unsgård, as well as susceptibility values manually extracted by the radiologist. This thesis aims to to establish a fully automated pipeline for segmenting and extracting susceptibility values from ROIs that have potential as imaging biomarkers for diagnostics of PD and other neurodegenerative diseases such as ALS.

One research objective of this thesis is to investigate the feasibility of using the novel DL-based tool SynthSeg to segment subcortical regions including the putamen, caudate, pallidum, thalamus and hippocampus with 7T MRI input data. The iron rich brain regions of the BG provide particularly good contrast in QSM images. It is of interest to investigate whether this effect can improve the automated segmentation of certain ROIs by using the QSM image instead of the more traditional choice of T1w for the segmentation. The performance of SynthSeg with QSM input data will be evaluated and compared to segments generated from T1w input data.

As SynthSeg at the writing moment is not trained to segment the SN and RN, this thesis uses a CNN U-network previously developed in the MR physics group at NTNU to segment the ROIs from QSM images. The thesis will extract susceptibility values from the automated segmentation masks and analyze and report the findings of susceptibility values in healthy volunteers. There will be a discussion of the variation expected between healthy subjects, age dependency and comparison of the susceptibility extracted by different segmentation methods. The susceptibility values will be discussed in the context of potential biomarkers for PD.

A final objective of this thesis is to investigate how well the automated segmentation can predict the susceptibility values extracted manually by a radiologist. In clinical practice a radiologist will usually not segment an entire ROI when extracting susceptibility values, instead placing smaller ROIs inside the structure of interest. For this reason, the mean susceptibility of the automated volume segments it is not expected to provide the exact values of the manual segmentations. This thesis performs a correlation analysis to investigate how well the automated segmentation of the SN and OMEGA can predict the manual values.

The main research questions investigated are stated below:

- **Q1** Is SynthSeg with 7T T1w input data feasible for segmenting sub-

cortical ROIs?

- **Q2** Is SynthSeg with 7T QSM input data feasible for segmenting subcortical ROIs?
- **Q3** What are the expected susceptibility values and variation in healthy volunteers for ROIs related to PD?
- **Q4** What is the potential of 7T QSM as a biomarker for PD?
- **Q5** How well can the susceptibility values manually extracted by a radiologist be predicted by automated segmentation?

Chapter 2

Theory

This chapter aims to provide the theoretical basis needed for understanding the topics discussed in this thesis. The material in this chapter is based on the theory of the specialization project initialized by the author the autumn of 2022. Firstly, there is an introduction to magnetic susceptibility and the magnetic properties of biological tissue. A section concerning degenerative diseases and the relevance of QSM as a tool in clinical diagnostics of PD and ALS then follows. The physics behind the nuclear magnetic resonance (NMR) phenomenon and the MRI signal are presented, followed by an introduction to the image reconstruction method of QSM. A brief explanation of the theory of the DL-based segmentation tool SynthSeg is included, as well as an introduction to co-registration. Finally, an introduction to the concepts of DL and CNNs will be presented.

2.1 Magnetic susceptibility of tissue

When a magnetic material is placed in an external magnetic field, the material will induce a magnetization $\mathbf{M}$ given by

$$\mathbf{M} = \chi \mathbf{H}, \tag{2.1}$$

where χ is the magnetic susceptibility defining the magnetic response of the material. $\mathbf{H}$ is the magnetic field strength and is related to the magnetic field $\mathbf{B_0}$ by the magnetic permeability of free space μ_0 and the magnetic permeability of the magnetized material μ_r by the following equation

$$\mathbf{B_0} = \mu_0 \mu_r \mathbf{H}. \tag{2.2}$$

The induced field $\mathbf{M}$ is referred to as the bulk magnetization and it should be noted that this is a macroscopic property arising from averaging over many fields, not a description of the field of a single molecule. When only considering the z-direction, Equation (2.1) can be rewritten as

$$\mathbf{M_z}(\mathbf{r}) = \chi(\mathbf{r})\frac{B_0}{\mu_0\mu_r} = \chi(\mathbf{r})\frac{B_0}{\mu_0(1-\chi(\mathbf{r})}, \tag{2.3}$$

where $\mathbf{B_0} = B_0\,\hat{\mathbf{z}}$. Assuming $\chi << 1$ results in the following approximation

$$\mathbf{M_z}(\mathbf{r}) \approx \chi(\mathbf{r})\frac{B_0}{\mu_0}. \tag{2.4}$$

2.1.1 Paramagnetism, Diamagnetism and Ferromagnetism

The magnetic properties of materials are commonly classified in three categories; diamagnetic, paramagnetic and ferromagnetic, depending on their magnetic susceptibility. This section gives a brief explanation of the physics behind the three types of magnetic materials.

The diamagnetic effect arises from the electron orbitals in response to an applied magnetic field. All materials have diamagnetic properties, caused by the motion of the charges of the electrons, generally inducing a field opposing the direction of an external magnetic field. Paramagnetic properties of unpaired spins will generally be much stronger than the diamagnetic effect. Atoms with even electron numbers therefore typically display a susceptibility close to zero or a weak diamagnetic effect. Note that the diamagnetic properties will depend on the structure of the molecule, for example the arrangement of the water molecule are known to produce an diamagnetic effect with a weak susceptibility $\chi < 0$ [12].

Due to the Pauli exclusion principle, atoms with even electron numbers will occupy energy states with two electrons of opposite spin quantum numbers, causing a cancellation of the total spin magnetic moment. For atoms with an unpaired electron, the single electron spin gives rise to a magnetic moment and can be considered as an magnetic dipole. In the presence of an external magnetic field the orientation of the dipole tends to align with the external field like a bar magnet. Materials magnetizing in the same direction as the applied field is referred to as paramagnetic with a susceptibility value $\chi > 0$.

The materials categorized as ferromagnetic are characterized by the ability to keep their magnetization after the external field is removed. Ferromagnetic materials consist of atoms with unpaired electrons, but in contrast to paramagnetic materials, the spins of ferromagnetic materials have a tendency to align with each other in groups. When applying an external field to an ferromagnetic material, groups of spins will align with the field and combine as the field strength increases. The groups of spins will keep their alignment when the external field is removed, thus creating a permanent magnet. Ferromagnetic materials such as iron have high positive

susceptibility and will induce especially strong magnetic fields when placed in an external field.

2.1.2 Magnetic susceptibility in brain tissue

A large fraction of biological tissue consists of water. In human brain tissue the water fraction is estimated to be 70-85% [13]. Water is weakly diamagnetic with a susceptibility of approximately -9 ppm depending on the temperature [14], and the susceptibility of general brain tissue is expected to be in this approximate range [13], but varies with myelin and iron concentrations [15]. Myelin is slightly diamagnetic compared to average brain tissue due to the diamagnetic properties of the lipids and will cause variations in susceptibility between gray matter and myelinated white matter. Studying the synthesis and degradation of myelin can be of interest for example in relation to neurodegenerative diseases. Biological compounds containing iron are ferromagnetic and provides high contrast to the diamagnetic brain tissue. Iron also plays a role in several biological processes and can be used to trace the functionality and activity of processes such as oxygen transfer, myelin synthesis and neurotransmitter metabolism. Pathology can disrupt the normal iron concentrations in specific brain areas and cause measurable susceptibility changes. Iron accumulation is associated with degenerative diseases such as PD and multiple Sclerosis and can be used as biomarkers in medical diagnostics [15].

2.1.3 Susceptibility in MRI

In MRI the bulk magnetization M_b caused by the susceptibility of the imaged object does not contribute directly to the MRI signal like the nuclear magnetization described in section 2.3. Instead it alters the local magnetic field experienced by the spins and their precessional frequencies. By shifting the precessional frequency, susceptibility sources can create artifacts like geometric distortions by disturbing the spatial encoding. The magnetic field inhomogeneities caused by the susceptibility can also lead to signal loss due to dephasing of spins. Some imaging techniques such as Susceptibility Weighted Imaging (SWI) and QSM use the susceptibility proprieties of tissue to contrast images and provide important information for biomedical practice. The following section continues the discussion of the application of susceptibility as an imaging contrast.

2.2 Neurodegenerative diseases

Neurodegenerative diseases are diseases of the nervous system characterised by a progressive loss of function or structures of nerve cells in the

brain, the spinal cord and peripheral nerves. Neurodegenerative diseases often result in cell death and can affect a variety of functions such as movement, memory and language. Motor neuron diseases are a classification of neurodegenerative diseases where the degeneration primarily affects motor neurons and preserves the other functions of the nervous system. Some of the most common neurodegenerative diseases are Alzheimer's disease, PD and ALS. Imaging techniques such as Computer Tomography and MR is commonly used to provide information in diagnostics of neurodegenerative diseases.

2.2.1 Parkinson's Disease

PD is a neurodegenerative movement disorder characterised by loss of dopaminergic neurons in the Substantia Nigra (SN). To be diagnosed with PD a patient must display two of the three following symptoms; tremor, rigidity or bradyakinesia, where bradykinesia describes a slowness of movement and the difficulty to initiate movements [3]. The causes of PD are not fully understood, but the cell damage are thought to be linked to the appearance of a protein aggregate called Lewy bodies in the Striatum pathway [16]. The loss of neurons in dopamine producing areas such as the SN cause a dopamine lack disrupting the functions of the Basal Ganglia (BG), which are thought to be the root of the symptoms of PD.

The BG is a group of sub-cortical nuclei that are primarily connected to the function of motor control, and includes the brain structures of the caudate nucleus, putamen, pallidum, SN and Subthalamic Nucleus (STN). Patients with PD will suffer significant damage to the neurons of the SN, as well as loss of neurons in the STN. During the progression of the disease, increased brain iron levels are observed, specifically in the Nigrostriatal pathway [17], which facilitates communication between the SN and the caudate and the putamen of the striatum. Increased iron deposition has also been reported in the RN and the STN for patients suffering from PD [18].

As iron is ferromagnetic, accumulation of iron will cause increased magnetic susceptibility in the relevant brain areas. The increased susceptibility can be measured by QSM and used as an indicator of the rate of progression of the disease. In this regard it is of interest to establish a range of normal susceptibility values of healthy individuals to aid the diagnostics of PD. QSM is also used to observe the progression of neuron loss. The areas of the STN, SN and RN are of relatively high susceptibility value and for this reason QSM is favorable to provide good imaging contrast.

2.2.2 Amyotrophic Lateral Sclerosis

ALS is a rare motor neuron disease characterised by progressive degeneration of both upper and lower motor neurons. The patients experience increasing paralysis and weakening of the musculature during the progression of the fatal disease. Patients suffering from ALS have a median survival time of 2.5 years after developing symptoms [19]. The prevalence of ALS in Norway is approximately 4-7 per 100.000 inhabitants[20], increasing with age. There is currently no known cure for ALS, but there are medications documented to slow down the progression of the disease [20]. An early diagnosis is thus critical for patients suffering from ALS. The clinical diagnostics face challenges related to identifying specific biomarkers and overlap with mimic syndromes. Measurements of the susceptibility distribution with the MR imaging technique of QSM have shown promising diagnostic potential of ALS [21].

The causes of ALS are not fully understood, but the progression of the disease is characterized by damage to the nerve cells responsible for transmitting signals from the brain and spinal chord to the muscles to facilitate will-controlled movement. The primary motor cortex (PMC) is located in the Precentral Gyrus and is populated with Betz cells. The degeneration of these cells are one of the hallmarks of ALS [22]. Increased susceptibility values in the PMC due to iron accumulation have shown correlation with disease progression [21], as well as potential as a biomarker for separating ALS from mimicking diseases. ALS affects specific regions of the PMC and the diagnostic feasibility of QSM measurements of ROIs such as the arm and face homunculus is investigated [21]. Other studies investigate correlation of susceptibility in ROIs such as the thalamus with disease progression [23].

2.3 Nuclear Magnetic Resonance

The phenomena of NMR arises when a nucleus with intrinsic spin is placed in an external magnetic field $\boldsymbol{B_0}$. Due to the Spin Quantum Numbers of the nucleus, a quantization of energy levels occurs, where the energy difference between the spin states is specific to the nucleus. By introducing an electromagnetic pulse with photon energy matching the energy gap, it is possible to excite and manipulate the spin states. A characteristic signal is produced, as the nuclei are returning to the equilibrium state. Hydrogen atoms (H) are commonly the nucleus of choice in MRI, with two possible Spin Quantum Numbers $m = +-1/2$. In the following discussion the terms spin, nucleus and H are used interchangeably.

2.3.1 Classical precession

The precessional motion of a spin in a magnetic field can be explained by a classical analogy. As nuclei are charged particles, a rotation such as the nuclear spin $\boldsymbol{I}$ gives rise to a magnetic moment $\boldsymbol{\mu} = \gamma \boldsymbol{I}$, where γ is the nuclei specific Gyromagnetic ratio. Analogous to the behavior of a rotating mass in a gravitational field, the magnetic moment will experience a torque $\boldsymbol{\tau} = \boldsymbol{\mu} \times \boldsymbol{B_0}$ in the presence of an external magnetic field $\boldsymbol{B_0}$. From Newtons second law of motion we have $\boldsymbol{\tau} = d\boldsymbol{L}/dt$, where $\boldsymbol{L}$ is the angular momentum. As $\boldsymbol{I}$ is the nuclear angular momentum and is linearly related to $\boldsymbol{\mu}$, we have the differential equation

$$\frac{d\boldsymbol{\mu}}{dt} = \gamma \boldsymbol{\mu} \times \boldsymbol{B_0}. \tag{2.5}$$

Solving Equation (2.5) results in a precessional motion of μ with Larmor frequency

$$\boldsymbol{\omega_0} = \gamma \boldsymbol{B_0} \tag{2.6}$$

around the axis of $\boldsymbol{B_0}$. It should be noted that this picture describes the expectation value of the spin and is only applicable to the macroscopic quantity arising from an ensemble of spins.

2.3.2 The net Magnetization vector

In the quantum description, a single spin will upon measuring be either parallel or anti-parallel with $\boldsymbol{B_0}$, inhabiting one of the two energy eigenstates $E = -m\gamma\hbar B_0$, where $\hbar$ is Planck's reduced constant and $m = \pm\frac{1}{2}$. The MRI procedure does not measure single spins, but an ensemble of spins inhabiting superpositions consisting of a linear combination of the eigenstates. This ensemble can be described in terms of Schrödinger's wave equation, and we find that the expectation value of the wave function behaves in accordance to the classical picture described in the previous section. The Boltzmann's distribution gives the ratio between the two populations of spin states when the system is in equilibrium

$$\frac{n_\uparrow}{n_\downarrow} = e^{-\Delta E/kT}, \tag{2.7}$$

where T denotes the temperature, k is Boltzman's constant and $\Delta E = \hbar\gamma_0$ is the energy difference between the quantum states. The net magnetization vector $\boldsymbol{M}$ per volume V can be found by averaging over over all possible quantum states.

$$\boldsymbol{M} = \frac{1}{V}\int \langle \boldsymbol{\mu} \rangle dV. \tag{2.8}$$

In equilibrium, only the longitudinal components of the magnetic moments contribute to $\boldsymbol{M}$, as the random phase of the transversal component tends to cancel each other out. The magnitude of the equilibrium value $\boldsymbol{M_0} = M_0\hat{\boldsymbol{z}}$ is given by

$$M_0 = \frac{\rho\gamma^2\hbar^2 B_0}{4kT}, \tag{2.9}$$

where ρ is the proton density.

2.3.3 Relaxation

The net magnetization vector $\boldsymbol{M}$ will undergo two separate relaxation processes, T_1-relaxation and the T_2-destruction process. The T_2-process is the destruction of the transversal component of the signal, due to increasing incoherence in the phase of the spins. T_2 is defined as the time it takes for the transversal component $\boldsymbol{M}_\perp$ to decrease to approximately 37% of its original value. The pure T_2-effect is due to interactions between neighbouring spins, slowing down or speeding up the precessional frequency of individual spins. As it is caused by random spin-spin interactions, this process is irreversible. In addition local inhomogeneities in the magnetic field due to susceptibility effects or deviations in the external field will add to the dephasing process, resulting in an observed effective relaxation time known as T_2^*

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'},$$

where T_2' denotes the relaxation time of the dephasing due to field inhomogeneities. The static field inhomogeneities causes this process to be reversible, in contrast to the T_2-dephasing. The T_1-relaxation process describes the relaxation of $\boldsymbol{M}$ in the longitudinal direction, where $\boldsymbol{M_z}$ exponentially increases towards the equilibrium value of M_0. T_1 is defined as the time it takes for $\boldsymbol{M_z}$ to reach approximately 63% of M_0. This is a process where the spins relax to the lower energy state described in 2.3.2 until the state of equilibrium given by Boltzmann's formula is reached. In a coordinate system rotating with the precessional frequency of $\boldsymbol{M}$, the decomposed relaxation processes of $\boldsymbol{M}$ can be described as a function of time t

$$\begin{aligned} \boldsymbol{M}_\perp(t) &= \boldsymbol{M}_\perp(t=0)e^{-t/T_2^*} \\ \boldsymbol{M_z}(t) &= \boldsymbol{M_0}(1-e^{-t/T_1}). \end{aligned} \tag{2.10}$$

2.3.4 Measuring the MRI signal

The MRI signal is measured by tipping a component of the magnetization vector previously described to the transversal plain. This is achieved by introducing a circularly polarized magnetic field $\boldsymbol{B_1}$ perpendicular to the external field $\boldsymbol{B_0}$. Depending on the time t the flipping the net magnetization $\boldsymbol{M}$ to an flip angle $\alpha = \gamma \boldsymbol{B_1} t$. When removing the $\boldsymbol{B_1}$-pulse, $\boldsymbol{M}_\perp$ will rotate in the xy-plane. By placing coils perpendicular to the external field B_0, the rotating $\boldsymbol{M}_\perp$ will induce a current in the coils, creating a measurable signal containing information about $\boldsymbol{M}$. This signal is the Free Induction Decay (FID) caused by the T_2^*-relaxation, describing the damped oscillations of $\boldsymbol{M}_\perp$. The signal describing the sinusoidal motion of $\boldsymbol{M}$ is complex, and is measured either by acquiring signal from coils in both x and y direction simultaneously, or by phase shifting one of the signal components with $\frac{\pi}{2}$ to acquire both the real and imaginary part. The precessional frequency's linear dependence on the magnetic field makes it possible to spatially encode the MRI signal and produce images by applying the theory of Fourier Transformations.

2.3.5 MGRE

The Multi Gradient Echo (MGRE) sequence is an acquisition method commonly used to acquire T_2^*-weighted MR images. The Gradient Echo (GRE) signal is produced by applying a magnetic field gradient forcing dephasing and re-phasing of the spins to create a signal echo. The gradient alters the precessional frequency and causes spins to acquire different phase shifts at different locations. By applying a gradient of equal magnitude and opposite sign for the same amount of time, the dephasing due to the gradients can be reversed, and an echo can be measured at the echo time TE, as illustrated in Section 2.3.5. As spins along the axis of the gradient will experience different magnetic fields, their precessional frequency can be encoded to their location. The MRI signal is recorded in k-space, and using the Fourier Transform and the encoded relation of the spatial placement and frequency of the signal, one can find the spatial distribution of the signal intensities. The MGRE apply repeated de-phasing and re-phasing gradients to generate multiple echos from one RF-excitation. Each echo can be spatially encoded by shifting the phase of the spins before applying the gradients.

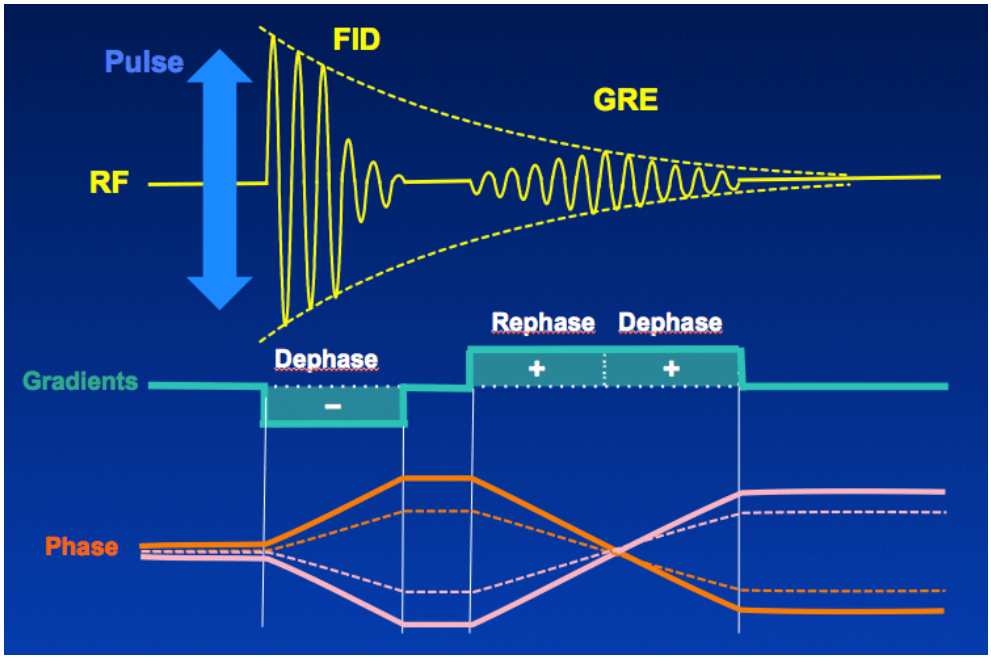


Figure 2.1: GRE acquisition sequence. The lower lines show how the phase of spins at different locations rephase and create the GRE due to the rephase gradient. The figure is retrieved from MRIquestions.com, courtesy of Allen D. Elster.

2.4 Quantitative Susceptibility Mapping

QSM is an MRI reconstruction method that provides the spatial distribution of susceptibility values. The QSM method has the advantage of extracting the quantitative susceptibility values from a single measurement, distinguishing it from other imaging techniques such as SWI and T_2^*- weighted images, where the susceptibility is one of several factors contributing to the image-contrast. As discussed in Section 2.1, the susceptibility distribution can be of interest when assessing diseased tissue and shows great promise as a tool for diagnostics of several neurodegenerative diseases. The potential contribution of QSM in clinical application is still explored. The QSM reconstruction method makes use of phase information from GRE acquisitions to assess the local magnetic field inhomogeneities. After extracting a mask of the ROI and removing the background field, the susceptibility map can be extracted by solving an ill-posed inversion problem, describing the relation between the field inhomogeneities and the susceptibility distribution. This section provides the theoretical background of the inversion problem, as well as a description of the QSM reconstruction process, with focus on the integrative Total General variation method [24].

2.4.1 Magnetic field inhomogeneities and phase

The measured phase of a MRI GRE acquisition is related to the local magnetic B-field inhomogeneities ΔB. The local field experienced by the spins can be expressed as

$$B_l = B_0 + \Delta B, \tag{2.11}$$

where B_l is the total local magnetic field and B_0 is the applied field in the $\hat{\boldsymbol{z}}$-direction. The inhomogeneities in B_l will cause a local change in the angular Larmour frequency of the spins $\Delta\omega = \gamma\Delta B$. After an echo time TE the spins will have acquired a change in phase of $\Delta\phi = \Delta\omega$TE, resulting in a total phase of

$$\phi = \phi_0 + \Delta\omega TE = \phi_0 + \gamma\Delta B \cdot TE, \tag{2.12}$$

where ϕ_0 is the phase at t = 0 and γ is the Gyromagnetic ratio. ΔB can then be calculated as

$$\Delta B = \frac{\Delta\phi}{\gamma TE} \tag{2.13}$$

The MRI measurements only detects the phase data $\phi \in [-\pi, \pi]$, so an unwrapping of the phase may be needed to extract the true phase. Unwrapping of the phase can be performed by adding and subtracting integer multiples n(r) of 2π as follows

$$\phi = \phi_w + n(r) \cdot 2\pi, \tag{2.14}$$

where ϕ_w is the wrapped phase. There are several methods of deciding n(r), relying on either the temporal evolution of the phase or the spatial phase distribution [25].

Another effect to consider is individual coil offset. The software at the 7T MAGNETOM TERRA (Siemens, Munchen, Germany) system at the St. Olav's Hospital currently does not combine the phase data of all coils, so one have to take into account the effect of each coil having an individual coil phase offset when extracting the phase data. To account for this offset the algorithm of A Simple Phase Image Reconstruction for multi-Echodata (ASPIRE) [26] is implemented in the scanners image reconstructor, allowing the phase data to be applied directly. The method of ASPIRE measures the phase at two different echo times to calculate the true phase offset. Given the condition that the phase difference is scaled by an integer, it has been shown that phase unwrapping is not necessary after applying ASPIRE. This condition is fulfilled by choosing echo times TE_i and TE_j satisfying $m \cdot TE_i = (m+1)TE_j$. A detailed description of the method is presented in the paper by Eckstein et al. [26].

2.4.2 Background field removal and Brain Extraction

The QSM reconstruction process involves removal of the background field. In addition to the susceptibility response of the ROI, B-field inhomogeneities can be caused by susceptibility sources outside the surface of the ROI. Susceptibility edges such as tissue-air interfaces can distort the measured data and create artifacts in the image. There are several methods for removing the background field, such as the Sophisticated Harmonic Artifact Reduction for Phase data (SHARP) method, which assumes no susceptibility sources close to boundaries of the ROI [27]. Generally, methods for background field removal are dependent on choosing good regularization parameters and are prone to errors and loss of information around the edges of the ROI. The background removal requires a mask separating the area of interest from the background. For brain imaging this mask can be generated using the Brain Extraction Tool (BET) from the FSL, an analytic software for brain-image data [28]. A visualization of the brain extraction process can be found in Figure 2.2. BET takes a T_2^*-weighted magnitude image as input and estimates the size and centre of mass of the brain. This information is used to initialize a triangular tesselation of a spherical surface. The model is deformed iteratively, subjected to forces and regularization terms until it reaches a certain threshold and estimates smooth edges of the brain [29]. The process can be re-run until a sufficiently smooth surface is achieved.

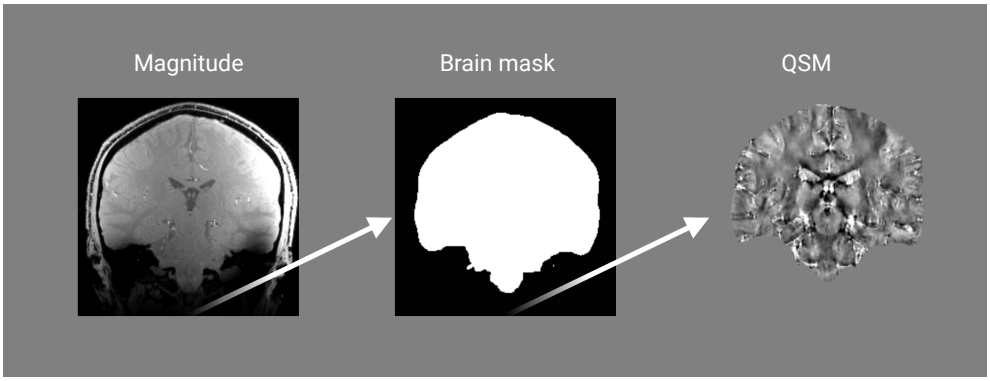


Figure 2.2: Illustration of brain extraction. The BET tool can be used to extract a brain mask from a magnitude image, to be used further in the QSM reconstruction process to separate the relevant data from the background. The figure was created by Sofie Vorren using bioRender.com.

2.4.3 Dipole field inversion

The goal of QSM reconstruction is to obtain the susceptibility distribution $\chi(\boldsymbol{r})$ from the measured field inhomogeneities $\Delta\boldsymbol{B}(\boldsymbol{r})$ retrieved from the

phase data. This relation can be derived by solving a forward problem and expressing the magnetization $\boldsymbol{M}(\boldsymbol{r})$ through $\Delta\boldsymbol{B}(\boldsymbol{r})$. The problem can be investigated by modelling the individual spins as point dipole sources. Note that $\boldsymbol{M}(\boldsymbol{r})$ describes a macroscopic quantity and is not directly describing the individual spins. The relation between the local field inhomogeneities experienced by the spins $\Delta\boldsymbol{B}(\boldsymbol{r})$ and the induced magnetization $\boldsymbol{M}(\boldsymbol{r})$ is therefore not straightforward. The distribution of susceptibility, orientation and geometry of the object will affect the local field of the individual spin. To take into account the discontinuous nature of the magnetization field, the Lorentz approximation is commonly introduced. The Lorentz sphere separates the description of the magnetization close to a dipole source for far field contributions, as illustrated in Figure 2.3.

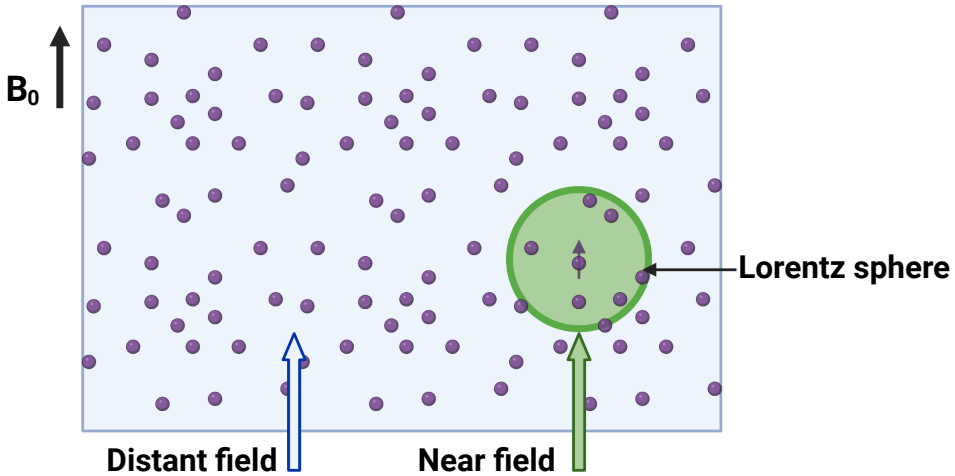


Figure 2.3: Visualization of the Lorentz sphere approximation. The figure illustrates how the contribution to the induced magnetic field $\boldsymbol{M}(\boldsymbol{r})$ is divided into distant field contributions and near field contributions, by a Lorentz sphere centered at position $\boldsymbol{r}$. The purple circles represents the individual spins. The figure was created by Sofie Vorren using bioRender.com.

The area inside the Lorentz sphere accounts for the local susceptibility effects by dealing with individual dipole sources, while $\boldsymbol{M}(\boldsymbol{r})$ is described by continuum theory in the area outside of the sphere.

Assume that the external magnetic field $\boldsymbol{B_0}$ is oriented along the $\hat{\boldsymbol{z}}$-direction and that contributions to the magnetic environment from chemical shifts can be neglected. The Maxwell's equations of static magnetism

$$\nabla \cdot \boldsymbol{B} = 0 \tag{2.15}$$

$$\nabla \times \boldsymbol{B} = \mu_0 \nabla \times \boldsymbol{M}, \tag{2.16}$$

can then be used together with the Lorentz approximation to express the field variation $\Delta \boldsymbol{B}(\boldsymbol{r})$ from $\boldsymbol{M}(\boldsymbol{r})$ and the point-dipole response $d_z(\boldsymbol{r})$ by the following relation [25]

$$\Delta B_z(\boldsymbol{r}) = \mu_0 \int M_z(\boldsymbol{r}) \cdot d_z(\boldsymbol{r} - \boldsymbol{r}')\, d^3\boldsymbol{r}'. \tag{2.17}$$

$d_z(\boldsymbol{r})$ is the dipole kernel in the $\hat{\boldsymbol{z}}$-direction and is a function of position given by

$$d_z(\boldsymbol{r}) = \frac{1}{4\pi} \frac{3\cos\theta - 1}{|\boldsymbol{r}|^3}, \tag{2.18}$$

where θ is the angle between the position vector $\boldsymbol{r}$ and the z-axis. Expressed in k-space with coordinates $\boldsymbol{k}$, the Fourier transformation of $d_z(\boldsymbol{r})$ can be written as

$$d_z(\boldsymbol{k}) = \frac{1}{3} - \frac{k_z^2}{\boldsymbol{k}^2}. \tag{2.19}$$

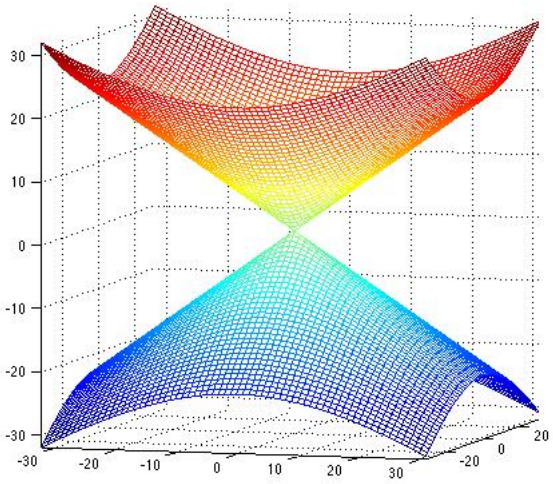


Figure 2.4: Visualization of the zero surface of the dipole kernel in the Fourier domain. The figure was created by Ericl. liu. CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=12809660.

Assuming a susceptibility << 1 and using the expression from Equation (2.4) derived in Section 2.1 to replace $M_z(\boldsymbol{r})$ with $\chi(\boldsymbol{r})$, Equation 2.17 can be re-expressed as

$$\frac{\Delta B_z(\boldsymbol{r})}{B_0} = \int \chi(\boldsymbol{r}') \cdot d_z(\boldsymbol{r} - \boldsymbol{r}')\, d^3\boldsymbol{r}'. \tag{2.20}$$

This expression is recognizable as the convolution between $\chi(\boldsymbol{r})$ and the dipole response $d_z(\boldsymbol{r})$. Upon Fourier transformation Equation 2.20 becomes a simple multiplication in the Fourier domain

$$\mathcal{F}\left\{\frac{\Delta B_z(\boldsymbol{r})}{B_0}\right\} = \chi(\boldsymbol{k}) \cdot d_z(\boldsymbol{k}), \tag{2.21}$$

where $\chi(\boldsymbol{k})$ and $d_z(\boldsymbol{k})$ is the Fourier transform of the susceptibility map and the dipole response, respectively. Equation (2.21) is known as the inversion problem of QSM. It should be noted that $d_z(\boldsymbol{r})$ is not defined for the point $\boldsymbol{r} = \boldsymbol{r}'$, which makes the problem ill-posed. In the Fourier domain the dipole kernel is zero on a double cone surface which is visualized in Figure 2.4, causing a non-unique solution to the inversion problem.

2.4.4 Total General Variation

The dipole field inversion problem can be solved with various methods to reconstruct the QSM image from gradient echo data. One implementation is the integrative Total Generalized Variation (TGV) based QSM reconstruction method (TGV-QSM) presented in the paper by Langkammer et al. in 2015 [24]. The method has the advantage of eliminating propagation of errors, as it combines the phase unwrapping, background field removal and dipole inversion in one single step. The integration of steps also reduces the number of parameters introduced in the reconstruction process, which may be a benefit in practical applications, which often requires a fine tuning of parameters for the different employments. The integrative TGV method produces robust solutions, allowing high quality QSM images to be reconstructed from gradient echo phase data. The method increases the reconstruction quality for low SNR data, allowing a reduction of the scan time without image quality cost compared to other reconstruction methods.

2.4.4.1 Regularization terms

QSM reconstruction often make use of regularization terms to improve the solution of the noise sensitive inversion problem. Some methods incorporate a measure of the variation of the solution in the minimization function to favor smooth solutions. The semi-norm Total Variation (TV) is a commonly used measure of the variation and is given as

$$TV(\chi) = \|\nabla\chi\|_M, \tag{2.22}$$

were ∇ denotes the gradient and $\|\cdot\|_M$ is the Radon norm. The problem of the TV is that it only considers the first derivatives of the solution, which causes staircase artifacts to appear in images that are not piecewise constant. To enforce a higher degree of smoothness on the solution, the TGV regularization term is introduced. The TGV is a semi-norm on a Banach space and the second order TGV^2 can be expressed as

$$TGV^2_{\alpha_{10}}(\chi) = min_w \alpha_1 \|\nabla\chi - w\|_M + \alpha_0 \|SW\|, \tag{2.23}$$

where $\alpha_1 i$ and α_0 are weighing factors and W represents all vector fields. It should be noted that an ideal ratio between the weighing factors has been identified [24], resulting in only one effective weighting parameter. The symmetrized derivative is denoted by S, and applied on a 2D vector $w = (w_1, w_2)^T$, is defined as

$$Sw = \begin{pmatrix} \dfrac{\partial w_1}{\partial x} & \dfrac{1}{2}\left(\dfrac{\partial w_1}{\partial x} + \dfrac{\partial w_2}{\partial x}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial w_1}{\partial x} + \dfrac{\partial w_2}{\partial x}\right) & \dfrac{\partial w_2}{\partial x} \end{pmatrix}.$$

2.4.4.2 Integrative QSM reconstruction method

Traditionally, solutions to the dipole field inversion problem rely on numerical stabilization strategies [27], involving either inverse filtering or iterative methods. The integrative TGV-QSM method uses the iterative algorithm described in (Bredies, 2014; Chambolle and Pock, 2010) [30] to solve an optimization problem by finding the saddle points to a convex-concave problem. A convex-concave problem is an optimization problem where the objective and constraint functions involve both convex and concave terms. The optimization problem of the dipole field inversion is described in the following discussion.

The TGV-QSM method makes use of wrapped phase data ϕ_w directly, through the Laplacian, here denoted by ∇^2, of the unwrapped phase

$$\nabla^2\phi = Im(\nabla^2 e^{i\phi_w} \cdot e^{-i\phi_w}). \tag{2.24}$$

The method of obtaining the $\nabla^2\phi$ from the wrapped phase data is described in the paper by Schofield and Zhu [31]. Using the relation of the phase and the magnetic field inhomogeneities described in Equation (2.13), the Laplacian of the phase can be implemented in the dipole inversion problem of Equation (2.21) as follows

$$\frac{1}{3}\left(\frac{\partial^2\chi}{\partial x^2} + \frac{\partial^2\chi}{\partial y^2} + \frac{\partial^2\chi}{\partial z^2}\right) = \frac{1}{2\pi TE\gamma B_0}\nabla^2\phi \tag{2.25}$$

Further details of the implementation of the dipole inversion on the Laplacian of the phase can be found in the paper by Li et al. [32]. The background field removal is accounted for by introducing an auxiliary variable ψ. ψ is defined as the Laplacian of the difference between Equation (2.25) an a brain mask Ω. The brain mask can be generated from the Brain Extraction Tool described in section 2.4.2. The optimization problem penalizes the ψ variable by a squared L^2-norm and incorporates the TGV regularization term defined in Equation (2.23). The problem can be expressed as a minimization over the susceptibility mapχ and the auxiliary variable ψ

$$\begin{aligned} \min_{\chi\psi} &= \int |\psi|^2 dx + TGV_\alpha^2 \text{ subject to} \\ \nabla^2\psi &= \frac{1}{3}\left(\frac{\partial^2\chi}{\partial x^2} + \frac{\partial^2\chi}{\partial y^2} + \frac{\partial^2\chi}{\partial z^2}\right) - \frac{1}{2\pi TE\gamma B_0}\nabla^2\phi \ in\ \Omega. \end{aligned} \tag{2.26}$$

The dipole inversion is solved by the minimization with respect to χ, while the minimization of ψ corresponds to removal of the harmonic background field. There are several advantages of applying an integrative method, one of them being the avoidance of introducing additional parameters. Background field removal often includes regularization parameters, such as the threshold parameter of the SHARP method. Addition of regularization terms might introduce accosted errors in the solution. In addition to removing propagation of errors this provides a robustness to the TGV-QSM method.

2.5 SynthSeg

SynthSeg is a DL-based tool for brain segmentation. The software is the first segmentation tool to use Convolutional Neural Networks (CNNs) to produce segmentations agnostic to different contrast and resolutions of the input images [33]. A challenge in CNN-based segmentation is that the network will be specific to the resolution and contrast of the training data. SynthSeg overcomes this issue by generating synthetic training data of random variations of contrast and resolution. The increased variations in the training data makes it possible to segment images from different modalities and pre-processing. The software is also robust with respect to healthy and diseased subjects of varying ages. Currently, T1w acquisitions is widely used in neuroimaging studies and for SynthSeg segmentations of MR acquisitions [34]. There is currently little literature on the quality of SynthSeg segmentation using QSM images as input. It is of interest to investigate if QSM contrast could potentially improve segmentation quality of ROIs with

high susceptibility contrast to surrounding tissue. Using the QSM image as basis for automated segmentation might also be favorable for automated extraction of susceptibility values, as the extraction pipeline would not depend on a secondary acquisition for generating the segmentation.

2.6 Linear Co-registration

Co-registration is often used in medical imaging to align multiple images to the same anatomical space. For example co-registering different individuals to the same MNI template [35] allows for direct comparison of voxel values at specific brain regions placed at the same coordinates. The registering is performed by fitting a transformation T to be applied on the original image I_{in} such that the difference of the transformed image $I_C = T \cdot I_{in}$ and the reference image I_{ref} is minimized. Registration is also used to align images of different modalities or contrasts of the same individual. The type of transformation T and degrees of freedom must be considered for different application of registration, and the transformation can be both linear or non-linear. Linear transformations can be categorized in rigid and affine transformations. Rigid transformations includes operations of translations and rotations, and does not alter or morph the geometry of the image. Affine transformations is a type of linear transformations that, like rigid transformations, are linear, but includes the operations of zooming and sheering. This means an affine transformation have the potential to morph the image geometry, but parallel lines are still preserved after the transformation. The choice of transformation will influence the accuracy of the co-registration, and should be chosen in accordance with the expected warping or movement between I_{in} and I_{ref}. Should the images be of different individuals, a morphing of the image would be expected, and a more complex transformation is needed to describe the transition between the spaces. Co-registration between images of the same individual will likely involve movement that can be described by a rigid transformation.

The registration problem can be described mathematically as an optimization problem with a cost function on the form

$$T^* = argmin_{T \in S_T} C(I_{ref}, I_{in} \cdot T),$$

where S_T is the space of transformation of the chosen form. When solving the optimalization problem, it is important to find the global minimum of the cost function, and not a local minimum. Searching through high resolution images is of high computational cost, and often a technique of down-sampling is applied. The linear registration software FLIRT by FSL [36] [37] applies this method of first searching for a global minimum for a low-resolution version of the problem. The optimization process is then

repeated with basis in the previous solution for higher resolutions, until a global solution for the original resolution is found. This way the computational speed can be drastically reduced, while the transformation giving the highest similarity between the input image and the reference image is identified.

2.7 Deep Learning

Recently, the impact of Artificial Intelligence (AI) in a variety of fields have been rapidly increasing, including in the medical sciences. Machine learning is a branch of AI where statistical algorithms are used to let machines improve through experience by recognizing patterns in data. DL is a subset of machine learning, and is characterised by the implementation of artificial neural networks, inspired by the function of the biological neurons found in the brain. Deep Neural Networks have proven particularly useful for tasks involving the extraction of complex features, such as image classification, speech recognition and text generation.

2.7.1 The Artificial Neuron

The main building block of artificial neural networks is the artificial neuron, commonly referred to as the perceptron. A simple neuron can be modeled as a function taking in a sum of input values x_n, and returning the sum of the input values scaled with the weighing factors w_n and with an added bias b as follows

$$\sum_{n=1}^{N} \sigma(w_n x_n + b). \tag{2.27}$$

σ denotes the activation function, a central concept in neural networks. The activation function introduces non-linearity in the perceptron, and in this way allows for more complexity in the neural network. The activation function is inspired by the role of the action potential we find in biological systems, as it gives a relation of how the input to the neuron relates to the output signal. In biological neurons a certain threshold value in terms of the cell membrane potential must be reached for the neuron to fire and forward information to other neurons [38]. Depending on the purpose of the neural network, the activation function can be chosen to have specific characteristics such as thresholding. The activation function can also be used to control the range of output values of the perceptron. The output range can be adapted to the specific task of the network. For example if the goal is to determine the probability that the input belongs to a classification

label, output values between 0-1 representing a probability might be suitable. A traditional choice for the activation function in artificial neurons is the Sigmoid function

$$g(x) = \frac{1}{1 + e^{-x}}, \tag{2.28}$$

a differentiable function that restricts the output values to the range between 0 and 1. Other widely used activation functions are the hyperbolic tangent, the Rectified Linear Unit (ReLu) and the Exponential Linear Unit (ELU) functions, each characterized by different advantages and challenges.

2.7.2 Artificial Neural Networks

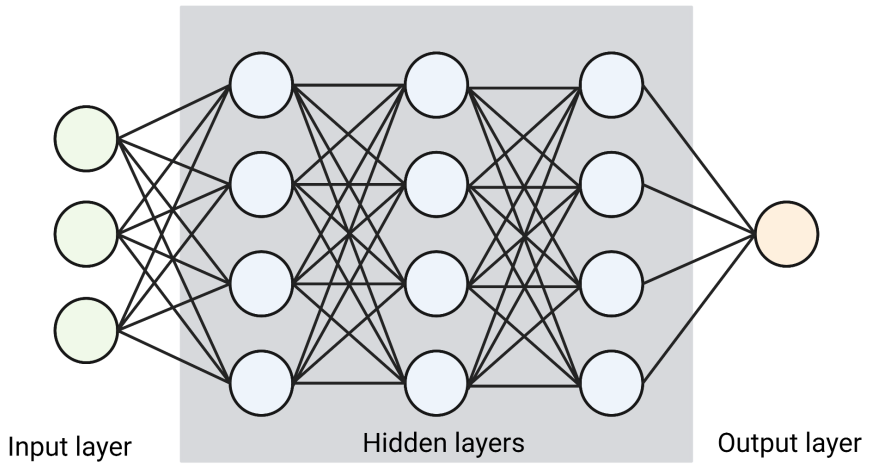


Figure 2.5: Example of an artificial neural network with 3 hidden layers. The artificial neuron is represented by a single circle. The figure was created by Sofie Vorren using BioRender.com.

The perception described in Section 2.7.1 is a single layer neural network. By letting multiple neurons communicate by forwarding the output of one layer to another, complex networks can be formed. These multi layer networks are the basis of DL. An example of a simple neural network consisting of five layers is displayed in Figure 2.5. Here the input layer is connect to three hidden layers consisting of four perceptions each. The output is obtained by combining the outputs of the third hidden layer.

A neural network can be trained to optimize the weights and biases making up the parameters $\theta = \{w_n, b\}$ to minimize a loss function $L(x, \theta)$,

where x is the input. The loss function is used as a measure of how well the output of the network corresponds with a correct solution, and can be defined in different ways, depending on the specific application of the network. In machine learning, supervised learning refers to the use of a training dataset with a labeled ground truth. Deep neural networks often use supervised learning to minimize the loss function, representing the difference between the predicted output and the correct label. The training becomes an optimization problem, with the loss function as the objective function to be minimized. The problem is solved by iterating the training data through the network, calculating the loss function and updating the parameters θ. The process is repeated until the loss function is minimized to a certain threshold of acceptable difference. One efficient method is the gradient decent technique. The method updates the parameters by taking the partial derivatives of $L(x, \theta)$ with respect to each parameter, and adjust the parameter in the opposite direction of the gradient with a distance d. Using this method, the next iteration with the updated weights is likely to have a lower loss function. Note that it is often not preferable to minimize the loss function too much, as this might cause the network to be too specifically tuned for the training data and possibly reduce the performance on new data.

The choice of loss function depends on task of the network and the data type to be processed. In the case of image segmentation, the DS is a widely used metric for measuring the similarity between images. The DS is defined as

$$DS = \frac{2 \cdot |P \cap T|}{|P| + |T|}, \tag{2.29}$$

where P is the predicted labels by the U-net, T is the ground truth, and $P \cap T$ denotes the intersection of the two. The DS ranges from 0 to 1, 1 indicates perfect similarity.

2.7.3 K-fold cross validation

When training a neural network, the available labelled data is often divided into a training set and a validation set. By not using the entire dataset for training, one can evaluate how well the network performs on new data. Another method of evaluation is k-fold cross validation, often used when the training data is sparse to allow the network to learn from all the available labeled data. In k-fold cross validation, the labeled data is divided into batches. For example, if one have a small dataset of 15 images, one can create 15 batches of data, each excluding one image that is used for validation. The network is then fit to one of the batches, and validated on the excluded image. The evaluation score from each validation is saved,

and the process is repeated for the remaining batches. The final evaluation of the performance of the network can be retrieved from the saved evaluation scores. In this way, the entire dataset is used for the training and validation, in contrast to separating out a validation set that are not involved in the training process.

2.7.4 Convolutional Neural Networks and U-nets

The CNN is a type of deep neural network commonly used for problems where spatial information is important, such as in image classification tasks. CNNs are particularly good at recognising and extracting complex features of data, and are defined by the use of convolutional layers. In a convolutional layer, the input data is convolved with a filter, also known as a kernel.For a 2D input image, the filter is typically a 3x3 matrix containing the parameters to be adjusted during the training of the network. The filters can be trained to recognise specific features in the data that are relevant for the task given to the network.

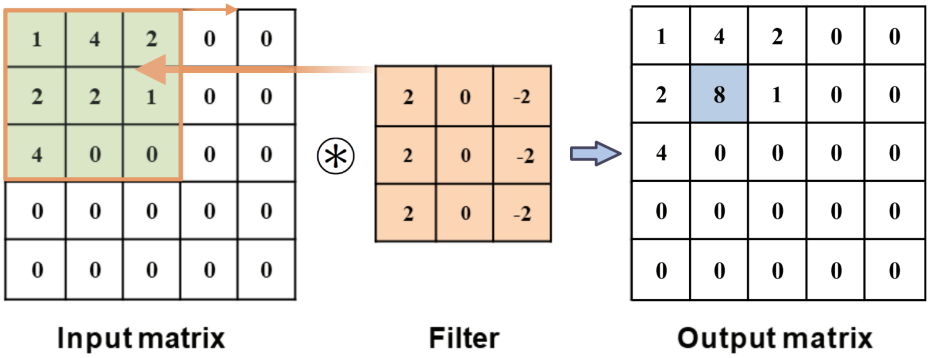


Figure 2.6: Visualisation of a filter operation applied on a 5x5 input matrix in a convolutional layer. The filter is slid through the pixels of the input matrix to obtain the full output matrix. The figure was created by Sofie Vorren using BioRender.com.

Figure 2.6 illustrates an example of the operation of a convolutional layer. The convolution of the filter with the input matrix is performed by sliding the filter over the pixels of the input image. At each position, the dot product between the filter and the overlapping input matrix is calculated and saved in the output matrix, also known as a feature map. In this way, the spatial information in the input image is considered, making the method suitable for image classification problems. If for example the filter is tuned to detect sharp edges, the output matrix will have increased values at positions where the input image contains edges.

A pooling layer is often added after convolutional layers, to reduce the resolution of the feature map. This will increase the feature information density to be passed on to subsequent layers of the network. In a CNN, the first convolutional layers will typically identify simple features like horizontal or vertical edges. Deeper layers usually extracts more complex features and shapes, building on the information extracted in the prior layers. The information is then combined in fully connected layers, that combines the feature information extracted in the convolutional layers to a final output classification. The CNN can in this way learn to recognise and extract abstract properties from the data.

The U-net is a type of CNN structure where the layers are connected in a U-shape, and is a typical choice for image segmentation tasks. The U-shaped architecture allows for an encoding and a decoding path. The encoding path is responsible for detecting and abstracting features from the input data, while the decoding path will upsample the data to match the resolution of the input image in a final segmentation map.

Chapter 3

Method

This section describes the methods applied in the thesis. Some of the material in this section are based on the specialization project written by the author the autumn of 2022. There will be a description of the dataset, including demographics, acquisition parameters and the QSM processing pipeline. The segmentation pipelines investigated will be presented, as well as the method applied for evaluating the segmentation quality. Finally, the method of the segmentation of the SN, RN and OMEGA by the CNN U-net and the correlation analysis to manual susceptibility values will be described.

3.1 Data acquisition

The raw MRI data used in this study were acquired at the 7T MR Centre at NTNU during previous projects [10] [11], and as a part of a currently ongoing study by radiologist Runa Unsgård with REK approval number 108066. The total dataset consisted of 29 healty volunteers, in the age range of 20-41, with a mean age of 27.97 ± 5.73 years. The data were acquired on a 7T MAGNETOM TERRA (Siemens, Munchen, Germany) with a 32-channel phased array-receiving coil. The T1w images were acquired with a MP2RAGE sequence with a spatial resolution of $(0.75mm)^3$ and a matrix size of $224 \times 300 \times 320$ using a GRAPPA coil mode. The acquisition parameters for the TR, TE, TI and FA were 4300 ms, 1.99 ms, 840 ms and 5°, respectively.

For the QSM acquisition, the imaging sequence performed was a Multi-echo GRE with TR = 31 ms, acquiring 4 echoes with 4.7 ms spacing, starting at echo time TE = 2.54 ms. The flip angle was 12° with a FoV = 23 cm and a spatial resolution of $(0.75mm)^3$. The number of axial slices acquired was 224. The algorithm of ASPIRE, described in section Section 2.4.1, was implemented on the scanners image reconstructor by the vendor and was initially performed on the phase data.

3.2 QSM reconstruction

The QSM images were reconstructed by performing background field removal and dipole inversion using the integrative TGV-QSM method [24] described in Section 2.4.4.2. As ASPIRE was already performed on the phase data, it was not necessary to unwrap the phase data. Additionally, the integrative TGV-QSM method performs the phase unwrapping as a part of the one-step method. The FSL library and TGV-QSM software are not compatible with the DICOM format, so phase and magnitude images were first converted to nifti format using the MRIcron conversion tool dcm2nii. A brain mask was then generated using the FSL library. From *Fslutils*, *fslroi* was used to separate the brain and skull from the background, then the BET tool [28] described in section Section 2.4.2 was applied to extract the brain from the skull. Using the generated mask and phase data, the TGV-QSM method was applied for each of the 3 last echoes. This step performed the dipole inversion with the background field removal incorporated as described in Section 2.4.4.2. The final QSM image was calculated using *fslmaths* as the median of the three single echo QSM images. Figure 3.1 illustrates the steps of the process in a flow chart.

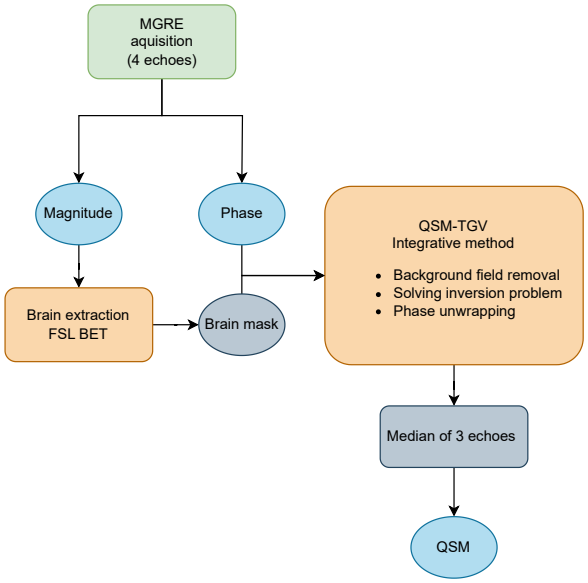


Figure 3.1: QSM reconstruction pipeline. Phase and magnitude data is acquired from a MGRE sequence of 4 echoes. The magnitude data is used to create a brain mask using the BET of FSL. Then TGV-QSM software uses the mask and phase data to solve the inversion problem for 3 separate QSM. In the final step, the 3 QSM images are combined in a single QSM image by taking the median.

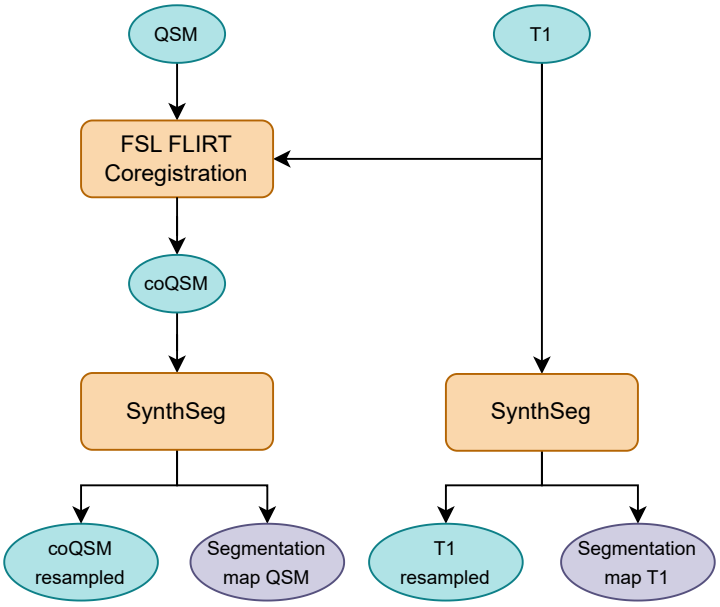


Figure 3.2: The segmentation pipelines of SynthSeg with QSM input and SynthSeg with T1w input data.

3.3 SynthSeg segmentation

In this thesis, two segmentation pipelines were evaluated for the 29 subjects, both using the DL-based brain segmentation software SynthSeg described in Section 2.5. The first pipeline uses the QSM image as input to SynthSeg to generate the segmentation map, while the other generates the segmentation map from T1w input data and uses the FLIRT co-registration from the FSL library to co-register the QSM image to the T1 space. Section 3.3.1 describes further details of the co-registration performed. The two pipelines are illustrated in Figure 3.2. Note that the resulting QSM and T1 segmentation maps are in the same T1 space, to allow for a quantitative comparison between the segmentations. The co-registering of the QSM image to the T1 image in the QSM segmentation pipeline would not be necessary if the QSM segmentation is found reliable and used for the final susceptibility extraction pipeline.

SynthSeg was run with the robust option to obtain optimal segmentation quality. The flags 'qc' and 'vol' were included to capture informa-

tion about the predicted quality scores of the segments and the segmentation volumes for further analysis. 38 different ROIs were segmented by SynthSeg, but a limited number was chosen for the analysis of this thesis. Specifically the right and left thalamus, caudate, putamen, pallidum, and hippocampus, as well as the general CSF. Note that SynthSeg always produce segmentation maps with 1 mm isotropic resolution. A resampled input image can be saved by including the 'resample' flag. As identical matrix size was needed to correctly extract the susceptibility from the ROIs, the resampled QSM image was used for extracting the susceptibility values.

3.3.1 Co-registration

In order to extract QSM values from the segmentations generated from T1 images, a co-registration of the QSM images to the reference T1 images is necessary. The linear registration tool FLIRT from FSL [37] was applied, using a rigid transformation with 6 parameters. As the QSM and T1 images were of the same individual and from the same scanning session, the rigid transformation with 6 parameters was thought to be the most suitable to describe the movement. The co-registered QSM images were all inspected in the image viewing software MRIcron [39]. For 3 subjects a more careful visual analysis was performed with the analytic image tool 3DSlicer [40], to ensure that the co-registered QSM image was moving compared to the original, and that the co-registered image was overlapping with the anatomy of the T1 image.

3.3.2 Evaluation of segmentation quality

The segmentation quality of both the QSM and T1 segmentation was evaluated in terms of a qualitative analysis, the segmentation volumes and the QC scores. Additionally, the segmentation pipelines were compared and a quantitative analysis was performed by calculating the DS of the whole segmentation map and the average DS for each ROI separately. The qualitative analysis was performed using the imaging viewing software MRIcron [39]. The segmentation masks were overlapped on the QSM and T1 images and inspected visually. Representative example images were generated for 3 of the subjects to be included in a more detailed analysis. The visual analysis investigated the correspondence with the QSM and T1 contrast and the expected anatomy in reference to the brain atlases of *7.0 Tesla MRI Brain Atlas* [41] and the online atlas of IMAIOS. It should be noted that this analysis was not performed by a trained professional, and is therefore limited. Still, the analysis were able to detect obvious mislabelling and deviations of the segmentation boarders.

SynthSeg implements a regressor trained to predict the quality of the

segmentation in the form of QC scores. Segments where QC scores below 0.65 are predicted are rejected. Studies have documented that the QC scores correlates well with manually discarded segments [34]. The QC scores are predicted only for certain groups of ROIs, specifically for the putamen and pallidum, the hippocampus and amygdala and separately for the thalamus. QC scores for the caudate is not predicted by SynthSeg, as the regressor is only implemented for some representative groups of ROIs to indicate the general segmentation quality of the whole map. The QC scores of each group is reported averaged over all subjects. The average volumes of each segmented ROI were also calculated and compared with the expected volumes from literature. Considering the purpose of the segmentation, the extracted susceptibility values were analysed and compared between the segmentation pipelines, as discussed in Section 3.3.3.

3.3.3 Analysis and extraction of susceptibility values

The mean and median susceptibility were extracted from the resampled QSM image from the resulting masks of the QSM and T1 SynthSeg segmentation. The extraction was performed with Python, using the numpy, nibabel and pandas libraries. For each ROI, the susceptibility averaged over all subjects was calculated, as well as the inter-subject standard deviation (SD). The intra-ROI SD was also recorded and averaged over all subjects, and a two-tailed t-test were performed to investigate whether the intra-ROI SD changed significantly with the segmentation pipelines. Violin plots were used to compare the distribution of the mean susceptibility from the QSM and T1 masks for each ROI. The difference in susceptibility of the two segmentation pipelines is also reported in terms of the magnitude difference of both the mean and median in ppb and the p-values from a two-tailed t-test. A normal distribution of the data was assumed.

The susceptibility extracted from the masks of the T1 SynthSeg segmentation were further analyzed with regard to age dependency, by plotting the raw susceptibility to the subject age. Susceptibility values of the manually segmented centrum semiovale (CS) were available for 19 of the subjects. The CS normalized values were additionally plotted to the subject age, to investigate the effects of normalization on the age. The age dependency analysis were also performed for the susceptibility extracted from the U-net segmentations of the SN, RN and OMEGA.

3.4 U-net segmentation of the SN, RN and OMEGA

3.4.1 U-net and choice of weights

The automated segmentation of the SN, RN and OMEGA were performed with a CNN created and trained by a previous master student [10] in the MR physics group at NTNU. The CNN is a 2D U-net, trained separately for each ROI on whole ROI segmentations drawn manually by radiologist Runa Unsgård. The OMEGA segment is a sub-segment of the motor cortex. The training data of the U-net consisted of 15 of the 29 subjects included in the dataset of this thesis. The U-net have an initial 16 filters in the first layer, and 256 filters in the lowest layer, and is described in further detail in the thesis by Eivind Lysheim [10]. The performance of the U-net was evaluated in the mentioned thesis by a 15-fold cross validation, resulting in a DS of 0.90 ± 0.01 for the SN, 0.94 ± 0.01 for the RN and 0.86 ± 0.02 for the OMEGA. As a saved set of final weights tuned from this 15-fold cross validation was not found in the available code, the best weights from a 5-fold cross-validation for the SN was chosen for the segmentation performed in this project. The best weights from a 5-fold cross-validation was also selected for the U-net segmentation of the RN. As a ground truth of the SN and RN segmentation were available for the 15 first subjects, a new evaluation of the DS of the predicted masks were performed before deciding on the chosen set of weights used to predict the segmentation of the full dataset. For the OMEGA, the saved weights from the 15-fold analysis were used for the 15 subjects in the training set, and one of the weights was chosen from the 15-fold evaluation, the 13th, to be used for all the new 19 subjects segmented by the U-net. The reason for this method was that the only available weights for the OMEGA was the 15 k-fold weights, and by using the 15-fold weights for the old dataset it was possible to avoid segmenting subjects using weights trained on the same subject. The mean susceptibility value and inter-subject SD were then extracted from the QSM images using Python libraries.

3.4.2 Correlation of predicted and manual susceptibility values in the SN and OMEGA

It is of interest to investigate if the automated segmentations can be used to predict the susceptibility values that are currently extracted manually by radiologists in clinical practice. For 19 of the subjects in the dataset, with mean age 27.94 ± 6.00, susceptibility values of the SN and OMEGA were extracted manually by radiologist Runa Unsgård. The method used by the radiologist was to manually place circular ROIs of minimum 2 mm diameter in one slice of the QSM image, and measuring the mean suscept-

ibility of this area. For the OMEGA the manual ROI was placed at the top part of the OMEGA. It should be noted that as the manual values were extracted from much smaller ROIs than the whole structure segmentation performed by the automated segmentation, this might affect the correlation of the two. It should also be noted that the manual method avoided areas containing blood vessels, characterised by increased intra-ROI SD. Both the manual and automated susceptibility values used in the following correlation analysis are the raw susceptibility values from the QSM image, and are not adjusted to a reference region. The correlation of the mean of the automated values to the manual values adjusted to the manually measured CS by the radiologist were tested by linear regression, but no changes in the correlation were observed.

A linear regression was performed in python, fitting the mean susceptibility extracted from the volume segmentations predicted by the U-net described in Section 3.4.1 to the values manually extracted by the radiologist from the smaller circular ROIs. The goal was to see how well the automated values would be able to predict the manual values, considering the manual values as the gold standard in the analysis. To evaluate the correlation, the R^2-value, the MAE (Mean Absolute Error) and the MSE (Mean Squared Error) between the predicted and the manual values were taken into account. Assuming a normal distribution of the values, a two-tailed t-test was performed to investigate the statistical significance. An exploratory method was used to investigate if other statistics had a better correlation to the values extracted manually than the mean susceptibility value of the automated segments. The same linear regression and calculation of p-values were performed for each statistic to the manual values for both the SN and OMEGA. Specifically, the statistics investigated were the median, max value and the 80, 90, 95 and 98 percentile values of susceptibility in the corresponding automated volume segments.

Chapter 4

Results

This section presents the results of the thesis in three parts. The first part concerns the SynthSeg segmentation of the caudate, pallidum, putamen, thalamus and hippocampus. The segmentation quality is reported in terms of a qualitative analysis, the ROI volumes and the QC scores predicted by SynthSeg. The difference between the segmentation maps generated with SynthSeg from QSM and T1 images are analysed quantitatively with regard to the DS and the extracted susceptibility values. The raw susceptibility values extracted from the masks are presented with the mean, median, inter-subject SD and intra-ROI SD. The second part concerns the U-net segmentation of the RN, SN and OMEGA, and reports the raw susceptibility values extracted from the automated segmentation with the inter-subject SD. An analysis of the correlation of the automated SN and OMEGA susceptibility values to manual values extracted by a radiologist is presented. The final part analyses the extracted susceptibility values further in terms of age dependency and normalization.

4.1 SynthSeg segmentation of the caudate, pallidum, putamen, thalamus and hippocampus

4.1.1 Qualitative analysis

4.1.1.1 Qualitative analysis: QSM input

This section presents visual examples of the masks generated by SynthSeg using 7T QSM images as input. The examples are all of the same individual, subject A. Figure 4.1 shows the segmentation of the left and right caudate, putamen, pallidum and thalamus overlapped on the input QSM image for one subject in the axial plane. The figure shows that the segments are placed in the expected locations anatomically, and that the segmentations seems to be continuous. It is notable that the right caudate does not seem

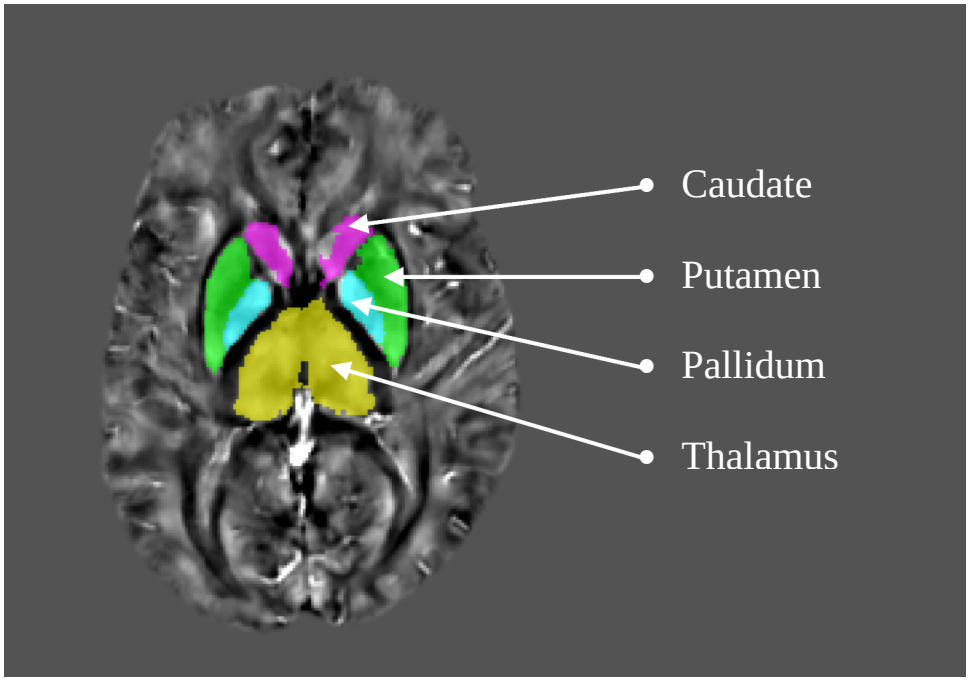


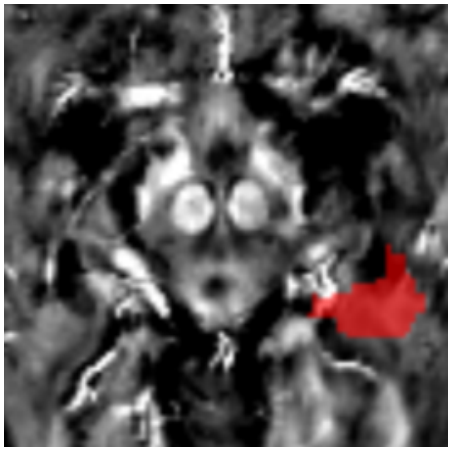
Figure 4.1: Segmentation of the caudate, putamen, pallidum and thalamus overlapped on the input QSM image in the axial plane. The segments are generated with SynthSeg with the 7T QSM image as input.

to overlap exactly with the visible anatomy of the QSM image.

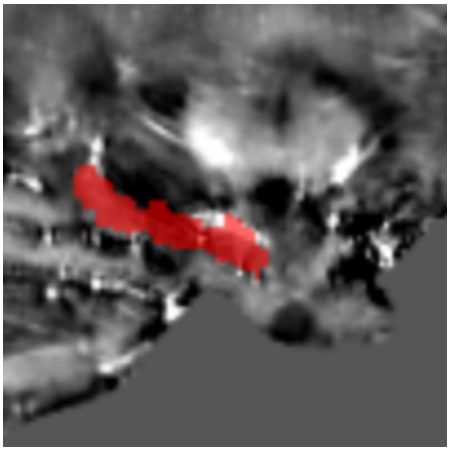
Figure 4.2 shows an example of the segmentation of the right hippocampus in the axial and sagittal plane. In the sagittal plane, the uneven boarders of the segment are particularly visible. The segmentation of the thalamus is displayed in Figure 4.3, and seems to generally correspond with the visible anatomy. Some over-labeling in the left direction is observed in the axial plane, and in the sagittal plane some parts of the segment seems to be missing towards the lower left part compared to the visible anatomy. An example of segmentation of the right putamen is displayed in Figure 4.4, and seems to correspond with the anatomy of the QSM image. Some likely mislabeling towards the right in the coronal plane is observed. The segmentation of the pallidum seems to correspond with the visible anatomy in Figure 4.5. No apparent mislabeling is observed for the caudate in Figure 4.6.

4.1.1.2 Qualitative analysis: T1w input

This section presents visual examples of the segmentations generated by SynthSeg using 7T T1w images as input. The examples presented here are all from the same subject A presented in Section 4.1.1.1, in the examples

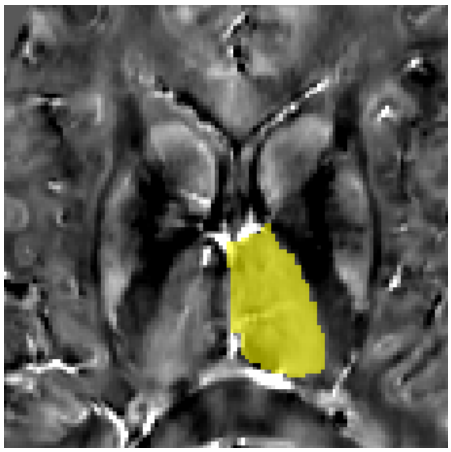


(a) Axial

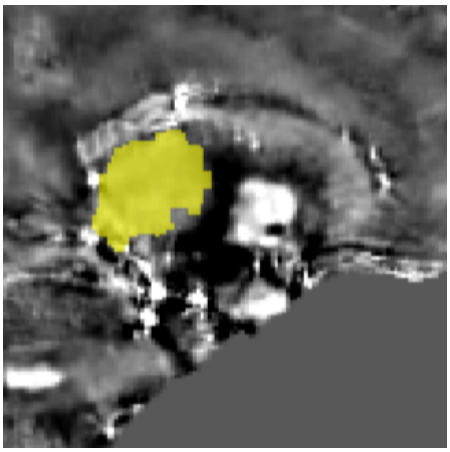


(b) Sagittal

Figure 4.2: Segmentation of the **hippocampus**. The segments are generated with SynthSeg with a 7T QSM image as input.

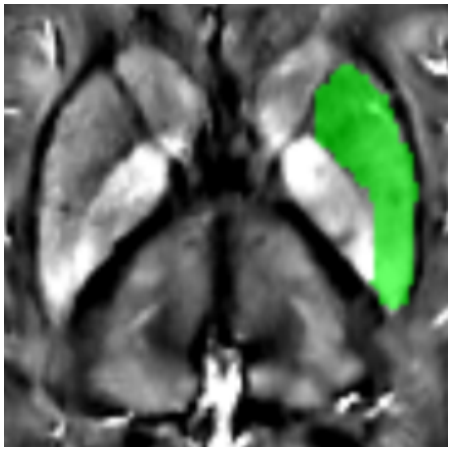


(a) Axial

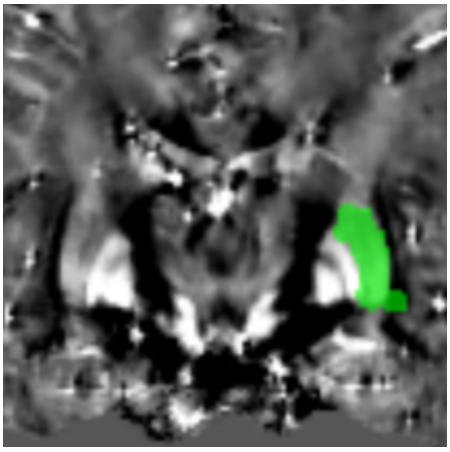


(b) Sagittal

Figure 4.3: Segmentation of the **thalamus**. The segments are generated with SynthSeg with a 7T QSM image as input.

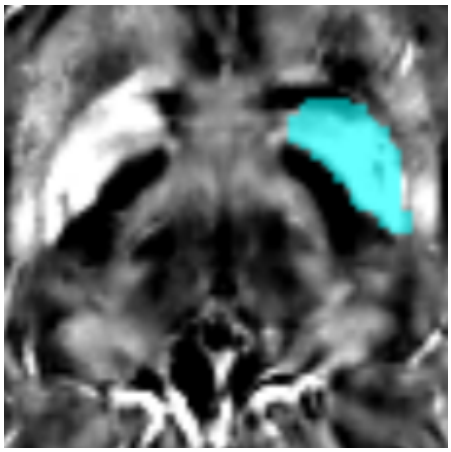


(a) Axial

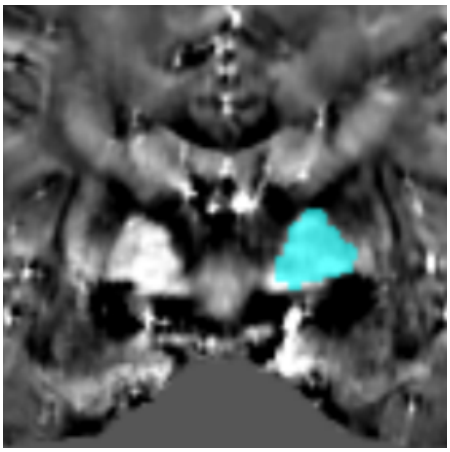


(b) Coronal

Figure 4.4: Segmentation of the **putamen**. The segments are generated with SynthSeg with a 7T QSM image as input.

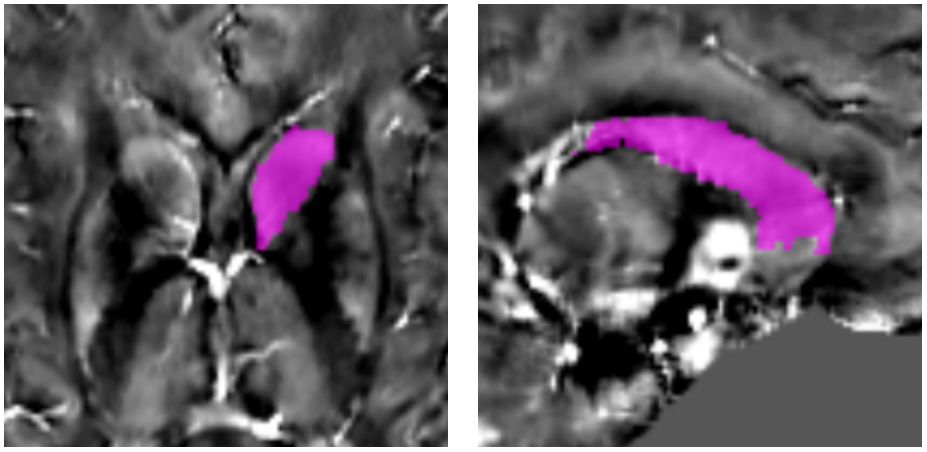


(a) Axial



(b) Coronal

Figure 4.5: Segmentation of the **pallidum**. The segments are generated with SynthSeg with a 7T QSM image as input.



(a) Axial (b) Sagittal

Figure 4.6: Segmentation of the **caudate**. The segments are generated with SynthSeg with a 7T QSM image as input.

for the QSM segmentation. Figure 4.7 shows the T1 segmentation of the left and right caudate, putamen, pallidum and thalamus overlapped on the QSM image for one subject in the axial plane. The figure shows that the segments seem to correspond well with expected anatomical placement, and that the segmentations seems to be continuous. It is notable that the right caudate does not seem to overlap exactly with the visible anatomy of the QSM image.

Figure 4.8 shows an example of segmentation of the right hippocampus in the axial and sagittal plane. In the sagittal plane particularly, we observe that the segmentation boarders are much smoother compared to the segmentation of the same individual generated from the QSM image in Figure 4.2. The T1 segmentation of the thalamus is displayed in Figure 4.9, and seems to correspond well with the visible anatomy. The under-labeling observed in the lower light part of the segment for the QSM segmentation in the sagittal plane in Figure 4.3 is not found for the T1 segmentation. Figure 4.10 displays the T1 segmentantion of the right putamen, and seems to correspond well with the visible anatomy. Inspecting the examples of the segmentation of the pallidum in Figure 4.11 and of the caudate in Figure 4.12 does not indicate any apparent deviations from expected anatomical shape.

4.1.1.3 Qualitative analysis: Comparison

This section presents a visual comparison of the SynthSeg segmentation with T1w and QSM input data. The T1 segments are visualized in blue and the QSM segments in red. The examples are all of the same individual, sub-

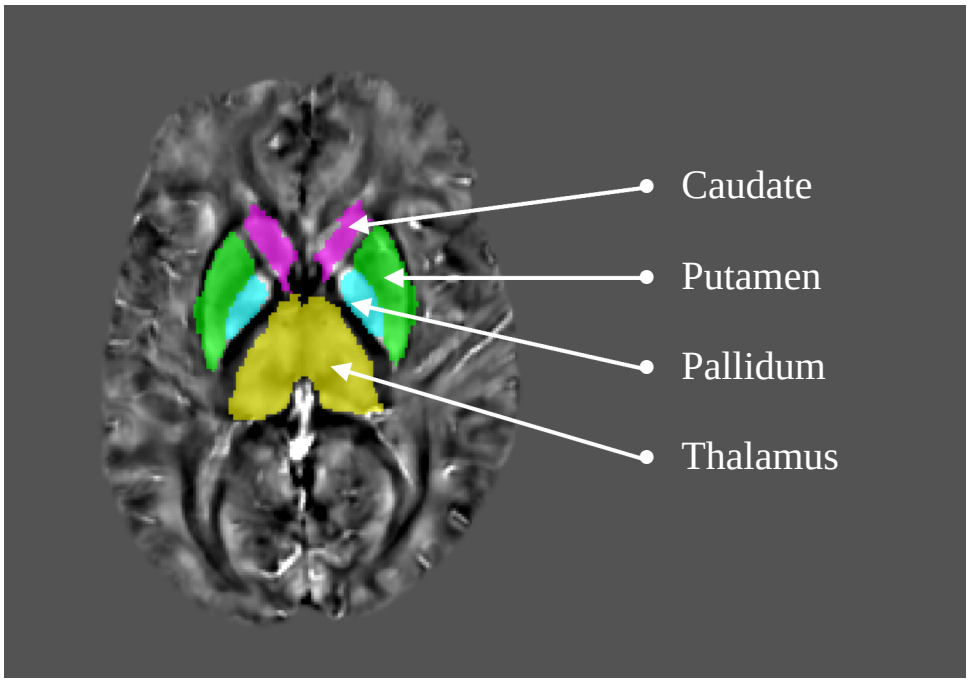
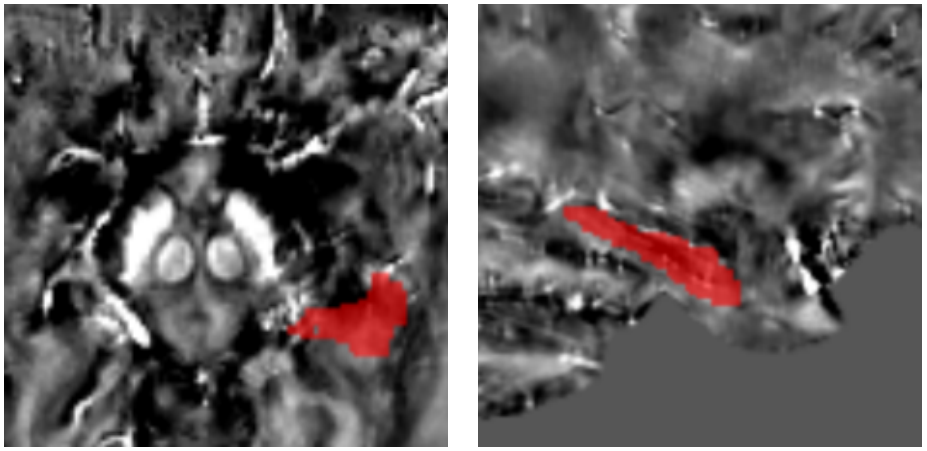
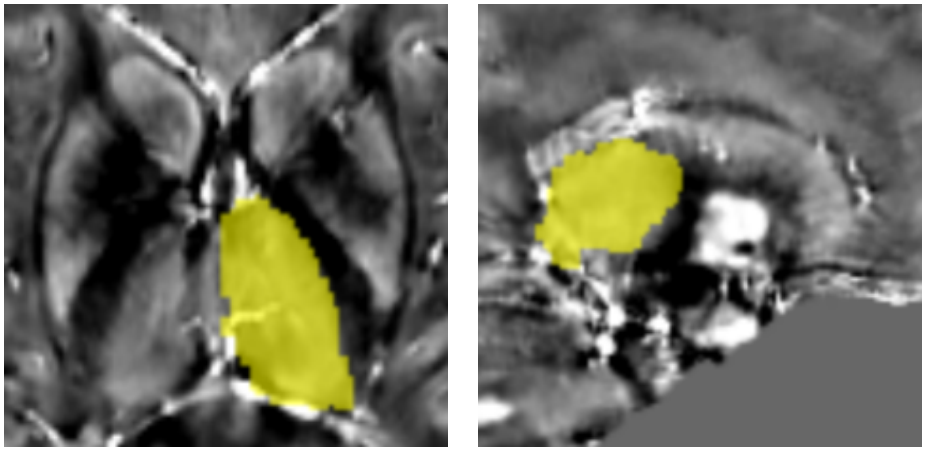


Figure 4.7: Segmentation of the caudate, putamen, pallidum and thalamus. The segments are generated with SynthSeg with a 7T T1w image as input.



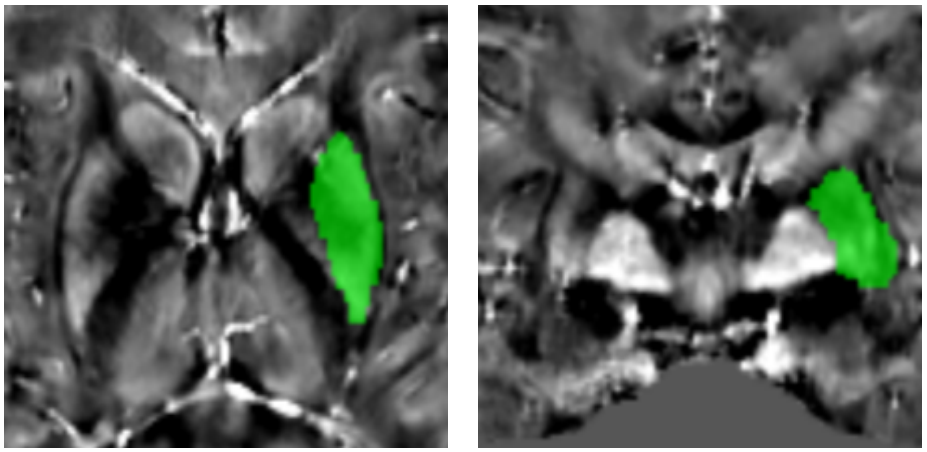
(a) Axial (b) Sagittal

Figure 4.8: Segmentation of the **hippocampus**. The segments are generated with SynthSeg with a 7T T1w image as input.



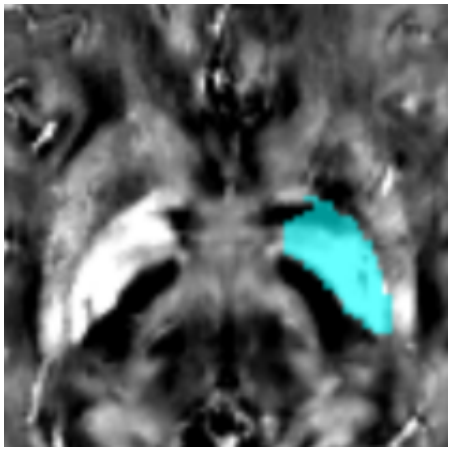
(a) Axial (b) Sagittal

Figure 4.9: Segmentation of the **thalamus**. The segments are generated with SynthSeg with a 7T T1w image as input.

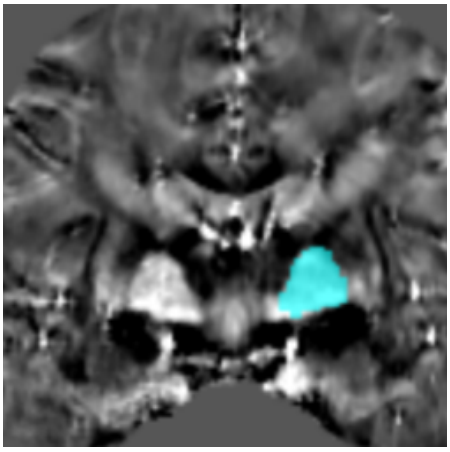


(a) Axial (b) Coronal

Figure 4.10: Segmentation of the **putamen**. The segments are generated with SynthSeg with a 7T T1w image as input.

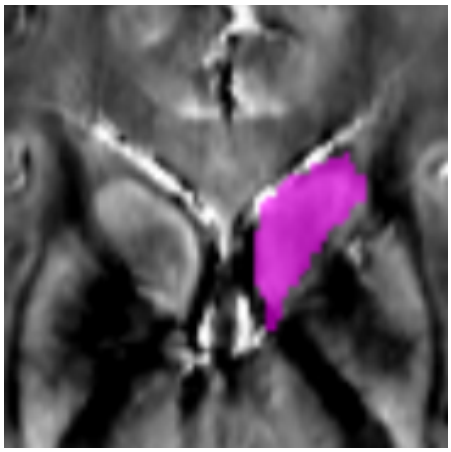


(a) Axial

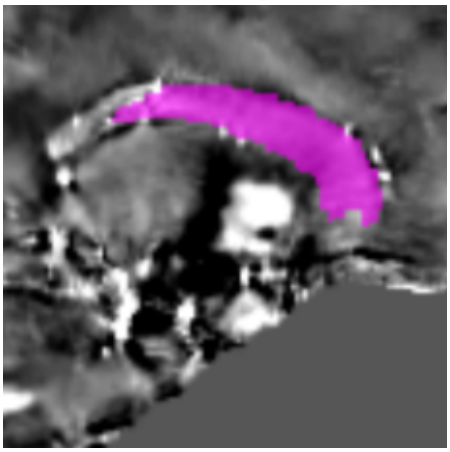


(b) Coronal

Figure 4.11: Segmentation of the **pallidum**. The segments are generated with SynthSeg with a 7T T1w image as input.



(a) Axial

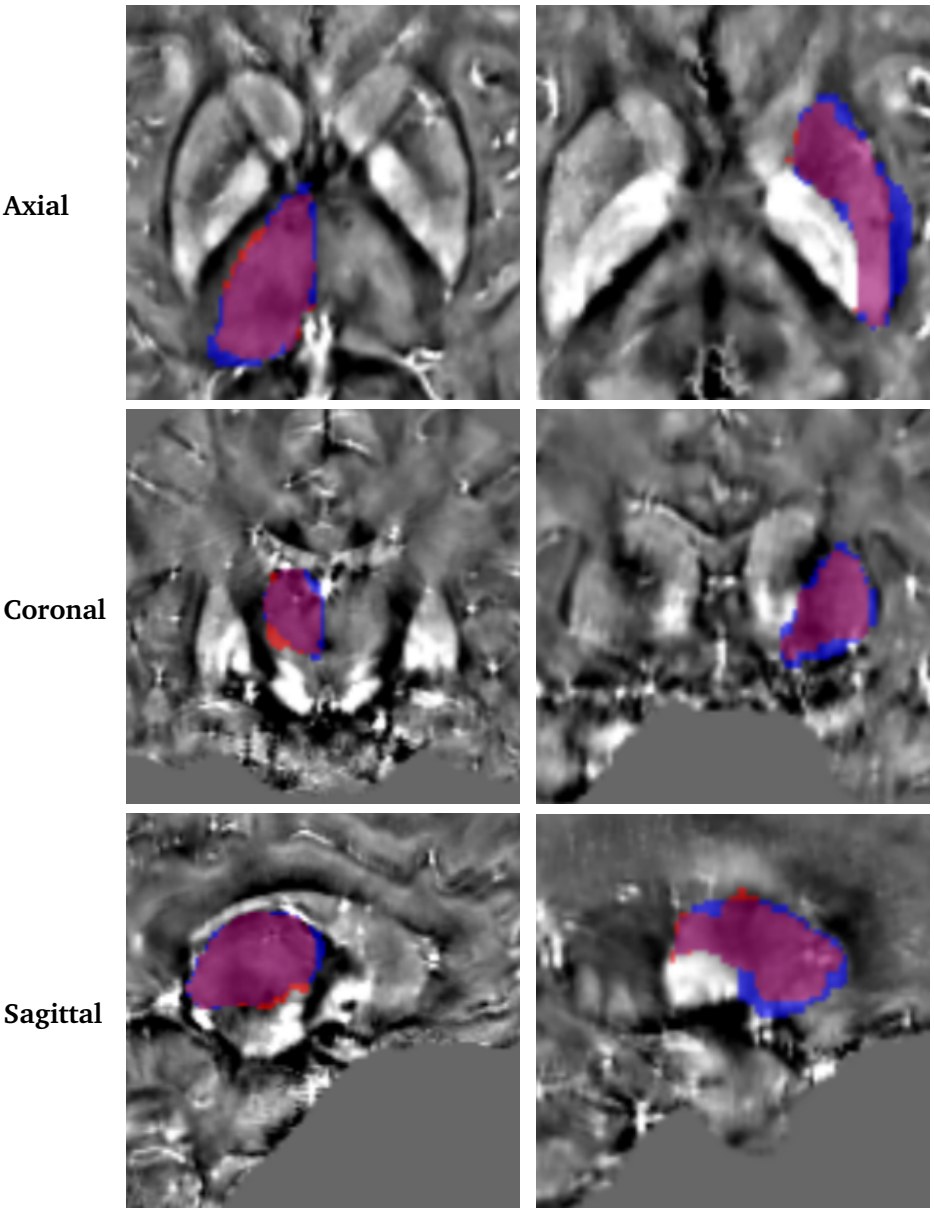


(b) Sagittal

Figure 4.12: Segmentation of the **caudate**. The segments are generated with SynthSeg with a 7T T1w image as input.

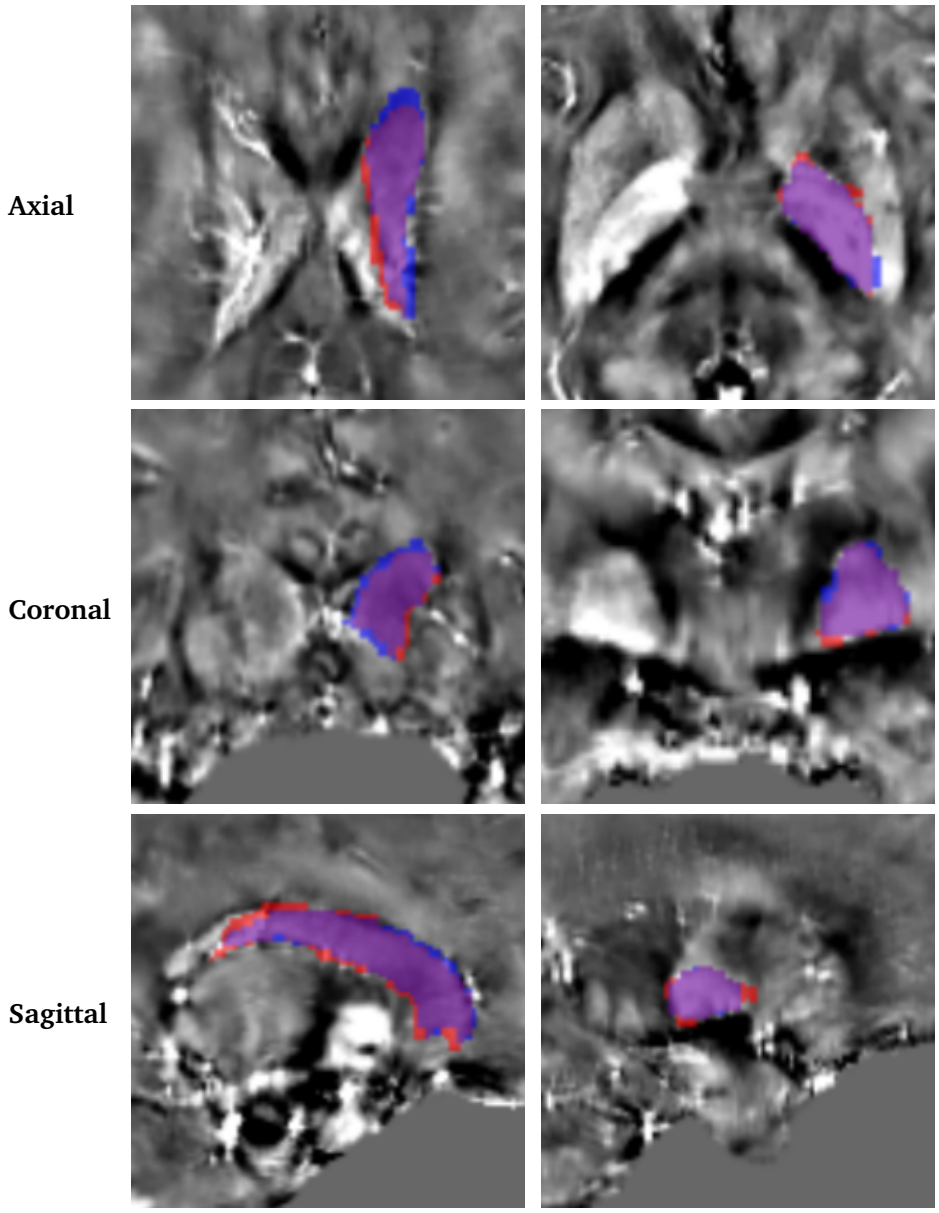
ject A, if not specified otherwise. The comparison of the segmentation of the left thalamus in Figure 4.13 shows that the T1 and QSM segmentation generally agrees in their labeling. However, the T1 segments seems to include more labels around the edges, while the QSM segment includes more labels towards the left in the coronal plane. Figure 4.14 (a) displays a general agreement for the segmentation of the right caudate. In the coronal plane a slight shift of the QSM segment towards the right is observed, while a slight shift is observed in the left direction in the axial plane. The segmentation of the right pallidum in Figure 4.14 shows a high agreement between the two segmentation inputs, though minor differences in labeling are observed around the edges in all planes. The putamen displayed in Figure 4.13 shows greater deviations around the edges. The T1 segmentation tends to label a larger volume, particularly in the lateral direction, as seen in the axial plane. Figure 4.15 displays the T1 and QSM masks of the putamen for two additional subjects, subject B and subject C. The trend of less labeling in the lateral direction by the QSM segmentation observed for subject A, can be seen even more clearly in subject B and C.

The largest difference between the two segmentation pipelines was observed for the hippocampus, visualised in Figure 4.16. The QSM input results in a segment much smaller than the T1 segmentation. Particularly in the lateral and ventral direction of the brain, the QSM input did not identify the parts labeled by the T1 input as the hippocampus, as displayed in Figure 4.16. Inspecting the QSM and T1 segmentation of the hippocampus for two additional subjects, the same trend of missing portions of the QSM segmentation was seen.



(a) Thalamus (b) Putamen

Figure 4.13: Difference in segmentation of the thalamus (a) and the putamen (b) in the axial, coronal and sagittal plane. The segment in red is generated with SynthSeg with a QSM image as input and the blue segment are generated from a 7T T1w image



(a) Caudate (b) Pallidum

Figure 4.14: Difference in segmentation of the caudate (a) and the pallidum (b) in the axial, coronal and sagittal plane. The segment in red is generated with SynthSeg with a QSM image as input and the blue segment are generated from a 7T T1w image

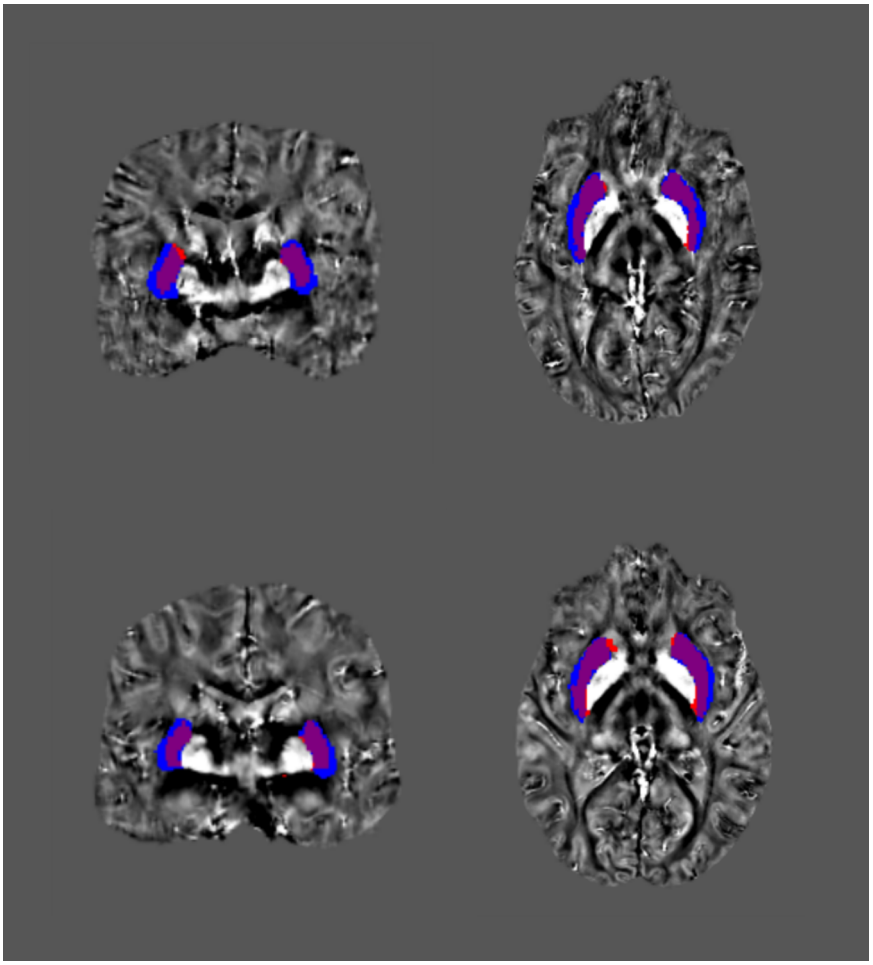


Figure 4.15: Segmentation of the putamen in the coronal (left) and axial (right) plane. The upper row displays subject B, and the lower row subject C. The segments in red are generated from QSM input and the segments in blue are generated from T1w input with SynthSeg.

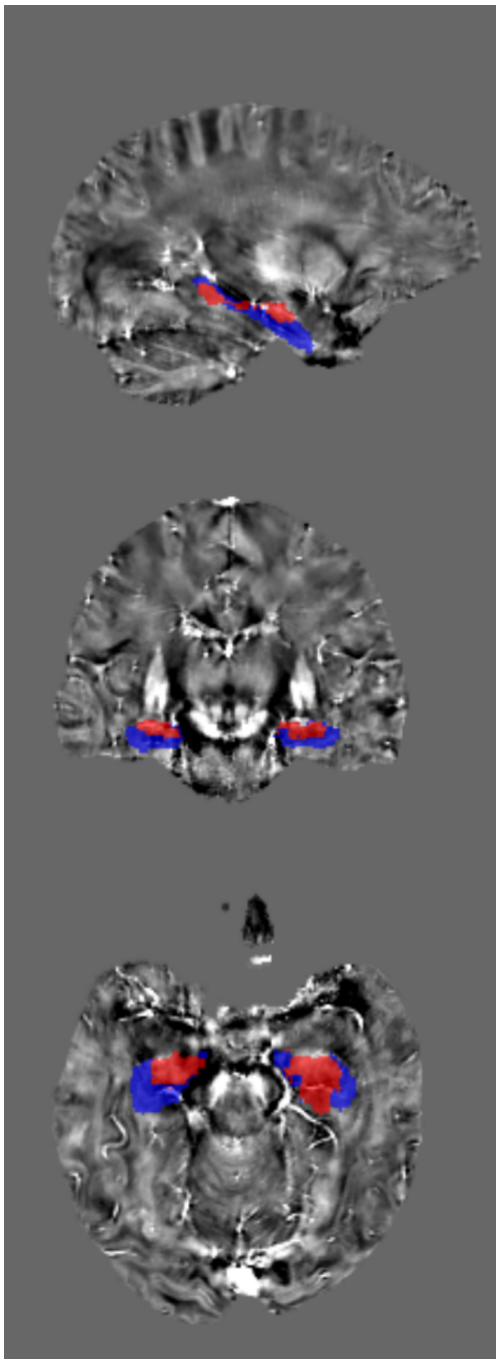


Figure 4.16: Difference in segmentation of the hippocampus in the sagittal, coronal and axial plane. The segment in red is generated with SynthSeg with a QSM image as input and the blue segment are generated from a 7T T1w image.

4.1.2 Segmentation volumes

The volumes of the segments generated from QSM images are reported in Table 4.1. The values are reported both in cm^3 and in % of the Total Intracranial Volume (ICV). The ICV is also segmented by SynthSeg from the QSM input image. As variability in brain volumes between individuals is expected, the relative volume of the ROIs might be the best indicator of whether the automated segmentation produces segments corresponding to expected volumes.

Table 4.1: Average volume for ROIs generated by SynthSeg from 7T QSM input for 29 healthy volunteers. The values are reported both in cm^3 and in % relative to the ICV. The values are presented with the inter-subject SD in the format mean (SD).

ROI	Mean volume (SD) cm^3	Mean volume (SD) % ICV
Total intracranial	1157.92 (111.67)	
CSF	124.05 (11.92)	10.73 (0.55)
Left thalamus	6.72 (0.67)	0.58 (0.05)
Right thalamus	7.10 (0.74)	0.61 (0.05)
Left caudate	4.05 (0.41)	0.35 (0.04)
Right caudate	4.29 (0.42)	0.37 (0.04)
Left putamen	4.07 (0.45)	0.35 (0.03)
Right putamen	4.58 (0.52)	0.40 (0.04)
Left pallidum	1.79 (0.16)	0.16 (0.01)
Right pallidum	1.80 (0.21)	0.16 (0.02)
Left hippocampus	2.14 (0.58)	0.18 (0.04)
Right hippocampus	2.47 (0.58)	0.21 (0.05)

The volumes of the segments generated from T1w images are reported in Table 4.2. It should be noted that the segmented volumes are generally larger for the T1 segments than for the QSM segments in Table 4.1. Particularly, the average total intracranial volume is much higher for the T1 segmentation. This causes the relative volumes to be smaller for most of the ROIs. The hippocampus in contrast increased significantly both in absolute and relative volume for the T1 segmentation, with p-values from a two-tailed t-test calculated to be < 0.001. This indicates a large deviation in the segmentation of the hippocampus dependent on the input image-contrast to SynthSeg. Furthermore, it is observed that substantially more of the CSF is included in the T1 segmentation compared to the QSM segmentation.

Table 4.2: Average volume for ROIs generated by SynthSeg from 7T T1w images for 29 healthy volunteers. The values are reported both in cm^3 and in % relative to the ICV. The values are presented with the inter-subject SD in the format mean (SD).

ROI	Mean volume (SD) cm^3	Mean volume (SD) % ICV
Total intracranial	1526.79 (129.95)	
CSF	211.76 (31.20)	13.84 (1.39)
Left thalamus	7.91 (0.57)	0.52 (0.03)
Right thalamus	8.26 (0.62)	0.54 (0.03)
Left caudate	4.16 (0.47)	0.27 (0.03)
Right caudate	4.37 (0.45)	0.29 (0.03)
Left putamen	5.95 (0.41)	0.39 (0.03)
Right putamen	5.84 (0.40)	0.38 (0.02)
Left pallidum	1.74 (0.16)	0.11 (0.01)
Right pallidum	1.86 (0.14)	0.12 (0.01)
Left hippocampus	4.19 (0.32)	0.28 (0.02)
Right hippocampus	4.03 (0.32)	0.27 (0.02)

4.1.3 Quality Control scores

The QC scores are automatically generated by SynthSeg and reflects the predicted reliability of the segmentation. The QC scores of both the QSM segmentation and T1 segmentation is presented in Table 4.3. The QC scores of the QSM ROIs are in the range of 0.72-0.81, above the discard limit defined by SynthSeg at 0.65. Still, the score of 0.73 ± 0.05 for the amygdala and hippocampus is only slightly above the discard limit. The quality of the hippocampus segments are therefore expected to be lower than for the other ROIs. The QC scores of the T1 segmented ROIs are in the range of 0.84-0.89. Compared to the values from the QSM segmentation, a statistically significant improvement is observed for all ROI groups. Particularly for the hippocampus and amygdala the QC score of 0.85 ± 0.01 indicates that the T1-input data produces adequate segmentation quality of the ROIs.

Table 4.3: Quality Control scores predicted by SynthSeg for segmentations generated from 7T QSM and T1w input data. The values are presented averaged over 29 subjects with the inter-subject SD on the format mean (SD). The significance of the difference in QC score between the segmentation methods is evaluated in terms of the p-value from a two-tailed t-test, assuming a normal distribution of the data.

ROI	**QC score QSM**	**QC score T1w**	**p-value**
Thalamus	0.81 (0.03)	0.85 (0.02)	< 0.0001
Putamen and Pallidum	0.81 (0.03)	0.89 (0.02)	< 0.0001
Hippocampus and Amygdala	0.73 (0.05)	0.85 (0.01)	< 0.0001
Average	0.78 (0.03)	0.86 (0.01)	< 0.0001
General CSF	0.72 (0.07)	0.76 (0.06)	< 0.05

4.1.4 Dice Scores

The DS between the SynthSeg segmentations generated from T1 and QSM input data are presented in Table 4.4. The DS were calculated both for the whole segmentation map, and for each ROI mask. The values are averaged over the 29 subjects, and given with the inter-subject SD. Generally, the DS indicates relatively high similarity between the segmentations. For the thalamus, caudate and pallidum, the averaged DS are in the range 0.83-0.87. The putamen scored slightly lower with 0.74 ± 0.04 and 0.80 ± 0.03 for the left and right putamen, respectively, where the error is the inter-subject SD. Significantly lower values were found for the hippocampus, the left hippocampus scoring as low as 0.51 ± 0.13, and the right 0.57 ± 0.11. The low DS implies that one of the two segmentation methods failed to segment the hippocampus correctly. The low DS of the CSF should also be noted at only 0.19 ± 0.03, showing a great variability in the segmentation of the CSF.

ROI	Mean DS
Left thalamus	0.85 (0.03)
Right thalamus	0.86 (0.03)
Left caudate	0.83 (0.03)
Right caudate	0.84 (0.02)
Left putamen	0.74 (0.04)
Right putamen	0.80 (0.03)
Left pallidum	0.87 (0.02)
Right pallidum	0.87 (0.02)
Left hippocampus	0.51 (0.13)
Right hippocampus	0.57 (0.11)
ROI average	0.77 (0.02)
Whole map	0.86 (0.01)
CSF	0.19 (0.03)

Table 4.4: DSs between SynthSeg segmentation generated from T1w and QSM input images.The DS was calculated for each ROI and averaged over 29 subjects, and are presented with the inter-subject SD on the format mean (SD).

4.1.5 Susceptibility values

Figure 4.17 displays the susceptibility values extracted from the QSM SynthSeg segments. The mean and median values of 29 healthy volunteers are presented with the inter-subject SD for the left and right ROIs. The susceptibility values extracted from the T1 segmentation is presented in Figure 4.18. For the T1 segmentation, the QSM images were first co-registered to the T1-images using the FSL FLIRT software [36]. The exact values and SD are

reported in tables in Appendix A, including the CSF susceptibility.

For the QSM segmentation, the CSF susceptibility was found to be -0.72 (1.07) and -0.06 (0.15) ppb for the mean and median, respectively. The parenthesized number is the inter-subject SD. The CSF was respectively measured to be 0.96 (0.55) and 0.00 (0.00) ppb for the mean and median with the T1 segmentation. The values displayed in Figure 4.17 and Figure 4.18 are not normalized to avoid introducing noise from a possible unstable CSF reference.

Generally, similar values were found for the mean and median susceptibility, with a tendency of the median value to be slightly higher. The inter-subject SD seems to increase as the mean value increases. The highest SD compared to the mean value were found in ROIs of lower susceptibility, like the thalamus and hippocampus. The caudate and pallidum show a lower variation between subjects compared to the mean value. Table 4.5 displays the average variation of susceptibility inside the ROIs, which can be used as an indicator of the uniformity of the segments. The intra-ROI SD were found to be in the range of 19-28 ppb for the QSM segmentation, and slightly higher at 19-32 ppb for the T1 segmentation. The p-values from a two-tailed t-test were calculated, and the caudate and the right putamen were the only segments showing no significant change in intra-ROI SD due to segmentation method.

Table 4.5: Intra-ROI variation of susceptibility averaged over 29 healthy volunteers for the QSM and T1 segmentation. The values are presented in the format mean (SD), where the uncertainty is the inter-subject SD for the intra-ROI variation. The significance of the change in mean variability due to the segmentation method is evaluated in terms of the p-value from a two-tailed t-test, assuming a normal distribution of the data.

ROI	**Intra-ROI SD (SD)** (ppb) **QSM input**	**Intra-ROI SD (SD)** (ppb) **T1w input**	**p-value**
Left thalamus	19.29 (2.17)	21.93 (2.70)	< 0.001
Right thalamus	20.44 (2.46)	21.83 (2.54)	< 0.05
Left caudate	18.83 (3.35)	19.08 (2.90)	0.762
Right caudate	18.67 (3.16)	19.33 (2.87)	0.409
Left putamen	27.14 (4.24)	24.20 (3.79)	< 0.05
Right putamen	25.11 (3.52)	23.63 (3.85)	0.132
Left pallidum	26.77 (5.72)	29.81 (5.34)	< 0.05
Right pallidum	27.99 (5.39)	31.92 (5.41)	< 0.05
Left hippocampus	22.74 (2.71)	19.93 (2.54)	< 0.001
Right hippocampus	20.58 (2.76)	18.88 (2.18)	< 0.05
CSF	22.33 (1.91)	16.29 (1.53)	< 0.001

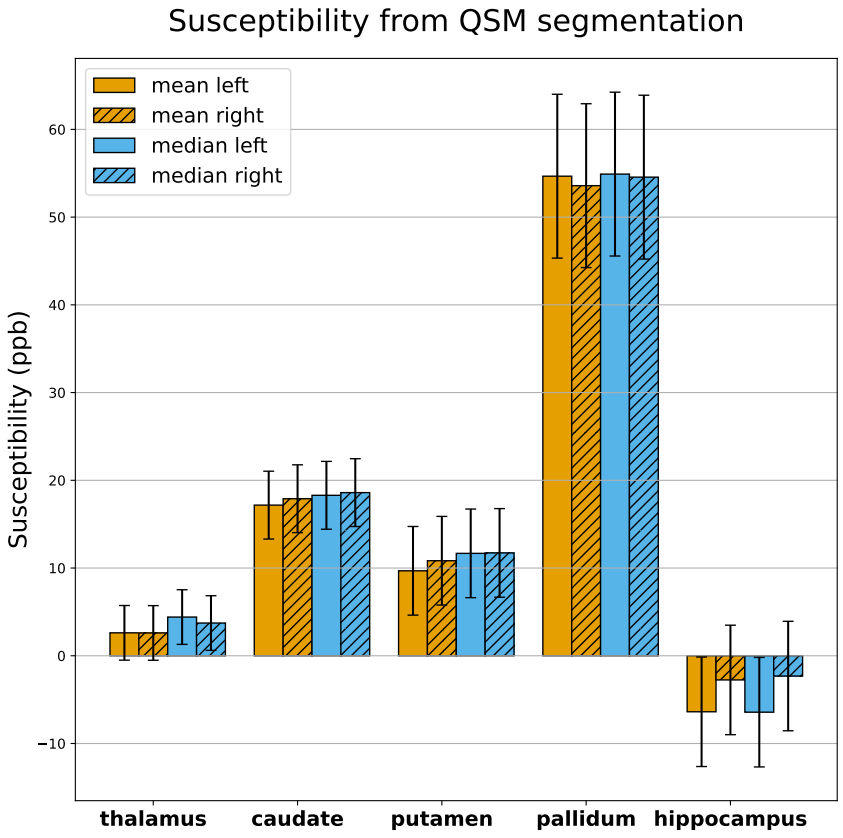


Figure 4.17: Susceptibility values extracted from the masks generated by SynthSeg with QSM input data. For each ROI, the diagram shows the mean and median susceptibility of the left and right segment averaged over 29 healthy subjects.

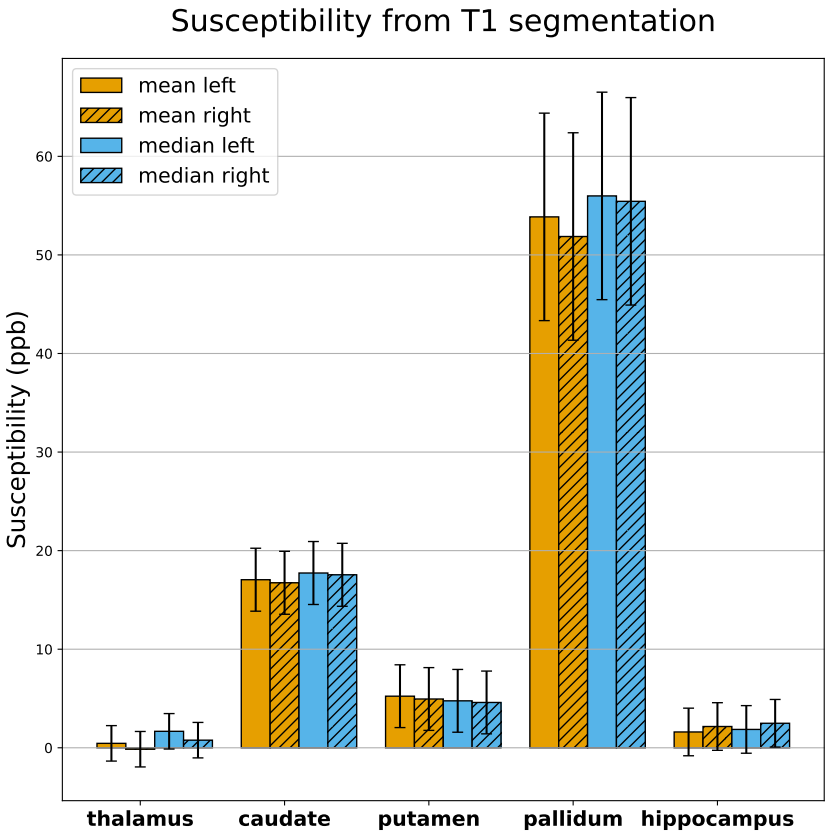


Figure 4.18: Susceptibility values extracted from the masks generated by SynthSeg with T1w input data. For each ROI, the diagram shows the mean and median susceptibility of the left and right segment averaged over 29 healthy subjects.

4.1.5.1 Comparison of Susceptibility values

Table 4.6 presents the difference in average susceptibility obtained for the ROIs when using the QSM segments compared to using the T1 segments. If the magnitude of difference is significantly less than the difference observed in pathology, it would indicate that the method of QSM segmentation could be useful in a diagnostic practice. However, a great variability in the difference of susceptibility is observed between the ROIs. The caudate and pallidum in particular shows a low difference in the averaged values between the two segmentation input data types of less than 2 ppb. The mean susceptibility value of the left and right putamen increases for the QSM segmentation with 4.45 ± 5.97 and 5.88 ± 6.40 ppb, respectively, with the error being the inter-subject SD. Considering that the mean value of the left and right pallidum for the T1 segmentation were found to be 5.23 ± 3.18 and 4.95 ± 3.41, the difference introduced by the QSM segmentation is notable. The statistical significance of the changes in terms of the p-value were calculated from a two-tailed t-test. The changes were found to be significant for the thalamus, putamen and hippocampus. The hippocampus stands out with a particularly high difference in the mean susceptibility, the left hippocampus decreasing with almost 8 ppb for the QSM segmentation.

The differences found in the median susceptibility values are of similar scales, and shows a statistical significance for the same ROIs as for the mean. The absolute value of the difference is increasing slightly for most segments in the median compared to the mean, while decreasing for the right pallidum, right caudate and right hippocampus. No significant changes were found for the caudate nor pallidum for neither the mean nor the median.

Figure 4.19 and Figure 4.20 displays the mean susceptibility values extracted from the QSM and T1 masks in violin plots. The distribution of mean susceptibility between the subjects in the left hippocampus is much more spread for the QSM than for the T1 segmentation. Generally, the T1 segmentation seems to produce more compact distribution of mean values, particularly for the hippocampus, thalamus and putamen.

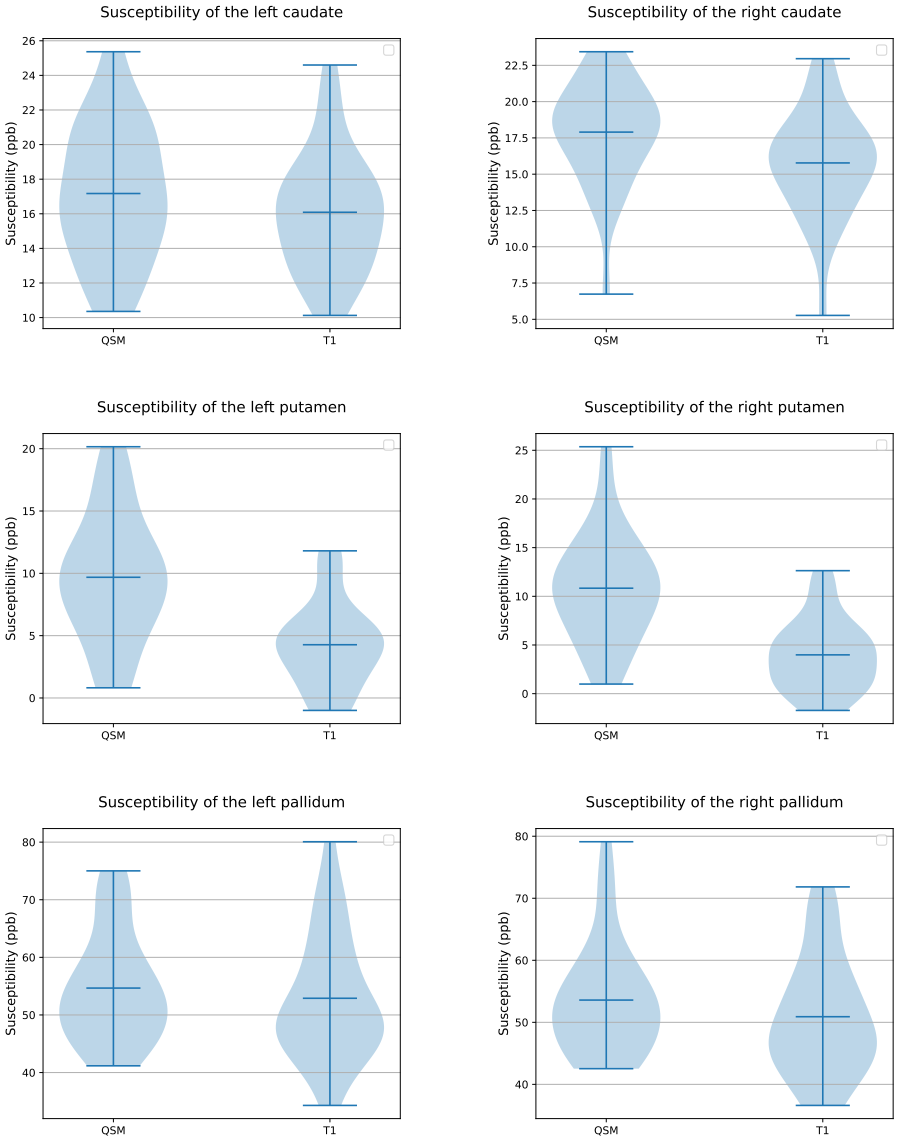


Figure 4.19: Violin plots comparing the distribution of mean susceptibility value in the left and right caudate, putamen and pallidum with QSM and T1 segmentation. N=29.

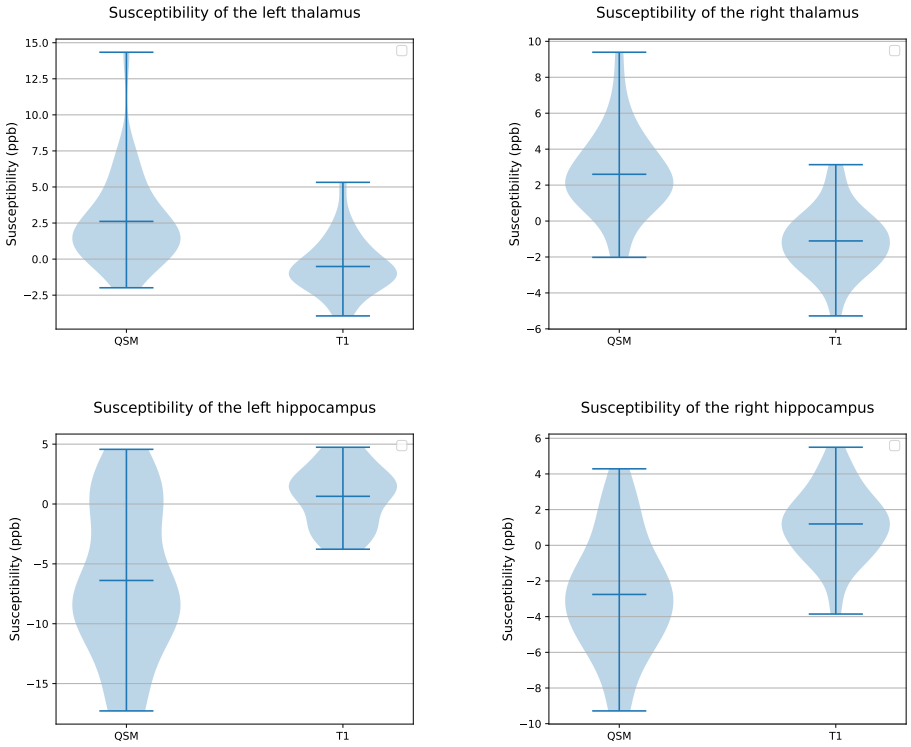


Figure 4.20: Violin plots comparing the distribution of mean susceptibility value in the left and right thalamus and hippocampus with QSM and T1 segmentation. N=29.

Table 4.6: This table displays the difference in magnitude of the mean and median susceptibility values between the ROIs generated with SynthSeg from T1 and QSM input. The values are the averaged susceptibility extracted from the QSM masks subtracted from the averaged susceptibility from the T1 masks. The table also reports p-values calculated from the mean and median for each ROI, using a two-tailed t-test and assuming a normal distribution of the data. The number in parenthesis is the inter-subject SD.

ROI	**Difference mean** (ppb)	**Difference median** (ppb)	**p-value mean**	**p-value median**
Left thalamus	-2.16 (3.59)	-2.74 (3.77)	< 0.005	< 0.0005
Right thalamus	-2.74 (2.89)	-2.96 (3.09)	< 0.0001	< 0.0001
Left caudate	-0.12 (5.02)	-0.56 (5.43)	0.898	0.581
Right caudate	-1.16 (4.92)	-1.05 (5.42)	0.210	0.302
Left putamen	-4.45 (5.97)	-6.91 (5.95)	< 0.0005	< 0.0001
Right putamen	-5.88 (6.40)	-7.12 (5.95)	< 0.0001	< 0.0001
Left pallidum	-0.81 (14.07)	1.09 (14.12)	0.758	0.679
Right pallidum	-1.72 (12.92)	0.88 (13.45)	0.476	0.726
Left hippocampus	7.99 (6.69)	8.29 (5.80)	< 0.0001	< 0.0001
Right hippocampus	4.91 (3.84)	4.80 (3.40)	< 0.0001	< 0.0001
CSF	1.68 (1.14)	-0.15 (0.15)	< 0.0001	< 0.05

4.2 U-net segmentation of the SN, RN and OMEGA

4.2.1 Susceptibility values of the SN, RN and OMEGA

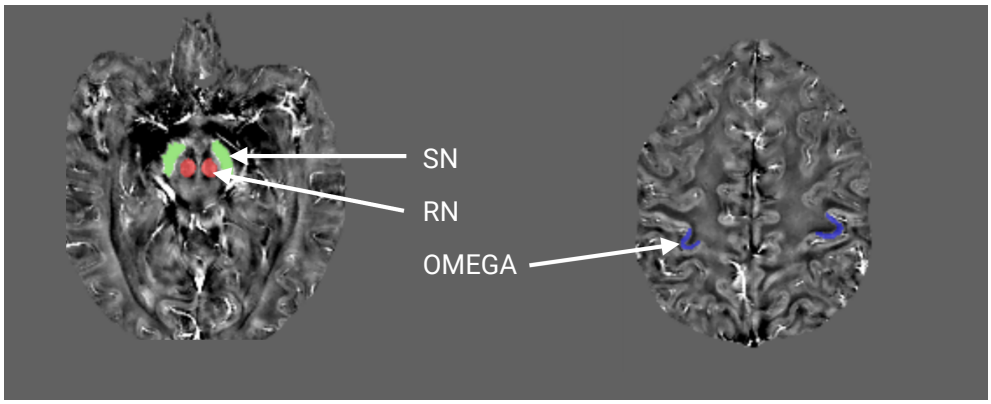


Figure 4.21: The figure displays an example of the segmented SN, RN and OMEGA generated by the CNN U-net in the axial plane.

This section presents the results of the susceptibility extraction from the U-net segmentation. An example of the resulting segments of the SN, RN and OMEGA is displayed in Figure 4.21. The chosen weights were first evaluated, using the manually segmented masks available for 15 of the subjects as a ground truth. The evaluation of the SN weights resulted in an average DS of 0.92 ± 0.026. For the RN, the average DS was 0.95 ± 0.017. The mean susceptibility extracted from the U-net masks of the SN, RN and OMEGA, are presented in table Table 4.7 with the inter-subject SD. The values are averaged over 29 healthy volunteers, and are not normalized with respect to a reference value, as the U-net is currently not trained to segment a suitable reference area. The mean susceptibility averaged over 15 of the healthy volunteers extracted from the manually segmented whole structure masks is included in Table 4.7, and shows a slight decrease in the mean susceptibility for the SN, RN and OMEGA compared to the automatically extracted values.

4.2.2 Correlation of predicted and manual susceptibility values in the SN and OMEGA

A linear regression of the mean susceptibility of the automatically segmented SN to the values manually extracted by a radiologist is presented in Figure 4.22. The R-squared value was 0.57, showing some correlation. The MAE and the MSE was calculated to be 13.77 and 0.33 ppb respectively. An exploratory method was used to investigate if other statistics of the susceptibility of the automated segment correlated stronger to the manu-

Table 4.7: Mean susceptibility of the SN, RN and Omega extracted from masks generated by the U-net. The values are averaged over 29 healthy volunteers with a mean age of 27.97 ± 5.73 years and presented with the inter-subject SD. The right column reports the mean susceptibility extracted from manually segmented masks averaged over 15 healthy volunteers with a mean age of 25.81 ± 4.08 years.

ROI	**Mean (SD) ppb U-net**	**Mean (SD) ppb Manual**
SN	70.16 (12.10)	68.82 (12.73)
RN	47.78 (11.75)	43.67 (10.56)
OMEGA	21.14 (4.50)	18.99 (3.95)

Table 4.8: Results of exploratory correlation analysis between U-net susceptibility and manually extracted susceptibility for the SN and OMEGA. The R^2-value of a linear regression to the manual values was calculated for different statistics of the susceptibility extracted from the automated segment, including the mean, median, max and percentile values.

ROI	**SN** R^2**-value**	**OMEGA** R^2**-value**
Statistic		
Mean	0.58	0.43
Median	0.50	0.40
Max	0.65	0.25
80-percentile	0.66	0.42
90-percentile	0.74	0.35
95-percentile	0.78	0.30
98-percentile	0.80	0.28

ally extracted values. A linear regression to the median, max, 80, 90, 95 and 98 percentile value were also performed, and the resulting R-squared values are presented in Table 4.8. The highest correlation in terms of the R-squared value was found for the 98 percentile value of the susceptibility, with $R^2 = 0.80$. Figure 4.23 displays the linear regression of the 98 percentile value to the manual values. The MAE and the MSE were calculated to be 9.86 and 0.15 ppb respectively, significantly lower than the inter-subject SD of the manual values of 28.45 ppb. The p-values from a two-tailed t-test were calculated between the manual values and predicted susceptibility from the linear regression of the automated values. The p-value was found to be < 0.0001 for the mean value and 0.089 for the 98-percentile prediction. The relatively high probability of finding the correlation due to chance for the 98-percentile value of this dataset should be noted.

The same exploratory correlation analysis was performed for the auto-

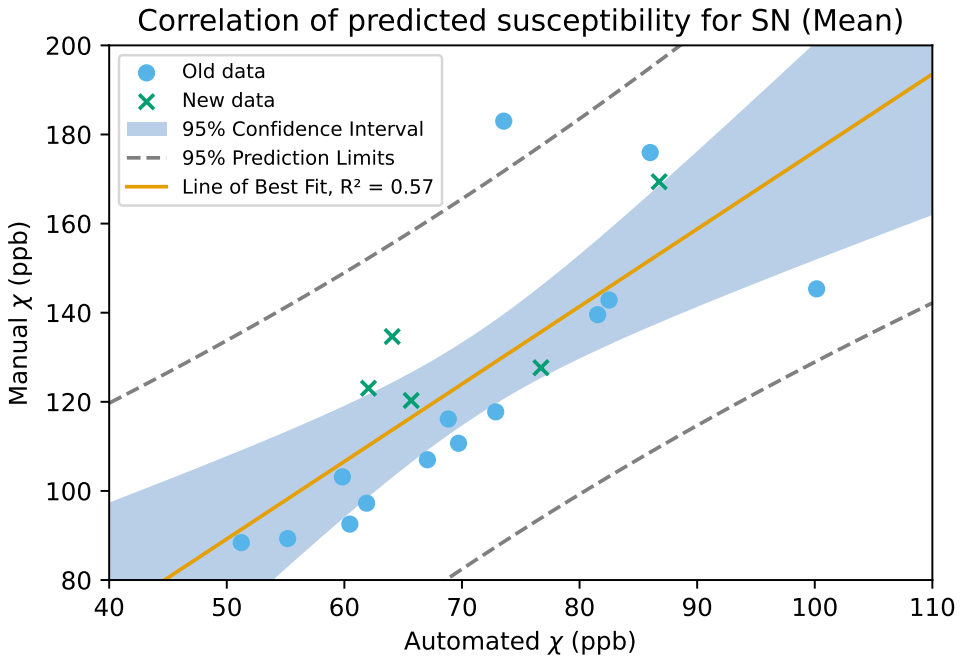


Figure 4.22: Correlation of mean susceptibility extracted from automated segmentation of the SN to susceptibility values manually extracted by a radiologist. The orange line shows the linear regression with equation $y = 1.74x + 0.0022$ and $R^2 = 0.57$. The blue dots are the manual values plotted to the predicted susceptibility values.

mated OMEGA susceptibility values, but found no pronounced correlation. The correlation of the mean susceptibility in the U-net segmented OMEGA to manual values are displayed in Figure 4.24. The R-squared value of 0.43 shows some correlation, but less than for the SN. It should be noted that as the correlation increased for higher percentile values for the SN, the correlation seemed to rather decrease to the higher susceptibility values in the OMEGA, as shown in Table 4.8. As some of the data in this study is the same data included in the training set of the U-net, Figure 4.22 and Figure 4.23 marks the data that has not been seen by the U-net. The U-net seems to be able to predict the manual values equally good for the new data as for the data included in the training of the net. The p-value from a two-tailed t-test was calculated, comparing the predicted values from the mean susceptibility in the automated OMEGA segment to the manual values. The p-value was found to be 0.284, showing no statistically significant correlation.

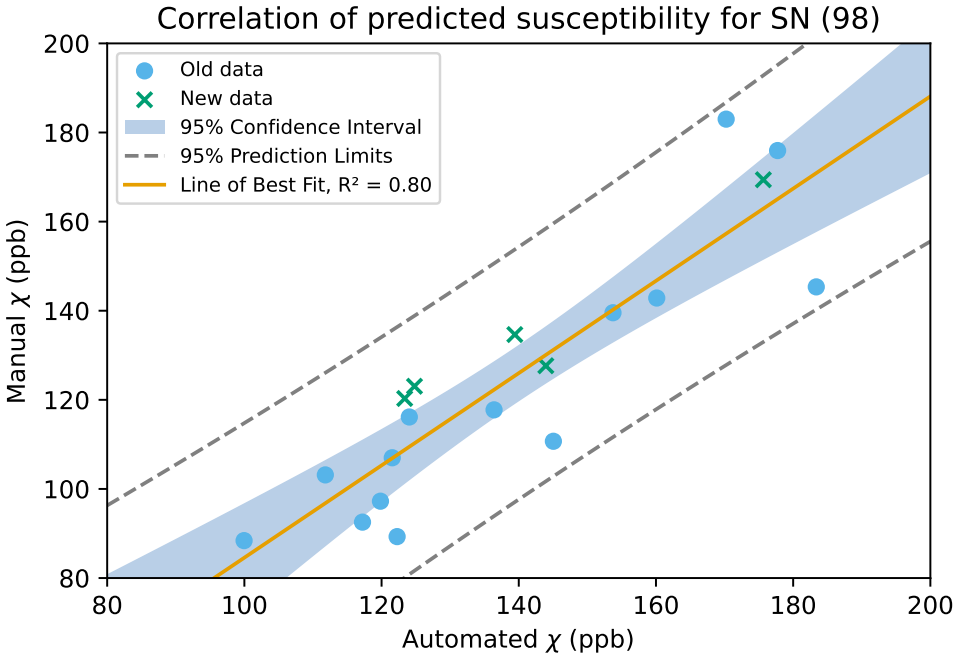


Figure 4.23: Correlation of the 98 percentile susceptibility extracted from automated segmentation of the SN to susceptibility values manually extracted by a radiologist. The orange line shows the linear regression with equation $y = 1.035x - 0.019$ and $R^2 = 0.80$. The blue dots are the manual values plotted to the predicted susceptibility values.

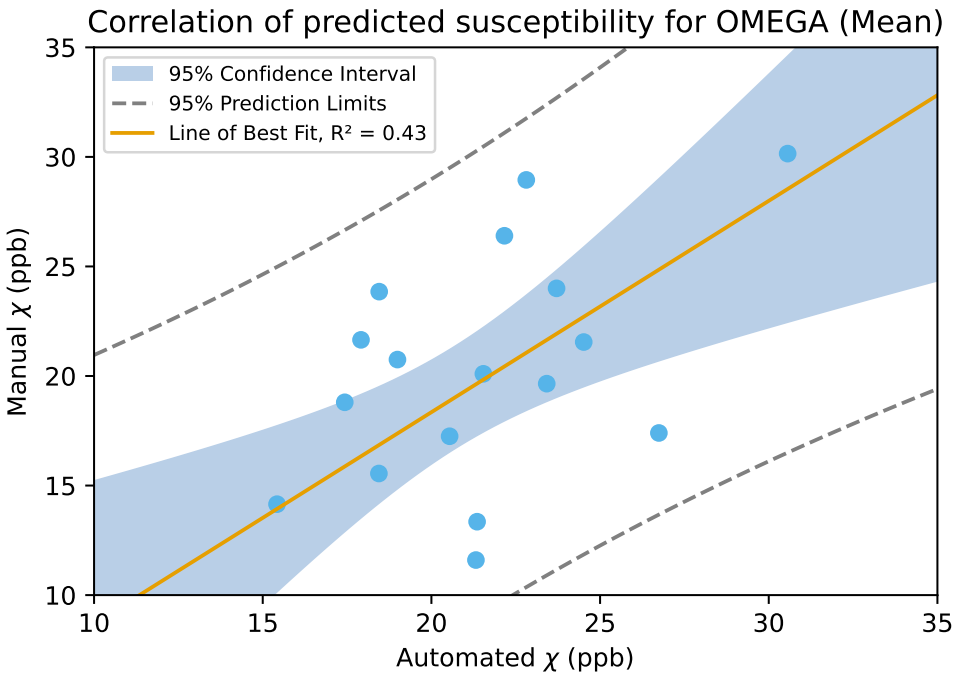


Figure 4.24: Correlation of mean susceptibility extracted from automated segmentation of the OMEGA to susceptibility values manually extracted by a radiologist. The orange line shows the linear regression with equation $y = 0.96x - 0.94$ and $R^2 = 0.43$. The blue dots are the manual values plotted to the predicted susceptibility values.

4.3 Variation of susceptibility with age

This section further analyses the susceptibility values extracted from the T1 SynthSeg masks, as well as the U-net segmented SN, RN and OMEGA. To investigate the age dependency of the data, the mean susceptibility are plotted to the subject age in Figure 4.26, Figure 4.27 and Figure 4.25. Both the raw susceptibility values, and the susceptibility values normalized to the centrum semiovale is plotted, to investigate the effects of normalization. Note that the manually segmented centrum semiovale were available only for 19 of the subjects, so the normalized age plots to the right have a smaller sample size than for the raw values. The plots shows that the increase with age is subtle compared to the subject variation, particularly for the raw susceptibility values.

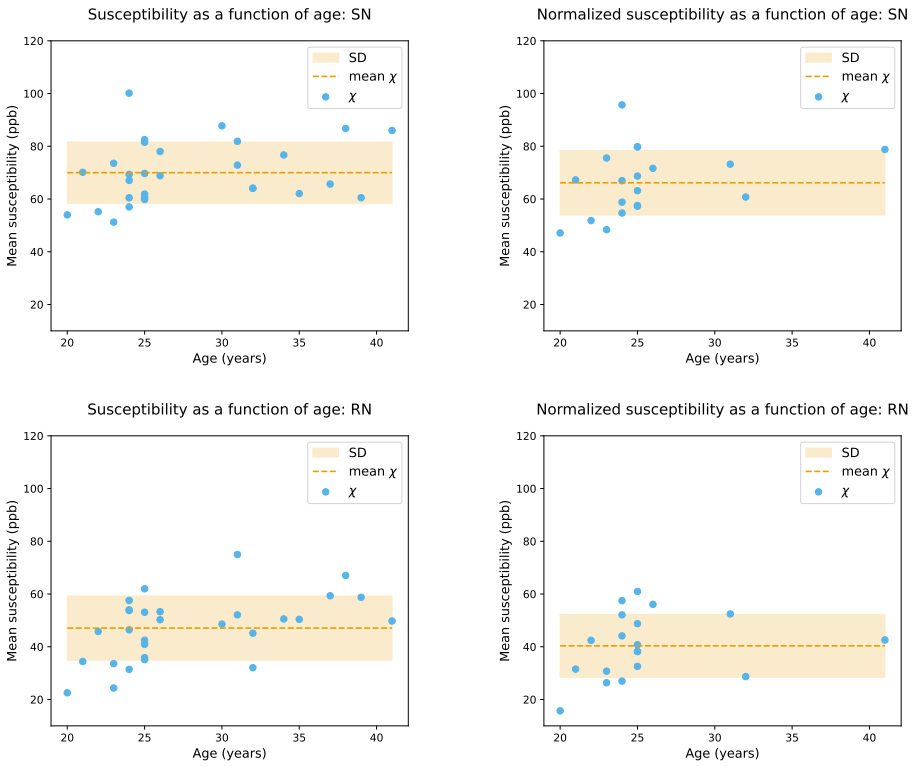


Figure 4.25: Susceptibility as a function of age in healthy subjects. The left column displays the mean susceptibility values extracted from the U-net segmentation of SN and RN for 29 subjects. The right column displays the same susceptibility values for 19 of the subjects normalized to the manually segmented centrum semiovale. The plots includes both the left and right ROIs, plotted as dots and stars respectively. The inter-subject SD is visualized with the shaded area

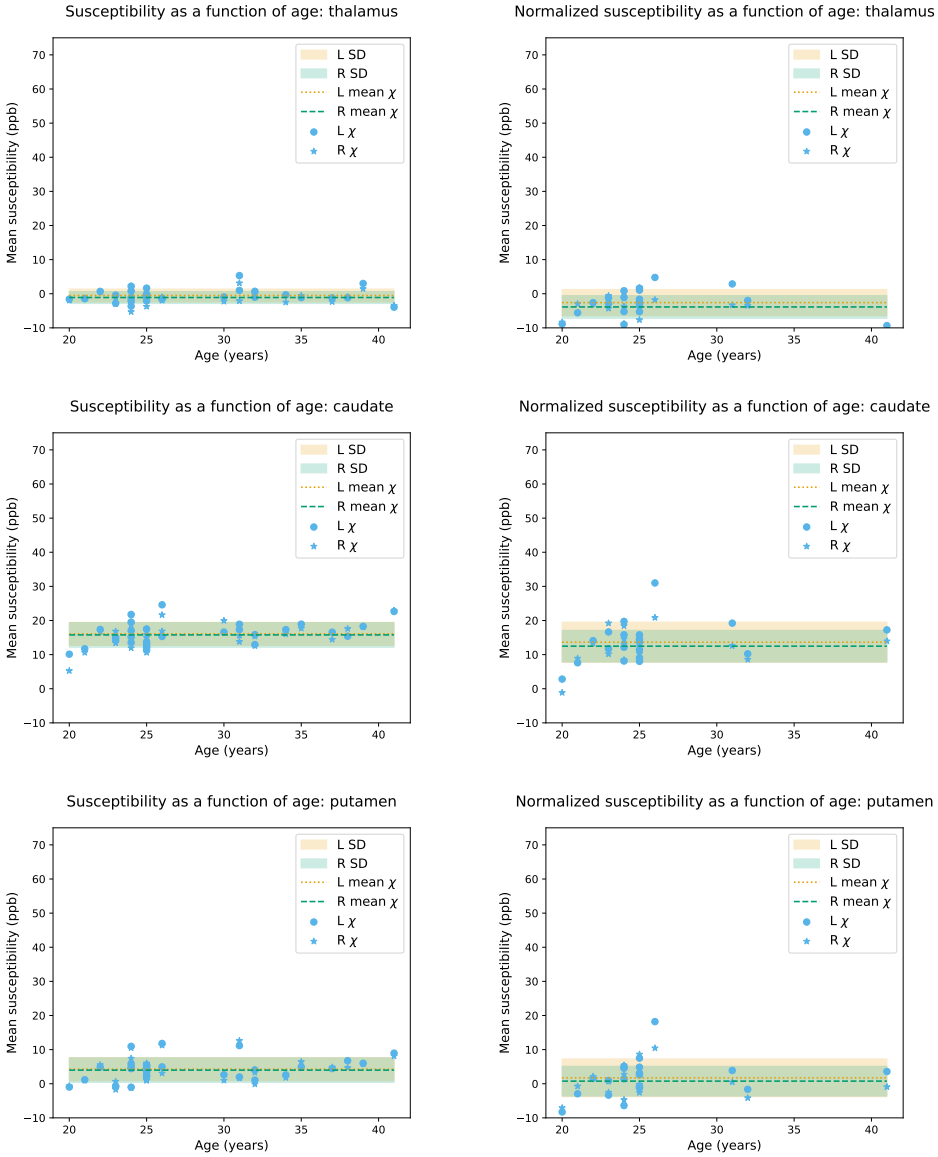


Figure 4.26: Susceptibility as a function of age in healthy subjects. The left column displays the mean susceptibility values extracted from SynthSeg segmentation based on T1 input data of 29 subjects. The right column displays the same susceptibility values for 19 of the subjects normalized to the manually segmented centrum semiovale. The upper row displays the thalamus, the middle row the caudate and the bottom row the putamen. The plots includes both the left and right ROIs, plotted as dots and stars respectively. The inter-subject SD is visualized with the shaded area.

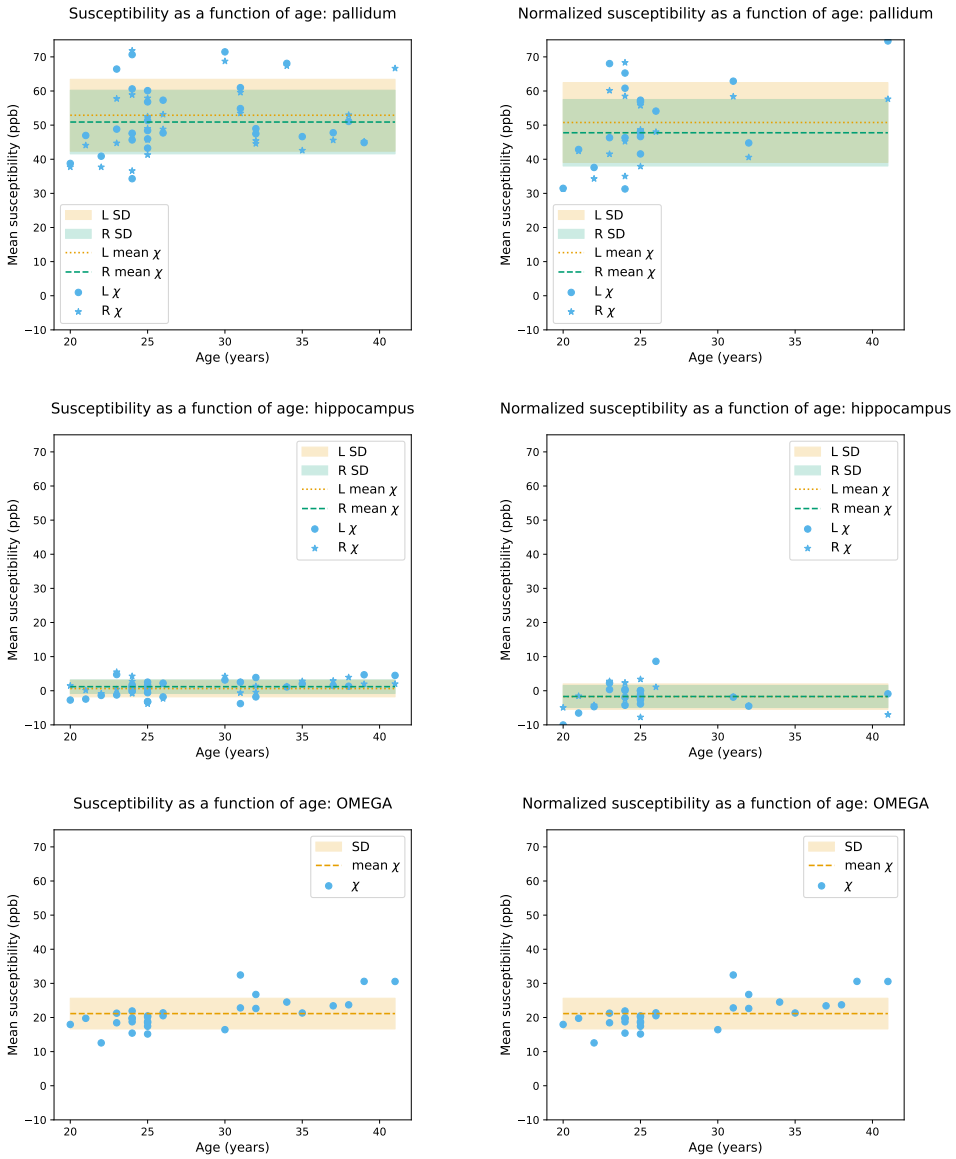


Figure 4.27: Susceptibility as a function of age in healthy subjects. The left column displays the mean susceptibility values extracted from SynthSeg segmentation based on T1 input data for the pallidum and the hippocampus, and from the U-net for the OMEGA for 29 healthy volunteers. The right column displays the same susceptibility values for 19 of the subjects normalized to the manually segmented centrum semiovale. The plots includes both the left and right ROIs, plotted as dots and stars respectively. The inter-subject SD is visualized with the shaded area

Chapter 5

Discussion

5.1 SynthSeg segmentation of the caudate, pallidum, putamen, thalamus and hippocampus

This section concerns the SynthSeg segmentation with T1w and QSM input data at ultra-high field strength. The section will first discuss the results of the three investigated indicators of segmentation quality, analysing the visual anatomy, segmentation volumes and QC scores. A comparison of the two segmentation pipelines will be performed, followed by an overall evaluation of the segmentation quality of the 7T T1w segmentation and QSM segmentation. The extracted susceptibility values will be discussed and compared to literature, and the potential of using the susceptibility values as biomarkers for PD will be considered. Finally, the section motivates a suggestion for a fully automated pipeline for extraction of susceptibility values for further exploring the diagnostic potential of voxel-based 7T QSM.

5.1.1 Segmentation quality

5.1.1.1 Qualitative analysis

This section will discuss the visual examples presented in Section 4.1.1. The analysis will mainly focus on the visible differences between the QSM and T1w segmentations, as the expertise of a radiologist is needed to provide a detailed assessment of the anatomical accuracy of the segmentation. However, it is noted that some of the segments seemed to clearly disagree with the outline of the structures visible in the anatomy of the QSM image. The segments were also evaluated to the expected anatomical shape, using the brain atlases *7.0 Tesla MRI Brain Atlas* [41] and the online atlas of IMAGIO as a reference. The imaging viewing software MRIcron [39] was used for the visual analysis, and examples from one representative subject

will be discussed in detail.

For the QSM segmentation of the thalamus displayed in Figure 4.3, a part of the ROI seems to be missing, while the T1 segmentation of the same subject in Figure 4.9 seems to include the missing labels and correspond better with the visible anatomy. The T1 and QSM based segmentation of the thalamus is displayed overlapped in Figure 4.13, and shows some disagreement around the edges, but a general overlap is noted. For the caudate and pallidum, no obvious mislabeling were observed for either the segmentation inputs. The segmentations of the caudate are displayed in Figure 4.12 and Figure 4.6, and for the pallidum in Figure 4.5 and Figure 4.11. Investigating the differences between the T1 and QSM segmentation masks displayed in Figure 4.14, the deviations around the segment boarders are slightly less apparent than for the thalamus. The visual analysis indicates that the SynthSeg segmentation of the ROIs are relatively agnostic to the two input data contrasts, particularly for the caudate and pallidum.

The example of segmentation of the putamen in Figure 4.4 displays an example of the QSM segmentation of the putamen. A small area to the left in the coronal plane is notably not corresponding with the expected anatomical shape and visible outlines in the QSM image. The same overlabeling is not found in the T1 segmentation of the same slice displayed in Figure 4.10. The difference between the T1 and QSM mask of the putamen is displayed in figure Figure 4.13 (b). The QSM segment is notably smaller than the T1 segment, and the surface of the T1 segment seems to correspond better with the anatomical outlines. The segmentation of the putamen is displayed for two other subjects in Figure 4.15. The deviations between the T1 and QSM masks are even more prominent in these examples. Note that the QSM segmentation seems to particularly fail at labeling the lateral parts of the putamen.

The largest deviation from expected anatomy was observed in the QSM segmentation of the hippocampus. The uneven boarders are visible in Figure 4.2, and in Figure 4.16 it is observed that large portions of the QSM segments are missing compared to the T1 segments. Inspecting the segments and visually comparing to brain atlases and the T1 weighted images of the subjects, the T1 segmentation seems to correspond with the expected the anatomy of the hippocampus, while large deviations are observed for the QSM segments.

5.1.1.2 Segmentation volumes

This section will discuss the volumes of the SynthSeg segments and compare to literature. Two large scale studies are considered in this analysis, and are presented in Table 5.1. For the purpose of this discussion, the study by Wang et al. [42] analysing the effects of age and sex on subcortical

volumes is labeled as Study 1, and the study of subcortical volumes across the lifespan of 18,605 healthy volunteers by Dima et al. [43] is labeled as Study 2.

Table 5.1: Overview of the parameters of the volumetric studies considered.. The studies are labeled as Study 1 and Study 2 for the purpose of the literature comparison of this thesis.

Study	**Study 1** [42] (2019)	**Study 2** [43] (2020)
Healthy subjects	563	18,605
Subject age	19-86 years	3-90 years
Segmentation method	VolBrain	Freesurfer
Image contrast	T1	T1

As the volumes of subcortical brain regions are known to vary with age and sex [42], it is necessary to compare the findings of this thesis to a similar demographic. Table 5.2 presents the volumes of the ROIs reported by Study 1 and 2 at an age of 30 years, which is close to the mean age of the cohort of this study of 27.97 ± 5.73. The given error is the inter-subject SD. The studies considered, including this study have an approximately even ratio between male and female subjects.

The T1 segmented volume of the thalamus seems to correspond relatively well with the literature values of Table 5.2, though for the right thalamus the average volume is notably higher than expected from literature. Considering the scale of variation found between the literature values, the deviations of this study falls in an expected range. With an inter-subject SD of the T1 segmented thalamus volume of 0.57 cm^3, the relatively high volume variations between subjects combined with the limited cohort size of this study might explain the variations observed. The QSM segmentation on the other hand produced smaller volumes than expected from literature. The volume of the left thalamus was found to be 1.3 cm^3 less than the reported volume from Study 1, while the right thalamus corresponded better with the literature. One possible explanation for this observation is that the cohort in this study tends to have larger volumes of the right thalamus than the left, compared to the larger cohort of Study 1 and 2, while the QSM SynthSeg segmentation systematically underestimates the size of the thalamus. It is also possible that error in the predicted masks occurs for both the T1 and QSM SynthSeg segmentation. The important observation from this discussion is the considerable difference in the thalamic volume predicted by SynthSeg from T1w and QSM input data, suggesting that the input data does not produce the same segmentation.

The volumes of the caudate was found to be similar for both the QSM and T1 segmentation, though the QSM segments tends to be slightly smal-

Table 5.2: Subcortical volumes in literature. The numbers in this table are based on the findings of study 1 [42] and 2 [43] in the age range of 20-40. Note that the numbers presented are approximate, as they are read by eyesight from the graphs presented in the studies. The details of study 1 and study 2 are presented in Table 5.1. The volumes reported in this study from the T1 and QSM SynthSeg segmentation is included for comparison, and are retrieved from Table 4.1 and Table 4.2. The numbers are given in cm^3, and L/R denotes the left and right ROI.

ROI	**T1w input**	**QSM input**	**Study 1**	**Study 2**
	L/R	L/R	L/R	L/R
Thalamus	7.9/8.3	6.7/7.1	8.0/7.2	7.7/7.5
Caudate	4.2/4.4	4.0/4.3	3.4/3.2	3.8/3.8
Putamen	6.0/5.8	4.1/4.6	5.0/4.8	6.0/6.0
Pallidum	1.7/1.9	1.8/1.8	1.4/1.4	1.8/1.7
Hippocampus	4.2/4.0	2.1/2.5	4.2/4.2	4.3/4.3

ler. Both the T1 and QSM volumes were slightly larger than the volumes reported by Study 1 and 2. It should be noted that the values in Table 5.2 is included for a general comparison and comes with a relatively high uncertainty in the decimal place, as the numbers are read from a graph by eyesight. The reported values are subject to many variables such as acquisition parameters, image contrast, resolution and automated segmentation method, in addition to subject variations in the datasets. With this in mind, the literature values serves as good indicators of whether the volumes found in this thesis is within a reasonable range. In this context, the volumes of the caudate is considered to indicate accurate segmentation. The segmented volumes of the pallidum from both T1w and QSM input data in this study correlates well with the findings of Study 2. Study 1 reports slightly smaller volumes, but as some variation is expected, the volumetric analysis indicates accurate segmentation also for the pallidum.

Looking at the volumes of the putamen, the T1 SynthSeg segmentation is almost identical to the findings of Study 2, although higher than Study 1. Considering the expected variation between studies, this serves as an indicator of successful T1 segmentation of the putamen. The QSM segmentation produced significantly lower volumes for the putamen compared to both Study 1 and 2, as well as the T1 segmentation. This indicates that SynthSeg significantly underestimates the labeling of the putamen from QSM input data. The same underestimation is observed for the QSM segmentation of the hippocampus, where only half of the volume is reported by the QSM segmentation compared to the literature and T1 segmentation. The T1 SynthSeg segmentation of the hippocampus seems to be in particularly good agreement with the literature volumes.

As a general trend, the T1 SynthSeg segmentation tends to overestimate the subcortical volumes compared to the considered literature, while the

QSM segmentation tends to underestimate the volumes for certain ROIs. Considering the volumetric analysis, the QSM segmentation seems to be robust for the pallidum and caudate, while significantly underestimating the volumes of the putamen and hippocampus. Generally, the T1 segmentation volumes fall within an acceptable range, considering the expected variation in reported values due to the many variables between different studies.

5.1.1.3 QC scores

The SynthSeg software includes a prediction of the segmentation quality in the form of QC scores. The scores are based on a regression task performed by fast CNNs [34], with a main purpose of identifying failed segmentations. In the context of the comparison of this thesis, it is not appropriate to use the QC scores alone to justify the evaluation of segmentation quality. Instead, the QC scores are viewed in the collected analysis as one of multiple indicators of the segmentation quality.

The limit set by SynthSeg for rejecting a segmentation is a QC score of 0.65. All the ROIs segmented in this study achieved QC scores above this limit, but for some of the ROIs, the QC scores were only slightly above the discard limit, specifically the QSM segmentation of the hippocampus. A statistically significant decrease in QC scores was observed for the QSM segmentation for all ROIs where the score was predicted, compared to the T1 segmentation. Particularly, the QC score of the hippocampus and amygdala decreased from 0.85 ± 0.01 to 0.73 ± 0.05, indicating that the QSM input significantly reduces the accuracy of the segmentation. It should be noted that the QSM segmentation of the thalamus and the putamen and pallidum resulted in relatively high QC scores of 0.81 ± 0.03, predicting adequate segmentation. Still, an significant increase was observed when using the T1w input, increasing the scores to 0.89 ± 0.02 for the putamen and pallidum and 0.85 ± 0.02 for the thalamus.

5.1.1.4 Dice scores

The segmentation masks generated from T1w and QSM input data were compared quantitatively, calculating the DS between the ROI masks. The results presented in Section 4.1.4 shows that there are some differences between the segmentations, though the DS of the whole-brain segmentation of 0.86 ± 0.01 indicates a generally high correlation. Note that the error given is the inter-subject SD. Investigating the DS of the individual ROIs, the DS of the hippocampus stands out as significantly lower than for the other ROIs. The DS of the left hippocampus was as low as 0.51 ± 0.13. It is clear that the QSM input did not reproduce the segmentation of the

hippocampus based on the T1w input. A general agreement in terms of DS is found for the other ROIs, with the second larges deviations found in the segmentation of the putamen. The DSs found for the thalamus, caudate and pallidum in the range 0.83-0.87 indicates a generally high correlation.

5.1.1.5 Evaluation of SynthSeg performance with 7T T1w input

As a ground truth is not available, the evaluation of the segmentation quality is based on the quantitative analysis in Section 5.1.1.1, the segment volumes discussed in Section 5.1.1.2 and the QC scores discussed in Section 5.1.1.3. SynthSeg is well documented to produce reliable segmentations with T1w input [34], however, sparse literature is found for automatic segmentation of 7T MRI [44]. Imaging at ultra-high field strength cause increased spatial field inhomogenities compared to 3T MRI, and the performance of automated segmentation tools can be restricted if they are not optimized for 7T MRI. One recent study by Wei et al. [45] found accurate performance of SynthSeg with 7T T1w input. The results of this thesis also indicates good segmentation quality at 7T for the subcortical regions investigated. With the limitation of not having access to manual segmentations to calculate the DS to a ground truth, all the investigated factors indicated that SynthSeg performs well with 7T T1w input data. Noted that the visual analysis was limited, as it was not performed by a clinical professional, no prominent deviations to expected anatomy was observed. The volumes corresponded generally well with literature, and the QC scores predicted good segmentation quality.

5.1.1.6 Evaluation of SynthSeg performance with 7T QSM input

This section will both evaluate the QSM segmentation quality based on the quantitative analysis in Section 5.1.1.1, the segment volumes discussed in Section 5.1.1.2 and the QC scores discussed in Section 5.1.1.3, and in comparison to the T1w segmentation. SynthSeg has been documented to perform well on a variety of input contrast and resolution [9] [34], but there is little documentation regarding QSM input data. Segmenting on QSM images is beneficial as the pipeline for extracting susceptibility values could be simplified, and would not be dependent on acquisition of a T1w image and co-registration. It is also of interest to test whether QSM input data can increase the segmentation quality of certain ROIs, as iron rich brain structures have high QSM contrast.

The analysis found that the performance of SynthSeg on QSM input varied for different ROIs. The segmentation quality indicators all suggested reliable segmentation of the caudate and pallidum, while showing signs of poor performance for the thalamus, putamen and hippocampus. Partic-

ularly the hippocampus deviated greatly from expected literature volumes and visible anatomy. Even though the QSM data seemed to segment the caudate and pallidum successfully, the results of the other ROIs indicates that SynthSeg does not perform reliably on QSM input data on a general basis.

A statistically significant decrease in QC-scores were found for the QSM segmentation compared to the T1 segmentation. Additionally, both the qualitative and the volumetric analysis found the QSM segmentation to perform worse or equal to the T1 segmentation. Referring to the discussion of Section 5.1.1.5, the segmentation of SynthSeg with 7T T1 input is assumed to be of good quality. Assuming the T1 segmentation as the ground truth, the calculated DSs indicates that the QSM input produce reliable segmentations of the caudate, thalamus and pallidum, while the QSM segmentation is limited for the putamen, and particularly failing to segment the hippocampus.

The limitations of the QSM segmentation might be caused by several factors. The qualitative analysis finds that the QSM segmentation seems to fail particularly in the lateral and ventral direction. This might be due to increased artifacts near the air filled cavities of the ears and nose. Edges of susceptibility contrast are found in the interfaces between air and brain tissue. Strong susceptibility interfaces are known to cause inhomogeneities in the magnetic environment and rapid signal loss. The rapid signal loss might cause missing information in the phase data used to construct the QSM images [46]. This is a possible explanation of why the QSM images failed to produce an accurate segmentation of brain regions near ear cavities, as seen with the putamen and in particular the hippocampus.

The under-performance of the QSM segmentation can also be due to properties of the ROI tissue, causing slightly different segment shapes in QSM contrast and T1 contrast. While the CNN of SynthSeg is trained on volume segments and shows reliable performance on a variety of input contrast [34], the QSM contrast of some ROIs might differ slightly from the anatomical outline.

The findings might also be explained by the variation in susceptibility contrast between the ROIs. While the caudate, putamen and pallidum is of significantly higher susceptibility than the general brain tissue, creating good QSM contrast, the hippocampus displays a more subtle QSM contrast. Most brain tissue have a susceptibility close to that of water at -9 ppm [13], and the susceptibility measured for the hippocampus in this study were in the range of 1.6-2.5 ppb for the T1 masks, and -6.4 to -2.8 ppb for the QSM masks. It should be noted that due to the relative nature of QSM values the susceptibility of water might not be measured at the correct value in this study, so the value might not be comparable to the extracted QSM values of this study. Still, it is a possibility that the low contrast might be contrib-

uting to the failure of the QSM segmentation. However, the thalamus was measured to be of even lower susceptibility, with the QSM segmentation corresponding reasonably well with the T1 segmentation. This might be explained by having several factors contributing to the segmentation performance, including the shape of the surface of the ROIs and differences in the difficulty of distinguishing the structure from neighbouring tissue.

5.1.2 Susceptibility values of the thalamus, caudate, putamen, pallidum and hippocampus

5.1.2.1 Susceptibility values in literature

This section will discuss the susceptibility values extracted from the thalamus, caudate, putamen, pallidum and hippocampus and compare to literature values. Comparing susceptibility values across studies is challenging, as QSM values are relative. Reported values are often normalized to a reference region that allows for comparison between different subjects. A number of factors including the specific acquisition parameters, scanner type, strength of B-field, as well as the reconstruction method applied and the quality of segmentation, highly influence the findings of different studies, and finding a stable reference region can be challenging. The inter-subject variation of mean susceptibility in literature is generally high, and the results of a study can thus be affected by the characteristics of the participants, particularly with regards to age [47]. The comparison to literature values carried out in this section will consider the three studies presented in Table 5.3. Information specifying the variables are included for the considered studies, as no literature was found reporting the susceptibility with the same acquisition parameters and TGV-QSM reconstruction pipeline used in this study. This study reports the mean and median susceptibility of 29 healthy subjects with mean age 28.0 and an inter-subject SD of 5.7, ranging from 20-41 years of age.

Normalization of susceptibility values to a reference region is important for analysing and comparing QSM data. Conventionally, whole-brain normalization or CSF normalization have been used, but have been found to be less stable due to variations in whole-brain susceptibility linked to pathology and non-uniformity of the CSF [48], possibly caused by movement. Other reference areas such as the centrum semiovale have been suggested, and might provide a more stable normalization [21]. The values of this study considered here are not with regard to a reference region, as a stable reference were not segmented. The QSM and T1 segmentation found the SynthSeg segmentation of the CSF to be -0.72 ± 1.07 ppb and 0.96 ± 0.55 bbp, respectively, where the error is the inter-subject SD. As a CSF normalization of the values found in this study would only change the

values slightly, the raw susceptibility values reported in Section 4.1.5 will be compared to CSF normalized values reported in literature. The comparison is presented visually in Figure 5.1.

Table 5.3: Mean susceptibility of the thalamus, caudate, putamen, pallidum and hippocampus in literature [49] [50] [51]. The values are presented in ppb in the format mean (SD).

Study	**L1** [49] Li et al. (2021)	**L2** [50] Feng et al. (2018)	**L3** [51] Li et al. (2016)
Field strength	1.5T	3T	7T
N	31	8	29
Age	61.9 (11.8)	24.1 (2.4)	57.9 (8.6)
Method	MEDI	HEIDI	Least square
Segmentation	Manual	Auto	Auto
Normalization	CSF	CSF	CSF

The mean value of both the QSM and T1 segmentation of the thalamus fall in the same range as the literature values. It is notable that the L3 study found a particularly negative susceptibility of the thalamus of -30 ppb, and that the normalization might not be comparable. The susceptibility values of the caudate corresponds with the expected range from literature. The susceptibility of the putamen was found to be lower when extracted from the T1 masks, approximately at 5 ppb, compared to the range of 18-61 ppb reported in literature. The QSM masks found a higher susceptibility of the putamen at approximately 10 ppb, still significantly lower than the literature values. The susceptibility of the putamen significantly increases with age, as found in the study by Treit et al. [47], as visualized in Figure 5.2. The younger age of the subjects investigated in this study might partly explain the deviations. The lower susceptibility found in this study can also be due to the segmentation quality. If the segmentation masks includes brain tissue of negative susceptibility outside of the ROI, the mean susceptibility can be decreased. This seems less likely, as we observed that the QSM segmentation generally labeled a smaller ROI than the T1 segmentation, while finding a more positive susceptibility.

The susceptibility found in the pallidum was in the range 51-55 ppb for both the QSM and T1 segmentation, notable lower than the literature values of study L2 and L3. We observe a high variation in reported literature, referring to the findings of study L1, with a susceptibility of only 25.9 ppb for the pallidum, and the relatively high inter-subject SD reported in literature. The mean susceptibility of the hippocampus extracted from the T1 segmentation is slightly higher than the reported literature values. For the QSM segmentation, the findings of this study is closer to the literat-

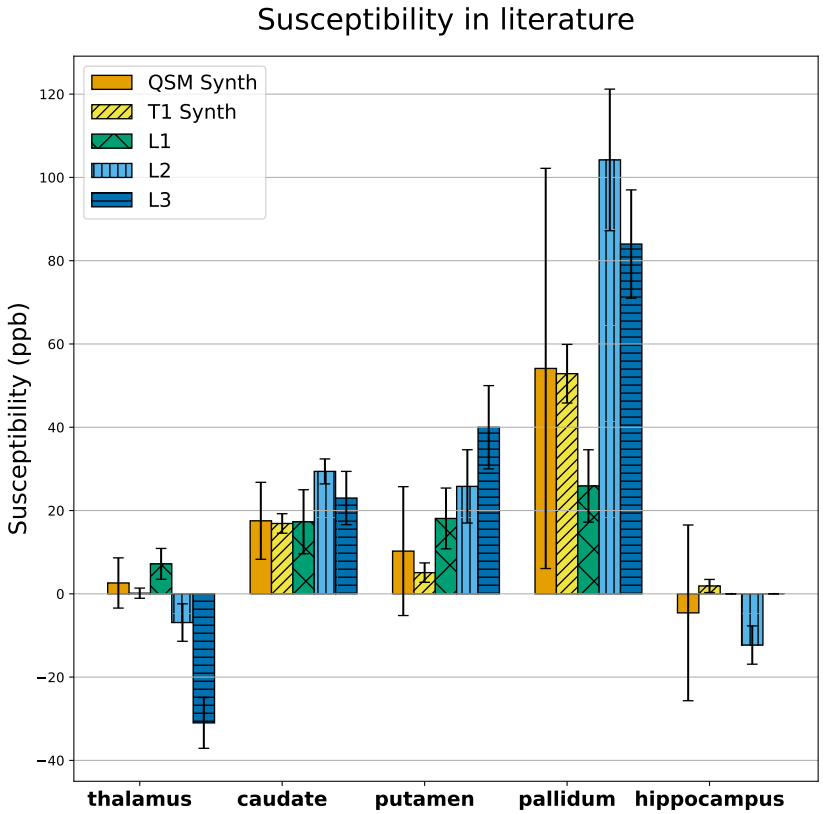


Figure 5.1: Susceptibility values reported in literature for the thalamus, caudate, putamen, pallidum and hippocampus. **L1** refers to the study by Li et al. (2021) [49], **L2** refers to the study by Feng et al. (2018) [50] and **L3** refers to Li et al. (2016) [51]. Further details of the studies is presented in Table 5.3. The values found in this study averaged over the left and right ROI are included for comparison. The error bars show the inter-subject SD.

ure, as more negative values were found. Considering the conclusions of the evaluation of the QSM segmentation in Section 5.1.1.6, the negative susceptibility found for the QSM segmentation is likely due to mislabeling and inclusion of brain tissue of negative susceptibility, and not due to a more accurate segmentation than the T1 segmentation. The deviations to literature is likely explained by the relative nature of the QSM values, and the multiple variables between the studies.

Compared to the literature values, a general underestimation of the mean susceptibility is observed for both the T1 and QSM segmentation, particularly for the putamen and pallidum. Some of the deviation to literature values might be due to the subject variation, particularly in small sample sizes, and it should be noted that study L2 only includes eight healthy controls. Study L1 and L3 includes subjects in a considerable higher age group than this study. The age dependency of susceptibility in the ROIs displayed in Figure 5.2, shows that a mean age of 28 might not be comparable to subjects in the age range above 60 years, particularly for the putamen and thalamus. The age dependency of this study was also examined further. The T1 segmented susceptibility in the ROIs as a function of age is presented in Figure 4.27 and Figure 4.26. A particular age dependency is not prominent for any of the ROIs. The age dependency was also investigated after normalization to the manually segmented centrum semiovale for 19 of the subjects. The thalamus and hippocampus seemed to show a slight increase with age for the normalized values. The result is expected considering the findings in Figure 5.2, as little variation is reported in the age range 20-40. An exception is found for the putamen, where Figure 5.2 shows a clear increasing trend for all ages, while the data of this study does not.

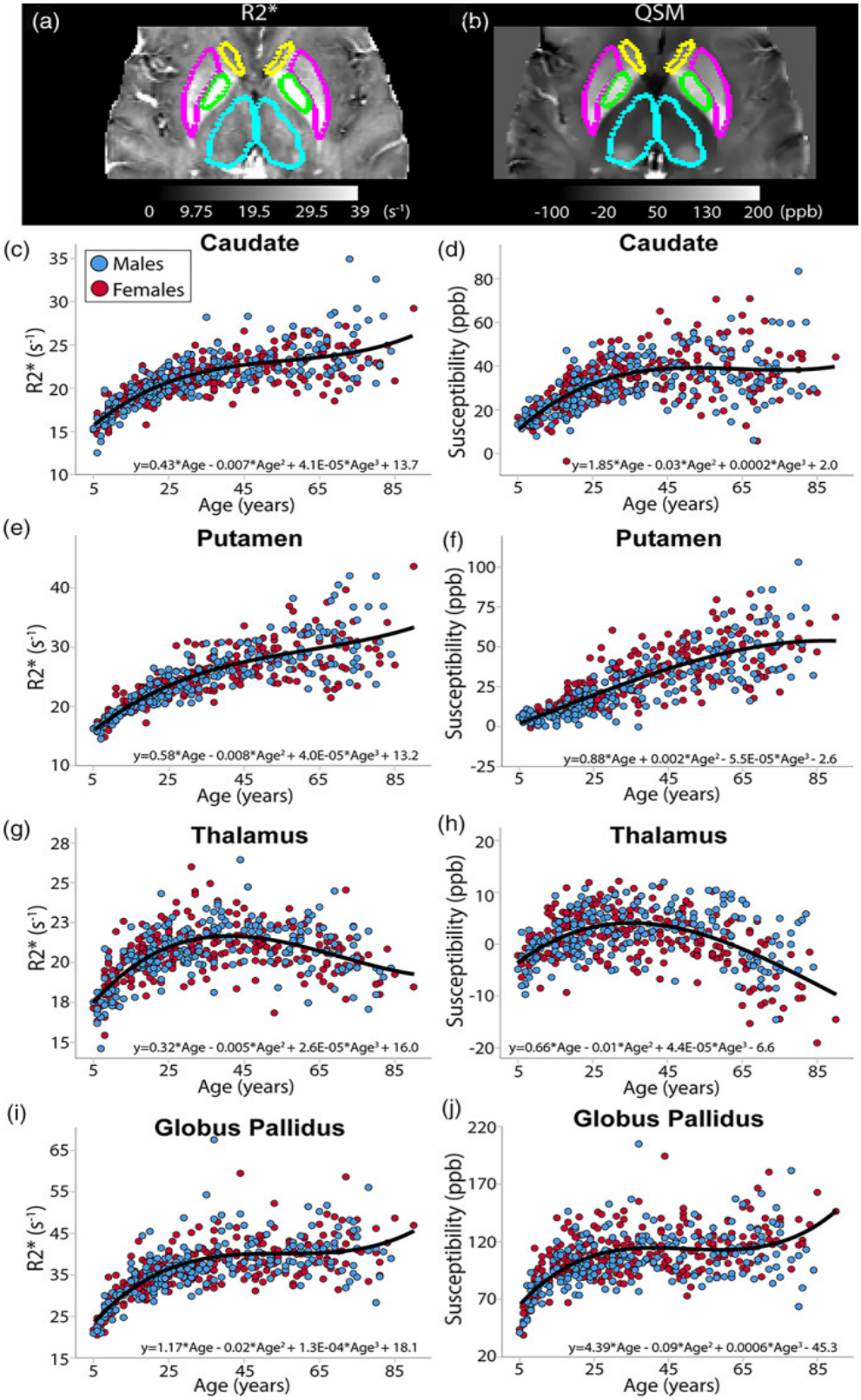
In addition to the subject variation, multiple factors might have contributed to the observed differences to literature values. The type of scanner, field strength and specific acquisition parameters introduce a large set of variables between studies. Instability in the CSF normalisation in the literature studies considered might also be a source of the found difference. A prior master thesis [11] investigated the dependency of the susceptibility values of different QSM reconstruction pipelines, as well as intra-subject variations across scans. The difference in susceptibility values from the single step TGV QSM reconstruction method implemented in this project and a threshold-based k-space division method [52] were found to be in the range of 1-4 ppb for the caudate, putamen and thalamus, 11 ppb for the pallidum and 6 ppb for the hippocampus. These magnitudes would not explain the deviations observed for the pallidum and putamen, but indicates that the different reconstruction methods in literature likely contributes to the variation in reported susceptibility. The master thesis [11] also reports intra-subject variations between different MRI scans in the range of

1-6 ppb, and shows that some variation is expected even with the same acquisition parameters and reconstruction method.

The extraction of susceptibility values is highly dependent on reliable segmentation. Sub-optimal segmentation quality might have caused the low susceptibility values found in this study. Overlabeling will generally decrease the mean value due to the negative susceptibility of general brain tissue. Still, it is difficult do draw any conclusion to the causes of the susceptibility differences. Likely, it is caused by a combination of multiple factors influencing the QSM values and the challenges of standardizing a normalization. Even with an underestimated susceptibility for some of the ROIs, the susceptibility ratios between the ROIs seemed to generally agree with literature Figure 5.2. The discussion emphasises the importance of controlling the variables in the susceptibility extraction pipeline, as the results are very sensitive to changes. A comparison of susceptibility values are more appropriate within a study, where the variables are controlled. This is particularly important when investigating the diagnostic potential of susceptibility values as biomarkers for neurodegenerative diseases like PD.

Variation and median value

A high variation in ROI susceptibility between subjects was found in this study, as well as in literature. The inter-subject SD of the mean ROI susceptibility values were found to increase slightly for the QSM segmentation compared to the T1 segmentation. An exception was found for the pallidum, where a slight decrease was observed for both the left and right region. The SD of the hippocampus, putamen and thalamus were found to increase with 1-4 ppb, indicating increased uncertainty for the QSM segmentation. The median susceptibility reported in this study was not found to differ significantly from the mean values. The inter-subject SD of the mean and median are approximately the same. The intra-ROI variation in the hippocampus and the left putamen were found to increase significantly ($p < 0.05$) for the QSM segmentation, and for the pallidum and thalamus the intra-ROI variation decreased significantly ($p < 0.05$), referring to Table 4.5. This might reflect the QSM segmentation's difficulty in segmenting the putamen and hippocampus discussed in Section 5.1.1.6.



© 2017, Treit et al.

Figure 5.2: Mean QSM value in ROIs as a function of age, in reference to the whole-brain. The figure is retrieved from a study of 498 healthy subjects published by Treit et al. in 2021 [47]. Licence: CC BY-NC-ND 4.0.

5.1.2.2 Susceptibility values as biomarkers for PD

It is of interest to establish a reliable automated method for extraction of QSM values, for the purpose of exploring the diagnostic potential of susceptibility measurements in subcortical brain regions. The analysis of segmentation pipelines in Section 5.1.1 found that SynthSeg seems to perform well with 7T T1w input data, and less reliably with 7T QSM input. As discussed in Section 5.1.2.1, QSM values are very sensitive to variables in the acquisition, reconstruction and post-processing of the data. For this reason the QSM extraction pipeline should be as simplified as possible, to limit the number of variables and post-processing steps that might introduce noise. Automated segmentation generated from QSM images might be beneficial, as it removes the involvement of a T1w acquisition and co-registration. The results showed that SynthSeg with QSM input seems to be reliable for some segments, the caudate and pallidum, while lacking in accuracy for the thalamus and putamen, and particularly failing to segment the hippocampus. The hypothesis of improved segmentation quality of iron rich subcortical nuclei does not correspond with the findings of decreased segmentation quality for the QSM segmentation compared to T1. It is still of interest to investigate how much the QSM segmentation deviates from the T1 segmentation in the mean susceptibility extracted. If the two segmentation pipelines reports susceptibility values that are sufficiently similar, some decrease in segmentation accuracy of the QSM segmentation might not be relevant for the purpose of identifying biomarkers for PD for certain ROIs.

For the caudate and pallidum no statistically significant difference were found in the mean susceptibility between the segmentation pipelines, with p-values in the range of 0.2-0.9. This indicates that the variation in measured susceptibility from the QSM segmentation is likely negligible for the purpose of exploring the diagnostic potential of susceptibility in the caudate and pallidum.

The susceptibility of the thalamus, putamen and hippocampus all changed significantly with p-values < 0.005. Whether this change will affect the identification of biomarkers depends on the expected deviation in susceptibility due to pathology. As this study only looks at healthy individuals, it is difficult to compare to an expected magnitude of change and draw any conclusions. Due to the relative nature of QSM values and the observed variation in literature, findings of the magnitude of increased susceptibility due to pathology can not be directly compared across studies. Still, the findings of literature will be considered in the following discussion, to indicate whether the error in the QSM segmentation could likely be neglected or not. Additionally, if the susceptibility of a ROI has the potential to be used as a biomarker for PD, the observed differences due to patho-

logy should be significant in comparison to the expected variation within healthy controls. The inter-subject SD can therefore provide information about what scale should be considered acceptable for the segmentation-dependent difference in extracted susceptibility. The smallest inter-subject variations were found for the T1 segmentation, and will be referred to in further discussion of the variation in this study.

The literature on susceptibility values in the basal ganglia as biomarkers for PD reports different findings of the statistical significance and the magnitude of difference expected from pathology [6] [5] [53] [54]. The systematic review of QSM in neurodegenerative diseases by Ravanfar et al. in 2021 [6] found for PD patients that 4 out of 22 studies found increased susceptibility in the putamen, 1 out of 18 studies found the same for the caudate, 6 out of 21 for the pallidum, 2 out of 4 for the hippocampus and 3 out of 11 studies for the thalamus. Langkammer et al. [54] reported an increase of 5 ppb for the thalamus in PD patients, with a p-value < 0.05. The inter-subject SD were found to be 1.80 and 1.66 ppb for the left and right thalamus in this study. As the difference in mean susceptibility between the T1 and QSM segmentation of the thalamus were slightly less than 3 ppb, it is likely that the QSM segmentation will limit the diagnostic potential of the method. The differences found for the hippocampus were as high as 8 ppb, approximately the same magnitude as the difference between PD patients and healthy controls found in the study by Li et al. in 2018 [55]. The difference of 8 ppb is additionally much higher than the inter-subject SD found in this study. The putamen shows a difference in susceptibility for the QSM segmentation and T1 segmentation of 4.2-5.9 ppb, a high value compared to the inter-subject variation of 3.2-3.4 ppb. A study by Chen et al. [56] found a pathological difference of 13 ppb. It should be noted that the mean susceptibility values of the putamen reported by Chen et al. were much higher than in this study, and that the susceptibility values are likely not comparable.

The magnitude difference found by the T1 and QSM segmentation for the left and right caudate were found to be -0.12 ± 5.02 and -1.16 ± 4.92 ppb, respectively, where the error is the variation between subjects in the differences between segmentation methods. The inter-subject SD of the mean susceptibility value of the left and right caudate were found to be 3.19 and 3.42 ppb. The difference due to segmentation method is considerably smaller than the variation between healthy subjects. A significant increase in susceptibility due to pathology would therefore likely be greater than the segmentation difference in extracted mean susceptibility. A recent study by Rong et al. [5] found a significant increase in susceptibility for the caudate ($p < 0.05$) and pallidum ($p < 0.005$) between healthy controls and late stage PD patients. It should be noted that no statistical significance were found in these regions for early stage PD, which might

partly explain the findings of no statistical significance for the caudate by some of the other studies [53] [54]. The study by Rong et al. [5] found similar mean susceptibility values as this study for the caudate at 21 ± 9 ppb, and an increase of 8 ppb for late stage PD patients. This indicates that the variation in measured susceptibility from the QSM segmentation is likely negligible for the purpose of exploring the diagnostic potential of susceptibility in the caudate.

There are several studies reporting findings of a statistically significant difference in susceptibility value for the pallidum in PD patients. The studies by Rong et al. [5], Shahmaei et al. [53] and Langkammer et al. [54] reports pathological differences in mean susceptibility of 28 ppb, 69 ppb and 14 ppb, respectively. The range of reported values demonstrates the limitations of comparison across studies, but does indicate that the variation of 1.72 ± 12.92 ppb and 0.81 ± 14.07 ppb due to the segmentation method of the pallidum are likely small compared to pathology. It should be noted that the range of susceptibility values reported in literature indicates that the QSM values of different studies are not only shifted by a term of reference, but also scaled in comparison to each other, as the magnitude differences between different ROIs are not constant, as can be seen in Figure 5.1.

Considering that the T1 segmentation seems to be generally robust and that the reliability of the QSM segmentation seems to differ greatly across the ROIs, it is suggested to use the T1 segmentation for the susceptibility extraction pipeline to investigate potential biomarkers for PD. Investigating the diagnostic power is out of the scope of this thesis, as only healthy individuals are included. For further work the pipeline should be evaluated in terms of the prediction of pathology. Another limitation of this project is the low age range of the subjects. Neurological diseases like PD and ALS are more prevalent in the older population. Considering the age dependency of the susceptibility in the ROIs, it is necessary to include a higher age range of healthy controls in further studies.

5.2 U-net segmentation of the SN and RN

5.2.1 Susceptibility values of the RN and SN

The mean susceptibility of the SN, RN and OMEGA were found to be 70.16 ± 12.10, 47.78 ± 11.75 and 21.14 ± 4.50 ppb, respectively, using the masks generated from the U-net described in Section 3.4.1. It should be noted that the 15 subjects making up the training set of the U-net are included in the dataset of this study. The same raw MRI data was used, but the reconstruction pipeline for the QSM training data was not identical to the one in this study. A ground truth of the segmentation for these 15 subjects were

available, referring to the manually drawn volume segments of the whole ROIs were used for the training. The 19 additional subjects segmented in this study were not seen by the U-net before.

For comparison, the mean susceptibility extracted from the manual volume segments for the 15 subjects where a ground truth were available are reported in Table 4.7. The mean susceptibility in the SN was found to be 68.82 ± 12.73 ppb, while the RN was found to be 43.67 ± 10.56 ppb, both slightly lower than the values extracted from the masks generated by the U-net. The OMEGA was also found to be of slightly lower susceptibility using the manual segmentation masks. If the deviation in the results were caused by inaccurate segmentation by the U-net, a decrease in the susceptibility extracted from the U-net masks would be expected instead of an increase, as the voxels outside the ROI would probably be of negative susceptibility. The slight increase is likely due to the inclusion of additional subjects, increasing the mean age from 25.81 ± 4.08 years to 27.97 ± 5.73. The susceptibility of the RN and SN is known to increase with age [57]. The investigation of age dependency in this study of the SN and RN in Figure 4.25 and the OMEGA in Figure 4.27, did not show a prominent increase with age compared to the inter-subject variation. Still it is notable that the lowest values occur for the lowest ages for all the ROIs. The slightly older cohort, as well as the large inter-subject variation, might explain the increase we observe in the susceptibility values compared to the manual segmentation. With basis in this discussion, the U-net segmentation appears to be reliable for extraction of susceptibility values in the SN, RN and OMEGA.

It is of interest to discuss whether the susceptibility values extracted from the U-net are suitable as predictors for PD, considering the relatively high variations found between healthy subjects. A study by Li et al. [51] in 2016 found statistically significant differences for both the SN and RN in patients with PD. The study found the susceptibility values of the SN to be 114 ± 40 ppb in patients with PD, and 90 ± 30 ppb in healthy controls, with a difference of 24 ± 50 ppb. With the assumption that the expected increase due to PD is around 25 ppb for the QSM pipeline in this study, the inter-subject SD found in this study using the automated segmentation at 12.20 ppb indicates that the U-net could be suitable for detecting pathology. It should be noted that the measured susceptibility values of the SN vary greatly in literature [58], and a variety of differences due to pathology are reported. The same study found the susceptibility in the RN to be 105 ± 40 ppb for PD patients, and 89 ± 30 ppb for healthy controls, with a statistical significance of $p = 0.001$. The inter-subject variation in this study at 11.75 ppb for the RN would be approximately of the same magnitude as the expected change due to pathology, assuming the literature difference [51] is representative for this study. The analysis suggests that the U-net

segmentation of the RN might not be able to detect pathology. Further studies should include patients with PD to investigate the prediction power of this specific QSM extraction pipeline.

5.2.2 Correlation of automated and manual susceptibility of the SN and OMEGA

Figure 4.22 displays a linear regression of the mean susceptibility extracted from the U-net masks of the SN to the manually extracted susceptibility of 19 healthy subjects. It should be noted that the manual values considered in the correlation analysis are not referring to the susceptibility extracted from the whole ROI manual segmentations considered in Section 5.2.1. The manual values used for the correlation analysis were acquired by a radiologist inspecting the QSM images and placing a circular ROI of approximately 2 mm in diameter within the SN and OMEGA structures. It is of interest to explore how the automatically extracted values correlate with the radiologists values, and how well they are able to predict the manual values.

A linear regression of the mean susceptibility of the automatically segmented SN masks found a R^2-value of 0.57, showing some correlation to the manual values. The analysis found a higher correlation to the 98-percentile susceptibility of the automated segments, with a R^2-value of 0.80. The correlation of other statistics to the manual values were investigated, specified in Table 4.8, but were not found to increase the correlation compared to the 98-percentile value. The linear regression of the 98-percentile value is displayed in Figure 4.23, and shows that the automated segmentation predicted values with high correlation to the values acquired manually. The p-values from a two-sided t-test were calculated between the predicted and manual values, resulting in $p = < 0.0001$ for the mean and $p = 0.089$ for the 98-percentile value. This suggest that the strong correlation found in the 98-percentile might be a characteristic of the dataset, and is not statistically significant. The correlation found for the mean value on the other hand shows a strong statistical significance. For this reason, the mean value of the automatically extracted susceptibility of the SN shows the greatest potential for prediction of the manual values.

The same correlation analysis was performed for the U-net segmentation of the OMEGA. The linear regression of the mean value found a more subtle correlation displayed in Figure 4.24, with a R^2-value of 0.43. The same parameters were explored as for the SN, but no increase in correlation was found compared to the mean value. The linear regression of the 98-percentile value resulted in a correlation of only 0.28, further suggesting that the 98-percentile value is not a stronger predictor of the manual values, and that the high correlation found for the SN data might be due to

chance. Assuming a normal distribution of the susceptibility, a two-sided t-test found the p-value to be 0.284, considering the correlation of the predicted values from the mean of the automated segment to the manual values. As the results are not statistically significant, the automated OMEGA susceptibility values does not seem to strongly predict the manual values.

It should be noted that an exploratory method was used to investigate whether the statistics specified in Table 4.8 showed a better correlation to the manual values than the mean value. The problem of multiple comparisons should be acknowledged here, as the chances of finding a correlation to one parameter increases as more parameters are investigated. This might be an explanation for the high correlation found of the 98-percentile value to the manual values for the SN. It would be of interest to test whether a new set of data would display a similar correlation. The sample size is a limitation of this analysis, and further analysis including a larger number of subjects would be of interest.

As the manual values were extracted from small parts of the SN, opposed to the whole structure masks the automated values are extracted from, it is likely that this affects the information in the susceptibility values extracted and interfere with the correlation. It is for this reason interesting to explore whether particular statistics of the automated segment such as higher susceptibility percentiles are more representative for the smaller regions of the manual ROIs. It should also be noted that the radiologist avoids inclusion of blood vessels, as this will increase the variation within the manual measurements. One limitation of the automated segmentation is that it is challenging to exclude voxels containing high signal from blood vessels within the ROI, which will increases the intra-ROI variation and might affect the mean susceptibility measured for the automated segmentation, as well as the ability to predict the manual values. It is important to mention that for the automated values to serve as possible biomarkers for neurodegenerative diseases such as PD or ALS, it is not necessarily required that they are able to exactly predict the manual values. Further work should include patients suffering from PD and ALS, and investigate the prediction power of the automated values of pathology compared to the manual values.

Chapter 6

Conclusion

This thesis investigated the automated segmentations of subcortical regions generated by the DL-based SynthSeg with T1w input data and QSM input data. The segmentation quality was evaluated in terms of a quantitative analysis, the segmented volumes and the predicted QC scores from SynthSeg. The thesis found a robust performance of SynthSeg on 7T T1w input images. The evaluation of SynthSeg with QSM input data found that the performance was reduced, particularly for segmentations of areas closer to the air filled cavities of the ears, specifically the hippocampus and putamen. This is thought to be caused by loss of information in the QSM phase data due to artifacts arising from the sharp susceptibility edges of air and tissue. The comparison of the QSM segmentation to the T1 segmentation found DS in the range of 0.83-0.87 for the thalamus, caudate and pallidum, while the left putamen and hippocampus scored lower, with DSs of 0.74 ± 0.04 and 0.51 ± 0.13, respectively. The difference in extracted susceptibility values between the QSM and T1 segmentation were not found to be significant for the caudate and pallidum, while a statistically significant change in susceptibility due to segmentation method was found for the thalamus, putamen and hippocampus ($p < 0.005$). Based on the findings of this thesis, the 7T T1w SynthSeg segmentation with FSL FLIRT co-registration of the QSM image is suggested to be used in the automated susceptibility extraction pipeline.

For the 29 healthy volunteers, the susceptibility extracted from the T1w segmented left thalamus, caudate, putamen, pallidum and hippocampus were found to be 0.45 (1.80), 17.05 (3.19), 5.23 (3.18), 53.86 (10.52) and 1.60 (2.41) ppb, respectively, presented with the inter-subject SD in parentheses. The raw susceptibility were found to be 70.16 (12.10), 47.78 (11.75) and 21.14 (4.50) ppb for the U-net segmentation of the SN, RN and OMEGA, respectively. A general underestimation of the susceptibility values compared to literature was found, but the problem of comparing QSM values across studies due to the relative nature of QSM values should

be noted, as well as the younger age group of the subjects of this study. The variations in ROI susceptibility between subjects were found to be in the range of 1.8-10.5 ppb, and reflects the variation found in literature. Further research including patients of pathology is needed to investigate the diagnostic power of 7T QSM for PD, as it is necessary to establish the expected change due to pathology of this method and compare to the relatively high variation found in healthy subjects.

A correlation was found for the mean susceptibility of the automatically segmented SN to values manually measured by a radiologist by performing a linear regression, with $R^2 = 0.57$. The p-value for the prediction was found to be significant ($p < 0.0001$). A correlation was found to the 98-percentile value in the automated SN segment, but with a p-value of 0.089, the R^2-value of 0.80 of the linear regression to the manual measurements was not found to be significant. The correlation analysis was also performed for the automatically segmented OMEGA, resulting in a R^2-value of 0.43 from a linear regression, but no statistical significance of the prediction was found. The correlation found for the mean susceptibility in the SN segmentation shows a potential for the automated segmentation to predict values currently extracted manually by radiologists, and might be feasible for identification of biomarkers of pathology. Further research is suggested including a larger dataset and patients of pathology to investigate the prediction power of the automated values. It is also suggested to include a higher age range of the healthy volunteers, as susceptibility values in the ROIs are known to vary with age and the prevalence of PD increases rapidly with age.

Bibliography

[1] B. Brakedal, L. Toker, Haugarvoll and K. et al, 'A nationwide study of the incidence, prevalence and mortality of parkinson's disease in the norwegian population,' *npj Parkinsons Disease*, vol. 8, Mar. 2022. DOI: 10.1038/s41531-022-00280-4.

[2] B. R. Bloem, M. S. Okun and C. Klein, 'Parkinson's disease,' *The Lancet*, vol. 397, no. 10291, pp. 2284–2303, 2021, ISSN: 0140-6736. DOI: https://doi.org/10.1016/S0140-6736(21)00218-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S014067362100218X.

[3] N. Helseinformatikk. 'Påvisning av parkinsons sykdom.' (2021), [Online]. Available: https://nhi.no/sykdommer/hjernenervesystem/parkinson/parkinson-diagnose/ (visited on 08/01/2023).

[4] M. Lawton, Y. Ben-Shlomo, M. T. May, F. Baig, T. R. Barber, J. C. Klein, D. M. A. Swallow, N. Malek, K. A. Grosset, N. Bajaj, R. A. Barker, N. Williams, D. J. Burn, T. Foltynie, H. R. Morris, N. W. Wood, D. G. Grosset and M. T. M. Hu, 'Developing and validating parkinson's disease subtypes and their motor and cognitive progression,' *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 89, no. 12, pp. 1279–1287, 2018, ISSN: 0022-3050. DOI: 10.1136/jnnp-2018-318337. eprint: https://jnnp.bmj.com/content/89/12/1279.full.pdf. [Online]. Available: https://jnnp.bmj.com/content/89/12/1279.

[5] Y. Rong, Z. Xu, Y. Zhu, X. Zhang, L. Lai, S. Sun, M. Gao, P. Guo, G. Zhang, Y. Geng, X. Ma, S. Wu, L. Yang, Z. Shen and J. Guan, 'Combination of quantitative susceptibility mapping and diffusion kurtosis imaging provides potential biomarkers for early-stage parkinson's disease,' *ACS Chemical Neuroscience*, vol. 13, no. 18, pp. 2699–2708, 2022, PMID: 36047877. DOI: 10.1021/acschemneuro.2c00321. eprint: https://doi.org/10.1021/acschemneuro.2c00321. [Online]. Available: https://doi.org/10.1021/acschemneuro.2c00321.

[6] P. Ravanfar, S. M. Loi, W. T. Syeda, T. E. Van Rheenen, A. I. Bush, P. Desmond, V. L. Cropley, D. J. R. Lane, C. M. Opazo, B. A. Moffat, D. Velakoulis and C. Pantelis, 'Systematic review: Quantitative susceptibility mapping (qsm) of brain iron profile in neurodegenerative diseases,' *Frontiers in Neuroscience*, vol. 15, 2021, ISSN: 1662-453X. DOI: 10.3389/fnins.2021.618435. [On-

line]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2021.618435.

[7] A. M. Dale, B. Fischl and M. I. Sereno, 'Cortical surface-based analysis: I. segmentation and surface reconstruction,' *NeuroImage*, vol. 9, no. 2, pp. 179–194, 1999, ISSN: 1053-8119. DOI: https://doi.org/10.1006/nimg.1998.0395. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811998903950.

[8] Y. Zhang, M. Brady and S. Smith, 'Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm,' *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001. DOI: 10.1109/42.906424.

[9] B. Billot, D. N. Greve, O. Puonti, A. Thielscher, K. Van Leemput, B. Fischl, A. V. Dalca and J. E. Iglesias, 'Synthseg: Domain Randomisation for Segmentation of Brain MRI Scans of any Contrast and Resolution,' *arXiv:2107.09559 [cs]*, 2021.

[10] E. Lysheim, *Analysis of quantitative susceptibility mapping in healthy volunteers at 3t and 7t*, 2021.

[11] M. R. Vik, *Quantitative susceptibility mapping reproducibility with 7t ultra-high field mri*, 2022.

[12] F. Gutiérrez-Mejía and J. Ruiz-Suárez, 'Ac magnetic susceptibility at medium frequencies suggests a paramagnetic behavior of pure water,' *Journal of Magnetism and Magnetic Materials*, vol. 324, no. 6, pp. 1129–1132, 2012, ISSN: 0304-8853. DOI: https://doi.org/10.1016/j.jmmm.2011.10.035. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S030488531100758X.

[13] J. H. Duyn and J. Schenck, 'Contributions to magnetic susceptibility of brain tissue,' *NMR in Biomedicine*, vol. 30, no. 4, e3546, 2017, e3546 NBM-15-0325.R1. DOI: https://doi.org/10.1002/nbm.3546. eprint: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.3546. [Online]. Available: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3546.

[14] J. Klohs and A. M. Hirt, 'Investigation of the magnetic susceptibility properties of fresh and fixed mouse heart, liver, skeletal muscle and brain tissue,' *Physica Medica*, vol. 88, pp. 37–44, 2021, ISSN: 1120-1797. DOI: https://doi.org/10.1016/j.ejmp.2021.06.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1120179721002362.

[15] H.-G. Shin, J. Lee and Y. H. Yun, '-separation: Magnetic susceptibility source separation toward iron and myelin mapping in the brain,' *Neuroimage*, vol. 240, Jul. 2021. DOI: 10.1016/j.neuroimage.2021.118371.

[16] W. Schulz-Schaeffer, 'The synaptic pathology of -synuclein aggregation in dementia with lewy bodies, parkinson's disease and parkinson's disease dementia,' *Acta neuropathologica*, vol. 120, pp. 131–43, Aug. 2010. DOI: 10.1007/s00401-010-0711-0.

[17] Q. Chen, Y. Chen, Y. Zhang, F. Wang, H. Yu, C. Zhang, Z. Jiang and W. Luo, 'Iron deposition in parkinson's disease by quantitative susceptibility mapping,' *BMC Neuroscience*, Apr. 2019. DOI: 10.1186/s12868-019-0505-9.

[18] Z. Jin, Y. Wang, M. Jokar, Y. Li, Z. Cheng and Y. Liu, 'Automatic detection of neuromelanin and iron in the midbrain nuclei using a magnetic resonance imaging-based brain template,' *Hum Brain Mapp*, Apr. 2022. DOI: 10.1002/hbm.25770.

[19] B. Vahsen, E. Gray and A. e. a. Thompson, 'Non-neuronal cells in amyotrophic lateral sclerosis — from pathogenesis to biomarkers,' *Nat Rev Neurol*, vol. 17, Jun. 2021. DOI: 10.1038/s41582-021-00487-8.

[20] H. Norge. 'Amyotrofisk lateral sklerose (als).' (2022), [Online]. Available: https://www.helsenorge.no/sykdom/hjerne-og-nerver/als/ (visited on 22/12/2022).

[21] K. E. Dean, B. Shen, G. Askin, A. D. Schweitzer, M. Shahbazi, Y. Wang, D. Lange and A. J. Tsiouris, 'A specific biomarker for amyotrophic lateral sclerosis: Quantitative susceptibility mapping,' en, *Clin. Imaging*, vol. 75, pp. 125–130, Jul. 2021.

[22] B. Genc, J. Jara, A. Lagrimas, P. Pytel, R. Roos, M. Mesulam, C. Geula, E. Bigio and H. Ozdinler, 'Apical dendrite degeneration, a novel cellular pathology for betz cells in als,' *Scientific Reports*, vol. 7, p. 41 765, Feb. 2017. DOI: 10.1038/srep41765.

[23] Q. Li, W. Zhu, X. Wen, Z. Zang, Y. Da and J. Lu, 'Beyond the motor cortex: Thalamic iron deposition accounts for disease severity in amyotrophic lateral sclerosis,' *Frontiers in neurology*, vol. 13, p. 791 300, 2022, ISSN: 1664-2295. DOI: 10.3389/fneur.2022.791300. [Online]. Available: https://europepmc.org/articles/PMC8907117.

[24] C. Langkammer, K. Bredies, B. A. Poser, M. Barth, G. Reishofer, A. P. Fan, B. Bilgic, F. Fazekas, C. Mainero and S. Ropele, 'Fast quantitative susceptibility mapping using 3d epi and total generalized variation,' *NeuroImage*, vol. 111, pp. 622–630, 2015, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2015.02.041. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811915001421.

[25] A. Deistung, F. Schweser and J. R. Reichenbach, 'Overview of quantitative susceptibility mapping,' *NMR in Biomedicine*, vol. 30, no. 4, 2017, e3569 NBM-15-0326.R2. DOI: https://doi.org/10.1002/nbm.3569. [Online]. Available: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3569.

[26] K. Eckstein, B. Dymerska, B. Bachrata, W. Bogner, K. Poljanc, S. Trattnig and S. Robinson, 'Computationally efficient combination of multi-channel phase data from multi-echo acquisitions (aspire),' *Magn. Reson. Med*, vol. 79, pp. 2996–3006, 2018. DOI: 10.1002/mrm.26963.

[27] W. Jung, S. Bollmann and J. Lee, 'Overview of quantitative susceptibility mapping using deep learning: Current status, challenges and opportunities,' *NMR in Biomedicine*, vol. 35, no. 4, e4292, 2022, e4292 NBM-19-0283.R1. DOI: https://doi.org/10.1002/nbm.4292. [Online]. Available: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4292.

[28] S. M. Smith, 'Fast robust automated brain extraction,' *Human Brain Mapping*, vol. 17, no. 3, pp. 143–155, 2002. DOI: https://doi.org/10.1002/hbm.10062. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.10062.

[29] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady and P. M. Matthews, 'Advances in functional and structural mr image analysis and implementation as fsl,' *NeuroImage*, vol. 23, S208–S219, 2004, Mathematics in Brain Imaging, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2004.07.051. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811904003933.

[30] K. Bredies, 'Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty,' *Lecture Notes in Computer Science*, vol. 8293, pp. 44–77, Jan. 2014. DOI: 10.1007/978-3-642-54774-4_3.

[31] M. A. Schofield and Y. Zhu, 'Fast phase unwrapping algorithm for interferometric applications,' *Opt. Lett.*, vol. 28, no. 14, pp. 1194–1196, Jul. 2003. DOI: 10.1364/OL.28.001194. [Online]. Available: https://opg.optica.org/ol/abstract.cfm?URI=ol-28-14-1194.

[32] W. Li, B. Wu and C. Liu, 'Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition,' *NeuroImage*, vol. 55, no. 4, pp. 1645–1656, 2011, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2010.11.088. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811911000164.

[33] B. Billot, M. Colin, S. Das and J. E. Iglesias, 'Robust Segmentation of Brain MRI in the Wild with Hierarchical CNNs and no Retraining,' in *MICCAI: Medical Image Computing and Computer Assisted Intervention*, 2022.

[34] B. Billot, C. Magdamo, S. Arnold, S. Das and J. Iglesias, 'Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain mri datasets,' Jan. 2023. DOI: 10.48550/arXiv.2209.02032.

[35] P. K. Mandal, R. Mahajan and, 'Structural brain atlases: Design, rationale, and applications in normal and pathological cohorts,' *Journal of Alzheimer's disease*, vol. 31, no. 3, pp. 169–188, 2012. DOI: https://doi.org/10.3233/JAD-2012-120412.

[36] 'A global optimisation method for robust affine registration of brain images,' *Medical Image Analysis*, vol. 5, no. 2, pp. 143–156, 2001, ISSN: 1361-8415. DOI: https://doi.org/10.1016/S1361-8415(01)00036-6. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841501000366.

[37] M. Jenkinson, P. Bannister, M. Brady and S. Smith, 'Improved optimization for the robust and accurate linear registration and motion correction of brain images,' *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002, ISSN: 1053-8119. DOI: https://doi.org/10.1006/nimg.2002.1132. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811902911328.

[38] D. Purves and et al., *Neuroscience*. New York: Sinauer Associates, Oxford University Press, 2018, ISBN: 978-1-605-35841-3.

[39] C. Rorden and M. Brett, 'Stereotaxic display of brain lesion,' *Behavioural neurology*, vol. 12, pp. 191–200, Feb. 2000. DOI: 10.1155/2000/421719.

[40] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper and R. Kikinis, '3d slicer as an image computing platform for the quantitative imaging network,' *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012, Quantitative Imaging in Cancer, ISSN: 0730-725X. DOI: https://doi.org/10.1016/j.mri.2012.05.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0730725X12001816.

[41] Z. Cho, *7.0 Tesla MRI Brain Atlas: In-vivo Atlas with Cryomacrotome Correlation*. Jan. 2015, ISBN: 978-3-642-54397-5. DOI: 10.1007/978-3-642-54398-2.

[42] Y. Wang, Q. Xu, J. Luo, M. Hu and C. Zuo, 'Effects of age and sex on subcortical volumes,' *Frontiers in Aging Neuroscience*, vol. 11, 2019, ISSN: 1663-4365. DOI: 10.3389/fnagi.2019.00259. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnagi.2019.00259.

[43] D. Dima, A. Modabbernia, E. Papachristou *et al.*, 'Subcortical volumes across the lifespan: Data from 18,605 healthy individuals aged 3–90years,' *Human Brain Mapping*, vol. 43, no. 1, pp. 452–469, 2022. DOI: https://doi.org/10.1002/hbm.25320. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.25320. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.25320.

[44] M. Svanera, S. Benini, D. Bontempi and L. Muckli, 'Cerebrum-7t: Fast and fully volumetric brain segmentation of 7 tesla mr volumes,' *Human Brain Mapping*, vol. 42, no. 17, pp. 5563–5580, 2021. DOI: https://doi.org/10.1002/hbm.25636. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.25636. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.25636.

[45] N. Wei, J. Jing, Y. Zhuo and Z. Zhang, 'Morphological characteristics of lenticulostriate arteries in a large age-span population: Results from 7t tof-mra,' *Frontiers in Neurology*, vol. 13, 2022, ISSN: 1664-2295. DOI: 10.3389/fneur.2022.944863. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fneur.2022.944863.

[46] Y. Wang and M. R. Prince, 'Quantitative susceptibility mapping is superior to t1-weighted imaging for detecting and measuring gadolinium,' *Radiology*, vol. 297, no. 1, pp. 151–153, 2020, PMID: 32697167. DOI: 10.1148/radiol.2020202801. eprint: https://doi.org/10.1148/radiol.2020202801. [Online]. Available: https://doi.org/10.1148/radiol.2020202801.

[47] S. Treit, N. Naji, P. Seres, J. Rickard, E. Stolz, A. H. Wilman and C. Beaulieu, 'R2* and quantitative susceptibility mapping in deep gray matter of 498 healthy controls from 5 to 90years,' *Human Brain Mapping*, vol. 42, no. 14, pp. 4597–4610, 2021. DOI: https://doi.org/10.1002/hbm.25569. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.25569.

[48] A. V. Dimov, T. D. Nguyen, P. Spincemaille, E. M. Sweeney, N. Zinger, I. Kovanlikaya, B. H. Kopell, S. A. Gauthier and Y. Wang, 'Global cerebrospinal fluid as a zero-reference regularization for brain quantitative susceptibility mapping,' *Journal of Neuroimaging*, vol. 32, no. 1, pp. 141–147, 2022. DOI: https://doi.org/10.1111/jon.12923. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jon.12923. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/jon.12923.

[49] X. Li, D. Jin, Y. Zhu, L. Liu, Y. Qiao, Y. Qian, J. Tian, B. Jiang, C. Hou, J. Geng, X. Li, X. Gao, Y. Ma, S. Wang, J. Zong and Y. Qin, 'Quantitative susceptibility mapping to evaluate brain iron deposition and its correlation with physiological parameters in hypertensive patients,' *Annals of Translational Medicine*, vol. 9, pp. 1582–1582, Oct. 2021. DOI: 10.21037/atm-21-5170.

[50] X. Feng, A. Deistung and J. R. Reichenbach, 'Quantitative susceptibility mapping (qsm) and r2* in the human brain at 3t: Evaluation of intra-scanner repeatability,' *Zeitschrift für Medizinische Physik*, vol. 28, no. 1, pp. 36–48, 2018, ISSN: 0939-3889. DOI: https://doi.org/10.1016/j.zemedi.2017.05.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0939388916301143.

[51] X. Li, R. P. Allen, C. J. Earley, H. Liu, T. E. Cruz, R. A. Edden, P. B. Barker and P. C. van Zijl, 'Brain iron deficiency in idiopathic restless legs syndrome measured by quantitative magnetic susceptibility at 7 tesla,' *Sleep Medicine*, vol. 22, pp. 75–82, 2016, ISSN: 1389-9457. DOI: https://doi.org/10.1016/j.sleep.2016.05.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389945716300478.

[52] S. Wharton, A. Schäfer and R. Bowtell, 'Susceptibility mapping in the human brain using threshold-based k-space division,' *Magnetic Resonance in Medicine*, vol. 63, no. 5, pp. 1292–1304, 2010. DOI: https://doi.org/10.1002/mrm.22334. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.22334. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.22334.

[53] V. Shahmaei, F. Faeghi, A. Mohammadbeigi, H. Hashemi and F. Ashrafi, 'Evaluation of iron deposition in brain basal ganglia of patients with parkinson's disease using quantitative susceptibility mapping,' *European Journal of Radiology Open*, vol. 6, pp. 169–174, 2019, ISSN: 2352-0477. DOI: https://doi.org/10.1016/j.ejro.2019.04.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352047719300206.

[54] L. C, P. L, S. S and D. A, 'Quantitative susceptibility mapping in parkinson's disease,' *PLoS One*, vol. 11, 2016. DOI: 10.1371/journal.pone.0162460. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/27598250/.

[55] D. T. Li, E. S. Hui, Q. Chan, N. Yao, S. Chua, G. M. McAlonan, S. Y. Pang, S. Ho and H. K. Mak, 'Quantitative susceptibility mapping as an indicator of subcortical and limbic iron abnormality in parkinson's disease with dementia,' *NeuroImage: Clinical*, vol. 20, pp. 365–373, 2018, ISSN: 2213-1582. DOI: https://doi.org/10.1016/j.nicl.2018.07.028. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158218302407.

[56] M. Costagli, G. Donatelli, L. Biagi, E. Caldarazzo Ienco, G. Siciliano, M. Tosetti and M. Cosottini, 'Magnetic susceptibility in the deep layers of the primary motor cortex in amyotrophic lateral sclerosis,' *NeuroImage: Clinical*, vol. 12, pp. 965–969, 2016, ISSN: 2213-1582. DOI: https://doi.org/10.1016/j.nicl.2016.04.011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158216300742.

[57] Y. Li, S. K. Sethi, C. Zhang, Y. Miao, K. K. Yerramsetty, V. K. Palutla, S. Gharabaghi, C. Wang, N. He, J. Cheng, F. Yan and E. M. Haacke, 'Iron content in deep gray matter as a function of age using quantitative susceptibility mapping: A multicenter study,' *Frontiers in Neuroscience*, vol. 14, 2021, ISSN: 1662-453X. DOI: 10.3389/fnins.2020.607705. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2020.607705.

[58] C. Tartaglia, P. Lei, J. Carlos, M. Castrillo, N. Pyatigorskaya, C. Sanz-Morère, R. Gaurav, E. Biondetti, R. Valabregue, M. Santin, L. Yahia-Cherif and S. Lehéricy, 'Systematic review iron imaging as a diagnostic tool for parkinson's disease: A systematic review and meta-analysis,' *Frontiers in Neurology*, vol. 11, May 2020. DOI: 10.3389/fneur.2020.00366.

Appendix A

Susceptibility values: Tables

This appendix presents the extracted susceptibility values in tables. Table A.1 reports the mean and median raw susceptibility value extracted from the masks generated with QSM input with SynthSeg. The inter-subject SD is also reported. Table A.2 reports the mean and median susceptibility value extracted from the masks generated with T1w input with SynthSeg from the 29 healthy volunteers.

Table A.1: Mean and median ROI susceptibility value averaged over 29 healthy volunteers. The susceptibility values were extracted from masks generated with SynthSeg from QSM images. The values are presented with the inter-subject SD and are not normalized to a reference.

ROI	**Mean (SD)** (ppb)	**Median (SD)** (ppb)
Left thalamus	2.61 (3.11)	4.41 (3.20)
Right thalamus	2.60 (2.37)	3.73 (2.48)
Left caudate	17.17 (3.87)	18.29 (3.96)
Right caudate	17.90 (3.54)	18.60 (3.87)
Left putamen	9.68 (5.05)	11.67 (4.85)
Right putamen	10.83 (5.42)	11.72 (4.75)
Left pallidum	54.67 (9.34)	54.90 (9.47)
Right pallidum	53.59 (8.91)	54.56 (9.29)
Left hippocampus	-6.39 (6.24)	-6.43 (5.45)
Right hippocampus	-2.75 (3.26)	-2.31 (2.79)
CSF	-0.72 (1.07)	-0.06 (0.15)

Table A.3 presents the mean susceptibility normalized to the CSF with the inter-subject SD for healthy volunteers. As QSM data is relative, it is necessary to use a reference area when comparing values from different studies. However, as the measured mean CSF susceptibility is close to 0 ppb, both the mean and SD is approximately the same as the raw values reported in Figure 4.18. Several studies have found that the susceptibility of subcortical regions vary with age [47]. To investigate the age dependency

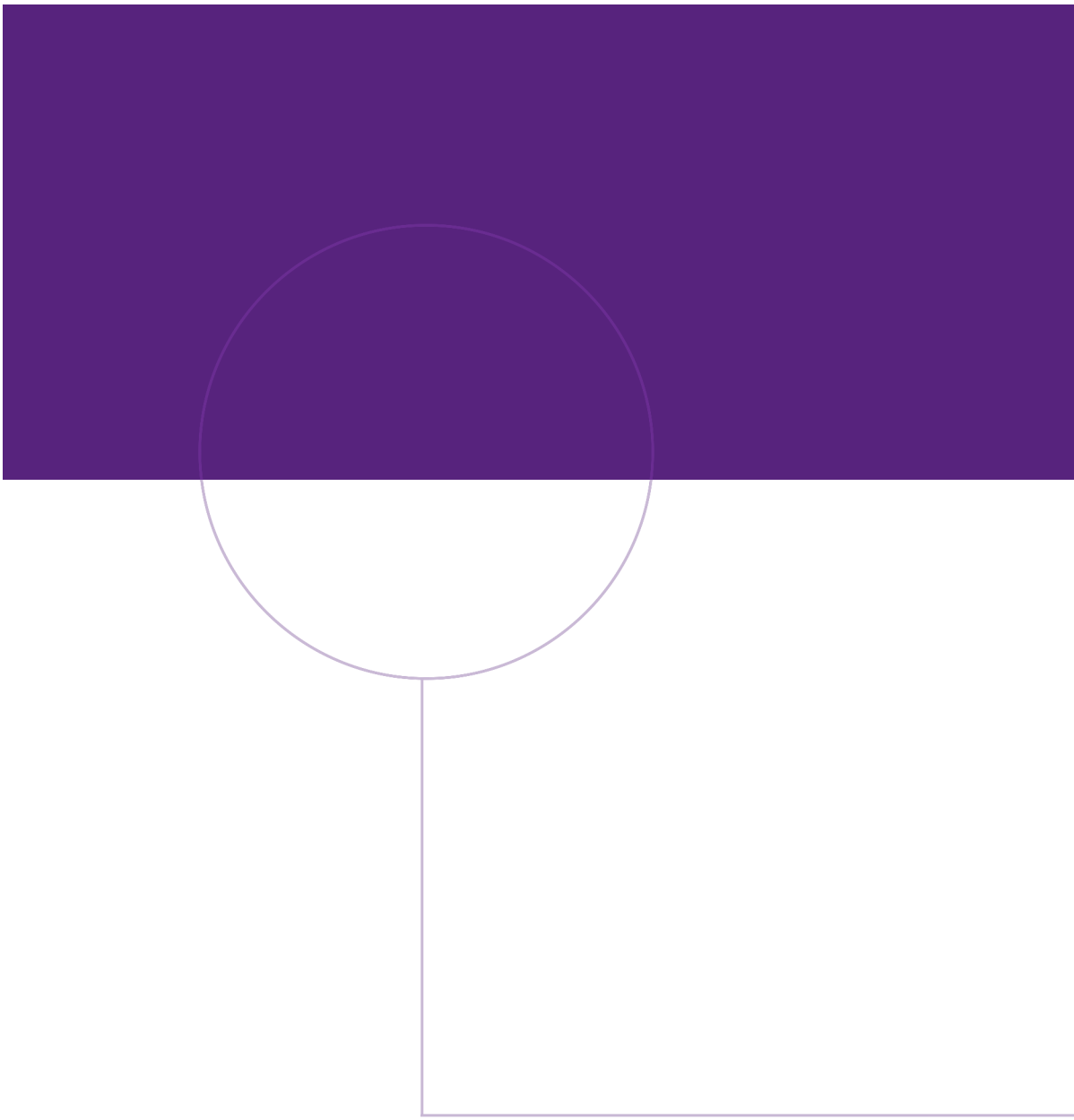
Table A.2: Mean and median ROI susceptibility value averaged over 29 healthy volunteers. The susceptibility values were extracted from masks generated with SynthSeg from T1w images. The values are presented with the inter-subject SD and are not normalized to a reference.

ROI	**Mean (SD)** (ppb)	**Median (SD)** (ppb)
Left thalamus	0.45 (1.80)	1.67 (2.00)
Right thalamus	-0.14 (1.66)	0.77 (1.84)
Left caudate	17.05 (3.19)	17.73 (3.71)
Right caudate	16.74 (3.42)	17.55 (3.80)
Left putamen	5.23 (3.18)	4.76 (3.45)
Right putamen	4.95 (3.41)	4.60 (3.58)
Left pallidum	53.86 (10.52)	55.99 (10.47)
Right pallidum	51.87 (9.35)	55.44 (9.72)
Left hippocampus	1.60 (2.41)	1.86 (1.99)
Right hippocampus	2.16 (2.02)	2.49 (1.94)
CSF	0.96 (0.55)	0.00 (0.00)

in this study, Table A.3 reports the CSF normalized susceptibility values divided in the age groups 20-41, 20-30 and 31-41. All of the ROIs display an increase of susceptibility for the older cohort, particularly the left caudate.

Table A.3: Mean ROI susceptibility value for age groups of 20-41, 20-30 and 31-41 and inter-subject standard deviation of ROIs. The susceptibility values are normalized to the CSF and were extracted from masks generated from T1 images with SynthSeg.

	Age 20-41 N = 29 **Mean (SD)**	**Age 20-30** N = 19 **Mean (SD)**	**Age 31-41** N = 10 **Mean (SD)**
ROI	(ppb)	(ppb)	(ppb)
Left thalamus	-0.51 (1.89)	-0.85 (1.42)	0.12 (2.44)
Right thalamus	-1.11 (1.76)	-1.33 (1.62)	-0.69 (1.92)
Left caudate	16.09 (3.38	15.39 (3.59)	17.42 (2.45)
Right caudate	15.78 (3.63)	15.44 (3.96)	16.42 (2.76)
Left putamen	4.27 (3.38)	3.78 (3.48)	5.20 (2.99)
Right putamen	3.99 (3.60)	3.50 (3.58)	4.90 (3.48)
Left pallidum	52.90 (10.53)	51.74 (10.20)	55.08 (10.81)
Right pallidum	50.90 (9.32)	50.15 (9.48)	52.34 (8.85)
Left hippocampus	0.64 (2.44)	0.14 (2.22)	1.60 (2.55)
Right hippocampus	1.20 (2.00)	0.90 (2.20)	1.76 (1.37)



NTNU

Norwegian University of
Science and Technology