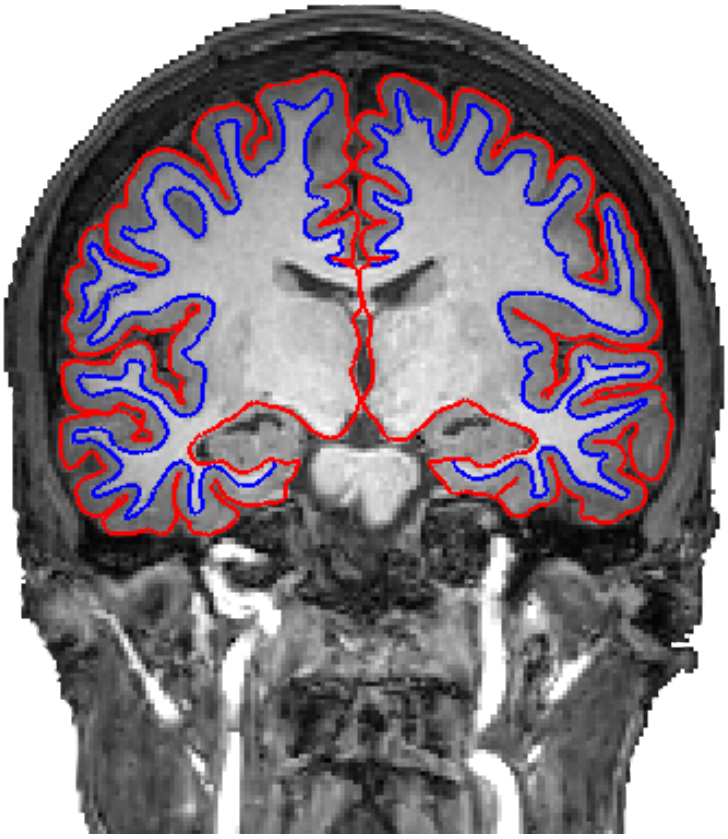


Brain Volume and Cortical Thickness in Type 2 Diabetes:

Software Implementation and Comparative Analysis

Master's thesis in Engineering Mathematics and Computational Science

ISAC STARK

DEPARTMENT OF MATHEMATICAL SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

Master's thesis 2024

Brain Volume and Cortical Thickness in Type 2 Diabetes:

Software Implementation and Comparative Analysis

ISAC STARK

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mathematical Sciences
Chalmers University of Technology
Gothenburg, Sweden 2024

Brain Volume and Cortical Thickness in Type 2 Diabetes:
Software Implementation and Comparative Analysis
ISAC STARK

Supervisor: Martin Schain and Marcus Wilander Björk, Antaros Medical AB
Examiner: Torbjörn Lundh, Department of Mathematical Sciences

Master's Thesis 2024
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: The figure shows one coronal slice of a T1-weighted MRI processed with FreeSurfer.

Typeset in LaTeX
Gothenburg, Sweden 2024

Brain Volume and Cortical Thickness in Type 2 Diabetes:
Software Implementation and Comparative Analysis
ISAC STARK
Department of Mathematical Sciences
Chalmers University of Technology

Abstract

This thesis aimed to investigate the effects of Type 2 Diabetes Mellitus (T2DM) on brain volumetrics, specifically in terms of changes in cortical thickness and grey matter volume. Using the surface based morphometry software FreeSurfer and FastSurfer, a comparative analysis was conducted to evaluate the reliability, effectiveness, and runtime of these software in assessing brain morphology in individuals with T2DM compared to healthy controls.

MRI data from two datasets, OASIS and MIND, were processed and analysed, focusing on regions of interest included in the Desikan–Killiany–Tourville atlas using both software. Intra-software reliability was assessed through three metrics, Pearson Correlation Coefficient, Intraclass Correlation Coefficient, and Test-Retest Variability. This was done to determine if both software are consistent in their estimates of brain volumetrics. The inter-software reliability was assessed through the Intraclass Correlation Coefficient. The inter-software reliability was performed to determine if FastSurfer gives a similar result to FreeSurfer, which is more extensively validated. Additionally, the runtime performance of FreeSurfer and FastSurfer was compared to determine their efficiency.

The results demonstrated high intra-software reliability for both FreeSurfer and FastSurfer in measuring brain volumes and cortical thickness. The two software also demonstrated high inter-software reliability, demonstrating that FastSurfer has a similar accuracy to FreeSurfer. FastSurfer also exhibited a significantly faster runtime. All of this combined highlights FastSurfer's potential for large-scale studies and clinical applications. However, contrary to expectations based on prior literature, significant differences in brain volumes between the T2DM group and healthy controls were not found.

In conclusion, while this study validates the use of FreeSurfer and FastSurfer for neuroimaging analysis, it also highlights the complexity of detecting brain volume changes associated with T2DM, pointing towards the necessity for further investigations. The improved runtime, as well as the high intra- and inter-software reliability of FastSurfer suggests it as a preferable software for this purpose.

Keywords: FreeSurfer, FastSurfer, Surface Based Morphometry, Type 2 Diabetes Mellitus, Alzheimer's disease

Acknowledgements

I want to start by thanking my two supervisors Martin Schain and Marcus Wilander Björk, whose guidance and support made this thesis a reality. I don't think one could hope for more involved supervisors and I've truly enjoyed our discussions during these last six months. I also want to thank the entirety of Antaros Medical and Carl Sjöberg in particular for providing me with the incredible opportunity to undertake this thesis project and for being so welcoming.

I would also like to take this moment to thank all of my dear friends, whose support has been a constant source of joy throughout this spring. A special note of appreciation goes to my friend Gabriel, for both proverbially and digitally lending me his brain during this thesis.

Also, thanks to my examiner Torbjörn Lundh for the help along the way.

Lastly, I want to thank my dear friend Adrian for all the memories over the years. Cheers.

The data used in this thesis were in part provided part by *The Open Access Series of Imaging Studies* (OASIS) [1].

Isac Stark, Gothenburg, June 2024

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

AD	Alzheimer's Disease
Aβ	Amyloid-β
ANCOVA	Analysis Of Covariance
CDR	Clinical Dementia Rating
CNN	Convolutional Neural Network
CSF	Cerebrospinal Fluid
CT	Cortical Thickness
DKT	Desikan-Killiany-Tourville
GMV	Grey Matter Volume
HC	Healthy Control
ICC	Intraclass Correlation Coefficient
MRI	Magnetic Resonance Imaging
MMSE	Mini-Mental State Examination
NFT	Neurofibrillary tangle
OASIS	Open Access Series of Imaging Studies
PCC	Pearson Correlation Coefficient
SA	Surface Area
SBM	Surface based morphometry
T2DM	Type 2 Diabetes Mellitus
TRV	Test-Retest Variability

Contents

List of Figures

List of Tables

1

Introduction

Obesity is a medical condition defined as the accumulation of excessive body fat, a condition closely related to other negative health conditions such as *Type 2 Diabetes mellitus* (T2DM) and cardiovascular diseases [2]. In turn, T2DM emerges as a notable risk factor for the development of neurodegenerative disorders and dementia later in life, yet the precise biological mechanisms that connects T2DM with neurodegenerative disorders remain unknown [3]. Due to the increased prevalence of both obesity and T2DM globally [4, 5], as well as a general increase in life expectancy, the prevalence of dementia disorders is expected to undergo a significant increase in coming decades [6]. This underscores the importance of increasing our understanding of how metabolic disorders contribute to neurodegeneration, and to identify biomarkers that can be used in intervention trials that aim to slow down or revert the disease progression.

Neurodegenerative disorders, such as *Alzheimer's Disease* (AD), involves structural changes in brain tissue [7], most notably cortical atrophy [8]. However, it remains unknown if T2DM is associated with similar structural changes in the brain, a critical consideration for assessing if a medication developed for, e.g., T2DM also influences the probability of developing neurodegenerative disorders [9, 10].

In order to characterise the brain morphometry of living people, there exists various methods utilising magnetic resonance imaging (MRI). One such method is that of surface-based morphometry (SBM). Surface-based morphometry typically uses high-resolution MRI with good contrast between white- and grey-matter to reconstruct the surface of the brain. In a preprocessing step, the intensity of the image is normalised and irrelevant tissue, such as the skull and eyes is removed [11]. Afterwards, the white matter is segmented and a white matter-surface is reconstructed. By expanding the white matter-surface outwards, the pial surface is reconstructed. The pial surface is the part where the cortex and the cerebrospinal fluid (CSF) meet. Finally, anatomical measures such as the cortical thickness and surface area can be calculated [12]. To account for inter-subject differences in anatomy when comparing the derived anatomical measures, subjects are compared after inflating the reconstructed surface measures to a unit sphere. The inflated surface is then mapped back onto a template brain on which inter-subject differences can be compared.

One commonly used SBM software is *FreeSurfer* [13]. However, the estimated time to reconstruct the cortex of a single subject in FreeSurfer is between 6 to 24 hours depending on the computer. This makes the software unsuitable for large datasets

without the proper infrastructure, such as access to a high-performance computing cluster. A recently developed software called *FastSurfer* aims to solve this issue by combining some of the methods included in FreeSurfer [13] with modern deep-learning techniques [14]. FastSurer has been shown to reduce the time for cortical reconstruction to less than one hour. However, the software is not as extensively tested as FreeSurfer, and requires further testing before it can be used be used in future drug intervention trials.

1.1 Aims

The objective of this master's thesis is to assess whether anatomical differences of the brain manifest in patients with T2DM and overweight individuals without a T2DM diagnosis, in comparison to healthy subjects. This will be done by leveraging MRI data obtained from Antaros Medical. Further, the thesis aims to validate and compare two different software that estimate brain morphometric measurements.

1.2 Scope and Limitations

The assessment of the potential difference in brain anatomy of T2DM patients will be limited to anatomical measures possible to derive from the surface-based morphometry methods FreeSurfer and FastSurfer. Further, to assess whether the potential difference even is reliable, the software will be validated in two ways. The first one is to test if the software themselves are reliable. That is, do they produce consistent anatomical estimates from input that in theory should produce identical estimates. For this, the comparison will be limited to the four anatomical measurements cortical thickness, grey matter volume, cortical surface area, and subcortical volumes. The second way the software will be validated is to test if they can detect a known difference in brain anatomy in a test dataset. This dataset will consist of patients with a clinical AD diagnosis and healthy controls. The statistical testing performed, both for assessing differences in AD patients and T2DM patients, will be limited to cortical thickness, grey matter volume, and subcortical volume.

Previous studies have been done on the topic of volumetric changes of the brain between T2DM, obese, and healthy controls [9]. The volumetric assessments performed in this thesis will therefore aim to replicate these previous findings in a new dataset. In addition, volumetric assessments in this dataset will enable a cross-comparison between structural changes and to metabolic changes, as metabolic imaging data have been acquired with positron emission tomography in the same individuals, thereby extending our understanding on how metabolic alterations in the periphery translates to brain pathology. Lastly, one important aspect of this work is to investigate if FastSurfer is a viable option to FreeSurfer in terms of improved processing time without disregarding sensitivity and reliability. Something that has not been tested in this setting to the best of the author's knowledge.

2
Theory

In this chapter the necessary theoretical background to understand this project will be presented. First, the necessary background to the anatomy of the brain and the pathology of neurodegenerative diseases will be presented. Afterwards, the basics of MRI and how these medical images can be used to estimate anatomical measures of the brain will be explained. Lastly, the necessary statistics to evaluate the result of the derived anatomic measures will be presented.

2.1 Anatomy of the Human Brain

The human brain is the bodily organ that is responsible for all human actions, both voluntary and involuntary [15]. The brain is commonly divided into four principle components: the brain steam, diencephalon, cerebrum, and the cerebellum. Involuntary actions, such as respiration is commonly attributed to the brain steam. Higher order functions on the other hand, such as memory and personality is usually attributed the cerebrum [16]. For this work, the cerebrum and related structures, such as the subcortical structures will be considered.

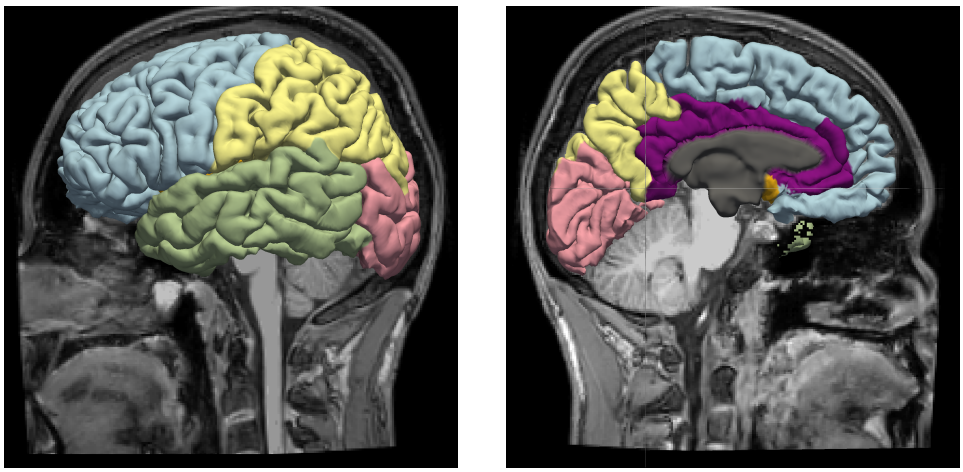
2.1.1 Anatomical Structure

In humans, the cerebrum is the largest part of the brain and accounts for approximately 80% of its weight [17]. The cerebrum consists primarily of two types of matter. The grey- and white matter. Grey matter is found in the outermost layer of the cerebrum, the *cerebral cortex*, as well as in the so called *basal ganglia*. The total space occupied by the brain, along with the surrounding structures that coexist within this cranial cavity, is collectively referred to as the intracranial volume (ICV) [18].

The cerebral cortex is only a few millimetres thick, on average, 2.5-3 mm [19]. During development the cerebral cortex is 'folded', which allows for a much larger surface area relative to the cranial volume. This is due to a process called *gyrification*. The peaks of these folds are called *gyri* (sing. Gyrus) and the valleys are called *sulci* (sing. Sulcus) [15]. Although the cortical thickness is quite consistent between individuals, regardless of intracranial volume, features such as grey matter volume and cortical surface area is highly correlated to the ICV [20]. Furthermore, both cortical thickness and grey matter volume have shown to be negatively correlated with age [21] and gender, with males generally having larger volumes of grey matter.

The white matter lies beneath the cerebral cortex [15]. This division between the grey and white matter is based on cellular composition. The grey matter is mainly composed of neuronal cell bodies [22], while the white matter is mainly composed of myelinated axons which are extensions of the neuronal cell bodies. [23]. Myelinated means that the axon is covered in a fatty material called myelin which helps transmit electrical impulses [24].

Besides this structural division, the cerebrum is furthered split into different regions. First, the cerebrum can be divided into two somewhat symmetrical halves called the left and right hemisphere. The hemispheres are separated by the cerebral fissure. The hemispheres are joined by a large white matter structure called the corpus callosum. Furthermore, the hemispheres are commonly divided into different 'lobes'. This division is usually made based on which cranial bone that covers the cerebrum [15]. The four lobes named after their corresponding cranial bone are the frontal, parietal, temporal and occipital lobes, see figure 2.1. Besides these four lobes, there is the *limbic* lobe. The limbic lobe is a ring shaped grey matter structure closest to the human 'midline' of each hemisphere [25].



(a) Sagittal view of the brain lobes

(b) Sagittal view of the midline of the brain lobes

Figure 2.1: The figure displays the five lobes of the brain, each shown in a different colour from two different viewpoints. The frontal lobe in blue, the temporal lobe in green, the parietal lobe in yellow, the occipital lobe in red, and the limbic lobe in purple. The limbic lobe is only visible in subfigure (**b**).

Another vital part of the anatomy of the brain are the broadly grouped subcortical structures. The subcortical structures include the hippocampus, amygdala, and basal ganglia, among others. When discussing the human ability to remember, both the hippocampus and amygdala are two very important structures. The hippocampus has shown to be highly related to both human learning, but also the human ability to remember [26]. The ventricular system, although not considered a subcortical structure, is another component located beneath and around the cerebral

cortex [27]. This system consists of a series of interconnected cavities within the brain that are filled with cerebrospinal fluid (CSF). CSF is also found surrounding the outside of the brain.

2.1.2 Pathology of Neurodegenerative Diseases

Neurodegenerative diseases includes a wide range of disease conditions that share the common characteristic of gradual loss of neuronal cell bodies. There exists multiple ways to classify neurodegenerative diseases, for example, based on the presence of abnormal protein aggregation, called proteinopathy [28]. Arguably, one of the most well-known, and also one of the most commonly occurring of these neurodegenerative diseases is Alzheimer's disease (AD) [29].

Clinically, AD is characterised by progressive loss of memory and a general cognitive decline [30]. While not unique for AD, common characteristics on the macroscopic scale are various degrees of regional cortical atrophy and enlargement of the lateral ventricles [30, 31]. Cortical atrophy often occurs in, but is not limited to, regions in the temporal, parietal and frontal lobe [31]. One specific region of the temporal lobe that is of interest in AD is the entorhinal cortex, which often is the first region of the brain to show signs of atrophy [32]. Atrophy is also prominent in subcortical structures. Most notably atrophy appears in the amygdala [33] and the hippocampus [34].

Seeing as cortical atrophy and ventricle enlargement is not unique for AD. AD is instead defined on the microscopic scale by the presence of amyloid plaques and neurofibrillary tangles (NFTs) [28]. Amyloid plaques are aggregates of the protein amyloid-β (Aβ) that have formed within the brain, and are believed to cause inflammation in the brain which in the long term affects the brain cells [35, 36]. Aβ are peptide chains consisting of around 40$\sim$ amino acids. The neurofibrillary tangles form when a different protein called *tau protein* form lumps in a neuron cell [37]. With progressing accumulation of NFTs, the host neuron cell eventually dies, leading to increased atrophy. Both of these pathologies have been showed to occur more often in people carrying the genetic variant APOE ε4 [38, 39].

2.1.3 Relation to Type 2 Diabetes Mellitus

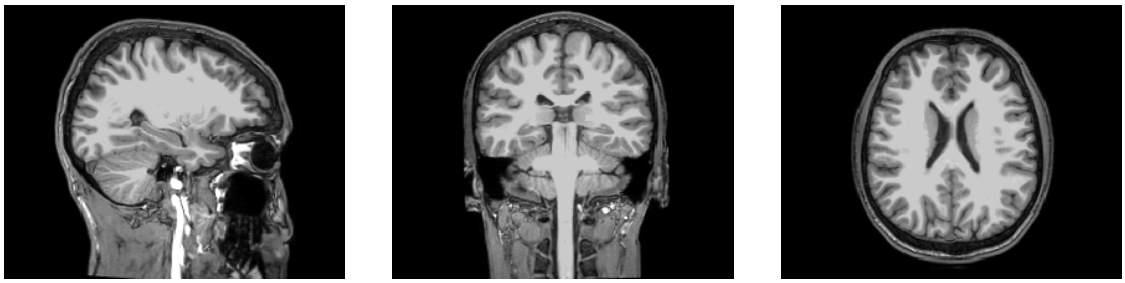
It has been shown that having T2DM increases the risk of getting AD later in life [40, 41]. These two conditions are theorised to be linked by various genetic and environmental factors. However, there exists multiple probable theories as to how exactly [42]. One possible mechanism is that the hyperinsulinemia (Increased insulin levels in the blood stream) caused by T2DM decreases the body's ability to remove accumulating amyloid plaques. Further, the same hyperinsulinemia may also promote the formation of neurofibrillary tangles [43]. There also exists evidence the same underlying reasons, namely increased amyloid plaques and increased neurofibrillary tangles are instead caused by insulin resistance in the brain [44].

2.2 Magnetic Resonance Imaging

One method to detect and characterise the volumetric changes in the brain caused by neurodegenerative disease is through *magnetic resonance imaging* (MRI). MRI is a non-invasive medical imaging technique utilising strong magnets. In short, conventional MRI works by aligning the hydrogen atoms in the body with the magnetic field produced by the magnet. The atoms are then excited via an induced radiofrequency pulse. When the atoms return to equilibrium after excitation a signal can be read which is then used to produce an image. Depending on how these pulses are applied, such as the time between the pulses and how the resulting signal is read out, different tissue contrasts can be achieved [45]. When using MRI to detect structural changes in the brain, it is common to use so called T1-weighted MRI [46, 47], as it provides good contrast between the white and grey matter, as well as the CSF. Clinically, MRI scans are commonly viewed as a collection of individual "slices" as illustrated in the Figure 2.2. This is referred to as a *tomographic* image. These slices can be viewed from different anatomical planes. There are three principal anatomical planes: sagittal, which divides the body into left and right halves; coronal, which divides the body into front and back sections; and transversal, which divides the body from top to bottom. An example of an MRI scan showcasing these three principal views can be seen in Figure 2.3.

Figure 2.2: Illustration of multiple slices from an MRI viewed coronally.

An example MRI showcasing the three principal views can be seen in figure 2.3 below. In practice, the magnetic field used in MRI is not perfectly homogeneous, which can lead to artefacts in the resulting images. This inhomogeneity can occur due to various reasons [48]. One reason is that certain parts of the body will naturally be located farther away from the centre of the magnetic coil, where the magnetic field is weaker. Additionally, different bodily tissues may have varying susceptibilities to the magnetic field, contributing to the inhomogeneity. The artefacts introduced by the MRI inhomogeneity often give rise to smooth and low-signal artefacts sometimes



(a) Sagittal view **(b)** Coronal view **(c)** Transversal view

Figure 2.3: Figure showing single slices of an MRI along the three principal anatomical planes: (a) Sagittal, (b) coronal and (c) transversal

rererred to as bias fields [49]. Bias fields can introduce challenges when attempting to process the MRI images digitally, if the intensity information of the MRI is used in the processing. However, there are algorithms developed to correct for these artefacts [49].

2.3 Brain Atlases

The aforementioned way of partitioning the brain into lobes and hemisphere is a quite rough approach, and typically not sufficiently detailed when comparing structural changes in the brain across individuals [50]. To address individual differences more accurately brain atlases can be created and used. A brain atlas is a template that, for example, specifies the boundaries between various brain structures in a standard space. This template can then be used to compare volumes and thicknesses of various brain structures across individuals. The initial brain atlases, including the well-known *Brodmann areas* [51], were developed based on the study of cellular structures observed in histological samples. Today, a plethora of brain atlases exists, designed in different ways and for different applications [52].

2.3.1 Desikan–Killiany–Tourville Atlas

One such atlas is the *Desikan–Killiany–Tourville* (DKT) atlas [53]. The DKT atlas is based on a previously developed atlas commonly referred to as the Desikan-Killiany (DK) atlas [54]. The DKT atlas was constructed using MRI data from 101 healthy individuals. The protocol used to label the cortical regions realises on information about the depth and curvature of the sulci. By mapping the brain to a common atlas like the DKT atlas it becomes easier to compare cortical regions between individuals, regardless of minor differences in regional anatomy.

In total, the DKT atlas contains 31 cortical regions of interests for each hemisphere. The full list of the cortical regions in the DKT atlas can be found in Table A.1 in Appendix A.1.

2.3.2 MNI305 Atlas

An other type of atlas is the MNI305 atlas. It is different from the DKT atlas in the sense that instead of dividing the cortex into different regions of interests, the MNI305 serves as a kind of coordinate system. The MNI305 was created by registering the brain MRI scans from 305 individuals to a template [55]. Image registration is when one image is transformed, or mapped onto another via a set of allowed operations so they occupy the same space after the transformation. The 305 registered scans were then averaged to create a template to which new scans can be registered.

By registering to the MNI305 atlas, we get a common space in which the same parts of the human brain from different individuals have roughly the same coordinates.

2.4 Surface Based Morphometry

Surface based morphometry (SBM) is a widely used class of methods in neuroimaging that estimates brain morphometric values from T1-weighted MRI of the brain. These methods have found a wide range of applications within research into neurological conditions such as Alzheimer's disease, autism spectrum disorders, and schizophrenia, among others.[56–59]. In short, SBMs work by reconstructing a mesh of the white matter surface of the brain as well as the '*pial*' surface, which is defined as the border between the grey matter and the CSF. These two surfaces then serve as the basis for calculating various morphometric features, such as cortical thickness and curvature, amongst others.

2.4.1 FreeSurfer

One common SBM software is FreeSurfer. FreeSurfer provides a complete pipeline for analysing structural MRI, meaning that it handles everything from preprocessing to calculating the morphometric features of interest. An overview of the FreeSurfer pipeline can be seen in Figure 2.4.

In total, the runtime of the entire pipeline can take up to several hours for a single subject, depending on the computer used in the processing.

2.4.1.1 Surface Reconstruction

In FreeSurfer, the mesh construction of the white and pial surface is done in seven steps. Using a T1-weighted MRI as input, the first step is to register the MRI to the MNI305 atlas. The coordinates of certain anatomical structures in the MNI305 atlas serves as basis for many of the later calculations [60]. As outlined in section 2.3.2, the MNI305 atlas serves as a coordinate system. This is used to generate initial 'seed points' as to where some cerebral structures are located. The registration of the input to the MNI305 space is done via an affine transformation. An affine transformation allows for a wide range of operations such as shear, translation, rotation and scaling [61].

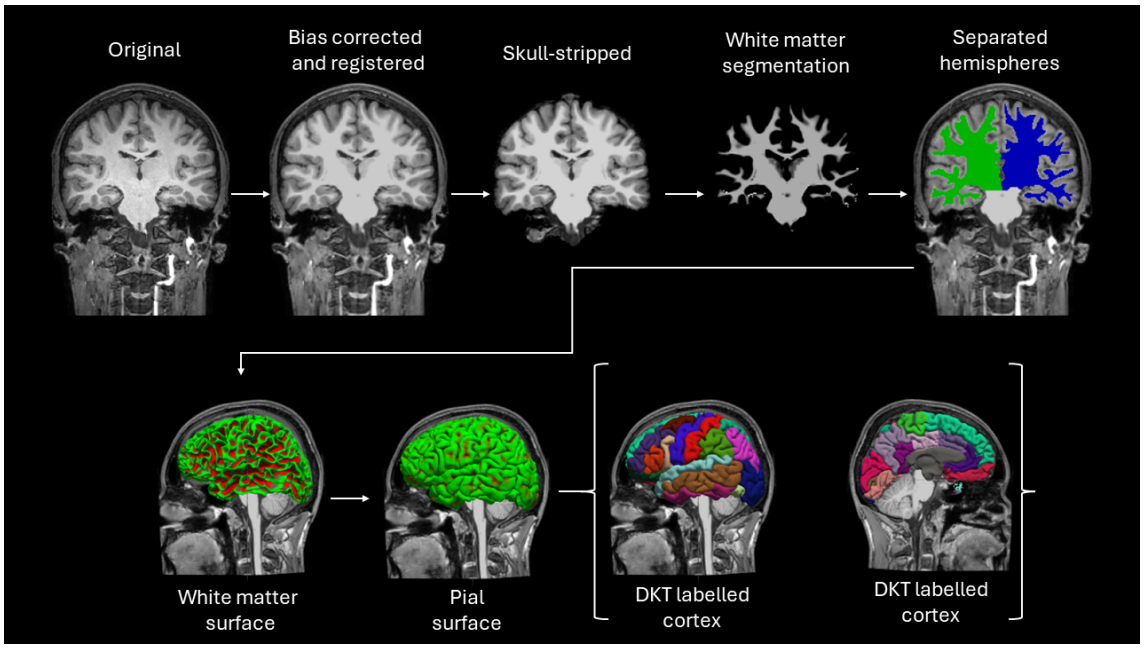


Figure 2.4: Schematic overview of the complete FreeSurfer surface reconstruction pipeline as outlined in Section 2.4.1.1, and the cortical labelling as outlined in Section 2.4.1.2

The second step in the FreeSurfer pipeline is to attempt to correct for the artefacts that can arise due to the magnetic field inhomogeneity mentioned in Section 2.2. The algorithm used to correct for the bias field can be read in full in the original paper by Dale, Fischl and Sereno [60]. In short, it utilises the fact that the input is a T1-weighted MRI brain scan, in which white matter should have the highest intensity (brightest colour). This, combined with the fact that white matter neighbouring each other should in theory have equal or at least similar intensities, is used to construct the bias field correction. By dividing the intensities in the MRI with the estimated bias field correction should ideally remove the artefacts. During the bias field correction, the intensities of the MRI is normalised to be in the range $[0, 255]$.

After correcting for the bias field, the third step in the FreeSurfer pipeline is to remove the skull so that only brain tissue remains. This process is accomplished by constructing a mesh of an ellipsoid and then updating the coordinate vector $\mathbf{x}_k$ of the nodes within the mesh according to

$$\mathbf{x}_k(t+1) = \mathbf{x}_k(t) + \mathbf{F}_S(t) + \mathbf{F}_M(t). \tag{2.1}$$

The term $\mathbf{F}_S$ in Equation (2.1) represents a smoothness force that incorporates the prior knowledge that the inner surface of the skull is a relatively smooth surface. The second force applied to the deformation of the ellipsoid is $\mathbf{F}_M$. This represents a MRI-based force that deforms the ellipsoid towards areas of low intensity values. As previously mentioned, white matter correspond to high intensity values in T1-weighted MRI, while CSF have low intensity values. The force $\mathbf{F}_M$ thus repels the

ellipsoid away from the brain. After these two forces have been iteratively applied to the mesh, everything outside the boundary of the mesh is removed. Ideally, this removes everything except the brain.

The fourth step is to use the skull-stripped brain to classify the tissue as either white matter or non white matter. This is done by using the normalised intensity information from the bias field correction and the output of the skull-stripping, I_0. The range on intensities in which we can find the different types of brain matter, white or grey are treated as fixed parameters. The lowest and highest expected intensity of the white matter is defined as $\text{WM\_LOW} = 90, \text{WM\_HIGH} = 140$ and the highest expected intensity of the grey matter is defined $\text{GREY\_HIGH} = 100$. Parts of the classification will be ambiguous as there is an overlap between the distribution of white- and grey matter. The first step in this process is to do a rough classification based solely on these expected intensities according to

$$I_1(\mathbf{x}) = \begin{cases} 1, & I_0(\mathbf{x}) \in [\text{WM\_LOW}, \text{WM\_HIGH}] \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

After the brain has been segmented into white matter and non-white matter the next step is to refine the estimate. This is done by creating three different sets of voxels that are ambiguous. A voxel can either be ambiguous due to it existing in the overlap between the expected white and grey matter intensities. It can also be ambiguous if it's classification is different from more than 20% of it's neighbours. Finally, it can be ambiguous if more than 60% of voxels lying in a plane minimising intensity variance have a different classification.

If a voxel exists in *all three* of these sets, the initial classification, $I_1(\mathbf{x})$ is inverted for these voxels. This constitutes the final classification of white matter.

The next two steps in the FreeSurfer is to divide the classified white matter from the previous steps into two hemispheres and then make sure that these two hemispheres constitute two fully connected regions. The division of the hemispheres is accomplished by utilising information obtained from the initial registration to the MNI305 atlas. The coordinates corresponding to the corpus callosum and the pons, a part of the brainstem, serve as an initial approximation for determining how to partition the white matter into the two hemispheres.

The final step in the FreeSurfer surface reconstruction pipeline is to use the newly created connected hemisphere volumes to construct the pial surface that was mentioned in section 2.4. This is done by first creating a discrete surface from the hemispherical volumes. This discretised surface consists of a triangular mesh. The triangular mesh is then refined by using the intensity difference between the white matter and grey matter. This refinement gives us the final white matter surface. Finally, this newly created surface is pushed outwards, guided by the intensity gradient were the grey matter meets the CSF. This gives us the pial surface.

The minimum distances between a vertex in the white matter surface and the pial

surface gives us the cortical thickness.

2.4.1.2 Cortical and Subcortical Labelling

The white matter and pial surfaces constructed by FreeSurfer serves as the basis for mapping the anatomical measures to an atlas, such as the DKT atlas mentioned in section 2.3.1. When the cortical surface is mapped to an atlas, this is called parcellation. However, the subcortical structures of the brain can also be mapped, and then it is called segmentation.

These two types of labelling uses the same algorithm, with minor differences [62, 63]. The first step of the labelling process is to register to a common coordinate system as intermediate. For the subcortical labelling the bias field corrected MRI is registered to previously mentioned MNI305 atlas is used. However, for the cortical parcellation the recreated surfaces are registered to a sphere. As the process to register to the MNI305 atlas has been brought up in section 2.4.1.1, this will focus on the spherical registration.

The spherical registration used by FreeSurfer to label the brain uses the information obtained from cortical reconstruction in section 2.4.1.1. As outlined in section 2.1.1, the cerebral cortex consists of peaks (gyri) and valleys (sulci). The topological information about these anatomical structures, such as their convexity is used for the spherical registration. More precise, the registration is done by minimising the total differences between the convexity of a specific subject, C compared to the average convexity of a set of subjects $\overline{C}$ [64].

After the registration to the common space, the labelling is done via a Bayesian algorithm. This is done by using two different kind of prior information. The first prior uses the information obtained from the spherical registration to the atlas and encodes the probability that each region of interest in the atlas exists at a given vertex. The second prior makes use of the regions of interests spatial relation to each other. That is, a vertex can only belong to a region of interest that makes sense anatomically given all the neighbouring vertices' region of interest [63].

2.4.2 FastSurfer

Another surface based morphometry software is FastSurfer which aims to replicate results from the FreeSurfer pipeline, with reduced runtime. FastSurfer is a hybrid method, combining traditional image analysis methods, as used in FreeSurfer, with modern deep learning methods [65]. The FastSurfer pipeline uses many of the same functions as developed by Fischl in FreeSurfer [60]. However, two primary novel methods are attributed to the reduced runtime and they will be covered in brief in this section. The full details of the method developed by Henschel et al can found in their original paper [65]. The first novel method compared to the FreeSurfer pipeline is the use of a convolutional neural network (CNN) to achieve the brain segmentation. In practice, the FastSurfer-CNN consists of independent CNNs trained separately on the three anatomical views sagittal, coronal and transversal.

The segmentation of brain tissue created by the CNN also removes the need for the iterative skull-stripping algorithm in the third step of the FreeSurfer pipeline. This is due to the fact that the accurate segmentation created by the CNN is used as the basis for the skull-stripping.

The second novel method introduced in FastSurfer is the use of a method called spectral embedding to perform the spherical registration used in the cortical labelling, see section 2.4.1.2. The full details of the inner workings spectral embedding is beyond the scope of this thesis, but the essence of the process is that it once again removes something that is an iterative process in FreeSurfer and replaces it with a single step.

2.5 Similarity Metrics

In an ideal scenario, when an individual undergoes multiple MRI scans within a short timeframe, the morphometric measurements estimated from these scans should yield identical or near-identical results. The underlying assumption is that the individual's anatomical structures, such as brain regions, remain unchanged between the scans. In reality, this is rarely the case. The cause of this can vary [66]. For example, the head tilt of the patient while in the scanner, the type of scanner as well as the type of coil used can have an impact on the end result. There also exists random noise in the scans [67]. Thus, when comparing different methods for deriving morphometric measures reliability is of utmost importance.

To assess the reliability of the different SBM software within themselves, a common approach is to re-scan the same patient twice and compare the output. This kind of reliability testing is usually called "test-retest" reliability. The variance in these repeated measurements will then reflect the uncertainty of a single measurement, and is a useful metric when assessing if a software is resistant to noise and consistent in the measurements it reports.

Test-rest reliability can also be used to determine if two or more different raters are consistent between each other [68].

3
Methods

In this chapter, the methodology employed in this thesis is outlined. The study utilised two datasets - one from an open access neuroimaging project called OASIS [1], and one propriety dataset called MIND provided through the courtesy of Antaros Medical [69]. Both datasets contain MRI scans of in-vivo human brains. The specific details regarding the acquisition of the MRI data and the demographic characteristics of the sample is provided in Section 3.1. Additionally, the analytical workflow used to process the data is presented, outlining the step-by-step procedures in Section 3.2. Finally, the statistical methods leveraged to analyse the reliability of the data and test the hypotheses that T2DM has a significant impact on brain atrophy are described in Sections 3.4 and 3.5.

3.1 Data Used

The specific details of the two aforementioned datasets are provided in this section.

3.1.1 OASIS

The Open Access Series of Imaging Studies (OASIS) is an open access project containing multiple T1-weighted brain MRI datasets. For this thesis the OASIS dataset was used. The OASIS dataset is a cross-sectional collection of brain MRI scans from right-handed individuals of various ages, including both men and women. The dataset comprises a total of 416 unique subjects. Among these, 100 individuals have been diagnosed with varying degrees of dementia, ranging from very mild to moderate. The diagnosis of dementia was determined using the Clinical Dementia Rating (CDR) protocol [70]. Additionally, 20 young, non-demented patients were scanned twice for test-retest purposes.

The OASIS MRI scans used were acquired using a 1.5T Siemens scanner with a voxel size of $1 \times 1 \times 1.25$ mm^3.

Due to time constraints processing all patients was not feasible. Therefore, a subset of the OASIS data was carefully selected to include all patients suffering from mild to moderate dementia, as well as healthy controls spanning all age groups. Furthermore, all of the test-retest scans were also included in the processed subset.

The purpose of using the OASIS dataset was two fold. The first was to test whether

the FreeSurfer and FastSurfer software could detect differences in brain morphometry between healthy controls (HC) and demented patients, as a fairly large difference is expected between patients with AD and HC. This was done to verify that the software implementations work as expected.The second purpose was to evaluate the reliability of the two software programs within themselves. For this purpose, the 20 patients who underwent repeated scans were used.

The demographic details of the processed subset are presented in table 3.1. In total, the dataset included 136 MRI scans, of which 20 were test-retest scans.

Table 3.1: Demographic of the processed subjects from the OASIS dataset

		Without Dementia			*With Dementia*		
Age Group	*Total patients*	*Patients*	*Male*	*Female*	*Patients*	*Male*	*Female*
<20	4	4	1	3	0	0	0
20s	37	37	14	23	0	0	0
30s	5	5	4	1	0	0	0
40s	8	8	2	6	0	0	0
50s	5	5	1	4	0	0	0
60s	12	9	1	8	3	0	3
70s	24	8	4	4	16	6	10
80s	19	9	4	5	10	4	6
90s	2	1	1	0	1	0	1
Total	**116**	**86**	**32**	**54**	**30**	**10**	**20**

3.1.2 MIND

MIND is a cross-sectional dataset proprietary to Antaros Medical, containing T1-weighted brain MRI scans from both lean and overweight non-diabetic patients as well as patients with T2DM. In total, the dataset consists of 38 individuals aged between 50 and 67. All of the scans were acquired on a 3T Siemens PET/MR scanner, with a voxel size of $0.5 \times 0.5 \times 1$ mm^3. The demographics of the patients in the MIND dataset, including disease status, can be seen in Table 3.2. All patients proceeded with a follow-up scan within 28 days. In this work, the follow-up scan was used for the purpose of reliability-testing the software. Thus, there were 76 MRI scans in total.

Table 3.2: Demographic of the processed subjects from the MIND data

		Lean Non-Diabetic			*Overweight Non-Diabetic*			*Type 2 Diabetes*		
Age Group	*Total patients*	*Patients*	*Male*	*Female*	*Patients*	*Male*	*Female*	*Patients*	*Male*	*Female*
50s	25	10	3	7	10	2	8	5	4	1
60s	13	5	1	4	3	1	2	5	3	2
Total	**38**	**15**	**4**	**11**	**13**	**3**	**10**	**10**	**7**	**3**

The MIND dataset included an array of other measurements in addition to the MRI acquisitions. This include for instance blood based biomarkers reflecting glucose control (e.g., plasma insulin and HbA1c) and systemic insulin resistance (HOMA-IR), CSF biomarkers reflecting amyloid and Tau deposition in the brain, and PET

measurements of radiolabelled glucose and fatty acid analogues, that reflect uptake of these substrates in the brain. Detailed analysis of the MR images will thus add the volumetric status in these subjects to this already rich dataset, which will allow for future detailed analyses that links together metabolic disturbances to brain function.

3.2 Data Processing

The initial step in the processing involved handling the input data. For the OASIS dataset, the input was an average of the 3-4 intra-session scans. The average scan used was created using the FreeSurfer function `mri-robust-template`. This creates an unbiased intra-subject template that is used in the following analysis [71]. This method has been demonstrated to increase reliability in test-retest data [71]. In contrast, for the MIND dataset, the native scan was used as the input due to the fact that the image-acquisition protocol was different, and only one scan was available.

Both datasets were processed on a computer running *Ubuntu 18.04.5 LTS*. The machine ran on a *Intel Xeon W-2133 CPU @ 3.60GHz* CPU as well as a *Nvidia GeForce RTX 2080 TI* GPU. For FreeSurfer, all samples were processed using version *7.4.1* and for FastSurfer version *2.2.0* was used. The type of machine used in the processing as well as the software version used has been shown to impact the end results [72, 73].

Even on modern machines, the full FreeSurfer pipeline, denoted `recon-all`, can take up to 20 hours [74]. To address this, a simple semaphore function was developed as a Bash shell script to enable parallel processing of multiple subjects for both FreeSurfer and FastSurfer. The shell script allows up to eight subjects to be processed in parallel, significantly reducing the overall computation time. The same semaphore function was used for the FastSurfer when processing in parallel as well.

By default, the FreeSurfer `recon-all` pipeline resamples the input MRI to a voxel size of $1 \times 1 \times 1$ mm^3, regardless if the input is natively of a higher resolution. However, FastSurfer allows for processing resolutions between $0.7 - 1$ mm^3. However, due to memory issues during the bias field correction step caused by the high resolution of the MIND data, and as to make comparison between the software more fair the flag `-vox-size 1` was used. This settings in FastSurfer resamples the MRI to a $1 \times 1 \times 1$ mm^3 voxel size, same as in the FreeSurfer recon-all pipeline.

For FreeSurfer, the entire processing pipeline was executed in parallel on the CPU. For FastSurfer however, the cortical segmentation step was run sequentially on the GPU to avoid potential memory issues due to only one GPU being available. All the later steps in FastSurfer were then also processed in parallel and on the CPU.

In the surface reconstruction process outlined in section 2.4.1.1, FreeSurfer (and FastSurfer) separates the brain into two hemispheres. Consequently, all measurements, including the average cortical thickness per region of interest in the atlas,

are divided into hemispheric values. This hemispheric division applies to a subset of subcortical structures as well. To prevent further fragmentation of the already limited dataset, which would result in separate datasets for each hemisphere, the average of all regions of interest spanning both hemispheres was calculated per subject and scan. It is worth mentioning that there is evidence that atrophy due to AD may be asymmetric [75, 76], but this was subsequently not investigated in this thesis.

All calculated cortical anatomical measurements were labelled to the DKT atlas. The full list of the cortical regions in the DKT atlas can be found in Table A.1 in Appendix A.1. All subcortical regions included in the analysis can be found in A.2 in Appendix A.1.

The calculated average measurements were used in all further analysis, including the similarity metrics and the statistical testing.

3.3 Quality Assurance

Due to lack of domain knowledge and time constraints, it was impractical to manually evaluate the quality of all 272 brain scans included in this thesis. However, a recent study suggests that manually correcting the boundaries between white matter and grey matter does not significantly improve the overall accuracy of cortical thickness, volume and surface area [77].

Due to this, only a small subset of the scans underwent visual inspection to ensure the segmentation process yielded reasonable results.

3.4 Similarity metrics

To compare how reliable FreeSurfer and FastSurfer are, three different metrics for similarity were implemented: *Pearson Correlation Coefficient*, *Intraclass Correlation Coefficient* and average variability. The output of FreeSurfer and FastSurfer, as well as the reliability testing was done in Python version 3.12.1.

3.4.1 Pearson Correlation Coefficient

The *Pearson Correlation Coefficient* (PCC) is a statistical measure that quantifies the degree of linear relationship between two sets of data [78]. The PCC between a set X and Y, denoted as $\rho_{X,Y}$, is defined by the following formula:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{3.1}$$

where $\text{cov}(X,Y)$ is the covariance of the two sets X and Y, and σ_X, σ_Y are their respective standard deviations. The PCC ranges from -1 to 1, where a positive PCC indicates a positive linear relationship, meaning that an increase in one data

set corresponds to an increase in the other. A negative PCC indicates a negative linear relationship, meaning that an increase in one data set corresponds to a decrease in the other. A PCC of 0 suggests no linear relationship between the two sets of data.

In the context of reliability testing, the ideal case would be for the PCC to be equal to 1. This would indicate that the measurements derived from the first scan are identical to the measurements derived from the second scan, demonstrating perfect intra-software reliability.

3.4.2 Intraclass Correlation Coefficient

Another common method to quantify reliability in the field of clinical statistics the so called "Intraclass correlation coefficient" (ICC) [79]. There exists multiple ways to calculate ICC, depending on what is being tested. For this thesis, two different ICC methods were used. The first one was used to test for intra-software reliability and the second was used to test for inter-software reliability.

In a test-retest setting it is recommended to use the ICC(2,k) measure [68]. This measure is suitable to use when trying for intra-rater reliability, that is, we want to measure the consistency of our rater. The rater in this case are the two software, FreeSurfer and FastSurfer. In this thesis the question is how well the software, either FreeSurfer or FastSurfer, replicate results that should be identical.

The ICC(2,k) is defined [68] as

$$\mathrm{ICC}(2,k) = \frac{MS_R - MS_E}{MS_R + \frac{MS_C - MS_E}{n}} \tag{3.2}$$

where MS_R represents the mean square of the rows, which captures the variability between subjects. This could, for instance, reflect the differences in cortical thickness within the same region across different individuals. MS_E denotes the mean square of the errors, capturing the variability between two measurements that should ideally be identical under perfect conditions. Lastly, MS_C stands for the mean square of the columns, reflecting the variability that different raters (in this case, the two different scans) may show systematic differences. For example, it is possible that one scan may overestimate anatomical measures due to some induced error, leading to a systematic bias.

The second form of ICC used was ICC(3,1). This was used to test for inter-software reliability, i.e., quantify how similar the results from two different raters are. In this case, the 'raters' are the two different software used, FreeSurfer and FastSurfer. This was mainly used to determine if FreeSurfer and FastSurfer differ widely in the anatomical measures they estimate. As FreeSurfer is a very commonly used software in structural MRI, and FastSurfer is not as extensively validated, it is of interest that they are reasonably similar. The ICC(3,1) is defined [68] as

$$\mathrm{ICC}(3,1) = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E}. \tag{3.3}$$

Where MS_R, MS_E are the same mean squares used in Equation (3.2) and k is the amount of raters, which for this case is equal to two.

According to [68] the reliability reported by the ICC can be categorised into four different ranges, less than 0.50, between 0.50 and 0.75, between 0.75 and 0.90 and above 0.90. These ranges correspond to poor, moderate, good and exceptional reliability respectively.

3.4.3 Test-retest Variability

The final similarity metric that will be used is the test-retest variability (TRV) metric [80]. The TRV is defined as the absolute average of two samples, divided by their common mean expressed as a percentage. In this case however, we will be using the average TRV. The only difference in this case is that we sum the TRV over all regions of interest, and divide by the number of samples as follows

$$\text{TRV} = \frac{1}{N}\sum_{i=1}^{N}\frac{|X_i - Y_i|}{\frac{(X_i+Y_i)}{2}} \times 100. \tag{3.4}$$

In the equation above X_i and Y_i are paired anatomical measurements from the test-retest scans on one specific region of interest.

The TRV is used to give an intuitive feel for how much a specific region of interest can vary on average. Ideally, a TRV of zero, or close to zero is preferable as this means the software evaluates the anatomical measures similarly between the scans.

3.5 Statistical Testing

In this section, the methodology employed to investigate potential volumetric differences between patients with Type 2 Diabetes Mellitus (T2DM) and healthy controls (HC) will be presented. The statistical testing was divided into three parts. Firstly, the data acquired from processing the MRI scans with FreeSurfer and FastSurfer underwent preprocessing. Subsequently, the multiple comparison problem was addressed by applying two distinct approaches, one tailored for the OASIS dataset and another for the MIND dataset. Finally, a linear model was utilised to determine if the diagnosis (AD for the OASIS dataset and T2DM for the MIND dataset) was a significant factor contributing to the observed differences in three anatomical measurements, cortical thickness, grey matter volume and subcortical volume.

3.5.1 Correction for Intracranial Volume

As outlined in section 2.1.1, the ICV is highly correlated with total grey matter volume and cortical surface area. Due to this, it is common in volumetric brain studies to adjust the regional volumes with the ICV. This can be done in various ways, the simplest being to divide the regional volume with the ICV to get the percentage of

the cranium that each volume occupies.

For this work, a method called the residual method was used [81]. It adjusts each regional volume for each subject according to the following formula

$$V_{\text{adj},i} = V_{\text{raw},i} - \beta(\text{ICV}_i - \overline{\text{ICV}}). \tag{3.5}$$

In equation 3.5, $V_{adj,i}$ represents the adjusted volume of a specific region in a specific subject. V_{raw} denotes the unadjusted volume of the same region in the same subject. The term β is the regression coefficient derived from linearly regressing the volume of a specific region of interest within the control group with the ICV. This allows for the determination of the specific linear relationship between that particular region and the ICV in a healthy population. Furthermore, ICV_i denotes the ICV of subject i, while $\overline{\text{ICV}}$ represents the average ICV within the control group.

The point of using only the control group as the basis of the linear regression coefficient β is to control for the association commonly found between ICV and the volumetric measure. That is, when comparing cross-sectional datasets we assume that the volumetric measure is effected by something other than the ICV, such as a neurodegenerative disorder.

3.5.2 Analysis of Covariance

To determine whether or not the disease status (AD for the OASIS dataset and T2DM for the MIND dataset) had a significant impact on cortical thickness and grey matter volume, an "Analysis of Covariance" (ANCOVA) model was used. ANCOVA is a useful technique when evaluating if the mean of a dependent variable differ between two or more categorical groups after correcting for independent covariates that may affect the dependent variable [82]. That is, ANCOVA tries to remove the effect of the covariates so the statistical effect of the categorical variable can be studied more clearly.

As outlined by Section 2.1.1, anatomical measurements such as cortical thickness and grey matter volume are influenced by the age and gender of the subject. As such, these were included as covariates. For both of the datasets, the model fitted was

$$\{CT, GMV, SCV\}_{\text{Adj}} \sim Age + Gender + Diagnosis, \tag{3.6}$$

where CT, GMV, and SCV represent the adjusted values of cortical thickness, grey matter volume, and subcortical structure volume, respectively. The categorical variable "*Diagnosis*" in Equation (3.6) varies depending on the dataset analysed. For the OASIS dataset, there are two categorical groups: AD and non-AD, as shown in Table 3.1. However, for the MIND dataset, there are three categories: lean, overweight, and T2DM, as indicated in Table 3.2. Although the overweight patient group was not of primary interest for this work, they were included in the ANCOVA model to provide additional data on how age and gender impact the anatomical measures.

In practice, a linear model was fitted for each region of interest according to equation 3.6 using the Python module `statsmodel`. The `OLS` function from the module `statsmodel` fits a linear model to the given data using ordinary least squares and returns the regression coefficients, their corresponding p-values, and other relevant information.

3.5.3 Bonferroni Correction

As mentioned in section 3.5, two forms of statistical correction were employed in this work. The first one involved the use of the Bonferroni correction [83], a method of statistical correction commonly utilised when performing multiple statistical tests to reduce the risk of spurious significance when the tests are independent of each other. Given the DKT atlas used in this thesis contained a large number of regions of interests (31), there existed a relatively high probability of at least one region of interest exhibiting significance due to pure chance. To reduce this likelihood, the desired significance level is divided by the number of tests being performed. That is, we reject the null hypothesis H_0 for the i-th test, T_i, if the p-value of the outcome is less than α/n, where α represents the commonly used critical value of 0.05, and n is equal to the number of regions of interest being tested simultaneously.

$$P(T_i \text{ passes} \mid H_0) \leq \frac{\alpha}{n}. \tag{3.7}$$

The p-value used here is the one calculated for the diagnosis from the linear model in Equation (3.6). The DKT atlas contains 31 regions of interest, and there are 22 non-zero subcortical structures. Consequently, n is either 31 or 22, depending on whether the cortical or subcortical structures are being tested.

It's worth mentioning that the anatomical measurements in the DKT atlas most definitely are dependant on each other, and not *independent* since they come from the same individual. This in turn runs the risk of over-correcting the statistical significance. However, since it's of interest to determine if regions most severely effected by AD is also effected by T2DM it is still a useful method for correction in this case.

3.5.4 A Priori Region Selection

Due to the MIND dataset being a small dataset and that the volumetric differences are assumed to be quite small as none of the patients have a clinical neurodegentrative disease diagnosis, the regions of interests were picked *a priori*.

What this means is that only regions that showed significant morphometric differences between AD and HC patients in the OASIS dataset, after the Bonferroni correction, were checked for significance in the MIND data. This method is recommended to use when analysing cross-sectional MRI data [84] and serves two purposes. The first is in an attempt to handle the multiple comparison problem, and the second is that we are only interested in morphometric changes *due* to the apparent connection between AD and T2DM.

4

Results

In this chapter, the results of this thesis will be presented. The first part of the result will focus on the performance, both in terms of runtime and reliability of FreeSurfer and FastSurfer. The second part of the results will first present the results related to the determination of which regions of interests were the most impacted by AD-related atrophy and if any of the same regions are effected by T2DM.

4.1 Software Comparison

The study compared two software tools, FreeSurfer and FastSurfer, based on three different metrics. The first metric evaluated the computational time required to process a T1-weighted MRI scan, which indicates the efficiency and processing speed of each software. The second metric assessed the reliability of the software, which was determined by the similarity measures presented in section 3.4, providing insights into the consistency and accuracy of the results generated by each tool. The third and final metric focused on the ability of the two software to detect the same morphometric signals, a crucial measure in determining whether the tools produce consistent and comparable results in terms of identifying relevant anatomical features.

4.1.1 Runtime

Table 4.1 presents the average runtime per subject, as well as the total runtime per software and dataset. The total runtime is defined as the sum of time taken to process all subjects individually. It is worth noting that the actual total runtime is lower for both software packages, as the subjects were processed in parallel as outlined in Section 3.2. However, due to user error, some subjects had to be manually rerun. This meant that some batches contained less than 8 subjects at a time. To enable comparison that ignores user error, the total run time and average run per subject is presented.

Table 4.1: Average and total runtime per software and dataset. The table shows the results for the two software FreeSurfer and FastSurfer per dataset.

Software	Dataset	Number of subjects	Average runtime per subject	Total run time
FastSurfer	MIND, Scan 1	38	02:02 hours	3 days 05:45
	MIND, Scan 2	38	01:48 hours	2 days 20:24
	OASIS	136	02:08 hours	12 days 04:05
FreeSurfer	MIND, Scan 1	38	06:48 hours	10 days 18:45
	MIND, Scan 2	38	06:46 hours	10 days 17:19
	OASIS	136	07:06 hours	40 days 07:13

4.1.2 Intra-software Reliability

In the following section, the results from the three different similarity metrics: Pearson correlation coefficient, intraclass correlation coeffiecient, and variability, described in Section 3.4 are presented for both datasets used. All the ICC calculations use the ICC$(2, k)$ metric, per equation (3.2). Three different anatomical measures were used to determine the similarity metrics: average cortical thickness, grey matter volume, and cortical surface area. They were calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately.

In Figures 4.1 and 4.5 we can see the results from the PCC calculated between the first and second MRI scans from the OASIS dataset. Generally, the PCC is high for both software and both datasets, with both averages being above 0.90. Furthermore, the PCC values for the cortical thickness is generally lower than that of the grey matter volume and cortical surface area. The same trend can be seen in the ICC similarity metric, shown in Figures 4.2 and 4.6. In Figures 4.3 and 4.7, the TRV is presented. For both software and datasets, the average variability is around 2%. Finally, all three of the similarity metrics for the subcortical volumes can be seen in Figures 4.4 and 4.8.

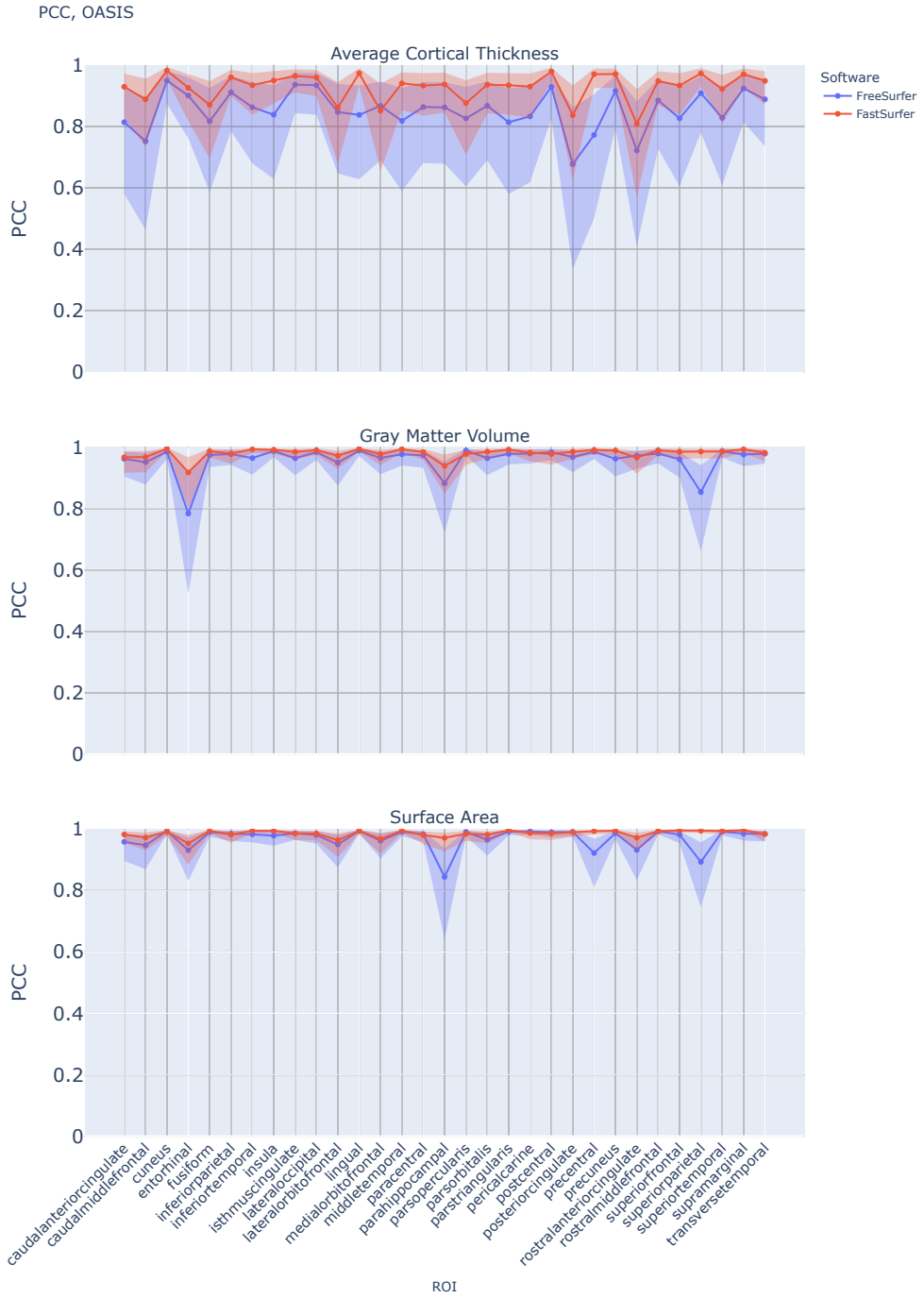


Figure 4.1: The test-retest reliability, quantified by the *Pearson correlation coefficient* (PCC), between the first and second MRI scans from the OASIS dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The PCC was calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the PCC values, while the transparent bands indicate the 95% confidence intervals around each PCC estimate. Details on specific values can be found in Tables B.1, B.2, B.3, B.4, B.5 and B.6.

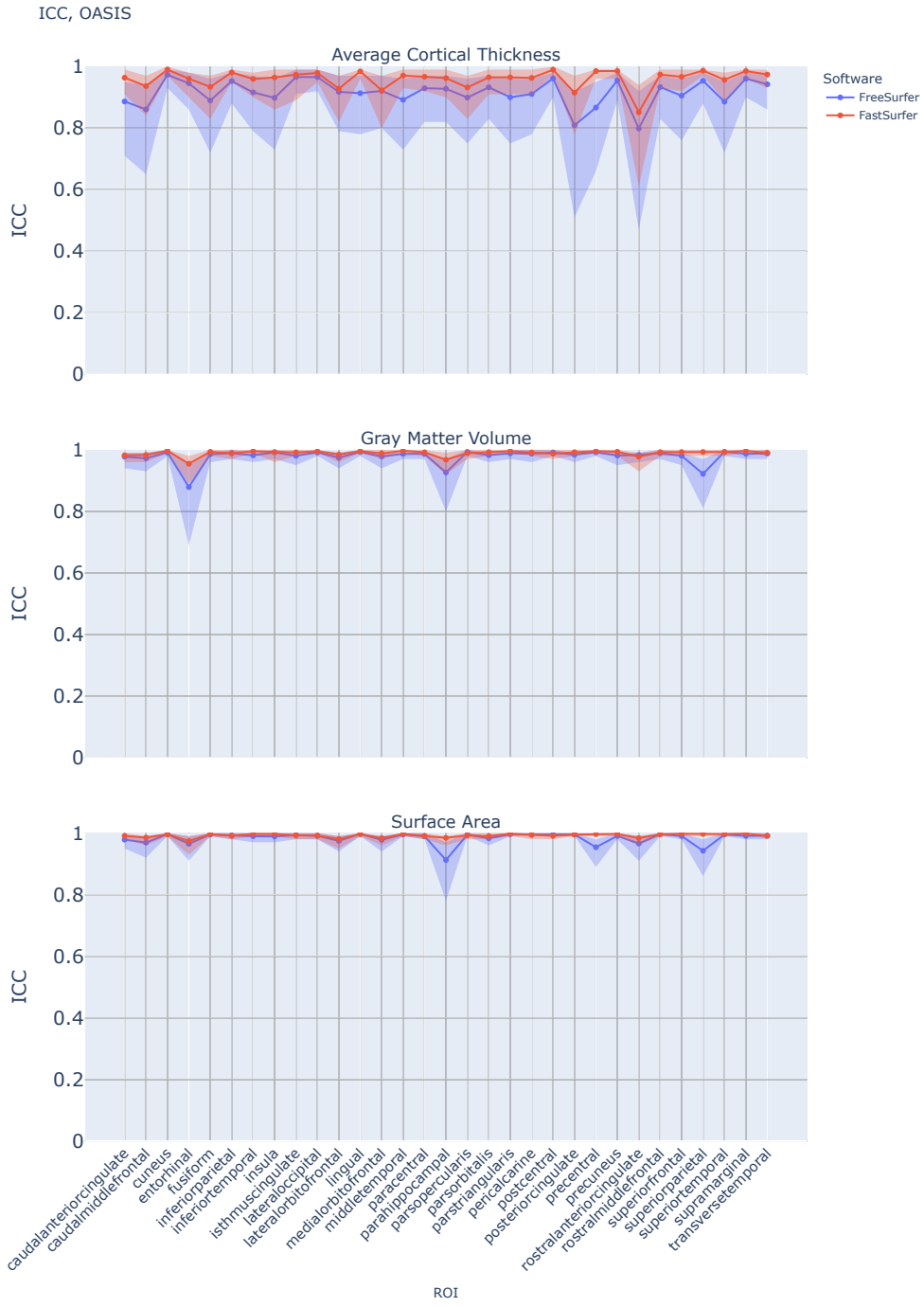


Figure 4.2: The test-retest reliability, quantified by the *Intraclass correlation coefficient* (ICC), between the first and second MRI scans from the OASIS dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The ICC was calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the ICC values, while the transparent bands indicate the 95% confidence intervals around each ICC estimate.

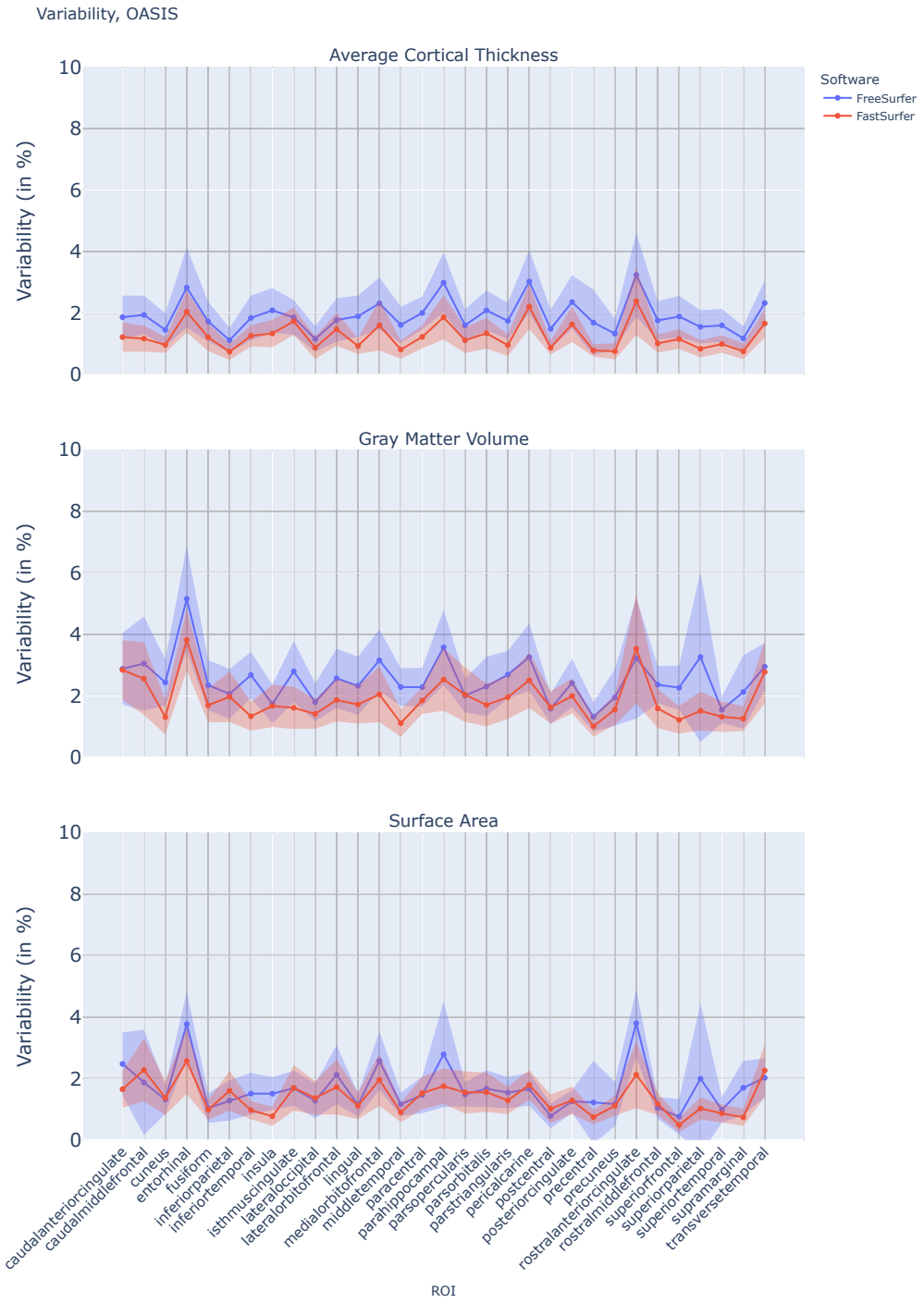


Figure 4.3: The *Test-retest variability* (TRV), between the first and second MRI scans from the OASIS dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The TRV was calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the TRV values, while the transparent bands indicate the 95% confidence intervals around each TRV estimate. Details on specific values can be found in Tables B.1, B.2, B.3, B.4, B.5 and B.6.

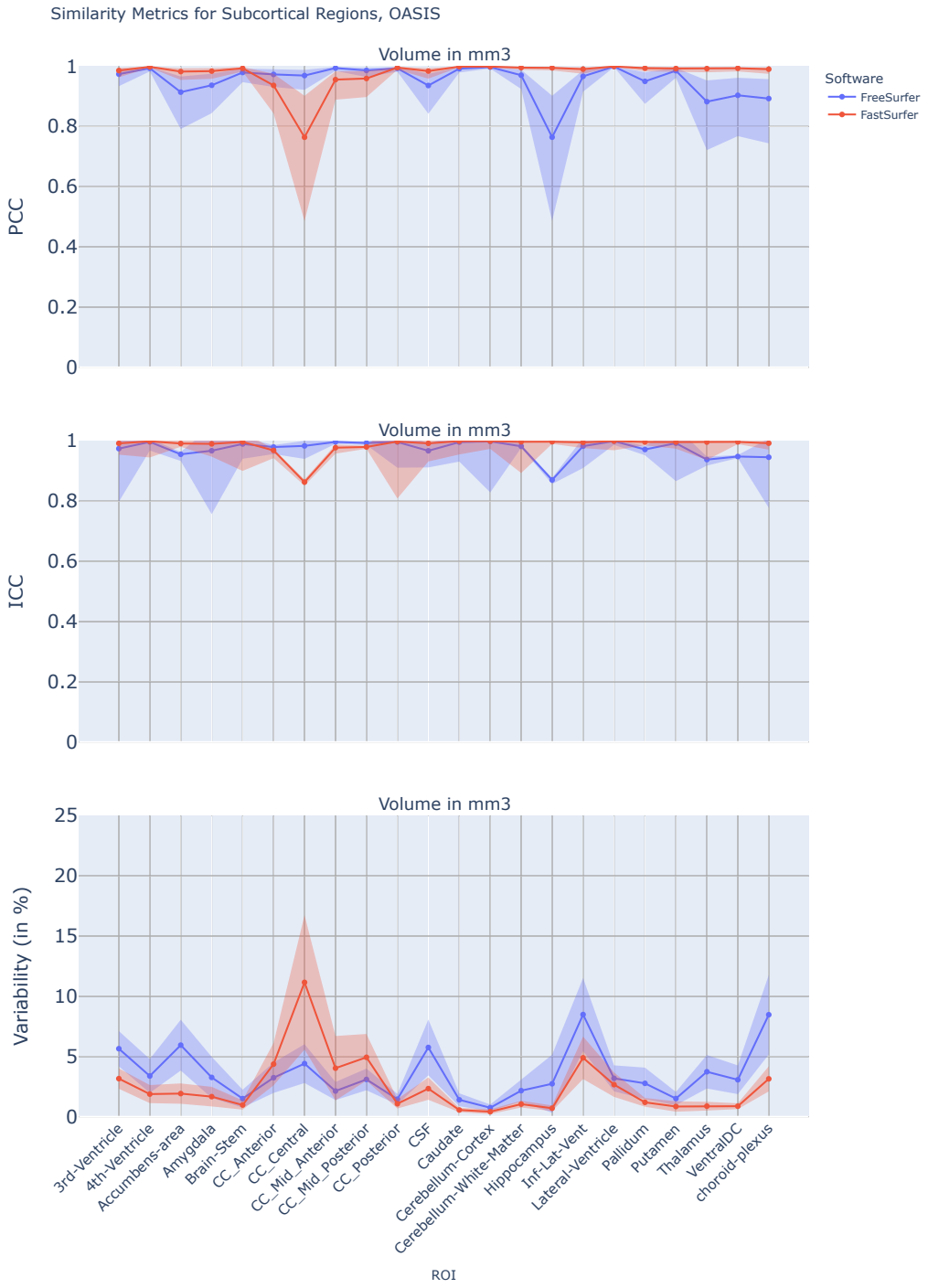


Figure 4.4: The test-retest reliability, quantified by the PCC, ICC and TRV, between the first and second MRI scans from the OASIS dataset for subcortical volume measurements. The plot at the top shows the PCC, the middle one the ICC and the bottom one the TRV. The metrics were calculated for all non-zero subcortical volumes, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the calculated values, while the transparent bands indicate the 95% confidence intervals around each estimate. Details on specific values can be found in Tables B.7 and B.8.

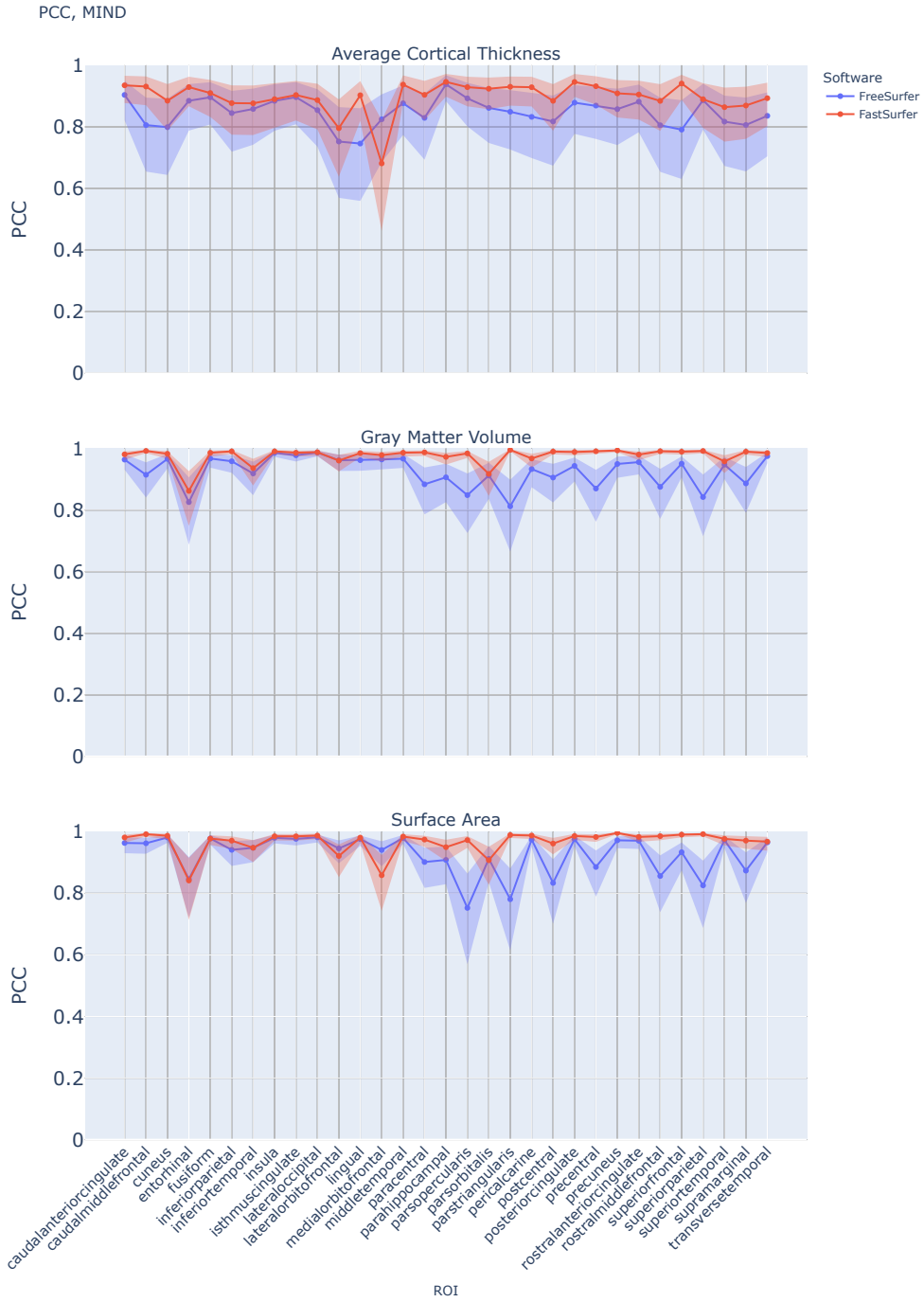


Figure 4.5: The test-retest reliability, quantified by the *Pearson correlation coefficient* (PCC), between the first and second MRI scans from the MIND dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The PCC was calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the PCC values, while the transparent bands indicate the 95% confidence intervals around each PCC estimate. Details on specific values can be found in Tables B.9, B.10, B.11, B.12, B.13 and B.14.

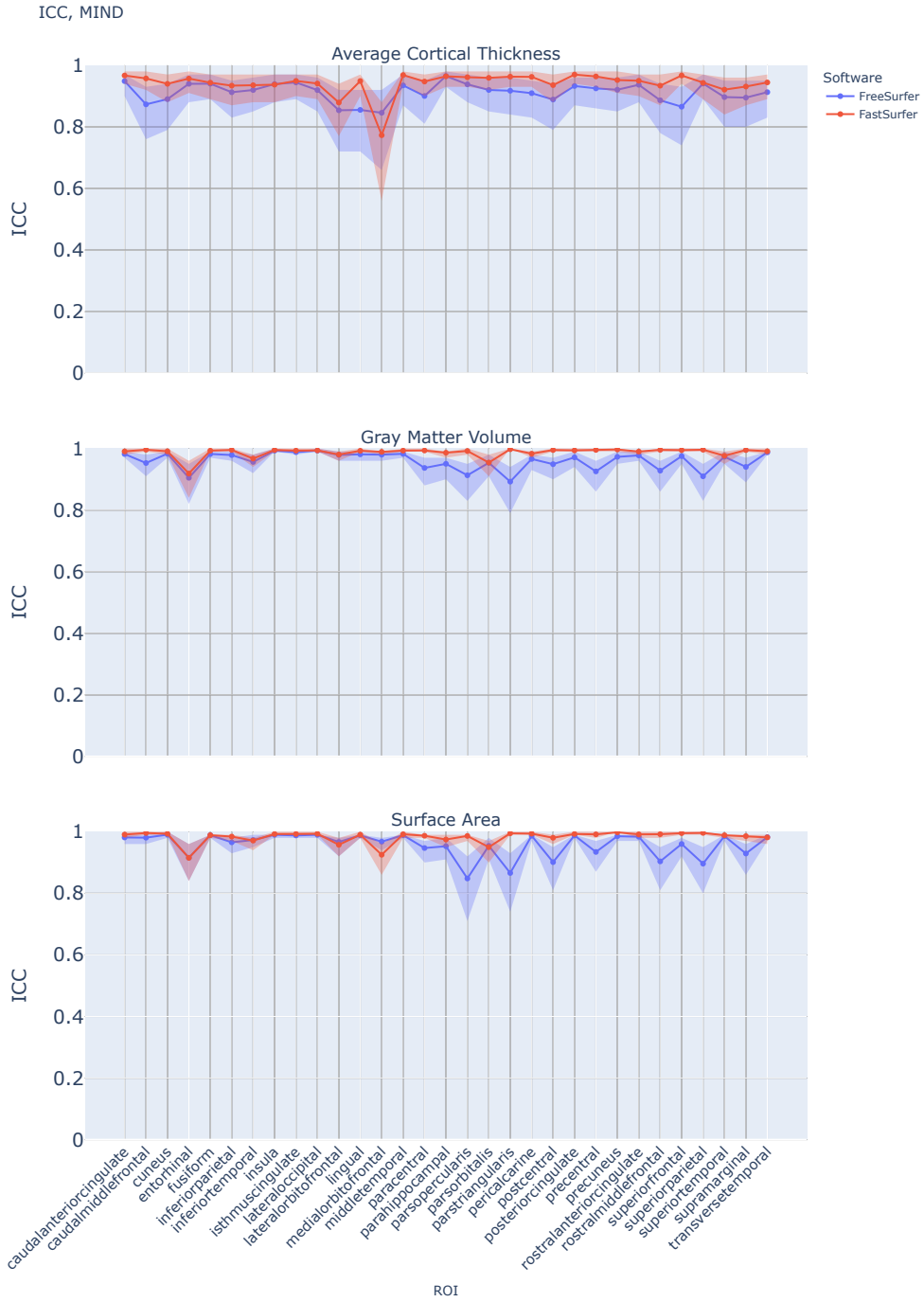


Figure 4.6: The test-retest reliability, quantified by the *Intraclass correlation coefficient* (ICC), between the first and second MRI scans from the MIND dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The ICC was calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the ICC values, while the transparent bands indicate the 95% confidence intervals around each ICC estimate. Details on specific values can be found in Tables B.9, B.10, B.11, B.12, B.13 and B.14.

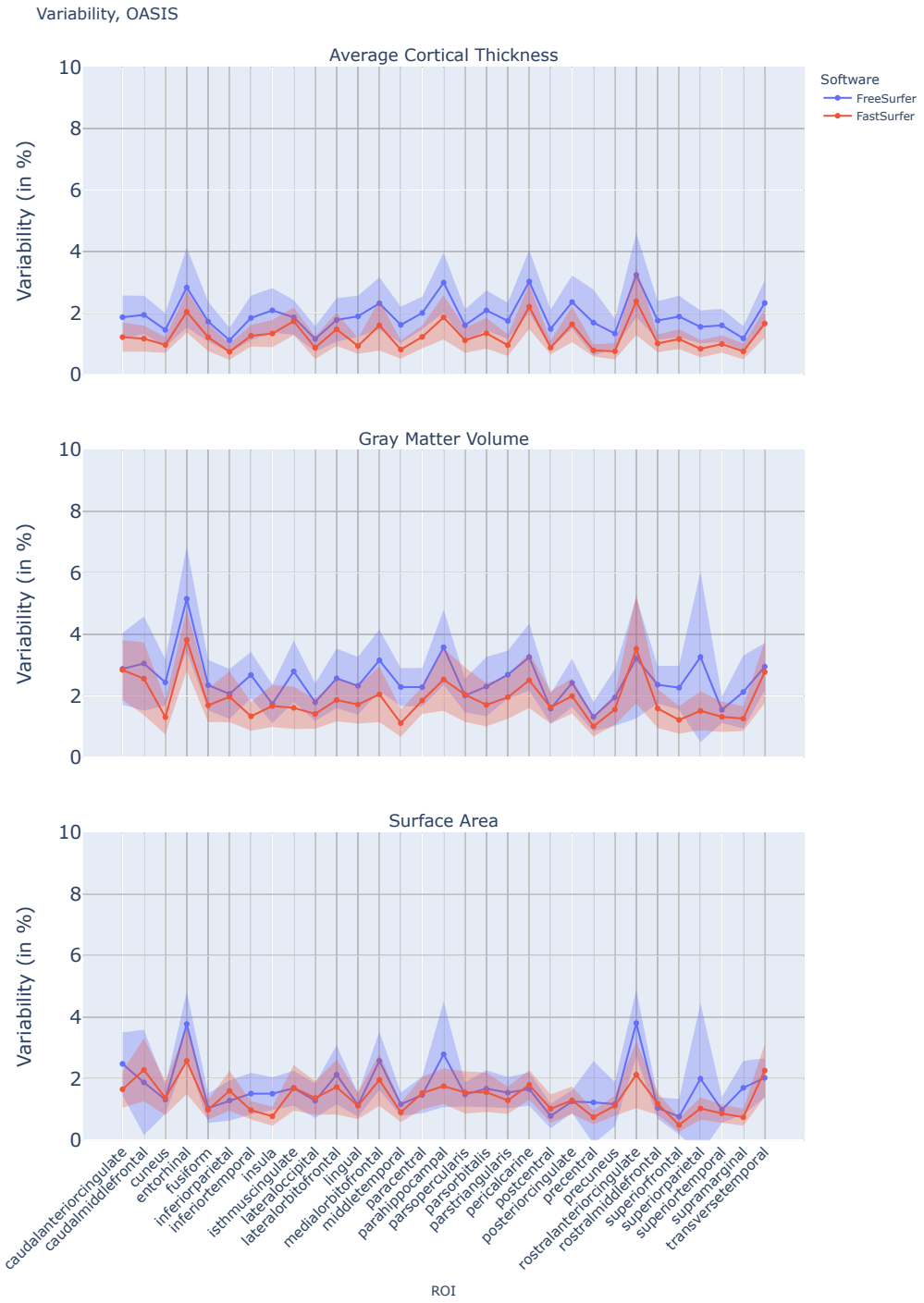


Figure 4.7: The *Test-retest variability* (TRV), between the first and second MRI scans from the MIND dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The TRV was calculated for each region of interest defined in the Desikan-Killiany-Tourville (DKT) atlas, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the TRV values, while the transparent bands indicate the 95% confidence intervals around each TRV estimate. Details on specific values can be found in Tables B.9, B.10, B.11, B.12, B.13 and B.14.

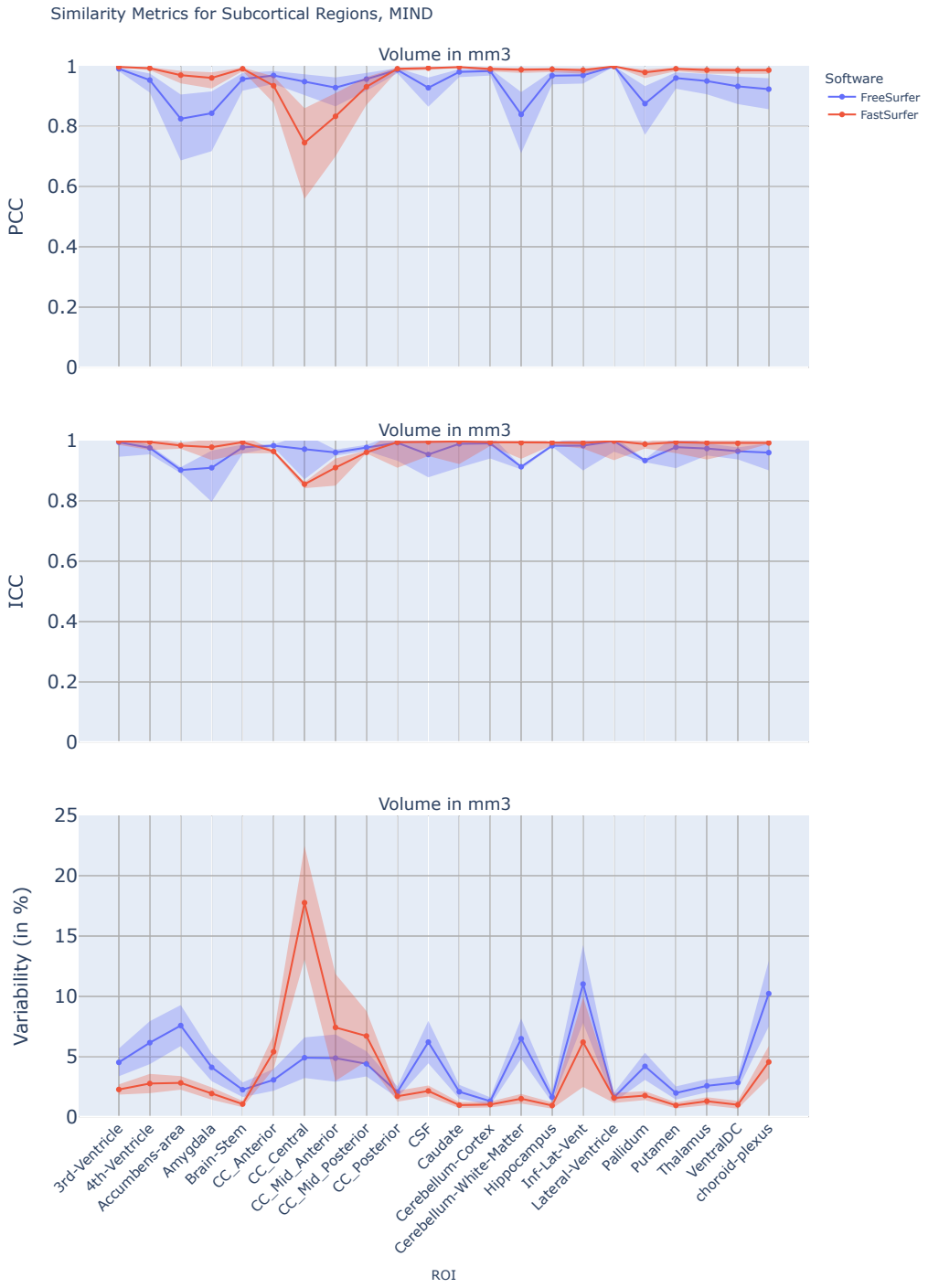


Figure 4.8: The test-retest reliability, quantified by the PCC, ICC and TRV, between the first and second MRI scans from the MIND dataset for subcortical volume measurements. The plot at the top shows the PCC, the middle one the ICC and the bottom one the TRV. The metrics were calculated for all non-zero subcortical volumes, using estimates from both the FreeSurfer and FastSurfer software separately. The solid lines represent the calculated values, while the transparent bands indicate the 95% confidence intervals around each estimate. Details on specific values can be found in Tables B.15 and B.16.

4.1.3 Inter-software Reliability

Here, the results from the inter-software reliability testing is presented. The ICC used to determine the inter-software reliability is the ICC$(3, 1)$ from Equation (3.3). The inter-software reliability was used to determine if FastSurfer and FreeSurfer differ in their estimates of anatomical measurements. The same three different anatomical measures as in the intra-software testing were used, namely average cortical thickness, grey matter volume, and cortical surface area. The calculations were performed separately for the two sets of scans available in the both datasets.

In Figures 4.9 and 4.10 the ICC can be seen for the OASIS dataset and the MIND dataset, respectively. For all three of the anatomical measurements and both scans, the average ICC is close to 0.93 with a few outlier regions.

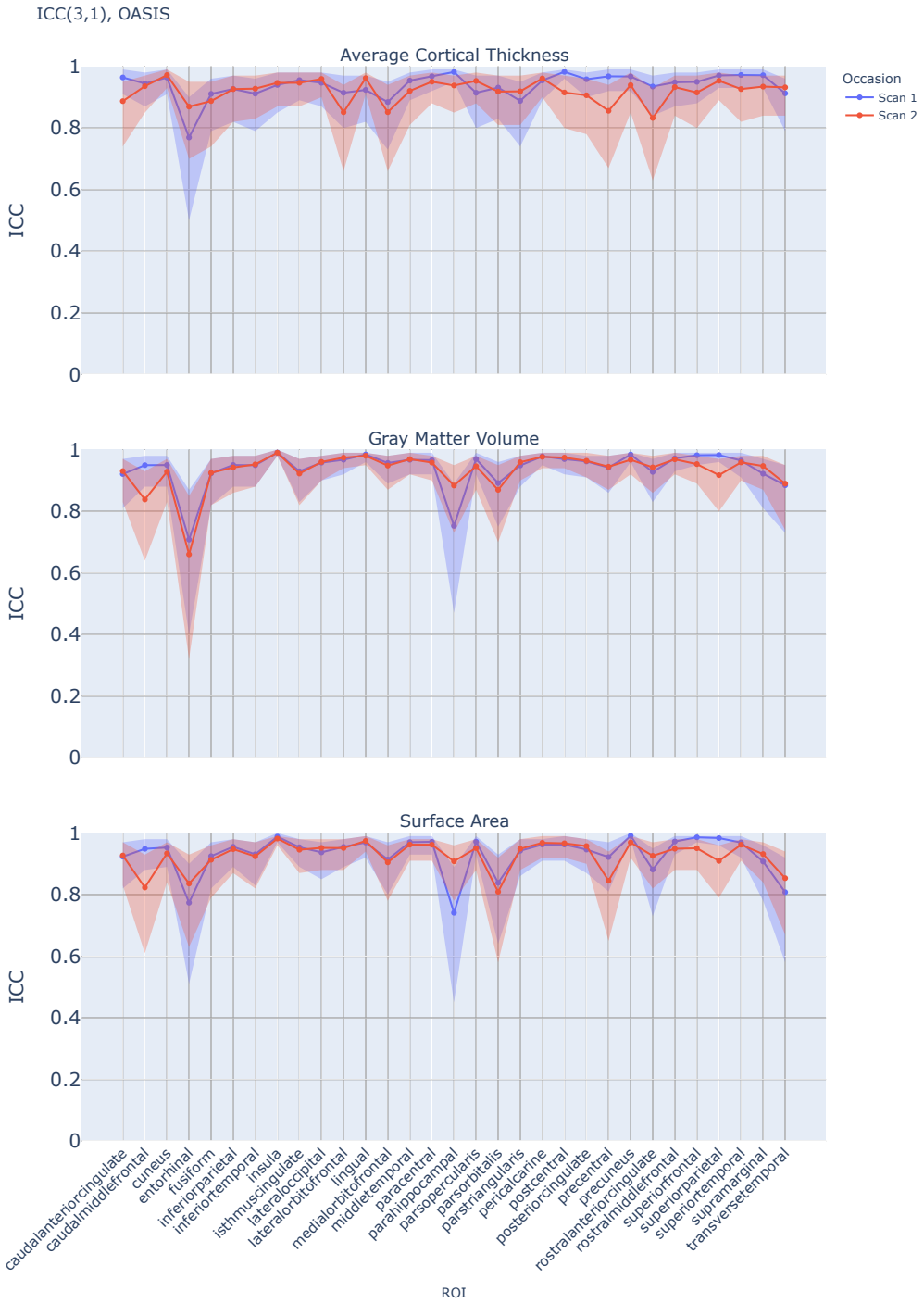


Figure 4.9: The inter-rater reliability, quantified by the ICC(3,1) metric, between the the two software FreeSurfer and FastSurfer using the OASIS dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The calculations were performed separately for the two sets of scans available in the dataset. The solid lines represent the calculated values, while the transparent bands indicate the 95% confidence intervals around each estimate.

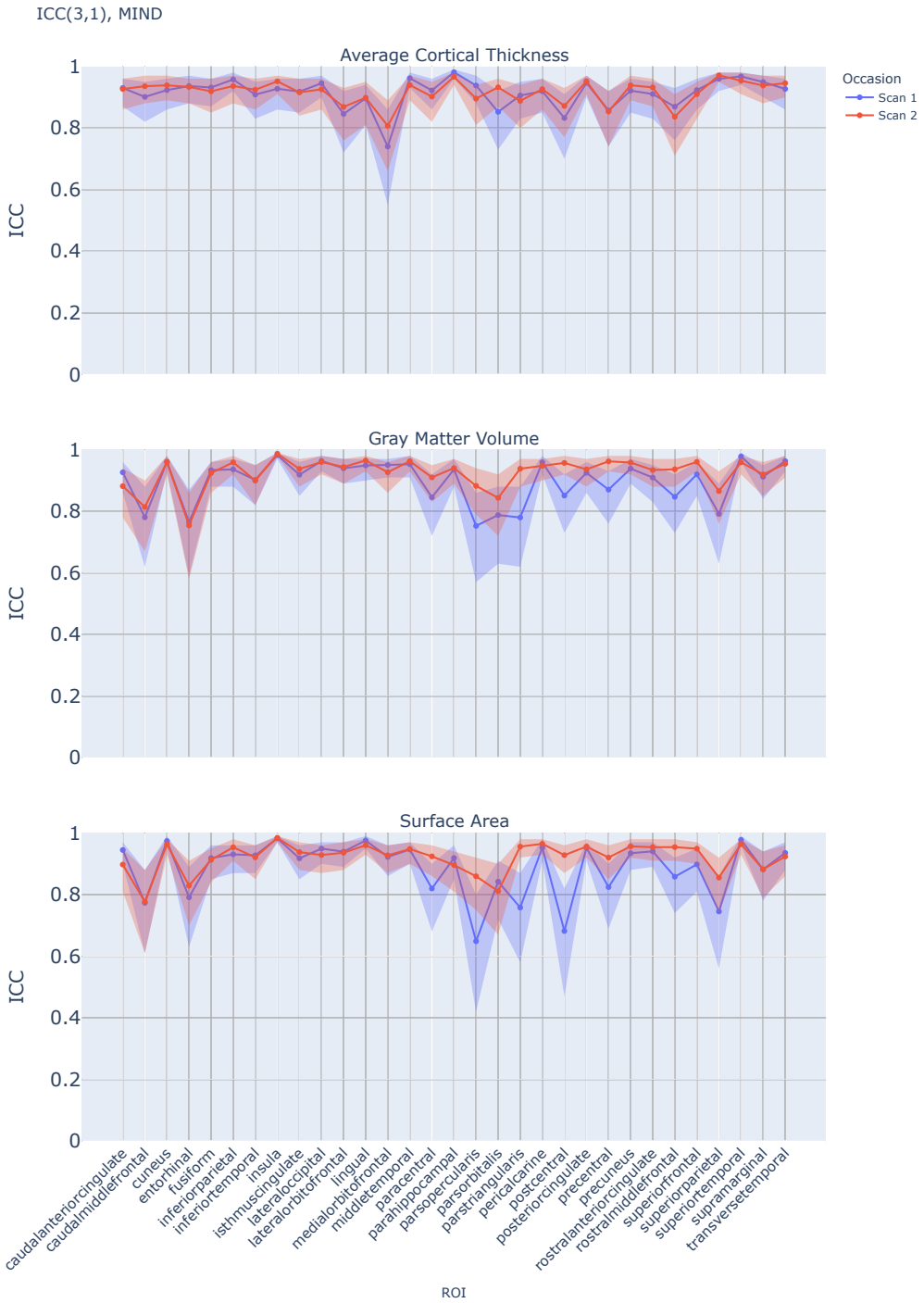


Figure 4.10: The inter-rater reliability, quantified by the ICC(3,1) metric, between the the two software FreeSurfer and FastSurfer using the MIND dataset. Three different anatomical measures are shown: average cortical thickness, grey matter volume, and cortical surface area. The calculations were performed separately for the two sets of scans available in the dataset. The solid lines represent the calculated values, while the transparent bands indicate the 95% confidence intervals around each estimate.

4.1.4 Visual Inspection

Here, an example of the difference between FreeSurfer and FastSurfer is presented by a single slice from a single subject in the MIND dataset.

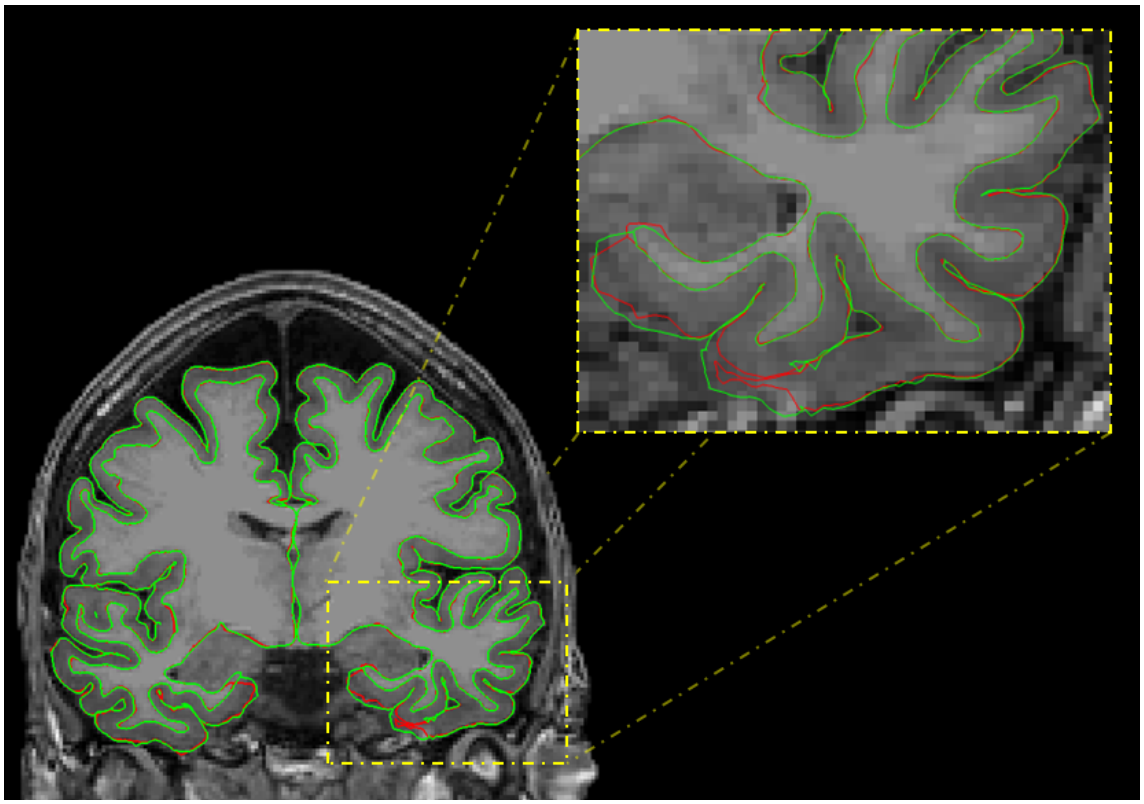


Figure 4.11: One slice of the cortical reconstruction in a subject from the MIND dataset. The colour of the lines determines the software used, with FreeSurfer represented in red and and FastSurfer in green. The figure also contains a zoomed in part of the right hemisphere of a part of the temporal lobe.

4.2 Assessment of Brain Volumetric Changes in Type 2 Diabetes

In this section, the results used to assess whether or not there exists morphometric changes in the patients with T2DM compared to the healthy controls are presented. This result is divided in to two parts. First, the results from the a priori region of interest selection are presented. Lastly, the results from the statistical difference between T2DM patients and the HC will be presented.

4.2.1 A Priori Region Selection

In this section, the most significantly impacted regions of interest in the DKT atlas from OASIS dataset will be presented. The significance was determined by fitting a linear model as outlined by Section 3.5 and the statistical correction was performed by dividing the significance with the number of regions. The regions had to be significant in both FreeSurfer and FastSurfer to be included. The significant regions in terms of average cortical thickness can be seen in Figure 4.12, the regions significantly effected in terms of grey matter volume can be seen in Figure 4.13 and finally, the most severely effected subcortical regions are presented in Figure 4.14.

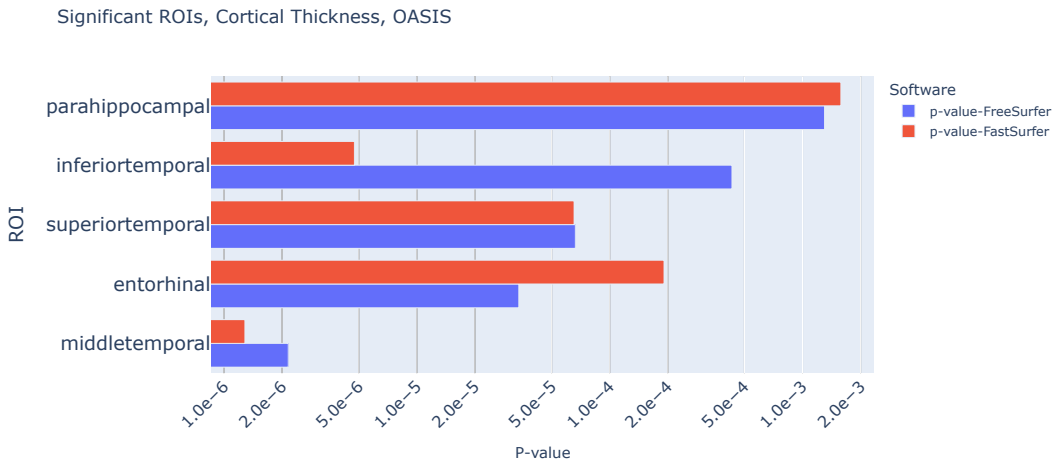


Figure 4.12: Regions of interest in the OASIS dataset where Alzheimer's disease diagnosis was identified as a statistically significant factor in both FreeSurfer and FastSurfer contributing to reduced cortical thickness at the significance level of $p < \frac{0.05}{31}$. This was based on a linear regression model that included age, gender, and Alzheimer's disease diagnosis as covariates.

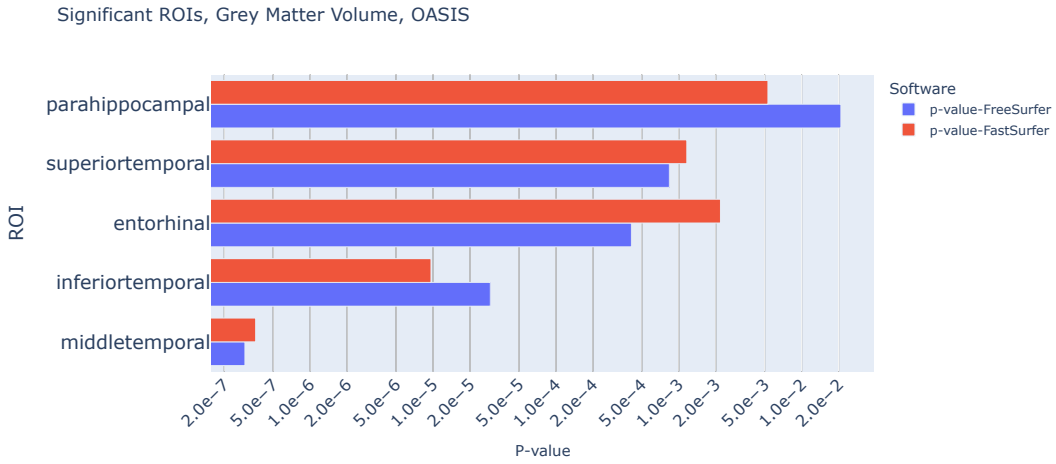


Figure 4.13: Regions of interest in the OASIS dataset where Alzheimer's disease diagnosis was identified as a statistically significant factor in both FreeSurfer and FastSurfer contributing to reduced grey matter volume at the significance level of $p < \frac{0.05}{31}$. This was based on a linear regression model that included age, gender, and Alzheimer's disease diagnosis as covariates.

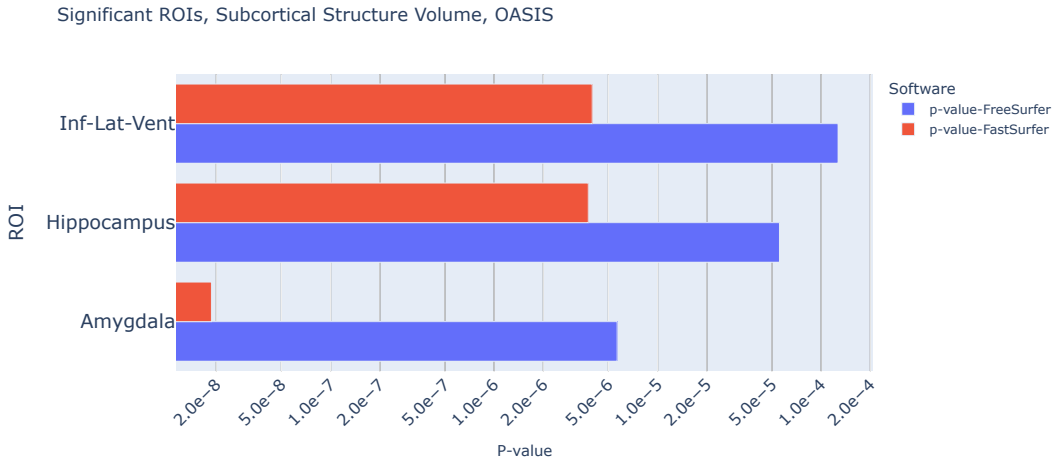


Figure 4.14: Regions of interest in the OASIS dataset where Alzheimer's disease diagnosis was identified as a statistically significant factor in both FreeSurfer and FastSurfer contributing to altered subcortical volume at the significance level of $p < \frac{0.05}{22}$. This was based on a linear regression model that included age, gender, and Alzheimer's disease diagnosis as covariates. Both the amygdala and the hippocampus exhibited negative coefficients, indicating atrophy. Conversely, the inferior lateral ventricle displayed a positive coefficient, suggesting an enlargement in its volume.

4.2.2 Statistical Difference Between T2DM and HC

In Table 4.2, the results from the statistical testing between the T2DM and HC patients from the MIND dataset is presented. All regions that were significant when comparing the cortical thickness, grey matter volume and subcortical volumes will be presented.

Table 4.2: Table presenting the regions in the DKT atlas that were significant in the MIND dataset at the 0.05 level as well the anatomical measure it was significant for.

Anatomical Measurement	Region of interest	P-value
Cortical Thickness	Transverse temporal gyrus	0.0077
	Middle frontal gyrus, *Rostral*	0.0304
Grey Matter Volume	Cingulate, *Rostral anterior*	0.0037
	Cingulate, *Isthmus*	0.0285
	Cingulate, *Caudal anterior*	0.0325
	Inferior temporal gyrus	0.0464
Subcortical Volume	Thalamus	0.0442

5
Discussion

This chapter present interpretations of the key findings reported in the chapter 4. The discussion is structured into five sections, each corresponding to a subsection of the results. In addition to discussing the implications of the current study, this chapter will explore potentials for future research questions. Both in terms of different methodological approaches that could be employed to further validate or extend the present findings. Secondly, it will consider the potential benefits of incorporating different datasets.

5.1 Software Comparison

Here, the results used to compare FreeSurfer to FastSurfer will be presented. The two software will be compared in terms of runtime, intra-software reliability and finally inter-software reliability.

5.1.1 Runtime

The results displayed in table 4.1 shows that processing the MR scan with FastSurfer provides a remarkable performance advantage over FreeSurfer. When using the semaphore Bash shell script outlined in section 3.2, FastSurfer finishes the equivalent full pipeline on the same dataset in less than a third of the time compared to FreeSurfer. It is important to note that the average runtime per subject and the total runtime presented in table 4.1 does not account for the parallel processing of multiple subjects. The actual time required to process the data is roughly one-eighth of the values shown in table 4.1 for both FreeSurfer and FastSurfer. However, due to the necessity of re-running certain subjects as a result of user error, the results were presented in their current format.

5.1.2 Intra-software Reliability

The primary objective of the intra-software reliability assessment was to determine the consistency of anatomical measurements estimated by the two software, FreeSurfer and FastSurfer, when comparing test-retest scans. The results show that, on average, both software have high correlation coefficients, both for Pearson's and intraclass. They also produce a relatively low average test-retest variability for the regions of interest in the DKT atlas. This pattern held true for both datasets and most regions. The cortical anatomical measurements demonstrate excellent reliability in terms of PCC and ICC. According to the grading ranges mentioned in Section

3.4, these measurements all fall into the excellent reliability category, with values above 0.90. However, there is one exception to this trend. FreeSurfer's estimates of the average cortical thickness had a slightly lower PCC value of around 0.85 for both datasets, placing it just outside the "excellent" reliability range.

One surprising finding was that FastSurfer generally outperformed FreeSurfer, demonstrating higher correlation coefficients and lower test-retest variability. This performance advantage was particularly evident in measurements with a larger range, such as the average cortical thickness.

However, there were certain regions where FastSurfer exhibited worse performance compared to FreeSurfer across both datasets. Examples of such regions include the medial orbitofrontal gyrus (*medialoribtofrontal*) and all the subcortical regions of interest related to the corpus callosum, except for the posterior corpus callosum, see Figures 4.4 and 4.8.

One notable result was that both FreeSurfer and FastSurfer exhibited poor performance when measuring the entorhinal cortex region. This observation is particularly concerning, as the entorhinal cortex is often one of the earliest brain regions to be affected by atrophy in Alzheimer's disease (AD), as mentioned in section 2.1.2.

Overall, FreeSurfer and FastSurfer exhibited marginally superior performance across all metrics when evaluated on the OASIS dataset. This enhanced performance can likely be attributed to the fact that the input data for the OASIS dataset consisted of an average of 3-4 intra-session scans. As discussed in Section 3.2, averaging multiple scans from the same session has been demonstrated to improve the reliability and accuracy of the results.

5.1.3 Inter-software Reliability

The purpose of the inter-software reliability test was to determine if the FastSurfer pipeline is generally able to recreate estimates obtained with FreeSurfer. Although the ground truth of the anatomical measurements in vivo are unknown, FreeSurfer is a widely used and accepted software in structural MRI, and thus, by measuring how similar measurements estimated by FastSurfer are to measurements estimated by FreeSurfer we get an idea if FastSurfer is a viable alternative.

From Figure 4.9 and Figure 4.10, we can see that almost all regions of interest in the DKT atlas have a high intraclass correlation coefficient. Most regions of interest have a coefficient around 0.9. Which regions that perform worse vary, some within the same dataset, depending on if the ICC was calculated for the first or the second scan. Overall, the inter-software reliability follow the same pattern as the intra-software ICC. For example, the enthorhinal cortex has lower ICC for all three measurements in the OASIS dataset, and worse for the grey matter volume and cortical surface area for the MIND dataset.

Regardless, considering that most of the regions fall into the category of excellent reliability, this indicates that FastSurfer is a viable option to FreeSurfer.

5.1.4 Visual Inspection

The visual inspection presented in Figure 4.11 shows that not only is the cortical reconstruction created by the white matter and pial surface quite good for this subject, it also shows that FreeSurfer and FastSurfer seem to give a similar result. In Figure 4.11, the red line shows the cortical reconstruction as created by FreeSurfer and the green line shows the same but for FastSurfer. When only one colour is visible it means that the two software are more or less identical.

The zoomed part of Figure 4.11 shows a close up inspection of a part of the right temporal lobe. Here, a relatively large difference between the segmentation of FreeSurfer and FastSurfer can be seen. The specific structure where the segmentation differs corresponds to the entorhinal cortex. This result is consistent with results from both the intra- and inter-software reliability tests performed, which indicated that this cortical region is difficult to estimate. Although the results presented in Section 4.1.4 only shows one slice from one subject, the shown slice serves as illustration for the rather abstract result presented in Sections 4.1.2 and 4.1.3.

5.2 Assessment of Brain Volumetric Changes in Type 2 Diabetes

Firstly, to determine in which regions to look for cortical atrophy in the MIND dataset, an *a priori* approach was adopted as outlined by Section 3.5.4. The results presented in Figures 4.12, 4.13 and 4.14 show that there were five cortical and three subcortical regions that were significant at the 0.05/31 and 0.05/22 level, respectively, when adjusted for ICV, age and gender in the OASIS dataset. For the cerebral cortex these were the parahippocampal gyrus, the superior, inferior and middle temporal gyrus as well as the enthorinal cortex. The subcortical regions that were significant with the same adjustments were the inferior lateral ventricle, the hippocampus and the amygdala.

These results agree with the literature. As mentioned in Section 2.1.2, the most commonly and severely affected regions of the brain due to AD related atrophy are the same ones we find in Figures 4.12, 4.13 and 4.14. As we expect these regions to show the most signs of atrophy in patients with AD, this also gives credence to FreeSurfer and FastSurfer as functional software as they are able to detect these changes.

Comparing Table 4.2 to Figures 4.12, 4.13, and 4.14, we can see that the only region in the MIND data that showed a significant effect at the 0.05 level, and also had a significant group-effect in the OASIS dataset, was the inferior temporal gyrus in regards to grey matter volume.

From literature, we expect to find larger differences in cerebral anatomy. Specifically in the hippocampus [85, 86]. We also expect to find it in the medial temporal lobe, with regions such as the entorhinal cortex showing the earliest signs of atrophy in AD patients.

One probable theory as to why a larger difference could not be detected is that the dataset used is rather small, only containing 10 subjects with T2DM and 15 HC. The initial atrophy in emerging neurodegeneration may not strictly adhere to a predefined spatial pattern. As such, it is possible that several of these patients has brain atrophy in different brain structures (or even in different hemispheres). Although the sample size included here did not allow us to detect any consistent changes across the brain, it is possible that there were indeed volumetric changes, but different subjects exhibited changes in different brain regions. This would effectively prevent us from detecting any volumetric changes in this small dataset. Of note, the primary reason for the small sample size is that the MIND-study was not primarily an MRI-study. MIND included PET imaging with two different radiotracers, and as PET imaging comes with radiation exposure, small sample sizes are often unavoidable. Nevertheless, another possible reason for why we did not detect any atrophy among T2DM group is that only participants that were cognitively unaffected were included in the MIND study. Having an Mini-Mental State Examination (MMSE) above or equal to 27 was an inclusion criteria in the MIND-study, which in essence mean that participants had no cognitive decline. It is therefore possible that the included T2DM subjects had no observable volumetric changes at the time of MR acquisition.

6
Conclusion

This thesis aimed to assess the potential differences in brain volumetrics between individuals with Type 2 Diabetes Mellitus and healthy controls using the two software FreeSurfer and FastSurfer. Despite previous literature indicating significant brain atrophy in T2DM patients, our findings did not reveal substantial differences in brain morphometry when comparing the two groups. Only one region, the inferior temporal gyrus, showed significant signs of atrophy.

The two software used, FreeSurfer and FastSurfer, demonstrated excellent intra-software reliability, with high values of Pearson's correlation coefficient and intra-class correlation coefficient. The two software also showed high inter-software reliability. This demonstrates that FastSurfer, which is not as extensively validated as FreeSurfer, can replicate the same results. This was of great importance to assess if FastSurfer is a valid alternative to FreeSurfer.

Furthermore, the regions of interest detected by both software as the most severely affected in the OASIS dataset correspond well with known literature on Alzheimer's disease-driven atrophy.

6.1 Future Work

The lack of detected differences between the T2DM patients and the healthy controls, when previous studies have detected it indicated the need for further studies in the field.

One potential improvement that could be implemented with the current dataset is the use of a higher resolution atlas, which contains more but smaller regions of interest. It is possible that the expected atrophy due to T2DM is more localised and does not significantly impact the average across the entire region in the DKT atlas used in the study. Alternatively, a vertex-to-vertex comparison could be employed, effectively serving as an extremely high-resolution atlas.

Another interesting future prospect the use of a longitudinal dataset instead of a cross-sectional dataset, as in this thesis. A longitudinal dataset would involve studying multiple patients over time, enabling individual assessment of anatomical

differences within the same time period. This approach would eliminate the need to correct for age, gender, and intracranial volume, as each patient would serve as their own baseline. As outlined in Section 2.1.1, age is negatively correlated with cortical thickness and grey matter volume, and it would be interesting to see if this decline is accelerated in patients with T2DM.

Furthermore, a longitudinal dataset would make it much easier to analyse anatomical differences, even if the atrophy is asymmetric.

If a larger dataset were acquired, and a clear difference in brain volumetrics could be established, the biomarker data from the MIND dataset would enable further research into understanding the metabolic factors driving brain atrophy.

Bibliography

[1] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, Sep. 2007, ISSN: 0898-929X. DOI: 10.1162/jocn.2007.19.9.1498. [Online]. Available: https://doi.org/10.1162/jocn.2007.19.9.1498 (visited on 01/18/2024).

[2] Y. Lim and J. Boster, "Obesity and Comorbid Conditions," eng, in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK574535/ (visited on 01/18/2024).

[3] M. Barbagallo and L. J. Dominguez, "Type 2 diabetes mellitus and Alzheimer's disease," *World Journal of Diabetes*, vol. 5, no. 6, pp. 889–893, Dec. 2014, ISSN: 1948-9358. DOI: 10.4239/wjd.v5.i6.889. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4265876/ (visited on 01/16/2024).

[4] P. T. James, R. Leach, E. Kalamara, and M. Shayeghi, "The worldwide obesity epidemic," *Obesity research*, vol. 9, no. S11, 228S–233S, 2001.

[5] M. A. B. Khan, M. J. Hashim, J. K. King, R. D. Govender, H. Mustafa, and J. Al Kaabi, "Epidemiology of type 2 diabetes – global burden of disease and forecasted trends," *Journal of Epidemiology and Global Health*, vol. 10, no. 1, p. 107, 2019, ISSN: 2210-6014. DOI: 10.2991/jegh.k.191028.001. [Online]. Available: http://dx.doi.org/10.2991/jegh.k.191028.001.

[6] "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the global burden of disease study 2019," *The Lancet. Public Health*, vol. 7, no. 2, e105–e125, Jan. 6, 2022, ISSN: 2468-2667. DOI: 10.1016/S2468-2667(21)00249-8. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8810394/ (visited on 06/04/2024).

[7] A. Serrano-Pozo, M. P. Frosch, E. Masliah, and B. T. Hyman, "Neuropathological Alterations in Alzheimer Disease," *Cold Spring Harbor Perspectives in Medicine:*, vol. 1, no. 1, a006189, Sep. 2011, ISSN: 2157-1422. DOI: 10.1101/cshperspect.a006189. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234452/ (visited on 01/16/2024).

[8] A. Bartos, D. Gregus, I. Ibrahim, and J. Tintěra, "Brain volumes and their ratios in Alzheimer´s disease on magnetic resonance imaging segmented using Freesurfer 6.0," *Psychiatry Research: Neuroimaging*, vol. 287, pp. 70–74, May 2019, ISSN: 0925-4927. DOI: 10.1016/j.pscychresns.2019.01.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925492718302324 (visited on 01/16/2024).

[9] G. Bernardes *et al.*, "Cortical and subcortical gray matter structural alterations in normoglycemic obese and type 2 diabetes patients: Relationship with adiposity, glucose, and insulin," en, *Metabolic Brain Disease*, vol. 33, no. 4, pp. 1211–1222, Aug. 2018, ISSN: 1573-7365. DOI: 10.1007/s11011-018-0223-5. [Online]. Available: https://doi.org/10.1007/s11011-018-0223-5 (visited on 01/22/2024).

[10] M. A. Espeland *et al.*, "Influence of Type 2 Diabetes on Brain Volumes and Changes in Brain Volumes: Results from the Women's Health Initiative Magnetic Resonance Imaging Studies," *Diabetes Care*, vol. 36, no. 1, pp. 90–97, Dec. 2012, ISSN: 0149-5992. DOI: 10.2337/dc12-0555. [Online]. Available: https://doi.org/10.2337/dc12-0555 (visited on 01/26/2024).

[11] 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, Feb. 1999, ISSN: 1053-8119. DOI: 10.1006/nimg.1998.0395. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811998903950 (visited on 01/22/2024).

[12] M. Goto *et al.*, "Advantages of Using Both Voxel- and Surface-based Morphometry in Cortical Morphology Analysis: A Review of Various Applications," *Magnetic Resonance in Medical Sciences*, vol. 21, no. 1, pp. 41–57, Feb. 2022, ISSN: 1347-3182. DOI: 10.2463/mrms.rev.2021-0096. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9199978/ (visited on 01/22/2024).

[13] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[14] L. Henschel, S. Conjeti, S. Estrada, K. Diers, B. Fischl, and M. Reuter, "Fastsurfer-a fast and accurate deep learning based neuroimaging pipeline," *NeuroImage*, vol. 219, p. 117 012, 2020.

[15] G. J. Tortora and B. H. Derrickson, *Introduction to the human body*. John Wiley & Sons, 2019.

[16] National Academy of Sciences, Institute of Medicine, and S. Ackerman, *Discovering the brain*. Washington, D.C., DC: National Academies Press, Jan. 1992.

[17] S. Herculano-Houzel, "The human brain in numbers: A linearly scaled-up primate brain," en, *Front Hum Neurosci*, vol. 3, p. 31, Nov. 2009.

[18] I. B. Malone *et al.*, "Accurate automatic estimation of total intracranial volume: A nuisance variable with less nuisance," *Neuroimage*, vol. 104, pp. 366–372, Jan. 2015, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2014.09.034. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4265726/ (visited on 05/22/2024).

[19] C. Hutton, E. De Vita, J. Ashburner, R. Deichmann, and R. Turner, "Voxel-based cortical thickness measurements in MRI," *Neuroimage*, vol. 40, no. 4, pp. 1701–1710, May 2008, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2008.01.027. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2330066/ (visited on 04/09/2024).

[20] K. Im, J.-M. Lee, O. Lyttelton, S. H. Kim, A. C. Evans, and S. I. Kim, "Brain Size and Cortical Structure in the Adult Human Brain," *Cerebral Cortex*, vol. 18, no. 9, pp. 2181–2191, Sep. 2008, ISSN: 1047-3211. DOI: 10.1093/

cercor/bhm244. [Online]. Available: https://doi.org/10.1093/cercor/bhm244 (visited on 05/16/2024).

[21] E. D. Gennatas *et al.*, "Age-Related Effects and Sex Differences in Gray Matter Density, Volume, Mass, and Cortical Thickness from Childhood to Young Adulthood," *The Journal of Neuroscience*, vol. 37, no. 20, pp. 5065–5073, May 2017, ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.3550-16.2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5444192/ (visited on 05/26/2024).

[22] A. A. Mercadante and P. Tadi, "Neuroanatomy, Gray Matter," eng, in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2024. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK553239/ (visited on 04/09/2024).

[23] K. Walhovd, H. Johansen-Berg, and R. Káradóttir, "Unraveling the secrets of white matter – Bridging the gap between cellular, animal and human imaging studies," *Neuroscience*, vol. 276, pp. 2–13, Sep. 2014, ISSN: 0306-4522. DOI: 10.1016/j.neuroscience.2014.06.058. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4155933/ (visited on 04/09/2024).

[24] P. Morell and R. H. Quarles, "The Myelin Sheath," en, in *Basic Neurochemistry: Molecular, Cellular and Medical Aspects. 6th edition*, Lippincott-Raven, 1999. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK27954/ (visited on 04/09/2024).

[25] L. Heimer and G. W. Van Hoesen, "The limbic lobe and its output channels: Implications for emotional functions and adaptive behavior," *Neuroscience & Biobehavioral Reviews*, The Limbic Brain: Structure and Function, vol. 30, no. 2, pp. 126–147, Jan. 2006, ISSN: 0149-7634. DOI: 10.1016/j.neubiorev.2005.06.006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0149763405001077 (visited on 05/16/2024).

[26] K. S. Anand and V. Dhikav, "Hippocampus in health and disease: An overview," *Annals of Indian Academy of Neurology*, vol. 15, no. 4, pp. 239–246, 2012, ISSN: 0972-2327. DOI: 10.4103/0972-2327.104323. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3548359/ (visited on 05/03/2024).

[27] C. Scelsi, T. Rahim, J. Morris, G. Kramer, B. Gilbert, and S. Forseen, "The Lateral Ventricles: A Detailed Review of Anatomy, Development, and Anatomic Variations," *AJNR: American Journal of Neuroradiology*, vol. 41, no. 4, pp. 566–572, Apr. 2020, ISSN: 0195-6108. DOI: 10.3174/ajnr.A6456. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7144651/ (visited on 05/03/2024).

[28] B. N. Dugger and D. W. Dickson, "Pathology of Neurodegenerative Diseases," *Cold Spring Harbor Perspectives in Biology*, vol. 9, no. 7, a028035, Jul. 2017, ISSN: 1943-0264. DOI: 10.1101/cshperspect.a028035. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5495060/ (visited on 05/13/2024).

[29] T. A. Yacoubian, "Chapter 1 - Neurodegenerative Disorders: Why Do We Need New Therapies?" In *Drug Discovery Approaches for the Treatment of Neurodegenerative Disorders*, A. Adejare, Ed., Academic Press, Jan. 2017, pp. 1–16,

ISBN: 978-0-12-802810-0. DOI: 10.1016/B978-0-12-802810-0.00001-5. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128028100000015 (visited on 05/13/2024).

[30] M. A. DeTure and D. W. Dickson, "The neuropathological diagnosis of Alzheimer's disease," *Molecular Neurodegeneration*, vol. 14, no. 1, p. 32, Aug. 2019, ISSN: 1750-1326. DOI: 10.1186/s13024-019-0333-5. [Online]. Available: https://doi.org/10.1186/s13024-019-0333-5 (visited on 05/13/2024).

[31] J. L. Whitwell, "Progression of Atrophy in Alzheimer's Disease and Related Disorders," en, *Neurotoxicity Research*, vol. 18, no. 3, pp. 339–346, Nov. 2010, ISSN: 1476-3524. DOI: 10.1007/s12640-010-9175-1. [Online]. Available: https://doi.org/10.1007/s12640-010-9175-1 (visited on 05/13/2024).

[32] K. M. Igarashi, "Entorhinal cortex dysfunction in Alzheimer's disease," *Trends in neurosciences*, vol. 46, no. 2, pp. 124–136, Feb. 2023, ISSN: 0166-2236. DOI: 10.1016/j.tins.2022.11.006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9877178/ (visited on 05/27/2024).

[33] S. P. Poulin, R. Dautoff, J. C. Morris, L. F. Barrett, and B. C. Dickerson, "Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity," *Psychiatry research*, vol. 194, no. 1, pp. 7–13, Oct. 2011, ISSN: 0165-1781. DOI: 10.1016/j.pscychresns.2011.06.014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3185127/ (visited on 05/17/2024).

[34] L. A. van de Pol, A. Hensel, F. Barkhof, H. J. Gertz, P. Scheltens, and W. M. van der Flier, "Hippocampal atrophy in Alzheimer disease: Age matters," *Neurology*, vol. 66, no. 2, pp. 236–238, Jan. 2006, Publisher: Wolters Kluwer. DOI: 10.1212/01.wnl.0000194240.47892.4d. [Online]. Available: https://www.neurology.org/doi/10.1212/01.wnl.0000194240.47892.4d (visited on 05/17/2024).

[35] L. C. Walker, "Aβ plaques," *Free Neuropathology*, vol. 1, pp. 1–31, ISSN: 2699-4445. DOI: 10.17879/freeneuropathology-2020-3025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7745791/ (visited on 05/17/2024).

[36] J. W. Kinney, S. M. Bemiller, A. S. Murtishaw, A. M. Leisgang, A. M. Salazar, and B. T. Lamb, "Inflammation as a central mechanism in Alzheimer's disease," *Alzheimer's & Dementia : Translational Research & Clinical Interventions*, vol. 4, pp. 575–590, Sep. 2018, ISSN: 2352-8737. DOI: 10.1016/j.trci.2018.06.014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6214864/ (visited on 06/02/2024).

[37] J. P. Brion *et al.*, "Neurofibrillary tangles and tau phosphorylation," eng, *Biochemical Society Symposium*, no. 67, pp. 81–88, 2001, ISSN: 0067-8694. DOI: 10.1042/bss0670081.

[38] J. M. Olichney *et al.*, "The apolipoprotein E epsilon 4 allele is associated with increased neuritic plaques and cerebral amyloid angiopathy in Alzheimer's disease and Lewy body variant," *Neurology*, vol. 47, no. 1, pp. 190–196, Jul. 1996, Publisher: Wolters Kluwer. DOI: 10.1212/WNL.47.1.190. [Online]. Available: https://www.neurology.org/doi/abs/10.1212/wnl.47.1.190 (visited on 05/17/2024).

[39] K. A. Josephs, Y. Tsuboi, N. Cookson, H. Watt, and D. W. Dickson, "Apolipoprotein E ϵ4 Is a Determinant for Alzheimer-Type Pathologic Features in Tauopathies, Synucleinopathies, and Frontotemporal Degeneration," *Archives of Neurology*, vol. 61, no. 10, pp. 1579–1584, Oct. 2004, ISSN: 0003-9942. DOI: 10.1001/archneur.61.10.1579. [Online]. Available: https://doi.org/10.1001/archneur.61.10.1579 (visited on 05/17/2024).

[40] J. Janson, T. Laedtke, J. E. Parisi, P. O'Brien, R. C. Petersen, and P. C. Butler, "Increased risk of type 2 diabetes in Alzheimer disease," eng, *Diabetes*, vol. 53, no. 2, pp. 474–481, Feb. 2004, ISSN: 0012-1797. DOI: 10.2337/diabetes.53.2.474.

[41] M. Barbagallo and L. J. Dominguez, "Type 2 diabetes mellitus and Alzheimer's disease," *World Journal of Diabetes*, vol. 5, no. 6, pp. 889–893, Dec. 2014, ISSN: 1948-9358. DOI: 10.4239/wjd.v5.i6.889. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4265876/ (visited on 05/26/2024).

[42] R. Hamzé *et al.*, "Type 2 Diabetes Mellitus and Alzheimer's Disease: Shared Molecular Mechanisms and Potential Common Therapeutic Targets," eng, *International Journal of Molecular Sciences*, vol. 23, no. 23, p. 15 287, Dec. 2022, ISSN: 1422-0067. DOI: 10.3390/ijms232315287.

[43] B. Cholerton, L. D. Baker, T. J. Montine, and S. Craft, "Type 2 Diabetes, Cognition, and Dementia in Older Adults: Toward a Precision Health Approach," *Diabetes Spectrum : A Publication of the American Diabetes Association*, vol. 29, no. 4, pp. 210–219, Nov. 2016, ISSN: 1040-9165. DOI: 10.2337/ds16-0041. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5111529/ (visited on 05/26/2024).

[44] S. M. de la Monte, "Insulin resistance and neurodegeneration: Progress towards the development of new therapeutics for alzheimer's disease," *Drugs*, vol. 77, no. 1, pp. 47–65, Jan. 2017, ISSN: 0012-6667. DOI: 10.1007/s40265-016-0674-0. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5575843/ (visited on 06/04/2024).

[45] G. Katti, S. A. Ara, and A. Shireen, "Magnetic resonance imaging (mri)–a review," *International journal of dental clinics*, vol. 3, no. 1, pp. 65–70, 2011.

[46] M. A. Rocca, G. Comi, and M. Filippi, "The Role of T1-Weighted Derived Measures of Neurodegeneration for Assessing Disability Progression in Multiple Sclerosis," English, *Frontiers in Neurology*, vol. 8, Sep. 2017, Publisher: Frontiers, ISSN: 1664-2295. DOI: 10.3389/fneur.2017.00433. [Online]. Available: https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2017.00433/full (visited on 05/15/2024).

[47] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," en, *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67–77, Feb. 2010, Publisher: Nature Publishing Group, ISSN: 1759-4766. DOI: 10.1038/nrneurol.2009.215. [Online]. Available: https://www.nature.com/articles/nrneurol.2009.215 (visited on 05/15/2024).

[48] E. N. Manson, A. N. Mumuni, S. Inkoom, and I. Shirazu, "Evaluation of the impact of magnetic field homogeneity on image quality in magnetic resonance imaging: A baseline quantitative study at 1.5 T," *Egyptian Journal*

of Radiology and Nuclear Medicine, vol. 54, no. 1, p. 146, Sep. 2023, ISSN: 2090-4762. DOI: 10.1186/s43055-023-01097-8. [Online]. Available: https://doi.org/10.1186/s43055-023-01097-8 (visited on 05/16/2024).

[49] J. Juntu, J. Sijbers, D. Dyck, and J. Gielen, "Bias Field Correction for MRI Images," Jan. 2005, pp. 543–551, ISBN: 978-3-540-25054-8. DOI: 10.1007/3-540-32390-2_64.

[50] M. Ono, S. Kubik, and C. Abernathey, *Atlas of the Cerebral Sulci* (Thieme Publishers Series). G. Thieme Verlag, 1990, ISBN: 978-3-13-732101-9. [Online]. Available: https://books.google.se/books?id=-MakzgEACAAJ.

[51] K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth, 1909.

[52] P. K. Mandal, R. Mahajan, and I. D. Dinov, "Structural Brain Atlases: Design, Rationale, and Applications in Normal and Pathological Cohorts," *Journal of Alzheimer's disease : JAD*, vol. 31, no. 0 3, S169–S188, 2012, ISSN: 1387-2877. DOI: 10.3233/JAD-2012-120412. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4324755/ (visited on 04/12/2024).

[53] A. Klein and J. Tourville, "101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol," English, *Frontiers in Neuroscience*, vol. 6, Dec. 2012, Publisher: Frontiers, ISSN: 1662-453X. DOI: 10.3389/fnins.2012.00171. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2012.00171/full (visited on 04/12/2024).

[54] R. S. Desikan *et al.*, "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," eng, *NeuroImage*, vol. 31, no. 3, pp. 968–980, Jul. 2006, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2006.01.021.

[55] A. Evans, D. Collins, S. Mills, E. Brown, R. Kelly, and T. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*, Oct. 1993, 1813–1817 vol.3. DOI: 10.1109/NSSMIC.1993.373602. [Online]. Available: https://ieeexplore.ieee.org/document/373602 (visited on 05/15/2024).

[56] C. W. Nordahl *et al.*, "Cortical folding abnormalities in autism revealed by surface-based morphometry," *Journal of Neuroscience*, vol. 27, no. 43, pp. 11 725–11 735, 2007.

[57] H.-J. Ahn *et al.*, "The cortical neuroanatomy of neuropsychological deficits in mild cognitive impairment and alzheimer's disease: A surface-based morphometric analysis," *Neuropsychologia*, vol. 49, no. 14, pp. 3931–3945, 2011.

[58] Z. Wu, Y. Peng, M. Hong, and Y. Zhang, "Gray matter deterioration pattern during alzheimer's disease progression: A regions-of-interest based surface morphometry study," *Frontiers in aging neuroscience*, vol. 13, p. 593 898, 2021.

[59] L. Palaniyappan and P. F. Liddle, "Aberrant cortical gyrification in schizophrenia: A surface-based morphometry study," *Journal of Psychiatry and Neuroscience*, vol. 37, no. 6, pp. 399–406, 2012.

[60] 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, Feb. 1999, ISSN: 1053-8119. DOI: 10.1006/nimg.1998.0395. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811998903950 (visited on 05/15/2024).

[61] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, no. 2, pp. 143–156, Jun. 2001, ISSN: 1361-8415. DOI: 10.1016/S1361-8415(01)00036-6. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841501000366 (visited on 05/15/2024).

[62] B. Fischl *et al.*, "Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain," eng, *Neuron*, vol. 33, no. 3, pp. 341–355, Jan. 2002, ISSN: 0896-6273. DOI: 10.1016/s0896-6273(02)00569-x.

[63] B. Fischl *et al.*, "Automatically parcellating the human cerebral cortex," eng, *Cerebral Cortex (New York, N.Y.: 1991)*, vol. 14, no. 1, pp. 11–22, Jan. 2004, ISSN: 1047-3211. DOI: 10.1093/cercor/bhg087.

[64] B. Fischl, M. I. Sereno, R. B. Tootell, and A. M. Dale, "High-resolution intersubject averaging and a coordinate system for the cortical surface," eng, *Human Brain Mapping*, vol. 8, no. 4, pp. 272–284, 1999, ISSN: 1065-9471. DOI: 10.1002/(sici)1097-0193(1999)8:4<272::aid-hbm10>3.0.co;2-4.

[65] L. Henschel, S. Conjeti, S. Estrada, K. Diers, B. Fischl, and M. Reuter, "FastSurfer - A fast and accurate deep learning based neuroimaging pipeline," *NeuroImage*, vol. 219, p. 117 012, Oct. 2020, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2020.117012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811920304985 (visited on 05/27/2024).

[66] E. P. Hedges *et al.*, "Reliability of structural MRI measurements: The effects of scan session, head tilt, inter-scan interval, acquisition sequence, FreeSurfer version and processing stream," *NeuroImage*, vol. 246, p. 118 751, Feb. 2022, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2021.118751. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811921010235 (visited on 04/09/2024).

[67] J. Veraart, E. Fieremans, and D. S. Novikov, "Diffusion MRI noise mapping using random matrix theory," en, *Magnetic Resonance in Medicine*, vol. 76, no. 5, pp. 1582–1593, 2016, ISSN: 1522-2594. DOI: 10.1002/mrm.26059. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.26059 (visited on 06/03/2024).

[68] T. K. Koo and M. Y. Li, "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research," *Journal of Chiropractic Medicine*, vol. 15, no. 2, pp. 155–163, Jun. 2016, ISSN: 1556-3707. DOI: 10.1016/j.jcm.2016.02.012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4913118/ (visited on 04/12/2024).

[69] M. Schain *et al.*, "59th EASD Annual Meeting of the European Association for the Study of Diabetes," *Diabetologia*, vol. 66, no. 1, S38, Sep. 2023, ISSN: 1432-0428. DOI: 10.1007/s00125-023-05969-6. [Online]. Available: https://doi.org/10.1007/s00125-023-05969-6.

[70] J. C. Morris, "The Clinical Dementia Rating (CDR): Current version and scoring rules," eng, *Neurology*, vol. 43, no. 11, pp. 2412–2414, Nov. 1993, ISSN: 0028-3878. DOI: 10.1212/wnl.43.11.2412-a.

[71] M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl, "Within-subject template estimation for unbiased longitudinal image analysis," *NeuroImage*, vol. 61, no. 4, pp. 1402–1418, Jul. 2012, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2012.02.084. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811912002765 (visited on 05/22/2024).

[72] E. Haddad *et al.*, "Multisite test–retest reliability and compatibility of brain metrics derived from FreeSurfer versions 7.1, 6.0, and 5.3," en, *Human Brain Mapping*, vol. 44, no. 4, pp. 1515–1532, 2023, ISSN: 1097-0193. DOI: 10.1002/hbm.26147. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.26147 (visited on 04/09/2024).

[73] E. H. B. M. Gronenschild *et al.*, "The Effects of FreeSurfer Version, Workstation Type, and Macintosh Operating System Version on Anatomical Volume and Cortical Thickness Measurements," en, *PLOS ONE*, vol. 7, no. 6, e38234, Jun. 2012, Publisher: Public Library of Science, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0038234. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0038234 (visited on 04/29/2024).

[74] R. Seiger, S. Ganger, G. S. Kranz, A. Hahn, and R. Lanzenberger, "Cortical Thickness Estimations of FreeSurfer and the CAT12 Toolbox in Patients with Alzheimer's Disease and Healthy Controls," en, *Journal of Neuroimaging*, vol. 28, no. 5, pp. 515–523, 2018, ISSN: 1552-6569. DOI: 10.1111/jon.12521. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/jon.12521 (visited on 05/02/2024).

[75] P. Giannakopoulos, P. R. Hof, and C. Bouras, "Alzheimer's disease with asymmetric atrophy of the cerebral hemispheres: Morphometric analysis of four cases," eng, *Acta Neuropathologica*, vol. 88, no. 5, pp. 440–447, 1994, ISSN: 0001-6322. DOI: 10.1007/BF00389496.

[76] O. Bugiani, J. Constantinidis, B. Ghetti, C. Bouras, and F. Tagliavini, "Asymmetrical cerebral atrophy in Alzheimer's disease," eng, *Clinical Neuropathology*, vol. 10, no. 2, pp. 55–60, 1991, ISSN: 0722-5091.

[77] V. Vahermaa *et al.*, "FreeSurfer 7 quality control: Key problem areas and importance of manual corrections," *NeuroImage*, vol. 279, p. 120 306, Oct. 2023, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2023.120306. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811923004573 (visited on 05/21/2024).

[78] J. J. Berman, "Chapter 4 - Understanding Your Data," in *Data Simplification*, J. J. Berman, Ed., Boston: Morgan Kaufmann, Jan. 2016, pp. 135–187, ISBN: 978-0-12-803781-2. DOI: 10.1016/B978-0-12-803781-2.00004-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128037812000047 (visited on 06/14/2024).

[79] K. Mcgraw and S. Wong, "Forming inferences about some intraclass correlation coefficients," *Psychological Methods*, vol. 1, pp. 30–46, Mar. 1996. DOI: 10.1037/1082-989X.1.1.30.

[80] T. Cole and D. Altman, "Statistics notes: What is a percentage difference?" *BMJ*, vol. 358, j3663, Aug. 2017. DOI: 10.1136/bmj.j3663.

[81] O. Voevodskaya *et al.*, "The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease," English, *Frontiers in Aging Neuroscience*, vol. 6, Oct. 2014, Publisher: Frontiers, ISSN: 1663-4365. DOI: 10.3389/fnagi.2014.00264. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnagi.2014.00264 (visited on 05/21/2024).

[82] A. Rutherford, *ANOVA and ANCOVA: a GLM approach*. John Wiley & Sons, 2012.

[83] E. W. Weisstein, "Bonferroni correction," *https://mathworld. wolfram. com/*, 2004.

[84] L. L. Backhausen, M. M. Herting, C. K. Tamnes, and N. C. Vetter, "Best Practices in Structural Neuroimaging of Neurodevelopmental Disorders," en, *Neuropsychology Review*, vol. 32, no. 2, pp. 400–418, Jun. 2022, ISSN: 1573-6660. DOI: 10.1007/s11065-021-09496-2. [Online]. Available: https://doi.org/10.1007/s11065-021-09496-2 (visited on 02/29/2024).

[85] C. Moran *et al.*, "Brain atrophy in type 2 diabetes," *Diabetes Care*, vol. 36, no. 12, pp. 4036–4042, Dec. 2013, ISSN: 0149-5992. DOI: 10.2337/dc13-0143. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3836136/ (visited on 06/04/2024).

[86] T. Zhang, M. Shaw, and N. Cherbuin, "Association between type 2 diabetes mellitus and brain atrophy: A meta-analysis," *Diabetes & Metabolism Journal*, vol. 46, no. 5, pp. 781–802, Sep. 2022, ISSN: 2233-6079. DOI: 10.4093/dmj.2021.0189. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9532183/ (visited on 06/04/2024).

A
Appendix

The Appendix contains the full table of all the ROIs in the DKT atlas.

A.1 Table of all cortical regions in the DKT atlas

Table A.1: Cortical regions in the DKT atlas [53] and their corresponding name in the output of FreeSurfer & FastSurfer.

Lobe	Cortical region	Output name
Temporal lobe	Entorhinal cortex	*entorhinal*
	Parahippocampal gyrus	*parahippocampal*
	Fusiform gyrus	*fusiform*
	Superior temporal gyrus	*superiortemporal*
	Middle temporal gyrus	*middletemporal*
	Inferior temporal gyrus	*inferiortemporal*
	Transverse temporal gyrus	*transversetemporal*
Frontal lobe	Superior frontal	*superiorfrontal*
	Middle frontal gyrus, *Rostral*	*rostralmiddlefrontal*
	Middle frontal gyrus, *Caudal*	*caudalmiddlefrontal*
	Inferior frontal gyrus, *Pars opercularis*	*parsopercularis*
	Inferior frontal gyrus, *Pars triangularis*	*parstriangularis*
	Inferior frontal gyrus, *Pars orbitalis*	*parsorbitalis*
	Orbitofrontal gyrus, *Lateral*	*lateralorbitofrontal*
	Orbitofrontal gyrus, *Medial*	*medialorbitofrontal*
	Precentral gyrus	*precentral*
	Paracentral lobule	*paracentral*
Parietal lobe	Postcentral gyrus	*postcentral*
	Supramarginal gyrus	*supramarginal*
	Superior parietal lobule	*superiorparietal*
	Inferior parietal lobule	*inferiorparietal*
	Precuneus	*precuneus*
Occipital lobe	Lingual gyrus	*lingual*
	Pericalcarine cortex	*pericalcarine*
	Cuneus cortex	*cuneus*
	Lateral occipital cortex	*lateraloccipital*
Limbic lobe	Cingulate, *Rostral anterior*	*rostralanteriorcingulate*
	Cingulate, *Caudal anterior*	*caudalanteriorcingulate*
	Cingulate, *Posterior*	*posteriorcingulate*
	Cingulate, *Isthmus*	*isthmuscingulate*
	Insula	*insula*

Table A.2: All subcortical regions included in the analysis and their corresponding name in the output of FreeSurfer & FastSurfer.

Subcortical region	Output name
3rd-Ventricle	3rd-Ventricle
4th-Ventricle	4th-Ventricle
Nucleus accumbens	Accumbens-area
Amygdala	Amygdala
Brain stem	Brain-Stem
Corpus callosum, *anterior*	CC_Anterior
Corpus callosum, *central*	CC_Central
Corpus callosum, *mid anterior*	CC_Mid_Anterior
Corpus callosum, *mid posterior*	CC_Mid_Posterior
Corpus callosum, *posterior*	CC_Posterior
Cerebrospinal fluid	CSF
Caudate nucleus	Caudate
Cerebellum, *cortex*	Cerebellum-Cortex
Cerebellum, *white matter*	Cerebellum-White-Matter
Hippocampus	Hippocampus
Lateral ventricle, *inferior*	Inf-Lat-Vent
Lateral ventricle	Lateral-Ventricle
Globus pallidus	Pallidum
Putamen	Putamen
Thalamus	Thalamus
Ventral diencephalon	VentralDC
Choroid plexus	choroid-plexus

B

Appendix

Appendix B contains all the data from the intra-software reliability testing. There is one table for each software and measurement (average cortical thickness, average grey matter volume, and surface area).

B.1 Intra-software reliability

Table B.1: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average cortical thickness measurements obtained from FreeSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.8140	0.5806	0.9238	0.8864	0.7100	0.9500	1.8658	1.1603	2.5712
caudalmiddlefrontal	0.7519	0.4636	0.8962	0.8607	0.6500	0.9400	1.9419	1.3190	2.5647
cuneus	0.9494	0.8742	0.9801	0.9732	0.9300	0.9900	1.4559	0.9131	1.9987
entorhinal	0.9006	0.7615	0.9604	0.9453	0.8600	0.9800	2.8328	1.5318	4.1337
fusiform	0.8170	0.5864	0.9250	0.8894	0.7200	0.9600	1.7245	1.0774	2.3715
inferiorparietal	0.9111	0.7852	0.9647	0.9525	0.8800	0.9800	1.1181	0.7171	1.5192
inferiortemporal	0.8627	0.6796	0.9446	0.9159	0.7900	0.9700	1.8422	1.1160	2.5685
insula	0.8383	0.6291	0.9342	0.8982	0.7300	0.9600	2.0895	1.3653	2.8137
isthmuscingulate	0.9364	0.8433	0.9749	0.9655	0.9100	0.9900	1.8607	1.2978	2.4235
lateraloccipital	0.9342	0.8383	0.9741	0.9658	0.9200	0.9900	1.1610	0.7491	1.5730
lateralorbitofrontal	0.8471	0.6472	0.9380	0.9176	0.7900	0.9700	1.7757	1.0680	2.4834
lingual	0.8378	0.6281	0.9340	0.9133	0.7800	0.9700	1.8963	1.2267	2.5659
medialorbitofrontal	0.8674	0.6896	0.9466	0.9207	0.8000	0.9700	2.3177	1.4729	3.1624
middletemporal	0.8184	0.5893	0.9257	0.8919	0.7300	0.9600	1.6171	1.0267	2.2075
paracentral	0.8634	0.6812	0.9449	0.9295	0.8200	0.9700	2.0040	1.4735	2.5346
parahippocampal	0.8622	0.6786	0.9444	0.9276	0.8200	0.9700	2.9905	2.0157	3.9654
parsopercularis	0.8261	0.6046	0.9290	0.8992	0.7500	0.9600	1.6032	1.0808	2.1256
parsorbitalis	0.8676	0.6900	0.9467	0.9324	0.8300	0.9700	2.0885	1.4435	2.7335
parstriangularis	0.8138	0.5802	0.9237	0.9001	0.7500	0.9600	1.7465	1.1541	2.3388
pericalcarine	0.8329	0.6183	0.9320	0.9103	0.7800	0.9600	3.0256	1.9996	4.0516
postcentral	0.9291	0.8263	0.9720	0.9612	0.9000	0.9800	1.4863	0.8588	2.1137
posteriorcingulate	0.6774	0.3354	0.8616	0.8090	0.5100	0.9200	2.3584	1.4884	3.2284
precentral	0.7723	0.5011	0.9054	0.8666	0.6600	0.9500	1.6904	0.6209	2.7600
precuneus	0.9160	0.7962	0.9667	0.9545	0.8900	0.9800	1.3360	0.8595	1.8126
rostralanteriorcingulate	0.7214	0.4096	0.8823	0.7986	0.4700	0.9200	3.2399	1.8746	4.6052
rostralmiddlefrontal	0.8849	0.7272	0.9539	0.9331	0.8300	0.9700	1.7597	1.1310	2.3884
superiorfrontal	0.8266	0.6055	0.9292	0.9052	0.7600	0.9600	1.8882	1.2135	2.5629
superiorparietal	0.9086	0.7795	0.9637	0.9536	0.8800	0.9800	1.5526	1.0177	2.0874
superiortemporal	0.8282	0.6088	0.9299	0.8856	0.7200	0.9500	1.6014	1.0647	2.1382
supramarginal	0.9238	0.8140	0.9698	0.9605	0.9000	0.9800	1.1790	0.7809	1.5771
transversetemporal	0.8886	0.7352	0.9554	0.9423	0.8600	0.9800	2.3261	1.5878	3.0644

Table B.2: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average grey matter volume measurements obtained from FreeSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9619	0.9043	0.9851	0.9776	0.9400	0.9900	2.8786	1.7125	4.0448
caudalmiddlefrontal	0.9510	0.8779	0.9808	0.9725	0.9300	0.9900	3.0514	1.5192	4.5836
cuneus	0.9849	0.9614	0.9941	0.9914	0.9800	1.0000	2.4386	1.6837	3.1935
entorhinal	0.7836	0.5221	0.9104	0.8790	0.6900	0.9500	5.1561	3.4441	6.8681
fusiform	0.9745	0.9353	0.9901	0.9858	0.9600	0.9900	2.3487	1.5321	3.1654
inferiorparietal	0.9773	0.9423	0.9912	0.9890	0.9700	1.0000	2.0633	1.2526	2.8740
inferiortemporal	0.9642	0.9100	0.9860	0.9820	0.9600	0.9900	2.6785	1.9272	3.4298
insula	0.9872	0.9672	0.9950	0.9914	0.9700	1.0000	1.7371	1.1120	2.3623
isthmuscingulate	0.9636	0.9084	0.9858	0.9807	0.9500	0.9900	2.7958	1.8023	3.7893
lateraloccipital	0.9837	0.9584	0.9937	0.9914	0.9800	1.0000	1.7936	1.1804	2.4069
lateralorbitofrontal	0.9495	0.8744	0.9802	0.9746	0.9400	0.9900	2.5701	1.6036	3.5366
lingual	0.9887	0.9709	0.9956	0.9928	0.9800	1.0000	2.3252	1.3814	3.2690
medialorbitofrontal	0.9649	0.9115	0.9863	0.9782	0.9400	0.9900	3.1565	2.1567	4.1563
middletemporal	0.9766	0.9405	0.9909	0.9868	0.9700	0.9900	2.2876	1.6738	2.9013
paracentral	0.9733	0.9323	0.9896	0.9867	0.9700	0.9900	2.2830	1.6552	2.9107
parahippocampal	0.8821	0.7209	0.9527	0.9263	0.8000	0.9700	3.5798	2.3562	4.8034
parsopercularis	0.9891	0.9720	0.9958	0.9933	0.9800	1.0000	2.0110	1.4625	2.5595
parsorbitalis	0.9638	0.9089	0.9858	0.9825	0.9600	0.9900	2.3031	1.3399	3.2663
parstriangularis	0.9779	0.9437	0.9914	0.9886	0.9700	1.0000	2.6896	1.9104	3.4687
pericalcarine	0.9790	0.9467	0.9919	0.9874	0.9600	1.0000	3.2568	2.1647	4.3490
postcentral	0.9835	0.9578	0.9936	0.9912	0.9800	1.0000	1.5845	1.0968	2.0723
posteriorcingulate	0.9677	0.9184	0.9874	0.9843	0.9600	0.9900	2.4203	1.6302	3.2104
precentral	0.9848	0.9613	0.9941	0.9919	0.9800	1.0000	1.3186	0.8443	1.7928
precuneus	0.9620	0.9046	0.9852	0.9815	0.9500	0.9900	1.9471	1.0272	2.8671
rostralanteriorcingulate	0.9722	0.9297	0.9892	0.9834	0.9600	0.9900	3.2226	1.2643	5.1809
rostralmiddlefrontal	0.9791	0.9468	0.9919	0.9888	0.9700	1.0000	2.3628	1.7525	2.9731
superiorfrontal	0.9603	0.9004	0.9845	0.9805	0.9500	0.9900	2.2659	1.5512	2.9806
superiorparietal	0.8535	0.6604	0.9407	0.9219	0.8100	0.9700	3.2641	0.4960	6.0323
superiortemporal	0.9867	0.9659	0.9948	0.9933	0.9800	1.0000	1.5459	1.1178	1.9741
supramarginal	0.9756	0.9381	0.9905	0.9870	0.9700	0.9900	2.1260	0.9357	3.3164
transversetemporal	0.9788	0.9461	0.9918	0.9878	0.9700	1.0000	2.9499	2.1851	3.7146

Table B.3: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average surface area measurements obtained from FreeSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9586	0.8963	0.9838	0.9787	0.9500	0.9900	2.4799	1.4567	3.5031
caudalmiddlefrontal	0.9475	0.8695	0.9794	0.9691	0.9200	0.9900	1.8773	0.1660	3.5886
cuneus	0.9939	0.9842	0.9976	0.9964	0.9900	1.0000	1.3182	0.8580	1.7784
entorhinal	0.9314	0.8318	0.9729	0.9653	0.9100	0.9900	3.7733	2.7189	4.8276
fusiform	0.9902	0.9749	0.9962	0.9953	0.9900	1.0000	1.0401	0.5618	1.5184
inferiorparietal	0.9851	0.9619	0.9942	0.9924	0.9800	1.0000	1.2874	0.6334	1.9414
inferiortemporal	0.9830	0.9566	0.9934	0.9899	0.9700	1.0000	1.5127	0.8366	2.1889
insula	0.9788	0.9461	0.9918	0.9891	0.9700	1.0000	1.5105	0.9768	2.0443
isthmuscingulate	0.9864	0.9652	0.9947	0.9934	0.9800	1.0000	1.6896	1.1201	2.2591
lateraloccipital	0.9813	0.9524	0.9927	0.9910	0.9800	1.0000	1.2805	0.7189	1.8421
lateralorbitofrontal	0.9499	0.8753	0.9803	0.9748	0.9400	0.9900	2.1273	1.1673	3.0873
lingual	0.9944	0.9857	0.9978	0.9973	0.9900	1.0000	1.1419	0.7608	1.5230
medialorbitofrontal	0.9621	0.9047	0.9852	0.9780	0.9400	0.9900	2.5775	1.6304	3.5246
middletemporal	0.9912	0.9774	0.9966	0.9955	0.9900	1.0000	1.1675	0.7877	1.5473
paracentral	0.9841	0.9594	0.9938	0.9900	0.9800	1.0000	1.4697	0.8728	2.0667
parahippocampal	0.8445	0.6418	0.9369	0.9126	0.7800	0.9700	2.7918	1.0778	4.5058
parsopercularis	0.9909	0.9767	0.9965	0.9953	0.9900	1.0000	1.4870	1.0898	1.8842
parsorbitalis	0.9658	0.9138	0.9866	0.9830	0.9600	0.9900	1.6729	1.0753	2.2706
parstriangularis	0.9921	0.9797	0.9969	0.9960	0.9900	1.0000	1.5446	1.0386	2.0505
pericalcarine	0.9928	0.9815	0.9972	0.9952	0.9900	1.0000	1.6599	1.1194	2.2004
postcentral	0.9897	0.9735	0.9960	0.9948	0.9900	1.0000	0.7850	0.3841	1.1860
posteriorcingulate	0.9910	0.9770	0.9965	0.9956	0.9900	1.0000	1.2575	0.8895	1.6256
precentral	0.9222	0.8104	0.9692	0.9541	0.8900	0.9800	1.2274	-0.1185	2.5734
precuneus	0.9870	0.9668	0.9950	0.9909	0.9800	1.0000	1.1735	0.4541	1.8929
rostralanteriorcingulate	0.9324	0.8341	0.9733	0.9662	0.9100	0.9900	3.8036	2.7145	4.8928
rostralmiddlefrontal	0.9934	0.9831	0.9975	0.9967	0.9900	1.0000	1.0477	0.6962	1.3992
superiorfrontal	0.9821	0.9543	0.9930	0.9904	0.9800	1.0000	0.7627	0.1844	1.3411
superiorparietal	0.8934	0.7457	0.9574	0.9427	0.8600	0.9800	1.9978	-0.4531	4.4486
superiortemporal	0.9922	0.9799	0.9970	0.9953	0.9900	1.0000	0.9864	0.5645	1.4082
supramarginal	0.9855	0.9630	0.9944	0.9909	0.9800	1.0000	1.7017	0.8266	2.5769
transversetemporal	0.9843	0.9599	0.9939	0.9913	0.9800	1.0000	2.0263	1.3929	2.6596

Table B.4: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average cortical thickness measurements obtained from FastSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9293	0.8268	0.9721	0.9636	0.9100	0.9900	1.2188	0.7414	1.6961
caudalmiddlefrontal	0.8883	0.7345	0.9553	0.9363	0.8400	0.9700	1.1688	0.7463	1.5913
cuneus	0.9820	0.9541	0.9930	0.9898	0.9700	1.0000	0.9648	0.7094	1.2202
entorhinal	0.9256	0.8182	0.9706	0.9602	0.9000	0.9800	2.0434	1.3691	2.7178
fusiform	0.8704	0.6959	0.9478	0.9337	0.8300	0.9700	1.2143	0.7656	1.6629
inferiorparietal	0.9602	0.9002	0.9844	0.9806	0.9500	0.9900	0.7429	0.4824	1.0035
inferiortemporal	0.9344	0.8387	0.9741	0.9595	0.9000	0.9800	1.2623	0.9139	1.6107
insula	0.9498	0.8750	0.9803	0.9644	0.8600	0.9900	1.3355	0.8910	1.7801
isthmuscingulate	0.9647	0.9111	0.9862	0.9736	0.8900	0.9900	1.7342	1.2869	2.1816
lateraloccipital	0.9596	0.8988	0.9842	0.9789	0.9500	0.9900	0.8757	0.5079	1.2435
lateralorbitofrontal	0.8612	0.6765	0.9440	0.9279	0.8200	0.9700	1.4738	0.9357	2.0118
lingual	0.9733	0.9324	0.9896	0.9841	0.9600	0.9900	0.9316	0.6744	1.1888
medialorbitofrontal	0.8512	0.6556	0.9397	0.9226	0.8000	0.9700	1.6027	0.7849	2.4204
middletemporal	0.9405	0.8531	0.9766	0.9705	0.9300	0.9900	0.8139	0.5177	1.1100
paracentral	0.9330	0.8354	0.9736	0.9665	0.9200	0.9900	1.2257	0.8540	1.5974
parahippocampal	0.9371	0.8450	0.9752	0.9622	0.9000	0.9900	1.8611	1.1416	2.5805
parsopercularis	0.8762	0.7083	0.9503	0.9319	0.8300	0.9700	1.1167	0.7043	1.5291
parsorbitalis	0.9361	0.8426	0.9748	0.9648	0.9100	0.9900	1.3435	0.8451	1.8419
parstriangularis	0.9343	0.8384	0.9741	0.9648	0.9100	0.9900	0.9562	0.6068	1.3055
pericalcarine	0.9300	0.8285	0.9724	0.9628	0.9100	0.9900	2.2103	1.4817	2.9388
postcentral	0.9787	0.9458	0.9917	0.9893	0.9700	1.0000	0.8793	0.6527	1.1060
posteriorcingulate	0.8366	0.6257	0.9335	0.9150	0.7800	0.9700	1.6375	1.0575	2.2174
precentral	0.9702	0.9247	0.9884	0.9852	0.9600	0.9900	0.7869	0.5891	0.9846
precuneus	0.9707	0.9258	0.9886	0.9854	0.9600	0.9900	0.7599	0.4962	1.0235
rostralanteriorcingulate	0.8087	0.5703	0.9215	0.8515	0.6100	0.9400	2.3852	1.2855	3.4849
rostralmiddlefrontal	0.9479	0.8704	0.9795	0.9738	0.9300	0.9900	1.0147	0.7240	1.3054
superiorfrontal	0.9335	0.8364	0.9737	0.9666	0.9200	0.9900	1.1558	0.8338	1.4777
superiorparietal	0.9726	0.9306	0.9893	0.9866	0.9700	0.9900	0.8397	0.5594	1.1199
superiortemporal	0.9218	0.8096	0.9690	0.9568	0.8900	0.9800	0.9947	0.7118	1.2775
supramarginal	0.9703	0.9250	0.9884	0.9853	0.9600	0.9900	0.7570	0.5017	1.0122
transversetemporal	0.9485	0.8720	0.9798	0.9739	0.9300	0.9900	1.6626	1.2049	2.1204

Table B.5: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average grey matter volume measurements obtained from FastSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9670	0.9167	0.9871	0.9821	0.9600	0.9900	2.8421	1.8738	3.8103
caudalmiddlefrontal	0.9674	0.9179	0.9873	0.9833	0.9600	0.9900	2.5577	1.3762	3.7392
cuneus	0.9950	0.9872	0.9981	0.9969	0.9900	1.0000	1.3098	0.7468	1.8728
entorhinal	0.9175	0.7997	0.9673	0.9547	0.8900	0.9800	3.8192	2.8237	4.8147
fusiform	0.9867	0.9658	0.9948	0.9935	0.9800	1.0000	1.6920	1.1456	2.2384
inferiorparietal	0.9789	0.9462	0.9918	0.9891	0.9700	1.0000	1.9709	1.1460	2.7957
inferiortemporal	0.9928	0.9814	0.9972	0.9952	0.9900	1.0000	1.3353	0.8701	1.8005
insula	0.9913	0.9777	0.9966	0.9925	0.9600	1.0000	1.6725	0.9882	2.3569
isthmuscingulate	0.9848	0.9612	0.9941	0.9916	0.9800	1.0000	1.6096	0.9230	2.2961
lateraloccipital	0.9901	0.9746	0.9962	0.9950	0.9900	1.0000	1.4141	0.9321	1.8962
lateralorbitofrontal	0.9719	0.9288	0.9890	0.9848	0.9600	0.9900	1.8640	1.1735	2.5545
lingual	0.9945	0.9859	0.9979	0.9956	0.9900	1.0000	1.7180	1.0990	2.3370
medialorbitofrontal	0.9768	0.9411	0.9910	0.9881	0.9700	1.0000	2.0479	1.1461	2.9498
middletemporal	0.9933	0.9828	0.9974	0.9968	0.9900	1.0000	1.1136	0.6684	1.5589
paracentral	0.9840	0.9591	0.9938	0.9918	0.9800	1.0000	1.8489	1.4152	2.2826
parahippocampal	0.9388	0.8489	0.9759	0.9677	0.9200	0.9900	2.5328	1.5130	3.5526
parsopercularis	0.9773	0.9423	0.9912	0.9891	0.9700	1.0000	2.0404	1.1538	2.9270
parsorbitalis	0.9854	0.9626	0.9943	0.9925	0.9800	1.0000	1.7059	1.0083	2.4036
parstriangularis	0.9919	0.9792	0.9969	0.9957	0.9900	1.0000	1.9568	1.2545	2.6591
pericalcarine	0.9820	0.9541	0.9930	0.9910	0.9800	1.0000	2.5065	1.6067	3.4063
postcentral	0.9776	0.9430	0.9913	0.9874	0.9700	1.0000	1.6245	1.0970	2.1519
posteriorcingulate	0.9850	0.9616	0.9942	0.9920	0.9800	1.0000	1.9933	1.4249	2.5618
precentral	0.9910	0.9769	0.9965	0.9954	0.9900	1.0000	1.0094	0.6783	1.3405
precuneus	0.9889	0.9714	0.9957	0.9938	0.9800	1.0000	1.5560	1.0585	2.0535
rostralanteriorcingulate	0.9656	0.9133	0.9866	0.9776	0.9300	0.9900	3.5319	1.7559	5.3079
rostralmiddlefrontal	0.9898	0.9737	0.9960	0.9924	0.9800	1.0000	1.5908	0.9551	2.2266
superiorfrontal	0.9849	0.9615	0.9942	0.9926	0.9800	1.0000	1.2182	0.7719	1.6646
superiorparietal	0.9855	0.9630	0.9944	0.9929	0.9800	1.0000	1.5119	0.8808	2.1431
superiortemporal	0.9857	0.9634	0.9944	0.9927	0.9800	1.0000	1.3197	0.8260	1.8133
supramarginal	0.9929	0.9817	0.9972	0.9963	0.9900	1.0000	1.2606	0.8588	1.6623
transversetemporal	0.9816	0.9531	0.9928	0.9902	0.9800	1.0000	2.7751	1.7609	3.7893

Table B.6: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average surface area measurements obtained from FastSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9827	0.9559	0.9933	0.9909	0.9800	1.0000	1.6543	1.0568	2.2518
caudalmiddlefrontal	0.9727	0.9309	0.9894	0.9855	0.9600	0.9900	2.2824	1.2611	3.3036
cuneus	0.9916	0.9784	0.9967	0.9960	0.9900	1.0000	1.3727	0.8181	1.9273
entorhinal	0.9537	0.8843	0.9818	0.9743	0.9300	0.9900	2.5746	1.4997	3.6496
fusiform	0.9945	0.9859	0.9979	0.9970	0.9900	1.0000	0.9867	0.6673	1.3062
inferiorparietal	0.9825	0.9554	0.9932	0.9913	0.9800	1.0000	1.6057	0.9552	2.2562
inferiortemporal	0.9948	0.9866	0.9980	0.9973	0.9900	1.0000	0.9744	0.6788	1.2699
insula	0.9946	0.9860	0.9979	0.9973	0.9900	1.0000	0.7766	0.4657	1.0875
isthmuscingulate	0.9864	0.9651	0.9947	0.9929	0.9800	1.0000	1.6997	0.9585	2.4408
lateraloccipital	0.9857	0.9635	0.9945	0.9921	0.9800	1.0000	1.3704	0.8247	1.9160
lateralorbitofrontal	0.9640	0.9094	0.9859	0.9809	0.9500	0.9900	1.7256	0.8365	2.6148
lingual	0.9940	0.9846	0.9977	0.9966	0.9900	1.0000	1.1172	0.6815	1.5529
medialorbitofrontal	0.9682	0.9198	0.9876	0.9837	0.9600	0.9900	1.9599	1.1168	2.8030
middletemporal	0.9950	0.9872	0.9981	0.9973	0.9900	1.0000	0.9028	0.5897	1.2160
paracentral	0.9809	0.9513	0.9926	0.9906	0.9800	1.0000	1.5319	1.0034	2.0603
parahippocampal	0.9715	0.9278	0.9889	0.9846	0.9600	0.9900	1.7550	1.1768	2.3331
parsopercularis	0.9851	0.9620	0.9942	0.9927	0.9800	1.0000	1.5513	0.8655	2.2370
parsorbitalis	0.9819	0.9537	0.9930	0.9901	0.9800	1.0000	1.5570	0.9168	2.1971
parstriangularis	0.9965	0.9910	0.9986	0.9980	0.9900	1.0000	1.2927	0.8523	1.7331
pericalcarine	0.9873	0.9675	0.9951	0.9939	0.9800	1.0000	1.7976	1.3176	2.2776
postcentral	0.9862	0.9647	0.9946	0.9911	0.9800	1.0000	1.0209	0.5581	1.4837
posteriorcingulate	0.9903	0.9751	0.9962	0.9951	0.9900	1.0000	1.2943	0.8442	1.7445
precentral	0.9936	0.9835	0.9975	0.9965	0.9900	1.0000	0.7480	0.5166	0.9793
precuneus	0.9947	0.9864	0.9980	0.9971	0.9900	1.0000	1.1196	0.7919	1.4473
rostralanteriorcingulate	0.9723	0.9298	0.9892	0.9837	0.9600	0.9900	2.1259	1.0414	3.2104
rostralmiddlefrontal	0.9935	0.9834	0.9975	0.9961	0.9900	1.0000	1.1840	0.8207	1.5473
superiorfrontal	0.9960	0.9896	0.9984	0.9979	0.9900	1.0000	0.4935	0.2601	0.7268
superiorparietal	0.9950	0.9872	0.9981	0.9972	0.9900	1.0000	1.0297	0.6691	1.3904
superiortemporal	0.9941	0.9848	0.9977	0.9967	0.9900	1.0000	0.8688	0.5786	1.1589
supramarginal	0.9966	0.9913	0.9987	0.9983	1.0000	1.0000	0.7432	0.4661	1.0204
transversetemporal	0.9846	0.9605	0.9940	0.9910	0.9800	1.0000	2.2633	1.4017	3.1249

Table B.7: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the volumetric measurements of the subcortical structures obtained from FreeSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
3rd-Ventricle	0.9738	0.9336	0.9898	0.9744	0.9100	0.9900	5.6505	4.1851	7.1158
4th-Ventricle	0.9933	0.9826	0.9974	0.9967	0.9900	1.0000	3.3860	1.9802	4.7917
Accumbens-area	0.9139	0.7915	0.9658	0.9556	0.8900	0.9800	5.9462	3.8375	8.0548
Amygdala	0.9367	0.8440	0.9751	0.9674	0.9200	0.9900	3.2670	1.5580	4.9760
Brain-Stem	0.9792	0.9470	0.9919	0.9898	0.9700	1.0000	1.5070	0.7666	2.2474
CC_Anterior	0.9728	0.9311	0.9894	0.9796	0.9500	0.9900	3.2427	1.9787	4.5068
CC_Central	0.9689	0.9216	0.9879	0.9835	0.9600	0.9900	4.4135	2.8028	6.0243
CC_Mid_Anterior	0.9939	0.9844	0.9977	0.9966	0.9900	1.0000	2.1366	1.3847	2.8886
CC_Mid_Posterior	0.9860	0.9641	0.9946	0.9933	0.9800	1.0000	3.0986	2.1990	3.9982
CC_Posterior	0.9929	0.9818	0.9973	0.9964	0.9900	1.0000	1.4611	1.0179	1.9044
CSF	0.9359	0.8422	0.9747	0.9672	0.9200	0.9900	5.7543	3.4358	8.0727
Caudate	0.9919	0.9791	0.9969	0.9959	0.9900	1.0000	1.4093	0.8704	1.9482
Cerebellum-Cortex	0.9970	0.9922	0.9988	0.9983	1.0000	1.0000	0.7493	0.4778	1.0207
Cerebellum-White-Matter	0.9700	0.9243	0.9883	0.9822	0.9500	0.9900	2.1695	1.2635	3.0754
Hippocampus	0.7643	0.4863	0.9018	0.8704	0.6700	0.9500	2.7392	0.3286	5.1498
Inf-Lat-Vent	0.9664	0.9153	0.9869	0.9825	0.9600	0.9900	8.4860	5.4526	11.5193
Lateral-Ventricle	0.9989	0.9970	0.9996	0.9993	1.0000	1.0000	3.1882	2.1070	4.2693
Pallidum	0.9497	0.8749	0.9803	0.9715	0.9300	0.9900	2.7745	1.4699	4.0791
Putamen	0.9850	0.9618	0.9942	0.9924	0.9800	1.0000	1.5116	0.9594	2.0638
Thalamus	0.8823	0.7214	0.9528	0.9379	0.8400	0.9800	3.7375	2.3469	5.1280
VentralDC	0.9034	0.7679	0.9615	0.9482	0.8700	0.9800	3.0724	1.8800	4.2648
choroid-plexus	0.8926	0.7438	0.9571	0.9458	0.8600	0.9800	8.4710	5.1714	11.7705

Table B.8: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the volumetric measurements of the subcortical structures obtained from FastSurfer using the OASIS dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
3rd-Ventricle	0.9856	0.9633	0.9944	0.9919	0.9800	1.0000	3.1667	2.3230	4.0105
4th-Ventricle	0.9980	0.9949	0.9992	0.9987	1.0000	1.0000	1.8822	1.1318	2.6326
Accumbens-area	0.9822	0.9546	0.9931	0.9914	0.9800	1.0000	1.9248	1.0801	2.7695
Amygdala	0.9837	0.9582	0.9937	0.9902	0.9800	1.0000	1.6713	0.8586	2.4840
Brain-Stem	0.9931	0.9822	0.9973	0.9967	0.9900	1.0000	0.9765	0.5932	1.3597
CC_Anterior	0.9363	0.8431	0.9749	0.9684	0.9200	0.9900	4.3718	2.6100	6.1337
CC_Central	0.7642	0.4860	0.9017	0.8632	0.6600	0.9500	11.1569	5.5794	16.7345
CC_Mid_Anterior	0.9555	0.8888	0.9826	0.9775	0.9400	0.9900	4.0336	1.3624	6.7048
CC_Mid_Posterior	0.9592	0.8976	0.9840	0.9799	0.9500	0.9900	4.9354	2.9973	6.8735
CC_Posterior	0.9954	0.9881	0.9982	0.9978	0.9900	1.0000	1.0791	0.6869	1.4713
CSF	0.9839	0.9588	0.9937	0.9919	0.9800	1.0000	2.3378	1.4092	3.2663
Caudate	0.9986	0.9965	0.9995	0.9993	1.0000	1.0000	0.5636	0.3823	0.7449
Cerebellum-Cortex	0.9987	0.9966	0.9995	0.9993	1.0000	1.0000	0.4216	0.2274	0.6157
Cerebellum-White-Matter	0.9953	0.9880	0.9982	0.9971	0.9900	1.0000	1.0508	0.7698	1.3319
Hippocampus	0.9946	0.9860	0.9979	0.9973	0.9900	1.0000	0.7075	0.4733	0.9417
Inf-Lat-Vent	0.9902	0.9749	0.9962	0.9952	0.9900	1.0000	4.8936	3.1322	6.6551
Lateral-Ventricle	0.9992	0.9980	0.9997	0.9996	1.0000	1.0000	2.6689	1.6526	3.6852
Pallidum	0.9935	0.9832	0.9975	0.9965	0.9900	1.0000	1.1979	0.8340	1.5619
Putamen	0.9925	0.9807	0.9971	0.9961	0.9900	1.0000	0.8507	0.4201	1.2813
Thalamus	0.9921	0.9798	0.9970	0.9962	0.9900	1.0000	0.8800	0.5165	1.2434
VentralDC	0.9933	0.9828	0.9974	0.9968	0.9900	1.0000	0.8779	0.6346	1.1212
choroid-plexus	0.9901	0.9747	0.9962	0.9924	0.9800	1.0000	3.1486	2.1071	4.1902

Table B.9: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average cortical thickness measurements obtained from FreeSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9037	0.8213	0.9492	0.9486	0.9000	0.9700	1.8431	1.3083	2.3778
caudalmiddlefrontal	0.8058	0.6548	0.8949	0.8731	0.7600	0.9300	1.7650	0.9272	2.6028
cuneus	0.7991	0.6439	0.8911	0.8900	0.7900	0.9400	2.7726	1.9879	3.5573
entorhinal	0.8847	0.7878	0.9389	0.9398	0.8800	0.9700	4.2283	3.2460	5.2106
fusiform	0.8960	0.8077	0.9450	0.9402	0.8900	0.9700	1.4971	1.0624	1.9318
inferiorparietal	0.8448	0.7194	0.9169	0.9125	0.8300	0.9500	1.7515	1.1980	2.3049
inferiortemporal	0.8581	0.7420	0.9242	0.9196	0.8500	0.9600	1.9029	1.3795	2.4263
insula	0.8844	0.7873	0.9387	0.9388	0.8800	0.9700	1.4932	1.1004	1.8860
isthmuscingulate	0.8964	0.8084	0.9452	0.9443	0.8900	0.9700	2.2346	1.7270	2.7422
lateraloccipital	0.8544	0.7356	0.9222	0.9194	0.8500	0.9600	1.9341	1.4206	2.4476
lateralorbitofrontal	0.7523	0.5696	0.8642	0.8539	0.7200	0.9200	1.9874	1.5579	2.4169
lingual	0.7459	0.5596	0.8604	0.8551	0.7200	0.9200	3.1667	2.2001	4.1332
medialorbitofrontal	0.8247	0.6858	0.9056	0.8457	0.6600	0.9200	2.1617	1.5642	2.7592
middletemporal	0.8766	0.7737	0.9344	0.9347	0.8700	0.9700	1.4498	0.9505	1.9491
paracentral	0.8292	0.6933	0.9081	0.9003	0.8100	0.9500	1.9103	1.2956	2.5249
parahippocampal	0.9382	0.8834	0.9676	0.9650	0.9300	0.9800	2.1829	1.5162	2.8496
parsopercularis	0.8925	0.8015	0.9431	0.9381	0.8800	0.9700	1.4559	1.1002	1.8116
parsorbitalis	0.8615	0.7477	0.9261	0.9198	0.8500	0.9600	2.0225	1.2287	2.8163
parstriangularis	0.8490	0.7265	0.9192	0.9173	0.8400	0.9600	1.6785	1.2576	2.0995
pericalcarine	0.8325	0.6988	0.9100	0.9094	0.8300	0.9500	4.5034	3.0712	5.9355
postcentral	0.8176	0.6741	0.9016	0.8888	0.7900	0.9400	1.9816	1.2443	2.7188
posteriorcingulate	0.8786	0.7772	0.9355	0.9327	0.8700	0.9600	2.0206	1.4525	2.5887
precentral	0.8691	0.7608	0.9303	0.9250	0.8600	0.9600	2.0906	1.3156	2.8657
precuneus	0.8575	0.7410	0.9239	0.9205	0.8500	0.9600	1.1807	0.7202	1.6412
rostralanteriorcingulate	0.8817	0.7826	0.9372	0.9367	0.8800	0.9700	2.1045	1.5208	2.6882
rostralmiddlefrontal	0.8053	0.6540	0.8947	0.8862	0.7800	0.9400	1.4928	1.0507	1.9350
superiorfrontal	0.7910	0.6308	0.8865	0.8654	0.7400	0.9300	1.5128	0.9454	2.0803
superiorparietal	0.8866	0.7912	0.9399	0.9407	0.8900	0.9700	1.3261	0.9256	1.7265
superiortemporal	0.8170	0.6731	0.9013	0.8967	0.8000	0.9500	1.7729	1.1794	2.3664
supramarginal	0.8061	0.6553	0.8951	0.8950	0.8000	0.9500	1.7492	1.1250	2.3735
transversetemporal	0.8358	0.7043	0.9119	0.9127	0.8300	0.9500	2.5920	1.8823	3.3017

Table B.10: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average grey matter volume measurements obtained from FreeSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9639	0.9311	0.9812	0.9820	0.9700	0.9900	3.1166	2.3175	3.9158
caudalmiddlefrontal	0.9147	0.8408	0.9551	0.9528	0.9100	0.9800	3.0403	0.7380	5.3427
cuneus	0.9662	0.9354	0.9824	0.9832	0.9700	0.9900	2.8827	2.0902	3.6752
entorhinal	0.8259	0.6877	0.9063	0.9048	0.8200	0.9500	8.6549	6.1522	11.1576
fusiform	0.9674	0.9377	0.9831	0.9822	0.9700	0.9900	2.7667	2.1441	3.3893
inferiorparietal	0.9585	0.9211	0.9784	0.9787	0.9600	0.9900	2.4108	1.2458	3.5759
inferiortemporal	0.9189	0.8484	0.9573	0.9563	0.9200	0.9800	3.3822	2.3852	4.3791
insula	0.9859	0.9728	0.9927	0.9930	0.9900	1.0000	1.2150	0.9243	1.5057
isthmuscingulate	0.9780	0.9577	0.9886	0.9881	0.9800	0.9900	2.3923	1.7952	2.9894
lateraloccipital	0.9868	0.9746	0.9932	0.9931	0.9900	1.0000	1.7877	1.3065	2.2689
lateralorbitofrontal	0.9617	0.9271	0.9801	0.9790	0.9600	0.9900	2.2436	1.5439	2.9434
lingual	0.9620	0.9276	0.9802	0.9811	0.9600	0.9900	2.8899	2.0745	3.7054
medialorbitofrontal	0.9644	0.9320	0.9815	0.9800	0.9600	0.9900	2.6099	1.9976	3.2222
middletemporal	0.9669	0.9368	0.9828	0.9828	0.9700	0.9900	2.1541	1.3095	2.9987
paracentral	0.8837	0.7861	0.9383	0.9369	0.8800	0.9700	3.5054	1.5643	5.4464
parahippocampal	0.9064	0.8260	0.9506	0.9502	0.9000	0.9700	3.6751	2.4917	4.8584
parsopercularis	0.8485	0.7256	0.9189	0.9130	0.8300	0.9500	3.2290	0.8231	5.6348
parsorbitalis	0.9123	0.8366	0.9538	0.9537	0.9100	0.9800	3.4123	2.2648	4.5598
parstriangularis	0.8123	0.6653	0.8986	0.8924	0.7900	0.9400	4.3106	1.4117	7.2095
pericalcarine	0.9332	0.8743	0.9650	0.9660	0.9300	0.9800	4.7621	3.2425	6.2818
postcentral	0.9056	0.8247	0.9502	0.9491	0.9000	0.9700	2.4940	1.0117	3.9763
posteriorcingulate	0.9435	0.8932	0.9705	0.9712	0.9400	0.9900	2.5016	1.4535	3.5496
precentral	0.8699	0.7623	0.9308	0.9256	0.8600	0.9600	2.7133	0.7207	4.7060
precuneus	0.9498	0.9048	0.9738	0.9728	0.9500	0.9900	1.9778	0.8558	3.0998
rostralanteriorcingulate	0.9557	0.9159	0.9769	0.9775	0.9600	0.9900	3.5222	2.5980	4.4465
rostralmiddlefrontal	0.8754	0.7717	0.9338	0.9280	0.8600	0.9600	3.2792	1.0692	5.4893
superiorfrontal	0.9507	0.9065	0.9743	0.9751	0.9500	0.9900	2.0281	1.0130	3.0433
superiorparietal	0.8423	0.7152	0.9155	0.9096	0.8300	0.9500	3.6906	0.9163	6.4650
superiortemporal	0.9475	0.9007	0.9726	0.9732	0.9500	0.9900	2.5777	1.9136	3.2417
supramarginal	0.8866	0.7911	0.9399	0.9403	0.8900	0.9700	3.0678	0.9058	5.2297
transversetemporal	0.9752	0.9524	0.9871	0.9875	0.9800	0.9900	3.1595	2.3209	3.9980

Table B.11: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average surface area measurements obtained from FreeSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9634	0.9302	0.9810	0.9813	0.9600	0.9900	2.6359	1.7958	3.4760
caudalmiddlefrontal	0.9625	0.9285	0.9805	0.9806	0.9600	0.9900	2.5811	1.6599	3.5023
cuneus	0.9812	0.9639	0.9903	0.9906	0.9800	1.0000	1.7822	1.2943	2.2701
entorhinal	0.8460	0.7214	0.9175	0.9156	0.8400	0.9600	7.2050	4.6991	9.7109
fusiform	0.9781	0.9580	0.9887	0.9889	0.9800	0.9900	2.0375	1.5817	2.4932
inferiorparietal	0.9411	0.8889	0.9692	0.9656	0.9300	0.9800	2.8182	1.5549	4.0814
inferiortemporal	0.9478	0.9011	0.9727	0.9730	0.9500	0.9900	2.6548	1.8213	3.4883
insula	0.9802	0.9620	0.9898	0.9900	0.9800	0.9900	1.2538	0.9089	1.5986
isthmuscingulate	0.9767	0.9554	0.9879	0.9879	0.9800	0.9900	2.4601	1.9084	3.0118
lateraloccipital	0.9819	0.9652	0.9906	0.9901	0.9800	0.9900	2.0049	1.3727	2.6371
lateralorbitofrontal	0.9457	0.8974	0.9717	0.9667	0.9200	0.9800	2.9397	2.1368	3.7427
lingual	0.9760	0.9540	0.9876	0.9880	0.9800	0.9900	2.0285	1.4360	2.6210
medialorbitofrontal	0.9405	0.8878	0.9689	0.9676	0.9400	0.9800	3.2464	2.4394	4.0533
middletemporal	0.9795	0.9607	0.9894	0.9887	0.9800	0.9900	1.8050	1.3177	2.2922
paracentral	0.9016	0.8176	0.9480	0.9470	0.9000	0.9700	2.6467	1.0991	4.1943
parahippocampal	0.9084	0.8296	0.9517	0.9528	0.9100	0.9800	2.7895	1.9779	3.6011
parsopercularis	0.7526	0.5701	0.8643	0.8482	0.7100	0.9200	3.1986	0.5072	5.8900
parsorbitalis	0.9119	0.8359	0.9536	0.9542	0.9100	0.9800	3.1892	2.1186	4.2597
parstriangularis	0.7809	0.6146	0.8807	0.8660	0.7400	0.9300	4.2860	1.1555	7.4165
pericalcarine	0.9744	0.9510	0.9867	0.9869	0.9700	0.9900	2.4633	1.8685	3.0582
postcentral	0.8337	0.7008	0.9107	0.9013	0.8100	0.9500	3.1199	1.1640	5.0758
posteriorcingulate	0.9757	0.9534	0.9874	0.9879	0.9800	0.9900	2.0404	1.3997	2.6811
precentral	0.8855	0.7892	0.9393	0.9343	0.8700	0.9700	2.6949	1.3704	4.0194
precuneus	0.9721	0.9465	0.9855	0.9855	0.9700	0.9900	1.5246	0.7128	2.3363
rostralanteriorcingulate	0.9708	0.9441	0.9848	0.9837	0.9700	0.9900	3.1178	2.3224	3.9133
rostralmiddlefrontal	0.8562	0.7387	0.9232	0.9037	0.8100	0.9500	3.6004	1.1709	6.0300
superiorfrontal	0.9335	0.8749	0.9651	0.9602	0.9200	0.9800	2.1051	1.0120	3.1983
superiorparietal	0.8257	0.6875	0.9062	0.8963	0.8000	0.9500	3.2724	0.7540	5.7908
superiortemporal	0.9725	0.9473	0.9857	0.9858	0.9700	0.9900	1.7468	1.2881	2.2054
supramarginal	0.8736	0.7686	0.9328	0.9293	0.8600	0.9600	3.6128	1.6856	5.5399
transversetemporal	0.9653	0.9338	0.9820	0.9825	0.9700	0.9900	3.1297	2.4477	3.8116

Table B.12: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average cortical thickness measurements obtained from FastSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9349	0.8774	0.9659	0.9667	0.9400	0.9800	1.5557	1.2068	1.9045
caudalmiddlefrontal	0.9309	0.8702	0.9638	0.9573	0.9200	0.9800	1.1371	0.8185	1.4557
cuneus	0.8849	0.7882	0.9390	0.9397	0.8800	0.9700	1.9244	1.4189	2.4299
entorhinal	0.9291	0.8670	0.9628	0.9567	0.9100	0.9800	3.2177	2.4629	3.9725
fusiform	0.9099	0.8322	0.9525	0.9436	0.8900	0.9700	1.3984	0.8796	1.9172
inferiorparietal	0.8771	0.7747	0.9347	0.9338	0.8700	0.9700	1.3692	0.8738	1.8646
inferiortemporal	0.8764	0.7735	0.9343	0.9352	0.8800	0.9700	1.7230	1.3206	2.1253
insula	0.8900	0.7971	0.9417	0.9371	0.8800	0.9700	1.3975	0.9243	1.8708
isthmuscingulate	0.9034	0.8208	0.9490	0.9492	0.9000	0.9700	1.9793	1.5388	2.4199
lateraloccipital	0.8869	0.7917	0.9401	0.9407	0.8900	0.9700	1.5772	1.0685	2.0860
lateralorbitofrontal	0.7957	0.6383	0.8892	0.8790	0.7700	0.9400	1.5800	1.1864	1.9735
lingual	0.9027	0.8195	0.9486	0.9488	0.9000	0.9700	2.2227	1.7405	2.7050
medialorbitofrontal	0.6814	0.4624	0.8220	0.7729	0.5600	0.8800	2.4020	1.7433	3.0607
middletemporal	0.9373	0.8818	0.9672	0.9683	0.9400	0.9800	1.1721	0.8794	1.4647
paracentral	0.9040	0.8218	0.9493	0.9472	0.9000	0.9700	1.3001	0.8898	1.7103
parahippocampal	0.9461	0.8980	0.9718	0.9642	0.9300	0.9800	2.2629	1.5131	3.0128
parsopercularis	0.9289	0.8666	0.9627	0.9612	0.9300	0.9800	1.0649	0.7736	1.3561
parsorbitalis	0.9238	0.8573	0.9600	0.9592	0.9200	0.9800	1.4434	1.0868	1.8000
parstriangularis	0.9302	0.8689	0.9634	0.9629	0.9300	0.9800	1.1216	0.8547	1.3886
pericalcarine	0.9287	0.8662	0.9626	0.9623	0.9300	0.9800	3.6591	2.9067	4.4116
postcentral	0.8846	0.7876	0.9388	0.9356	0.8800	0.9700	1.5663	1.0356	2.0970
posteriorcingulate	0.9458	0.8974	0.9717	0.9697	0.9400	0.9800	1.3268	0.9678	1.6859
precentral	0.9319	0.8720	0.9643	0.9637	0.9300	0.9800	1.4205	1.0387	1.8023
precuneus	0.9089	0.8305	0.9520	0.9520	0.9100	0.9800	1.0030	0.6484	1.3576
rostralanteriorcingulate	0.9050	0.8237	0.9499	0.9498	0.9000	0.9700	1.8040	1.3435	2.2646
rostralmiddlefrontal	0.8843	0.7871	0.9386	0.9338	0.8700	0.9700	1.2617	0.9684	1.5551
superiorfrontal	0.9407	0.8881	0.9690	0.9668	0.9400	0.9800	0.7612	0.5398	0.9827
superiorparietal	0.8898	0.7967	0.9416	0.9428	0.8900	0.9700	1.2207	0.7790	1.6625
superiortemporal	0.8640	0.7520	0.9275	0.9208	0.8400	0.9600	1.4919	0.9507	2.0331
supramarginal	0.8693	0.7611	0.9304	0.9309	0.8700	0.9600	1.4116	0.9300	1.8932
transversetemporal	0.8932	0.8027	0.9435	0.9443	0.8900	0.9700	2.2809	1.7592	2.8026

Table B.13: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average grey matter volume measurements obtained from FastSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9810	0.9634	0.9901	0.9906	0.9800	1.0000	2.2326	1.5843	2.8810
caudalmiddlefrontal	0.9921	0.9848	0.9959	0.9960	0.9900	1.0000	1.8351	1.4134	2.2568
cuneus	0.9828	0.9670	0.9911	0.9910	0.9800	1.0000	2.1767	1.6127	2.7406
entorhinal	0.8622	0.7490	0.9265	0.9187	0.8400	0.9600	6.1515	4.6172	7.6858
fusiform	0.9862	0.9734	0.9929	0.9930	0.9900	1.0000	1.5201	0.9658	2.0744
inferiorparietal	0.9907	0.9820	0.9952	0.9953	0.9900	1.0000	1.4614	1.0554	1.8673
inferiortemporal	0.9362	0.8799	0.9666	0.9672	0.9400	0.9800	2.5463	1.6552	3.4373
insula	0.9906	0.9819	0.9952	0.9951	0.9900	1.0000	1.0062	0.7120	1.3004
isthmuscingulate	0.9864	0.9738	0.9930	0.9932	0.9900	1.0000	1.9605	1.4825	2.4384
lateraloccipital	0.9878	0.9765	0.9937	0.9937	0.9900	1.0000	1.6126	1.0331	2.1920
lateralorbitofrontal	0.9607	0.9252	0.9796	0.9803	0.9600	0.9900	2.1786	1.6581	2.6991
lingual	0.9851	0.9713	0.9923	0.9926	0.9900	1.0000	1.8569	1.3511	2.3628
medialorbitofrontal	0.9784	0.9585	0.9888	0.9892	0.9800	0.9900	1.8302	1.3522	2.3083
middletemporal	0.9864	0.9737	0.9929	0.9931	0.9900	1.0000	1.5123	1.1162	1.9083
paracentral	0.9875	0.9759	0.9935	0.9935	0.9900	1.0000	1.7162	1.3210	2.1114
parahippocampal	0.9727	0.9477	0.9858	0.9864	0.9700	0.9900	1.9769	1.2471	2.7068
parsopercularis	0.9841	0.9694	0.9918	0.9920	0.9800	1.0000	1.9389	1.2310	2.6468
parsorbitalis	0.9176	0.8461	0.9566	0.9543	0.9100	0.9800	3.4411	1.9176	4.9645
parstriangularis	0.9954	0.9911	0.9976	0.9976	1.0000	1.0000	1.5390	1.1828	1.8952
pericalcarine	0.9676	0.9380	0.9831	0.9832	0.9700	0.9900	4.1214	3.2044	5.0384
postcentral	0.9901	0.9809	0.9949	0.9950	0.9900	1.0000	1.3121	0.9533	1.6710
posteriorcingulate	0.9887	0.9782	0.9941	0.9939	0.9900	1.0000	1.4775	1.0378	1.9171
precentral	0.9910	0.9826	0.9953	0.9954	0.9900	1.0000	1.2789	0.9188	1.6389
precuneus	0.9937	0.9878	0.9967	0.9968	0.9900	1.0000	1.0952	0.7754	1.4150
rostralanteriorcingulate	0.9802	0.9619	0.9897	0.9901	0.9800	0.9900	2.2634	1.6103	2.9165
rostralmiddlefrontal	0.9912	0.9829	0.9954	0.9956	0.9900	1.0000	1.6103	1.1605	2.0600
superiorfrontal	0.9896	0.9799	0.9946	0.9946	0.9900	1.0000	1.1072	0.6901	1.5244
superiorparietal	0.9919	0.9843	0.9958	0.9960	0.9900	1.0000	1.3938	0.9242	1.8634
superiortemporal	0.9581	0.9203	0.9782	0.9764	0.9500	0.9900	2.3971	1.7267	3.0676
supramarginal	0.9897	0.9801	0.9947	0.9949	0.9900	1.0000	1.4021	0.9629	1.8412
transversetemporal	0.9854	0.9718	0.9924	0.9910	0.9800	1.0000	2.9519	2.2242	3.6797

Table B.14: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the average surface area measurements obtained from FastSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
caudalanteriorcingulate	0.9813	0.9641	0.9903	0.9908	0.9800	1.0000	2.2410	1.7011	2.7809
caudalmiddlefrontal	0.9919	0.9844	0.9958	0.9959	0.9900	1.0000	1.5376	1.0939	1.9813
cuneus	0.9869	0.9747	0.9932	0.9935	0.9900	1.0000	1.6576	1.2755	2.0397
entorhinal	0.8416	0.7140	0.9151	0.9152	0.8400	0.9600	5.6972	4.0796	7.3148
fusiform	0.9776	0.9571	0.9884	0.9888	0.9800	0.9900	1.9323	1.4191	2.4455
inferiorparietal	0.9709	0.9444	0.9849	0.9842	0.9700	0.9900	2.2662	1.4220	3.1103
inferiortemporal	0.9479	0.9013	0.9728	0.9711	0.9400	0.9800	2.5488	1.8092	3.2884
insula	0.9852	0.9715	0.9923	0.9925	0.9900	1.0000	1.1361	0.8406	1.4316
isthmuscingulate	0.9850	0.9712	0.9923	0.9926	0.9900	1.0000	1.8295	1.3588	2.3002
lateraloccipital	0.9875	0.9759	0.9935	0.9934	0.9900	1.0000	1.6665	1.2267	2.1063
lateralorbitofrontal	0.9212	0.8525	0.9586	0.9575	0.9200	0.9800	3.0491	2.2020	3.8963
lingual	0.9808	0.9631	0.9901	0.9904	0.9800	1.0000	1.9558	1.4753	2.4363
medialorbitofrontal	0.8587	0.7430	0.9246	0.9252	0.8600	0.9600	4.3113	2.9191	5.7034
middletemporal	0.9843	0.9698	0.9919	0.9918	0.9800	1.0000	1.3783	0.9146	1.8421
paracentral	0.9745	0.9511	0.9868	0.9869	0.9800	0.9900	1.7698	1.1767	2.3629
parahippocampal	0.9498	0.9049	0.9738	0.9747	0.9500	0.9900	2.3012	1.8073	2.7950
parsopercularis	0.9728	0.9479	0.9859	0.9863	0.9700	0.9900	2.3178	1.6622	2.9734
parsorbitalis	0.9076	0.8281	0.9513	0.9498	0.9000	0.9700	3.2676	1.8917	4.6434
parstriangularis	0.9894	0.9796	0.9945	0.9948	0.9900	1.0000	2.2041	1.6551	2.7531
pericalcarine	0.9877	0.9763	0.9936	0.9937	0.9900	1.0000	1.8106	1.3031	2.3182
postcentral	0.9616	0.9268	0.9800	0.9803	0.9600	0.9900	2.0569	1.4280	2.6859
posteriorcingulate	0.9861	0.9733	0.9928	0.9926	0.9900	1.0000	1.7920	1.2691	2.3148
precentral	0.9828	0.9670	0.9911	0.9914	0.9800	1.0000	1.7942	1.3679	2.2205
precuneus	0.9960	0.9922	0.9979	0.9980	1.0000	1.0000	0.9803	0.7701	1.1905
rostralanteriorcingulate	0.9830	0.9673	0.9912	0.9915	0.9800	1.0000	2.3818	1.8118	2.9519
rostralmiddlefrontal	0.9856	0.9723	0.9926	0.9919	0.9800	1.0000	1.9613	1.3958	2.5269
superiorfrontal	0.9906	0.9818	0.9951	0.9953	0.9900	1.0000	1.1421	0.7620	1.5222
superiorparietal	0.9922	0.9849	0.9960	0.9961	0.9900	1.0000	1.1568	0.7630	1.5507
superiortemporal	0.9769	0.9556	0.9880	0.9886	0.9800	0.9900	1.5585	1.0690	2.0480
supramarginal	0.9711	0.9447	0.9850	0.9857	0.9700	0.9900	2.2101	1.5726	2.8477
transversetemporal	0.9678	0.9385	0.9833	0.9816	0.9600	0.9900	2.9449	2.0907	3.7991

Table B.15: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the volumetric measurements of the subcortical structures obtained from FreeSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
3rd-Ventricle	0.9914	0.9835	0.9956	0.9958	0.9900	1.0000	4.5151	3.3618	5.6683
4th-Ventricle	0.9539	0.9125	0.9760	0.9768	0.9600	0.9900	6.1504	4.3674	7.9334
Accumbens-area	0.8255	0.6871	0.9060	0.9033	0.8200	0.9500	7.5717	5.8680	9.2755
Amygdala	0.8439	0.7179	0.9164	0.9109	0.8300	0.9500	4.1040	2.9509	5.2571
Brain-Stem	0.9572	0.9186	0.9777	0.9787	0.9600	0.9900	2.2434	1.6371	2.8496
CC_Anterior	0.9692	0.9412	0.9840	0.9842	0.9700	0.9900	3.0632	2.1594	3.9669
CC_Central	0.9488	0.9030	0.9733	0.9726	0.9500	0.9900	4.9035	3.2178	6.5892
CC_Mid_Anterior	0.9290	0.8667	0.9627	0.9617	0.9300	0.9800	4.8691	2.9074	6.8308
CC_Mid_Posterior	0.9572	0.9186	0.9777	0.9780	0.9600	0.9900	4.3955	3.3482	5.4427
CC_Posterior	0.9877	0.9764	0.9937	0.9939	0.9900	1.0000	2.0514	1.5782	2.5246
CSF	0.9284	0.8656	0.9624	0.9548	0.9100	0.9800	6.1980	4.4342	7.9618
Caudate	0.9811	0.9637	0.9902	0.9904	0.9800	1.0000	2.0785	1.5039	2.6531
Cerebellum-Cortex	0.9839	0.9690	0.9917	0.9918	0.9800	1.0000	1.2929	0.9678	1.6179
Cerebellum-White-Matter	0.8400	0.7113	0.9142	0.9142	0.8400	0.9600	6.4795	4.8103	8.1488
Hippocampus	0.9685	0.9398	0.9837	0.9839	0.9700	0.9900	1.6345	1.1714	2.0977
Inf-Lat-Vent	0.9699	0.9425	0.9844	0.9842	0.9700	0.9900	11.0185	7.7874	14.2497
Lateral-Ventricle	0.9994	0.9988	0.9997	0.9997	1.0000	1.0000	1.6146	1.2185	2.0108
Pallidum	0.8757	0.7723	0.9340	0.9348	0.8700	0.9700	4.1859	3.0595	5.3122
Putamen	0.9607	0.9251	0.9795	0.9793	0.9600	0.9900	1.9751	1.4271	2.5230
Thalamus	0.9508	0.9067	0.9743	0.9748	0.9500	0.9900	2.5692	2.0249	3.1136
VentralDC	0.9332	0.8743	0.9650	0.9657	0.9300	0.9800	2.8464	2.2677	3.4252
choroid-plexus	0.9237	0.8572	0.9599	0.9612	0.9300	0.9800	10.2222	7.4924	12.9520

Table B.16: The PCC, ICC and variability, as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the volumetric measurements of the subcortical structures obtained from FastSurfer using the MIND dataset.

Region of interest	PCC	LL	UL	ICC	LL	UL	TRV	LL	UL
3rd-Ventricle	0.9977	0.9955	0.9988	0.9989	1.0000	1.0000	2.2635	1.8386	2.6885
4th-Ventricle	0.9933	0.9871	0.9966	0.9965	0.9900	1.0000	2.7597	1.9599	3.5596
Accumbens-area	0.9703	0.9432	0.9846	0.9847	0.9700	0.9900	2.8127	2.2493	3.3761
Amygdala	0.9612	0.9261	0.9798	0.9792	0.9600	0.9900	1.9272	1.4186	2.4359
Brain-Stem	0.9915	0.9836	0.9956	0.9958	0.9900	1.0000	1.0549	0.7957	1.3141
CC_Anterior	0.9352	0.8780	0.9661	0.9651	0.9300	0.9800	5.3972	3.9759	6.8184
CC_Central	0.7460	0.5599	0.8605	0.8563	0.7200	0.9300	17.7700	13.0688	22.4712
CC_Mid_Anterior	0.8337	0.7007	0.9106	0.9113	0.8300	0.9500	7.4165	2.9721	11.8610
CC_Mid_Posterior	0.9317	0.8716	0.9642	0.9621	0.9300	0.9800	6.7041	4.6305	8.7777
CC_Posterior	0.9916	0.9837	0.9957	0.9957	0.9900	1.0000	1.7031	1.2493	2.1568
CSF	0.9936	0.9876	0.9967	0.9966	0.9900	1.0000	2.1421	1.6862	2.5980
Caudate	0.9971	0.9944	0.9985	0.9983	1.0000	1.0000	0.9695	0.7217	1.2173
Cerebellum-Cortex	0.9910	0.9827	0.9954	0.9955	0.9900	1.0000	1.0165	0.7747	1.2582
Cerebellum-White-Matter	0.9884	0.9776	0.9940	0.9940	0.9900	1.0000	1.4988	1.1003	1.8974
Hippocampus	0.9900	0.9806	0.9948	0.9946	0.9900	1.0000	0.9371	0.6733	1.2008
Inf-Lat-Vent	0.9869	0.9748	0.9932	0.9933	0.9900	1.0000	6.1944	2.4763	9.9124
Lateral-Ventricle	0.9996	0.9991	0.9998	0.9998	1.0000	1.0000	1.5611	1.1441	1.9782
Pallidum	0.9795	0.9606	0.9894	0.9892	0.9800	0.9900	1.7584	1.3887	2.1281
Putamen	0.9915	0.9835	0.9956	0.9956	0.9900	1.0000	0.9473	0.7001	1.1944
Thalamus	0.9872	0.9753	0.9934	0.9934	0.9900	1.0000	1.2918	0.9607	1.6230
VentralDC	0.9869	0.9747	0.9932	0.9923	0.9800	1.0000	0.9961	0.6746	1.3176
choroid-plexus	0.9869	0.9747	0.9932	0.9933	0.9900	1.0000	4.5467	3.2196	5.8739

C

Appendix

C.1 Inter-software reliability

Table C.1: Table presenting the ICC(3,1) as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the three different anatomical measurements: cortical thickness (CT), grey matter volume (GMV), and surface area (SA), for the inter-software reliability testing of the first scan from the OASIS dataset.

Region of interest	ICC CT	LL	UL	ICC GMV	LL	UL	ICC SA	LL	UL
caudalanteriorcingulate	0.9641	0.9100	0.9900	0.9211	0.8100	0.9700	0.9233	0.8200	0.9700
caudalmiddlefrontal	0.9444	0.8700	0.9800	0.9499	0.8800	0.9800	0.9487	0.8800	0.9800
cuneus	0.9637	0.9100	0.9900	0.9503	0.8800	0.9800	0.9529	0.8900	0.9800
entorhinal	0.7694	0.5000	0.9000	0.7081	0.4000	0.8700	0.7739	0.5100	0.9000
fusiform	0.9108	0.7900	0.9600	0.9233	0.8200	0.9700	0.9250	0.8200	0.9700
inferiorparietal	0.9265	0.8200	0.9700	0.9504	0.8800	0.9800	0.9553	0.8900	0.9800
inferiortemporal	0.9115	0.7900	0.9600	0.9497	0.8800	0.9800	0.9309	0.8300	0.9700
insula	0.9398	0.8500	0.9800	0.9903	0.9800	1.0000	0.9881	0.9700	1.0000
isthmuscingulate	0.9554	0.8900	0.9800	0.9300	0.8300	0.9700	0.9538	0.8900	0.9800
lateraloccipital	0.9468	0.8700	0.9800	0.9581	0.9000	0.9800	0.9373	0.8500	0.9700
lateralorbitofrontal	0.9144	0.8000	0.9700	0.9673	0.9200	0.9900	0.9552	0.8900	0.9800
lingual	0.9236	0.8200	0.9700	0.9845	0.9600	0.9900	0.9693	0.9200	0.9900
medialorbitofrontal	0.8842	0.7300	0.9500	0.9569	0.8900	0.9800	0.9140	0.8000	0.9700
middletemporal	0.9539	0.8900	0.9800	0.9673	0.9200	0.9900	0.9703	0.9300	0.9900
paracentral	0.9679	0.9200	0.9900	0.9661	0.9200	0.9900	0.9716	0.9300	0.9900
parahippocampal	0.9817	0.9500	0.9900	0.7528	0.4700	0.8900	0.7412	0.4500	0.8900
parsopercularis	0.9144	0.8000	0.9700	0.9696	0.9200	0.9900	0.9719	0.9300	0.9900
parsorbitalis	0.9312	0.8300	0.9700	0.8917	0.7500	0.9600	0.8384	0.6400	0.9300
parstriangularis	0.8880	0.7400	0.9500	0.9488	0.8800	0.9800	0.9424	0.8600	0.9800
pericalcarine	0.9543	0.8900	0.9800	0.9784	0.9500	0.9900	0.9625	0.9100	0.9800
postcentral	0.9823	0.9600	0.9900	0.9695	0.9200	0.9900	0.9626	0.9100	0.9900
posteriorcingulate	0.9582	0.9000	0.9800	0.9620	0.9100	0.9800	0.9464	0.8700	0.9800
precentral	0.9677	0.9200	0.9900	0.9422	0.8600	0.9800	0.9219	0.8100	0.9700
precuneus	0.9672	0.9200	0.9900	0.9844	0.9600	0.9900	0.9912	0.9800	1.0000
rostralanteriorcingulate	0.9345	0.8400	0.9700	0.9280	0.8300	0.9700	0.8816	0.7300	0.9500
rostralmiddlefrontal	0.9483	0.8700	0.9800	0.9719	0.9300	0.9900	0.9725	0.9300	0.9900
superiorfrontal	0.9498	0.8800	0.9800	0.9819	0.9500	0.9900	0.9860	0.9700	0.9900
superiorparietal	0.9706	0.9300	0.9900	0.9827	0.9600	0.9900	0.9836	0.9600	0.9900
superiortemporal	0.9720	0.9300	0.9900	0.9650	0.9100	0.9900	0.9692	0.9200	0.9900
supramarginal	0.9712	0.9300	0.9900	0.9222	0.8100	0.9700	0.9076	0.7800	0.9600
transversetemporal	0.9125	0.7900	0.9600	0.8846	0.7300	0.9500	0.8080	0.5800	0.9200

Table C.2: Table presenting the ICC(3,1) as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the three different anatomical measurements: cortical thickness (CT), grey matter volume (GMV), and surface area (SA), for the inter-software reliability testing of the second scan from the OASIS dataset.

Region of interest	ICC CT	LL	UL	ICC GMV	LL	UL	ICC SA	LL	UL
caudalanteriorcingulate	0.8871	0.7400	0.9500	0.9306	0.8300	0.9700	0.9271	0.8300	0.9700
caudalmiddlefrontal	0.9358	0.8500	0.9700	0.8384	0.6400	0.9300	0.8231	0.6100	0.9300
cuneus	0.9718	0.9300	0.9900	0.9286	0.8300	0.9700	0.9338	0.8400	0.9700
entorhinal	0.8694	0.7000	0.9500	0.6602	0.3200	0.8500	0.8360	0.6300	0.9300
fusiform	0.8875	0.7400	0.9500	0.9250	0.8200	0.9700	0.9131	0.7900	0.9600
inferiorparietal	0.9263	0.8200	0.9700	0.9424	0.8600	0.9800	0.9477	0.8700	0.9800
inferiortemporal	0.9277	0.8300	0.9700	0.9524	0.8800	0.9800	0.9242	0.8200	0.9700
insula	0.9465	0.8700	0.9800	0.9902	0.9800	1.0000	0.9820	0.9600	0.9900
isthmuscingulate	0.9468	0.8700	0.9800	0.9226	0.8200	0.9700	0.9458	0.8700	0.9800
lateraloccipital	0.9591	0.9000	0.9800	0.9603	0.9000	0.9800	0.9518	0.8800	0.9800
lateralorbitofrontal	0.8511	0.6600	0.9400	0.9743	0.9400	0.9900	0.9513	0.8800	0.9800
lingual	0.9618	0.9100	0.9800	0.9803	0.9500	0.9900	0.9743	0.9400	0.9900
medialorbitofrontal	0.8514	0.6600	0.9400	0.9485	0.8700	0.9800	0.9048	0.7800	0.9600
middletemporal	0.9204	0.8100	0.9700	0.9694	0.9200	0.9900	0.9623	0.9100	0.9800
paracentral	0.9503	0.8800	0.9800	0.9583	0.9000	0.9800	0.9622	0.9100	0.9800
parahippocampal	0.9380	0.8500	0.9700	0.8837	0.7300	0.9500	0.9082	0.7800	0.9600
parsopercularis	0.9528	0.8800	0.9800	0.9462	0.8700	0.9800	0.9510	0.8800	0.9800
parsorbitalis	0.9182	0.8100	0.9700	0.8697	0.7000	0.9500	0.8092	0.5800	0.9200
parstriangularis	0.9190	0.8100	0.9700	0.9594	0.9000	0.9800	0.9488	0.8800	0.9800
pericalcarine	0.9609	0.9000	0.9800	0.9768	0.9400	0.9900	0.9682	0.9200	0.9900
postcentral	0.9152	0.8000	0.9700	0.9744	0.9400	0.9900	0.9666	0.9200	0.9900
posteriorcingulate	0.9062	0.7800	0.9600	0.9639	0.9100	0.9900	0.9577	0.9000	0.9800
precentral	0.8559	0.6700	0.9400	0.9448	0.8700	0.9800	0.8451	0.6500	0.9400
precuneus	0.9392	0.8500	0.9800	0.9673	0.9200	0.9900	0.9693	0.9200	0.9900
rostralanteriorcingulate	0.8327	0.6300	0.9300	0.9428	0.8600	0.9800	0.9260	0.8200	0.9700
rostralmiddlefrontal	0.9327	0.8400	0.9700	0.9687	0.9200	0.9900	0.9490	0.8800	0.9800
superiorfrontal	0.9152	0.8000	0.9700	0.9533	0.8900	0.9800	0.9505	0.8800	0.9800
superiorparietal	0.9534	0.8900	0.9800	0.9175	0.8000	0.9700	0.9094	0.7900	0.9600
superiortemporal	0.9263	0.8200	0.9700	0.9589	0.9000	0.9800	0.9620	0.9100	0.9800
supramarginal	0.9343	0.8400	0.9700	0.9471	0.8700	0.9800	0.9317	0.8400	0.9700
transversetemporal	0.9319	0.8400	0.9700	0.8899	0.7400	0.9500	0.8534	0.6700	0.9400

Table C.3: Table presenting the ICC(3,1) as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the three different anatomical measurements cortical thickness (CT), grey matter volume (GMV), and surface area (SA), for the inter-software reliability testing of the first scan from the MIND dataset.

Region of interest	ICC CT	LL	UL	ICC GMV	LL	UL	ICC SA	LL	UL
caudalanteriorcingulate	0.9299	0.8700	0.9600	0.9265	0.8600	0.9600	0.9449	0.9000	0.9700
caudalmiddlefrontal	0.9010	0.8200	0.9500	0.7807	0.6200	0.8800	0.7735	0.6100	0.8800
cuneus	0.9231	0.8600	0.9600	0.9609	0.9300	0.9800	0.9740	0.9500	0.9900
entorhinal	0.9363	0.8800	0.9700	0.7649	0.5900	0.8700	0.7912	0.6300	0.8900
fusiform	0.9315	0.8700	0.9600	0.9337	0.8800	0.9600	0.9187	0.8500	0.9600
inferiorparietal	0.9576	0.9200	0.9800	0.9362	0.8800	0.9700	0.9308	0.8700	0.9600
inferiortemporal	0.9085	0.8300	0.9500	0.9018	0.8200	0.9500	0.9276	0.8700	0.9600
insula	0.9274	0.8600	0.9600	0.9826	0.9700	0.9900	0.9827	0.9700	0.9900
isthmuscingulate	0.9182	0.8500	0.9600	0.9188	0.8500	0.9600	0.9178	0.8500	0.9600
lateraloccipital	0.9452	0.9000	0.9700	0.9633	0.9300	0.9800	0.9493	0.9000	0.9700
lateralorbitofrontal	0.8455	0.7200	0.9200	0.9396	0.8900	0.9700	0.9395	0.8900	0.9700
lingual	0.8965	0.8100	0.9400	0.9490	0.9000	0.9700	0.9758	0.9500	0.9900
medialorbitofrontal	0.7398	0.5500	0.8600	0.9504	0.9100	0.9700	0.9225	0.8600	0.9600
middletemporal	0.9618	0.9300	0.9800	0.9531	0.9100	0.9800	0.9464	0.9000	0.9700
paracentral	0.9220	0.8600	0.9600	0.8456	0.7200	0.9200	0.8199	0.6800	0.9000
parahippocampal	0.9812	0.9600	0.9900	0.9361	0.8800	0.9700	0.9189	0.8500	0.9600
parsopercularis	0.9380	0.8800	0.9700	0.7528	0.5700	0.8600	0.6488	0.4200	0.8000
parsorbitalis	0.8524	0.7300	0.9200	0.7880	0.6300	0.8800	0.8422	0.7200	0.9100
parstriangularis	0.9056	0.8300	0.9500	0.7799	0.6200	0.8800	0.7580	0.5800	0.8700
pericalcarine	0.9195	0.8500	0.9600	0.9592	0.9200	0.9800	0.9529	0.9100	0.9800
postcentral	0.8325	0.7000	0.9100	0.8512	0.7300	0.9200	0.6819	0.4700	0.8200
posteriorcingulate	0.9466	0.9000	0.9700	0.9250	0.8600	0.9600	0.9518	0.9100	0.9700
precentral	0.8569	0.7400	0.9200	0.8705	0.7600	0.9300	0.8245	0.6900	0.9000
precuneus	0.9213	0.8500	0.9600	0.9386	0.8900	0.9700	0.9345	0.8800	0.9700
rostralanteriorcingulate	0.9103	0.8300	0.9500	0.9097	0.8300	0.9500	0.9405	0.8900	0.9700
rostralmiddlefrontal	0.8692	0.7600	0.9300	0.8470	0.7300	0.9200	0.8581	0.7400	0.9200
superiorfrontal	0.9232	0.8600	0.9600	0.9201	0.8500	0.9600	0.8982	0.8100	0.9500
superiorparietal	0.9597	0.9200	0.9800	0.7917	0.6300	0.8900	0.7454	0.5600	0.8600
superiortemporal	0.9674	0.9400	0.9800	0.9782	0.9600	0.9900	0.9779	0.9600	0.9900
supramarginal	0.9493	0.9000	0.9700	0.9122	0.8400	0.9500	0.8818	0.7800	0.9400
transversetemporal	0.9269	0.8600	0.9600	0.9632	0.9300	0.9800	0.9356	0.8800	0.9700

Table C.4: Table presenting the ICC(3,1) as well as the upper- (UL) and lower limit (LL) of the 95% confidence interval for the three different anatomical measurements: cortical thickness (CT), grey matter volume (GMV), and surface area (SA), for the inter-software reliability testing of the second scan from the MIND dataset.

Region of interest	ICC CT	LL	UL	ICC GMV	LL	UL	ICC SA	LL	UL
caudalanteriorcingulate	0.9266	0.8600	0.9600	0.8816	0.7800	0.9400	0.8975	0.8100	0.9500
caudalmiddlefrontal	0.9357	0.8800	0.9700	0.8140	0.6700	0.9000	0.7769	0.6100	0.8800
cuneus	0.9386	0.8900	0.9700	0.9600	0.9200	0.9800	0.9637	0.9300	0.9800
entorhinal	0.9333	0.8800	0.9600	0.7544	0.5800	0.8600	0.8291	0.7000	0.9100
fusiform	0.9191	0.8500	0.9600	0.9243	0.8600	0.9600	0.9129	0.8400	0.9500
inferiorparietal	0.9363	0.8800	0.9700	0.9590	0.9200	0.9800	0.9537	0.9100	0.9800
inferiortemporal	0.9243	0.8600	0.9600	0.8997	0.8200	0.9500	0.9212	0.8500	0.9600
insula	0.9516	0.9100	0.9700	0.9870	0.9800	0.9900	0.9846	0.9700	0.9900
isthmuscingulate	0.9158	0.8400	0.9600	0.9376	0.8800	0.9700	0.9381	0.8800	0.9700
lateraloccipital	0.9255	0.8600	0.9600	0.9600	0.9200	0.9800	0.9286	0.8700	0.9600
lateralorbitofrontal	0.8685	0.7600	0.9300	0.9440	0.8900	0.9700	0.9361	0.8800	0.9700
lingual	0.8985	0.8100	0.9500	0.9651	0.9300	0.9800	0.9603	0.9300	0.9800
medialorbitofrontal	0.8069	0.6600	0.8900	0.9268	0.8600	0.9600	0.9275	0.8700	0.9600
middletemporal	0.9395	0.8900	0.9700	0.9627	0.9300	0.9800	0.9480	0.9000	0.9700
paracentral	0.9020	0.8200	0.9500	0.9101	0.8300	0.9500	0.9242	0.8600	0.9600
parahippocampal	0.9663	0.9400	0.9800	0.9405	0.8900	0.9700	0.8960	0.8100	0.9400
parsopercularis	0.8956	0.8100	0.9400	0.8826	0.7900	0.9400	0.8597	0.7500	0.9200
parsorbitalis	0.9320	0.8700	0.9600	0.8437	0.7200	0.9200	0.8112	0.6700	0.9000
parstriangularis	0.8880	0.8000	0.9400	0.9384	0.8800	0.9700	0.9559	0.9200	0.9800
pericalcarine	0.9264	0.8600	0.9600	0.9475	0.9000	0.9700	0.9650	0.9300	0.9800
postcentral	0.8716	0.7700	0.9300	0.9570	0.9200	0.9800	0.9283	0.8700	0.9600
posteriorcingulate	0.9521	0.9100	0.9700	0.9363	0.8800	0.9700	0.9560	0.9200	0.9800
precentral	0.8535	0.7400	0.9200	0.9622	0.9300	0.9800	0.9204	0.8500	0.9600
precuneus	0.9388	0.8900	0.9700	0.9583	0.9200	0.9800	0.9567	0.9200	0.9800
rostralanteriorcingulate	0.9319	0.8700	0.9600	0.9340	0.8800	0.9700	0.9539	0.9100	0.9800
rostralmiddlefrontal	0.8364	0.7100	0.9100	0.9365	0.8800	0.9700	0.9543	0.9100	0.9800
superiorfrontal	0.9101	0.8300	0.9500	0.9606	0.9300	0.9800	0.9492	0.9000	0.9700
superiorparietal	0.9711	0.9500	0.9800	0.8657	0.7600	0.9300	0.8553	0.7400	0.9200
superiortemporal	0.9537	0.9100	0.9800	0.9592	0.9200	0.9800	0.9640	0.9300	0.9800
supramarginal	0.9379	0.8800	0.9700	0.9199	0.8500	0.9600	0.8824	0.7900	0.9400
transversetemporal	0.9455	0.9000	0.9700	0.9541	0.9100	0.9800	0.9236	0.8600	0.9600

DEPARTMENT OF MATHEMATICAL SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY