Aus der Neurologischen Universitätsklinik Tübingen
Abteilung Neurologie mit Schwerpunkt Neurodegenerative Erkrankungen

MRI substrates of specific neuropsychological dysfunctions within and across FTD genotypes at the presymptomatic and symptomatic disease stage

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Wabersich-Flad, Dominik David

2021

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Professor Dr. M. Synofzik
2. Berichterstatter: Professorin Dr. B. Derntl

Tag der Disputation: 27.09.2021

Contents

Titlepage **I**

Table of contents **III**

List of Figures **VII**

List of Tables **XIII**

Abbreviations **XV**

1 Introduction **1**

2 General Material and Methods **7**

2.1 GENFI 7

2.2 GENFI subjects 8

2.3 Surface based analysis framework 11

2.3.1 MRI data acquisition 13

2.3.2 MRI data processing 13

2.4 Testing framework 14

2.4.1 Permutation and maximum statistic based test extension 16

2.5 HCP Parcellation for the results 17

3 FTLD **19**

3.1 Material and Methods 19

3.1.1 Specific area detection model and tests 19

3.2 Results 23

3.2.1 HCP Parcellation for the FTLD results 25

3.3 Discussion 27

4 Apraxia **29**

4.1 Material and Methods 29

4.1.1 Specific area detection models and tests 32

4.1.2 Comparison of the apraxia results to apraxia in stroke 32

4.2 Results 35

4.2.1 Area detection Analysis 35

4.2.2 Comparison of the apraxia results to apraxia in stroke 43

4.2.3 HCP Parcellation for the apraxia results 45

4.3 Discussion 47

5 TMT **51**

5.1 Material and Methods 51

5.1.1 Specific area detection model and tests 52

5.2 Results 53

5.2.1 TMTA analysis 53

5.2.2 HCP Parcellation for the TMTA results 55

5.2.3 TMTB analysis 56

5.2.4 HCP Parcellation for the TMTB results 58

5.3 Discussion 59

6 Presymptomatic cortical degeneration **61**

6.1 Methods 61

6.2 FTLD related areas 63

6.2.1 Methods 63

6.2.2 Results 63

6.3 Apraxia related areas 66

6.3.1 Methods 66

6.3.2 Results 66

6.4 TMT related areas 69

6.4.1 Methods 69

6.4.2 Results 69

6.5 Discussion 72

7 General Discussion 75

8 Summary 87

8.1 Abstract 87

8.2 Zusammenfassung 88

9 References 89

10 Authorship – Eigenanteil 95

10.1 Statement of Authorship 95

10.2 Erklärung zum Eigenanteil 95

Appendix A FreeSurfer citation 97

Appendix B R code for Wald Test with permutation based test distribution 99

List of Figures

1 The left plot shows the number of subjects within each genetic group; the right plot shows the number of subjects within each genetic group only for subjects from new families, that have not been present in the GENFI dataset of the first phase (which ended in the beginning of 2015). M- depicts the M- group; all other groups are depicted by their mutated gene. 9

2 This plot shows the family relations of the subjects: on the abscissa, every single family is shown; the ordinate shows the number of subjects. Every bar thus represents the number of subjects in one family. The smallest bar height corresponds to a family with one subject. 10

3 Illustration of a mesh put on a brain to create a two dimensional vertex grid of the cortical surface. The magnifying glass shows a close up of a vertex and its neighbouring vertices, with every node being a vertex. . . . 12

4 Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+S+ vs. M+S- 23

5 Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$ used on the GENFI dataset with $\theta_{1(j)}$ modelling the difference value between M+S+ vs. M+S-. 24

6 Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A}$, capturing the group difference between M+S+ and M+S-. In addition, an overlay of the human cortex parcellation is shown. 25

7 FTLD-CDR-SOB scores histograms for the four generated data groups - from left to right and top to bottom: the M+A+ group, the M-A+ group, the M+A- group and the M-A- group. 31

8 Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M-A- 36

9 Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M-A- 37

10 Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M+A- 38

11 Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M+A- 39

12 Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. 62M+A- 40

13 Lineup of the left (top) and right (bottom) hemisphere results for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. Shown are the results for the compared groups: M+A+ vs. M+A- (red) and M+A+ vs. 62M+A- (blue) 41

14 Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M+A- 42

15 Inflated left hemisphere surface of the fsaverage subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A0}$, capturing the group difference between M+A+ and M+A- (blue areas). In addition, an overlay of the unspecific apraxia stroke mask is shown (black stripes). 43

16 Inflated left hemisphere surface of the fsaverage subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A0}$, capturing the group difference between M+A+ and M+A- (blue areas). In addition, an overlay of the specific apraxia stroke masks are shown (black stripes): from left to right: pantomime, finger, hand. 44

17 Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between M+A+ and M+A-. In addition, an overlay of the human cortex parcellation is shown. 45

18 Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: TMTA|D+ vs. TMTA|D- 53

19 Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$ used on the TMTA dataset with $\theta_{1(j)}$ modelling the difference value between TMTA|D+ vs. TMTA|D-. 54

20 Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between TMTA|D+ and TMTA|D-. In addition, an overlay of the human cortex parcellation is shown. 55

21 Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: TMTB|D+ vs. TMTB|D- 56

22 Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$ used on the TMTB dataset with $\theta_{1(j)}$ modelling the difference value between TMTB|D+ vs. TMTB|D-. 57

23 Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between TMTB|D+ and TMTB|D-. In addition, an overlay of the human cortex parcellation is shown. 58

24 The ordinate shows the mean thickness in masked areas (from the previous results of the M+S+ vs. M+S- analysis); the abscissa shows the age of the subjects. Blue circles denote M-, red crosses denote M+S- and green stars denote M+S+. The lower left inlay shows again the M+S+ with the same ordinate and the YSO on the abscissa. The lines show the group trend of the group with the same color, respectively. 64

25 Scatterplot with the mean thickness of the masked areas (by the previous results from the M+S+ vs. M+S- comparison) as ordinate and the EYO as abscissa. Red crosses delineate M+, whereas blue circles delineate M-. 65

26 Similar scatterplot as shown in Figure 25, however, for M+S+ datapoints, the EYO variable has been corrected to show the actual YSO. 65

27 The ordinate shows the mean thickness in masked areas (from the previous results of the M+A+ vs. M+A- analysis); the abscissa shows the age of the subjects. Blue circles denote M-, red crosses denote M+S- and green stars denote M+S+. The lower left inlay shows again the M+S+ with the same ordinate and the YSO on the abscissa. The lines show the group trend of the group with the same color, respectively. 67

28 Scatterplot with the mean thickness of the masked areas (by the previous results from the M+A+ vs. M+A- comparison) as ordinate and the EYO as abscissa. Red crosses delineate M+, whereas blue circles delineate M-. 68

29 Similar scatterplot as shown in Figure 28, however, for M+S+ datapoints, the EYO variable has been corrected to show the actual YSO. 68

30 The ordinate shows the mean thickness in masked areas (from the previous results of the TMTB|D+ vs. TMTB|D- analysis); the abscissa shows the age of the subjects. Blue circles denote non-mutation carriers, red crosses denote presymptomatic mutation carriers and green stars denote symptomatic mutation carriers. The lower left inlay shows again the symptomatic mutation carriers with the same ordinate and the YSO on the abscissa. The lines show the group trend of the group with the same color, respectively. 70

31 Scatterplot with the mean thickness of the masked areas (by the previous results from the TMTB|D+ vs. TMTB|D- comparison) as ordinate and the EYO as abscissa. Red crosses delineate M+, whereas blue circles delineate M-. 71

32 Similar scatterplot as shown in Figure 31, however, for M+S+ datapoints, the EYO variable has been corrected to show the actual YSO. 71

List of Tables

1 Mean age (with standard deviation in brackets) and Gender of the subjects. 9

2 Rating scale for the clinical researcher to judge examination results. This rating scale was used to judge limb apraxia on the left and right, as well as many other examination results. 10

3 Descriptive statistics for group M+S+ and group M+S-. Descriptive statistics are means with standard deviations in brackets. N is a placeholder for the number of subjects. Male, female and symptomatic gives the count of subjects for the respective group. 19

4 Fourfould table for the datased used with detailed information. Descriptive statistics are means with standard deviations in brackets. Male, female and symptomatic gives the count of subjects for the respective group. The two columns segregate M+ and M-, the two rows segregate probands with apraxia (A+) and probands without apraxia (A-). 30

5 Descriptive statistics for group 62M+A-. Descriptive statistics are means with standard deviations in brackets. N is a placeholder for the number of subjects. Male, female and symptomatic gives the count of subjects for the respective group. 32

6 Descriptive statistics for the TMTA analysis: the TMTA|D+ group and the TMTA|D- group; as well as for the TMTB analysis: the TMTB|D+ group and the TMTB|D- group. Descriptive statistics are means with standard deviations in brackets. N is a placeholder for the number of subjects. Male, female and symptomatic gives the count of subjects for the respective group. 51

7 Predictions for cortical thickness of fixed effects in main analysis. Columns depict age given in first row, rows depict group given in first column. The thickness values are given in mm. 66

8 Predictions for cortical thickness of fixed effects in alternative analysis. Columns depict age given in first row, rows depict group given in first column. The thickness values are given in mm. 67

Abbreviations

62M+A- 62 mutation carriers without apraxia. VIII, XIII, 32, 35, 40, 41, 47

ALS Amyotrophic Lateral Sclerosis. 1, 9

ASL Arterial Spin Labelling. 13

bvFTD behavioral variant of Frontotemporal Dementia. 1, 77

C9ORF72 chromosome 9 open reading frame 72. 1, 9, 82, 83

CBDS Corticobasal Degeneration Syndrome. 1, 83, 84

CDR Clinical Dementia Rating Scale. 9

CSF cerebrospinal fluid. 8

DICOM Digital Imaging and Communications in Medicine. 13

DTI Diffusion Tensor Imaging. 13

EYO expected years of onset. X, XI, 62, 64, 65, 67, 68, 70–72, 80–83, 85, 86

fMRI functional Magnetic Resonance Imaging. 13

FTD Frontotemporal dementia. 1, 3

FTLD Frontotemporal Lobar Degeneration. 1–5, 7–9, 11, 19, 27, 29, 32, 47, 48, 51, 59, 61, 63, 72, 75–78, 84, 86

FTLD-CDR-SOB FTLD modified CDR SOB. VII, 10, 29, 31, 48, 78

FWHM Full-Width-Half-Max. 14, 82

GENFI Genetic Frontotemporal Dementia Initiative. VII, 7–10, 13, 19, 24, 29, 47–49, 51, 59, 61, 63, 66, 69, 76, 80–83, 85, 86

GERS GENFI Examination Rating Scale. 10, 29

GRN progranulin. 1, 9

HCP Human Connectome Project. VII, IX, X, 4, 17, 25, 45, 55, 58, 79

HPCC High-Performance Computing Cluster. 13

LMM linear mixed model. 61

LRT Likelihood-Ratio-Test. 61, 62

M+ mutation carriers. X, XI, XIII, 8, 9, 19, 29, 30, 51, 62, 64, 65, 67, 68, 70–72

M+A+ mutation carriers with apraxia. VII–X, 29, 31, 32, 35–45, 47, 66–68

M+A- mutation carriers without apraxia. VII–X, 29, 31, 32, 35, 38–45, 47, 66–68

M+S+ symptomatic mutation carriers. VII, X, XI, XIII, 8, 19, 20, 23–25, 27, 61–72

M+S- presymptomatic mutation carriers. VII, X, XIII, 8, 19, 23–25, 27, 61–67, 69, 72

M- non-mutation carriers. VII, X, XI, XIII, 8–10, 19, 29, 30, 48, 51, 52, 61–71

M-A+ non-mutation carriers with apraxia. VII, 29, 31, 48

M-A- non-mutation carriers without apraxia. VII, VIII, 29, 31, 32, 35–37, 47

MAPT microtubule-associated protein tau. 1, 9

MCA middle cerebral artery. 77

MLM multivariate linear model. 15, 16, 19, 35

MND Motor Neuron Disease. 1

MRI Magnetic Resonance Imaging. 2, 8, 11, 13, 17, 79

NIfTI Neuroimaging Informatics Technology Initiative. 13

PA Progressive Aphasia. 1

PNFA Progressive Nonfluent Aphasia. 84

PSP Progressive Supranuclear Palsy. 1

SD Semantic Dementia. 1

SOB sum of boxes. 9

SPECT single-photon emission computed tomography. 83

TBK1 TANK-binding kinase 1. 9

TMT Trail Making Test. 4, 5, 10, 51, 59, 60, 75, 86

TMTA TMT Subtest A. IX, XIII, 51, 52, 54, 59, 60, 82

TMTA|D+ deficiency in TMTA. IX, XIII, 51–55, 59

TMTA|D- no deficiency in TMTA. IX, XIII, 51–55, 59

TMTB TMT Subtest B. IX, XIII, 51, 52, 57, 59, 60

TMTB|D+ deficiency in TMTB. IX–XI, XIII, 51, 52, 56–59, 69–71

TMTB|D- no deficiency in TMTB. IX–XI, XIII, 51, 52, 56–59, 69–71

YSO years since onset. X, XI, 64, 65, 67, 68, 70–72

1 Introduction

Frontotemporal dementia (FTD) is a neurodegenerative disorder, that commonly affects men and women at the age of 45 to 65 years, lasts around six to eight years and ends with death (Neary, Snowden, & Mann, 2005). It has a prevalence rate of 15 among 100.000 in the 45-64 age group (Ratnavalli, Brayne, Dawson, & Hodges, 2002). Patients suffering from the disease show slowly progressing behavioral change, executive dysfunctions and language problems. Up to now, there seems to be no general consensus on the term FTD, however usually it refers to the most prominent of the Frontotemporal Lobar Degeneration (FTLD) spectrum disorders, more specifically called behavioral variant of Frontotemporal Dementia (bvFTD). Often it is used vicariously for all FTLD spectrum disorders, as they all share a similar brain pathology: the atrophy of the prefrontal and anterior temporal lobes (Neary et al., 2005; Hodges et al., 2004). Originally focusing only on three subtypes, bvFTD, Progressive Aphasia (PA) and Semantic Dementia (SD) (Neary et al., 1998), the recognition of other diseases with similar pathology introduced Corticobasal Degeneration Syndrome (CBDS), Progressive Supranuclear Palsy (PSP) and Motor Neuron Disease (MND)[1] to the FTLD spectrum diseases (Seelaar, Rohrer, Pijnenburg, Fox, & Van Swieten, 2010; Kertesz, McMonagle, Blair, Davidson, & Munoz, 2005). These terms all refer to disease syndromes, i.e. clinically defined entities that summarize the occurrence of specific symptoms to one syndrome. The underlying pathology for these diseases is referred to by the term FTLD. To avoid further confusion, I will use the term FTLD vicariously to refer to the whole group of neurodegenerative diseases, that are caused by FTLD.[2]

FTLD is highly heritable: approximately 30% - 50% of patients with FTLD have genetic FTLD (Rohrer & Warren, 2011), with three well identified genetic mutations in the following genes: chromosome 9 open reading frame 72 (C9ORF72), microtubule-associated protein tau (MAPT) and progranulin (GRN) (Boeve et al., 2012; Mahoney et al., 2012; Seelaar et al., 2010). All of these genetic mutations are autosomal dominant, and carriers will at some point in their life develop a disease (Chow, Miller, Hayashi, & Geschwind, 1999; Rohrer et al., 2009). Logically, genetic FTLD, sometimes also called familial FTLD, refers to FTLD cases, in which the impacted persons carry a genetic mutation, that is known to cause FTLD, including presymptomatic and symptomatic persons if not otherwise specified. The description presymptomatic/symptomatic refers to the clinical status of a person with known FTLD mutations. A person who fulfills all clinical diagnostic criteria for a FTLD spectrum disease would be labeled as symptomatic. In contrast, persons who do not fulfill the criteria would be considered as presymptomatic. Please note that although the status of a person might be presymptomatic, this person might show individual clinical symptoms but does not yet fulfill all diagnostic criteria.[3]

---

[1] Sometimes more specifically referred to Amyotrophic Lateral Sclerosis (ALS).

[2] It seemed logical to use the term FTLD as representative to describe all FTLD spectrum diseases, however strictly speaking this is also not correct, as FTLD in itself is not a disease or syndrome, it is just the pathological process underlying all of these diseases.

Over the years, not only clinical, but also histological and biochemical criteria for the FTLD spectrum diseases have been discussed (Englund et al., 1994). The discussion on valid and reliable diagnostic criteria is still ongoing (Kertesz, Hillis, & Munoz, 2003) and can be quite confusing, as different researchers might use the terms in different ways: while a pathologist might diagnose a FTLD spectrum disease based exclusively on histological criteria, a clinical researcher uses only clinical criteria for the diagnosis. A laboratory clinician might even use genetic mutations as basic reason for a diagnosis. Further complicating the issue is the fact that they might even find different diagnoses for the same individual using their respective criteria. The problem is that there seems to be no clear and easy relationship between the histologic pathology, genetics and clinical presentation of the diseases (Neary et al., 2005; Kertesz et al., 2005; Hodges et al., 2004).

While the objective pathological process in the brain can only fully be assessed through an autopsy, Magnetic Resonance Imaging (MRI) offers a technique to at least judge anatomical changes in the brain on a macro anatomical scale. Not only can anatomical changes in the brain, mainly cortical atrophy, be detected in the MRI, previous studies also identified anatomical differences between genetic groups and correlated them with various clinical measurements and neuropsychological tests (Whitwell et al., 2012; Rohrer, Ridgway, et al., 2010; Whitwell et al., 2009). The different genetic mutations that cause FTLD could lead to mutation-specific disease processes and thus to distinct anatomical brain changes during the course of the disease, which at last all end with widespread frontotemporal degeneration.

Because of a typically insidious beginning, it is reasonable to assume that the disease begins before any overt clinical symptoms could be detected. In fact, past MRI studies have shown, that prior to the noticeable onset of the disease, often also referred to as the disease being at a clinical stage, changes in the brain are already detectable (Bertrand et al., 2018; Papma et al., 2017; Rohrer et al., 2015). Thus, it does make sense to take a look not only at symptomatic (i.e. clinically affected) FTLD subjects, but also at subjects, that are presymptomatic.

---

[3] In the course of this work, I often speak about symptoms of presymptomatic persons, which seems to be a contradiction, but does make sense in the context of the given definition of a symptomatic person.

As a first step, I will establish the presented methods by replicating the general and well known atrophy pattern of FTLD, through comparing symptomatic with presymptomatic subjects. The identified areas of this analyses will then serve as a starting point for a subsequent presymptomatic analysis, presented further below.

Next, to improve our understanding of brain changes in a neurodegenerative disease, I would like to take a closer look at where these changes take place and with which neuropsychological functions these brain areas are related. The identification of neuropsychological dysfunction related areas could help to understand the progression of atrophy patterns as well as the clinical diagnosis in relation to the specific atrophy pattern in genetic FTLD. Different disease entities could be caused by the same or similar pathological process, but different progression of the atrophy pattern with the corresponding deficits in the patients. I chose to investigate the neuropsychological deficit apraxia and its brain substrates.

The decision to investigate apraxia within FTD came from the belief, that the relevance of this symptom is underestimated for FTLD patients and that it might be a symptom that is present frequently but often overlooked (Bertrand et al., 2018; Johnen et al., 2016). Apraxia can be described as dysfunctional actions, that are not caused by loss of motor strength or coordination problems. More specifically, apraxia can be divided into three subdomains: disturbed imitation of gestures, disturbed tool-use and disturbed gesture production on command (Goldenberg, 2014). In the following, I will use the term apraxia as an umbrella term for gesture imitation deficits, if not otherwise specified. I assume that apraxia has a neurological cause that should be detectable in the cortex, as suggested by previous research about apraxia in stroke: Buxbaum, Shapiro, and Coslett (2014) used voxel-based lesion-symptom mapping and found associations of gesture tasks with lesions in left inferior parietal and frontal regions. Regions that are usually affected in patients with FTLD as well. Moreover, studying anatomical brain correlates of apraxia (i.e. gesture imitation deficits) within genetic FTLD could also support apraxia research, as most studies about brain correlates of apraxia originate from patients who suffered a stroke (see, e.g. Buxbaum et al., 2014; Goldenberg, Hermsdörfer, Glindemann, Rorden, & Karnath, 2007; Goldenberg & Karnath, 2006).

Up until now, apraxia is not a symptom typically discussed in the context of FTLD. Executive dysfunctions, on the other hand, represent a typical feature of patients with FTLD (Neary et al., 2005). The performance of patients in the Trail Making Test (TMT) (Bowie & Harvey, 2006) is associated with executive dysfunction (Rohrer et al., 2015; Demakis, 2004; Reitan & Wolfson, 1995), which is why I chose to investigate TMT performance deficits and their brain substrates as another use case of the presented methods.

In addition to the above, I will show my results in light of the Human Connectome Project (HCP) based parcellation by Glasser et al. (2016). The HCP is a very recent, huge and promising approach to the mapping of the human connectome and providing its data to the scientific community (Van Essen et al., 2013, 2012). Glasser et al. (2016) is one study that, based on the HCP data, provided an excellent parcellation of the human cortex, based on the latest advances in technology, using several different modalities on a healthy and young cohort and additionally taking previous knowledge of brain areas into account. Showing my results on the basis of this parcellation could help creating a bridge to common brain areas and their function.

At last, I will look for presymptomatic changes in the brain. Rohrer et al. (2015) as well as Bertrand et al. (2018) have previously demonstrated, that presymptomatic changes in the brain can be detected for genetic FTLD subjects. I plan on using a different approach, yet also substantially trying to reveal presymptomatic changes: I will use the areas detected in the preceding analyses as masks. I suspect that the masked areas could underlie presymptomatic changes, which I will try to reveal in this subsequent analysis. Thus, this subsequent analysis will only concentrate on the masked brain areas and not the whole brain.

To sum up, I first establish my methods and replicate expected findings about genetic FTLD subjects: I hypothesize to find brain areas with reduced cortical thickness among symptomatic genetic FTLD subjects compared to presymptomatic genetic FTLD subjects. This analysis should in addition serve as a comparison of brain atrophy pattern, by providing the typical FTLD atrophy in the temporal and frontal cortex.

In the apraxia analysis, I hypothesize to find brain areas with reduced cortical thickness among (symptomatic and presymptomatic) genetic FTLD subjects with apraxia compared to non-mutation carriers (i.e. supposedly healthy controls) and find brain areas with reduced cortical thickness, that are specific to a group of (symptomatic and presymptomatic) genetic FTLD subjects with apraxia, compared to (symptomatic and presymptomatic) genetic FTLD subjects without apraxia and find different brain areas within the different genetic groups of FTLD. Also, I demonstrate my results in comparison to previous research of apraxia in stroke.

In a third analysis, I will again use the same methods and technique for another use case: TMT performance. I hypothesize to find brain areas with reduced cortical thickness, that are specific to a group of (symptomatic and presymptomatic) genetic FTLD subjects with TMT deficits, compared to (symptomatic and presymptomatic) genetic FTLD subjects without TMT deficits.

In a subsequent analysis of the three just proposed analyses, I will try to find presymptomatic changes in the identified brain regions, respectively. I hypothesize to find reduced cortical thickness prior to the symptomatic onset of the disease, in the previously detected brain areas for genetic FTLD subjects (for the three analyses and associated brain areas respectively).

After presenting the general material and methods, I will present the three different analyses (Apraxia / Symptomatic / TMT) and the presymptomatic analyses separately and discuss them altogether in the general discussion at the end.

2 General Material and Methods

The Genetic Frontotemporal Dementia Initiative (GENFI) study cohort consists of symptomatic genetic FTLD subjects and their relatives, who are at risk of developing FTLD within their life. The at-risk FTLD subjects divide thus into two subgroups: presymptomatic FTLD subjects, who carry a genetic mutation that is known to cause FTLD, but do not yet fulfill all clinical diagnostic criteria for a FTLD spectrum disease; and control subjects, who do not carry such a known genetic mutation and therefore will most likely not develop FTLD. A symptomatic person, i.e. a person affected by a FTLD spectrum disease, is a person who has been clinically diagnosed by a neurologist as having a disorder in the FTLD disease spectrum.

2.1 GENFI

GENFI is a consortium of 22 clinical research centers across Europe (UK, Netherlands, Belgium, France, Spain, Portugal, Italy, Germany, Sweden) and Canada. In each of these research centers, patients and relatives of patients meeting the inclusion criteria (described further below), got selected to be part of the GENFI study dataset. This large group of research centers is necessary, to create a dataset like the GENFI dataset, as FTLD is rare.

Subjects included in the GENFI dataset were either known genetic FTLD patients or first-degree relatives of known genetic FTLD patients, who have a 50% chance of carrying the genetic mutation as well and are therefore termed at-risk subjects.

The GENFI data acquisition process consisted of two phases: GENFI 1 and GENFI 2. The data from GENFI 1 was collected from 2012 to 2015 and represents the initial phase of the GENFI data collection, which is further described in Rohrer, Warren, Fox, and Rossor (2013). The first published work with parts of the GENFI 1 dataset is described in Rohrer et al. (2015). In 2015, the second phase began, with data collection running up until now. Compared to the first phase, the data of the second phase underwent some changes. This is why the collection of subjects was restarted and even though previous subjects were partially included, they had to redo all data acquisition examinations (which will be further described below). I used the data from GENFI 2, which started with data collection in March 2015 up until the third data freeze time point on the 30th of January

2017.[4] From now onwards, all references to the GENFI data will refer to the GENFI 2 data from March 2015 to the end of January 2017.

The GENFI data includes: clinical examination, cerebrospinal fluid (CSF) and blood biomarkers, neuropsychological testing, structural and functional MRI scans, and genetics. Of course, not all of this information is relevant for the here presented approach: I only used parts of the data from the clinical examination, the neuropsychological testing and some of the MRI scans.

After collection of the data at each research center, the data was pseudonymized and securely sent to the main research coordination center at the University College London. Every research center had approved consent by the local ethics committee as well as written informed consent by every subject of the study. This evidently includes the research center in Tübingen, where this work originates from. Note, that the data collection process was strictly separated from data analysing projects (as this one), which is why at no point during my work I was able to identify single individuals in the anonymized data, provided by the main research coordination center in London.

2.2 GENFI subjects

The original GENFI data consists of 533 different subjects from 224 different families. 123 of the subjects showed up a second time one year later for a followup, resulting in a total of 656 data rows. Of the 533 different subjects, 442 subjects (83%) are at-risk, i.e. without a diagnosis and not affected by FTLD. Of the 442 at-risk subjects, 237 (54%) are presymptomatic and therefore mutation carriers (M+) and 205 (46%) are control subjects and thus non-mutation carriers (M-). Logically, the 91 subjects, who are clinically affected by the disease are all M+. Also, the symptomatic M+ have been the starting point for the recruiting process of study participants: symptomatic M+ were asked for study participation in addition to the identification and study participation proposal of their first degree relatives.

I excluded subjects without T1-weighted MRI scans (46), T1-weighted MRI scans of insufficient quality for image processing (14), and one participant with wrong demographic data. The resulting sample of 472 subjects comprised 77 (16%) symptomatic mutation carriers (M+S+) and 217 (46%) presymptomatic mutation carriers (M+S-), as well as 178 (38%) M-. If a subject had two visits with usable data, the last visit was included, to avoid duplicate subjects.

---

[4] The collection of data is still running, but the data which was used in this work goes up until the described data freeze time point.

Figure 1 shows the distribution of genetic alterations among the subjects of the GENFI dataset. I created one plot with all families and another plot only with subjects from families that have not already been present in the first phase of the data collection. The distribution of genetic groups differs between the different phases of data collection. Also note, that there is a fourth genetic group: TANK-binding kinase 1 (TBK1), which was recently discovered to be relevant for FTLD (Freischmidt et al., 2015). However, there is only one subject of this group in the dataset (so far). To address the family relations of the subjects, see Figure 2. Table 1 shows the age and gender of the M- group and the different genetic M+ groups.[5]

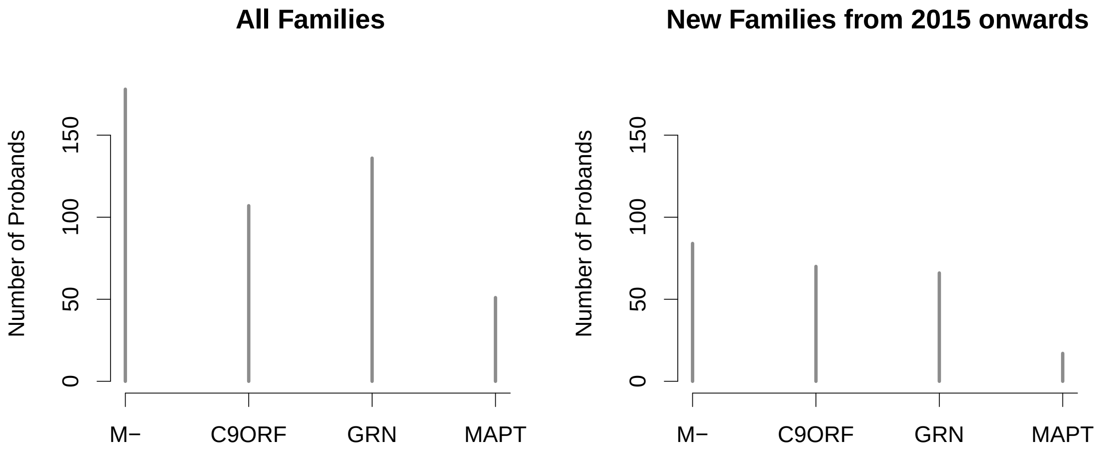

Figure 1: The left plot shows the number of subjects within each genetic group; the right plot shows the number of subjects within each genetic group only for subjects from new families, that have not been present in the GENFI dataset of the first phase (which ended in the beginning of 2015). M- depicts the M- group; all other groups are depicted by their mutated gene.

Table 1: Mean age (with standard deviation in brackets) and Gender of the subjects.

	M-	C9ORF72	GRN	MAPT
Age	46.3 (12.7)	52.0 (13.9)	50.3 (13.4)	45.9 (12.0)
Male/Female	78/100	47/60	51/85	22/29

Education of each subject was asked or, if asking was not possible, assessed by the clinical researcher and noted in years of education. Subjects also underwent a FTLD modified Clinical Dementia Rating Scale (CDR) and their score was noted as sum of boxes (SOB),

---

[5] The TBK1 group was deliberately omitted in the table, as it contains only one male subject with the age of 38, who is diagnosed with ALS.

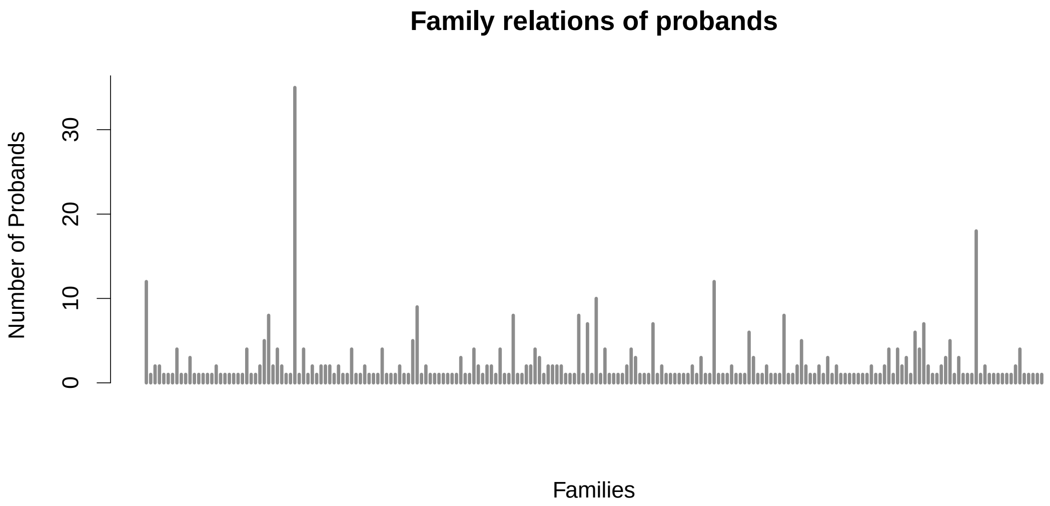

Figure 2: This plot shows the family relations of the subjects: on the abscissa, every single family is shown; the ordinate shows the number of subjects. Every bar thus represents the number of subjects in one family. The smallest bar height corresponds to a family with one subject.

later referenced as FTLD modified CDR SOB (FTLD-CDR-SOB). Apraxia was tested in the clinical examination by a trained neurologist for every left and right hand separately and judged on an ordinal scale from zero to three. Table 2 depicts the scores and their respective meaning of the rating scale used; this rating scale will from now onwards be referred to as GENFI Examination Rating Scale (GERS). The TMT was conducted by a trained psychologist and resulting test scores have been standardized relative to the M- scores in the GENFI cohort. Much more information about each single subject was collected, however only the variables and scores, that are relevant for this work were mentioned and described.

Table 2: Rating scale for the clinical researcher to judge examination results. This rating scale was used to judge limb apraxia on the left and right, as well as many other examination results.

Score	Meaning
0	Absent
0.5	Questionable / Very mild
1	Mild
2	Moderate
3	Severe

2.3 Surface based analysis framework

I decided to analyse the anatomical MRI data with a surface based approach: this means concentrating on data from the cortical surface and disregarding information outside of the cortex. This decision was made, as I was investigating changes in a neurodegenerative disorder and the main parameter of interest is the neurodegeneration of the cell bodies, which are located in the cortex. While completely neglecting the subcortical nuclei, I focused solely on the cortex surface: especially the thickness of the cortex should be a good measure to capture the atrophy, that takes places in a neurodegenerative disorder like FTLD. The used surface-based approach provides the thickness of the cortex at different locations of the cortical surface (more details on this will be described in section 2.3.2).

In order to investigate the surface of the brain, imagine putting a mesh on a brain that covers the whole surface of the brain, including all sulci and gyri. The mesh is basically a grid consisting of several nodes. Each node has a specific location on the brain surface. These nodes will be called vertices (singular: vertex). For an illustration, see Figure 3. Typically, in a surface-based approach, one uses these vertices to refer to a specific location of the cortex. A vertex contains information about its neighbouring vertices and additional properties. Cortical thickness is one of those additional properties a vertex supplies. Vertex maps thus supply a two dimensional map of the cortex, with additional information like cortical thickness at every vertex. In contrast to voxels, typically used in a volume-based approach, vertices only need two dimensions, not three. Furthermore, the brain surface can be inflated, like a balloon, so there are no more visible sulci and gyri. The inflation helps in visualizations, to see all areas of the brain, while keeping the surface structure consistent.

In order to use surface based data from the MRI images, I needed to process the original data with FreeSurfer (for a good overview of the software and what it does, see Fischl, 2012), which is freely available online (`http://surfer.nmr.mgh.harvard.edu/`), together with wiki pages documenting its usage and giving further references to the scientific literature.

I used FreeSurfer v6.0.0 on Linux for all presented analyses. A summary of the underlying steps is given below, while an extensive description of the FreeSurfer processing methods is given in Appendix A.

The FreeSurfer cortical reconstruction processes performed by the `recon-all` routine

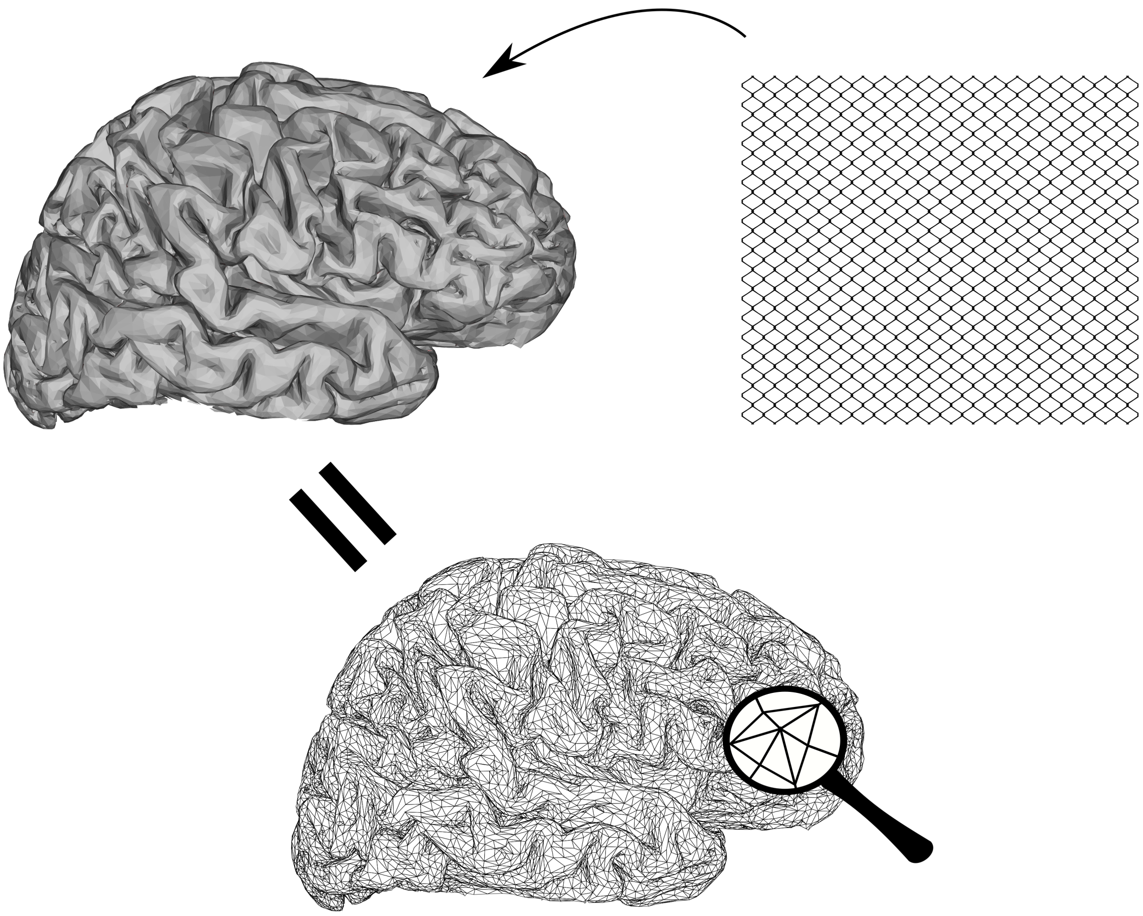

Figure 3: Illustration of a mesh put on a brain to create a two dimensional vertex grid of the cortical surface. The magnifying glass shows a close up of a vertex and its neighbouring vertices, with every node being a vertex.

include motion correction, skull stripping, intensity normalization, automated Talairach transformation, segmentation of (subcortical) grey and white matter structures, tessellation of (cortical) gray and white matter boundary, surface deformation following intensity gradients to optimize borders between tissue, registration to enable comparisons of cortical thickness across subjects, parcellation respecting gyral and sulcal structures, finally resulting in the creation of surface based data including cortical thickness values for every vertex. As described later, a smoothing algorithm is applied on the surface data, before using it for the upcoming analyses. Also note, that the resulting surface data is not restricted to the voxel resolution of the original data, which is why the detection of submillimeter differences between the groups is possible (Fischl & Dale, 2000). To visualize the data an inflated surface was used.

2.3.1 MRI data acquisition

At every site, MRI data was acquired according to the standardized GENFI 2 protocol. Images acquired include: T1-weighted, T2-weighted, Diffusion Tensor Imaging (DTI), resting state functional Magnetic Resonance Imaging (fMRI), FieldMap and Arterial Spin Labelling (ASL). All images were acquired on a 3T scanner. For the upcoming analysis only the anatomical T1-weighted and T2-weighted images were required.

The T1-weighted images were acquired with the following parameters: $TR = 2000ms$, $TE = 2.85ms$, $TI = 850ms$, $FoV = 282mm$, 208 sagittal slices, slice thickness of $1.1mm$, flip angle of 8 degree, acquisition time of 8 minutes and 32 seconds, phase encoding direction from anterior to posterior, base resolution of 256.

The T2-weighted images were acquired with the following parameters: $TR = 3200ms$, $TE = 401ms$, $FoV = 282mm$, 176 sagittal slices, slice thickness of $1.1mm$, acquisition time of 4 minutes and 46 seconds, phase encoding direction from anterior to posterior, base resolution of 256.

2.3.2 MRI data processing

The MRI images of the subjects were all in the Digital Imaging and Communications in Medicine (DICOM) standard format. As FreeSurfer processes original data in the Neuroimaging Informatics Technology Initiative (NIfTI) format, the first requirement was to convert the DICOM data to the NIfTI data format. For an overview, on the different formats and how to convert them, see Li, Morgan, Ashburner, Smith, and Rorden (2016). I used the `dcm2niix` command line tool, together with a python script to automatically call the conversion command for every file separately and write out useful log files to see if there occurred any relevant errors during conversion.

After the data was converted to the NIfTI file format, I processed the data with FreeSurfer. To do this, I used the command line tool `recon-all`, together with a python script to automatically call the processing routine on a High-Performance Computing Cluster (HPCC) for every subject separately and write out useful log files during this process.

For the FreeSurfer processing routine, T1-weighted MRI images have been used for every subject, in addition to T2-weighted images for better results (Lindroth et al., 2019), if present for the subject.

The FreeSurfer standard routine creates a brain surface for every subject, consisting of several vertices including their location with information about neighbouring vertices and additional properties like the thickness of the cortex at the vertex. However, as the brain surface is specifically created for a single subject, another step is necessary in order to compare subjects or groups of subjects with one another: resample the surface data to a common space. `fsaverage` is an average subject, that is included in FreeSurfer and can be used as a common space; its creation is described in Fischl, Sereno, Tootell, and Dale (1999). As the cortical thickness values represent the parameter of interest, I resampled the cortical thickness values of the subjects to the fsaverage common space. In addition, following standard conventions, I defined a Gaussian kernel Full-Width-Half-Max (FWHM) value of 10 mm, to smooth the thickness values.

All further references to thickness values (of subjects), will refer to cortical thickness values, that were smoothed with a FWHM value of 10 mm and resampled to the fsaverage common space. This was done for every subject and thickness values are given for every vertex of the common space. In accordance, all further references to vertices will (if not otherwise specified) refer to the vertices of the fsaverage common space.

Also note, that in surface-based analyses it is common to split the brain in two hemispheres: the left and the right hemisphere. Thus, all upcoming references to vertices and corresponding values of vertices like the cortical thickness are always vertices of one hemisphere.

2.4 Testing framework

All statistical analyses were performed with the free software environment for statistical computing R (R Core Team, 2018) and with the FreeSurfer tools. For the later presented permutation and maximum statistic based approach, an analysis with the FreeSurfer tools was not possible, as this method was newly developed inside R. Thus it could only be conducted with R.

For the statistical analyses, I used a multivariate linear model (MLM): see Equation 1.

$$\underset{(I\times J)}{\boldsymbol{Y}} = \underset{(I\times K)}{\boldsymbol{X}} \cdot \underset{(K\times J)}{\boldsymbol{B}} + \underset{(I\times J)}{\boldsymbol{E}} = \begin{bmatrix} y_{1(1)} & y_{1(2)} & \cdots & \cdots & \cdots & y_{1(J)} \\ y_{2(1)} & y_{2(2)} & \cdots & \cdots & \cdots & y_{2(J)} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & y_{i(j)} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ y_{I(1)} & y_{I(2)} & \cdots & \cdots & \cdots & y_{I(J)} \end{bmatrix} \tag{1}$$

$\boldsymbol{Y}$ denotes the dependent variable, which is a matrix with $i = 1, \ldots, I$ rows and $j = 1, \ldots, J$ columns, where i denotes the subject index and j denotes the vertex index. I and J denote the absolute and complete number of subjects and vertices respectively. $\boldsymbol{X}$ is the design matrix with I rows and K columns, where K denotes the absolute and complete number of regression parameters. $\boldsymbol{B}$ is the regression matrix with $k = 1, \ldots, K$ rows and J columns. It therefore consists of a set of regression parameters for each vertex respectively. At last, the variable $\boldsymbol{E}$ denotes the individual error for every subject at every vertex respectively and thus is a $I \times J$ matrix. This model further assumes that the errors $\varepsilon_{i(j)}$ follow a normal distribution as shown in Equation 2.

$$\boldsymbol{E} = \begin{bmatrix} \varepsilon_{1(1)} & \varepsilon_{1(2)} & \cdots & \varepsilon_{1(J)} \\ \varepsilon_{2(1)} & \varepsilon_{2(2)} & \cdots & \varepsilon_{2(J)} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{I(1)} & \varepsilon_{I(2)} & \cdots & \varepsilon_{I(J)} \end{bmatrix}, \ \varepsilon_{i(j)} \sim N(0, \sigma) \tag{2}$$

As the dependent variable, I used the cortical thickness values (of one hemisphere), so $y_{i(j)}$ denoted the cortical thickness value of vertex j for the i-th subject. The normality assumption for the errors has been visually examined by looking at scatterplots and histograms of randomly chosen subject and vertex samples of the data.

The parameters of the model can be tested with Wald tests to check if they are different from zero. The test statistic of the Wald test is shown in Equation 3. When testing a parameter, the test statistic is compared to a theoretical distribution and, based on its location in the theoretical distribution, the test decision is made. However, by looking at tests for single parameters repeatedly for every vertex, the test theoretical basis for deciding if a result is significant is no longer given: due to the accumulation of type one errors, a correction for multiple testing is necessary. Thus, this method cannot be used to decide which parameters at which vertices do have a significant role, without correcting it for multiple comparisons. However, I still used the uncorrected test as a first method to identify vertices, that *might* have a significant role, but then used a test that corrects for multiple comparisons in order to see if the parameters significantly differed from zero and thus have a significant role in the model.

$$T_{Wald} = \frac{\hat{\theta} - \theta_0}{se(\hat{\theta})} \tag{3}$$

To solve the problem of multiple testing, the general solution offered by FreeSurfer is to use a permutation based cluster-wise correction (see, e.g. Hyatt, Haney-Caron, & Stevens, 2012). I wanted to use a vertex-wise approach, that does not directly take into account the size of a significant cluster, i.e. a coherent significant area, as I feared that this could lead to small significant areas being neglected. Therefore, I used the solution presented next.

2.4.1 Permutation and maximum statistic based test extension

I now present a method, based on the testing framework of Blair and Karniski (1993), to test single parameters of a model. With the presented method, the problem of multiple comparisons gets solved, by using the absolute maximum statistic value of several permutations to create an empirical distribution for the test decision.

The test extends the previously presented Wald test (with the test statistic shown in Equation 3), with a correction for multiple comparisons: rather than comparing the test statistic to the theoretical distribution (which would be a normal distribution for the shown test statistic), I compare it to an empirical distribution of absolute maximum test statistics. The empirical distribution of absolute maximum test statistics is obtained by permutation of the grouping variable for every subject and estimating a new MLM for every permutation. From this newly estimated model, based on the data with permutated grouping labels, I extract the value for the Wald statistic for every parameter. To create the empirical distribution of absolute maximum test statistics, I take the absolute maximum value of the statistic for every parameter across all vertices. This means, that the empirical distribution of absolute maximum test statistics looks the same for one parameter at every vertex. Now, I can compare the value of the Wald test statistic for one parameter of the actual model estimated for the real data, with the just described empirical distribution of absolute maximum test statistics. If all possible permutations were performed and used to create the empirical distribution, the obtained p-value would be an exact p-value. However, if the number of subjects is high, calculating all possible permutations is practically not feasible. Therefore, approximate p-values can be computed, by creating an empirical distribution which is based on a high number of random permutations of the grouping variable (and the previously described subsequent steps). The here reported p-values based on this approach are always approximate p-values due to a high number of subjects. This test is therefore an approximate permutation test.

To further clarify the permutation approach, imagine the following: for a desired error rate of 5%, at least 20 permutations are necessary. Assuming one false positive among the 20 permutations results in a 5% error rate. This small sample would, however, not provide a confident test decision, due to a high amount of uncertainty about the small sample. On the other hand, calculating all possible permutations would be the ideal way, as no approximation is thus necessary. However, this was not possible for the later presented analysis, as it would have taken too long, while using lots of computational resources. This is why a reasonable big number of random permutations was chosen, somewhere in between the two presented extremes: the number of random permutations used to create the approximate p-values was 5000, which seemed reasonable after trying out different values. With an error rate of 5%, 250 out of the 5000 samples are false positives.

For technical details, see Appendix B for the actual R code of the described extended Wald test and a simulated usage example.

2.5 HCP Parcellation for the results

Glasser et al. (2016) present a parcellation of the human cortex, which is based on a healthy and young cohort with MRI data from different modalities. It is a fairly new parcellation of the cortex, based on HCP data (Van Essen et al., 2013, 2012). This is why I decided to use this parcellation to visualize my results. This should lead to a clearer view of which brain areas are actually affected.

The parcellation by Glasser et al. (2016) was made for the HCP 32k fs_LR subject. The later presented p-map overlays of my data were registered to the general fsaverage subject from FreeSurfer in 164k. It had to be resampled to the HCP 32k fs_LR subject with the workbench tool `wb_command`.

The images of the results were created using the connectome workbench viewer `wb_view`.

3 FTLD

In this analysis I will replicate the general and well known atrophy pattern of FTLD (Neary et al., 2005; Hodges et al., 2004).

3.1 Material and Methods

For this subset of the GENFI data, I selected all 294 M+ and divided them into two groups: the M+S+ group and the M+S- group. Note, that M- were excluded for this analysis. Table 3 provides descriptive statistics about this dataset.

Table 3: Descriptive statistics for group M+S+ and group M+S-. Descriptive statistics are means with standard deviations in brackets. N is a placeholder for the number of subjects. Male, female and symptomatic gives the count of subjects for the respective group.

	N (male/female)	**Age**	**Education**	**Symptomatic**
M+S+	77 (43/34)	63.09 (8.41)	12.73 (3.86)	77
M+S-	217 (77/140)	45.55 (11.83)	14.55 (3.25)	0

3.1.1 Specific area detection model and tests

In this section I will describe the actual specific MLM, based on the earlier described basic MLM, that I used for this analysis, together with the tests that have been conducted on the model parameters.

The model $\mathcal{M}_{A1}$ is shown in Equation 4. Remember, that $y_{i(j)}$ denotes the cortical thickness value of subject i at vertex j of one hemisphere. The model will be used on both, the left and the right hemisphere separately.

μ denotes the grand mean and θ denotes regression parameters of discrete variables, which are dummy coded in the design matrix. β denotes the regression parameter for a continuous variable.

The design matrix contains always a 1 for μ, thus modelling the overall mean. The design matrix contains a 1 for the $\theta_{1(j)}$ parameter, if the subject is in the M+S+ group, and it contains a 0 otherwise. Therefore, the $\theta_{1(j)}$ parameter models the difference between the groups, i.e. the difference value for the M+S+ group. As there are J parameters for $\theta_{1(j)}$, the difference between the groups is given for every single vertex. The $\theta_{2(j)}$

parameter takes a value of 1 if the subject is male and 0 if the subject is female, thus modelling the mean difference value for male subjects. The $\beta_{1(j)}$ parameter models the regression coefficient for age, while the $\beta_{2(j)}$ parameter models the regression coefficient for education.

$$\boldsymbol{Y} = \begin{bmatrix} 1 & 0 & 1 & age_i & edu_i \\ 1 & 1 & 0 & age_i & edu_i \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} \mu_{(1)} & \cdots & \mu_{(J)} \\ \theta_{1(1)} & \cdots & \theta_{1(J)} \\ \theta_{2(1)} & \cdots & \theta_{2(J)} \\ \beta_{1(1)} & \cdots & \beta_{1(J)} \\ \beta_{2(1)} & \cdots & \beta_{2(J)} \end{bmatrix} + \boldsymbol{E} \tag{4}$$

Every parameter of the model got tested whether it is different from zero with a Wald test (uncorrected for multiple comparisons), and the p-values for every vertex of this test were then written to a mgh file[6], which can be used as an overlay for the inflated surface (or any other surface) of the fsaverage subject. The resulting p-map, visualized with FreeSurfers *Freeview* tool, was thresholded with $p = 0.0001$ depicted by colored areas, the brightly colored areas with a threshold of $p = 0.00001$. Colored areas consequently mark areas with cortical thickness changes, i.e. reduced cortical thickness.

It is of interest to identify brain areas, that underlie significant changes, i.e. significantly reduced cortical thickness, in the M+S+ group. Apart from $\theta_{1(j)}$, the other parameters were included in the model to ensure that no confounding variable, like presumably an age, an education or a gender difference between the groups, is responsible for possible group differences of the M+S+ group compared to the other group. The inclusion of these parameters reduces the unexplained variance in the model thus providing a more sensitive test statistic for $\theta_{1(j)}$. Consequently, as the parameter $\theta_{1(j)}$ contains the interesting group difference, the test, if this parameter is significantly different from 0 would substantially test for group differences:

$$\mathcal{H}_0 : \theta_{1(j)} = 0$$

To test the hypothesis $\mathcal{H}_0$, I conducted the previously described permutation and maximum statistic based Wald test. The p-value for every vertex of this test was then written to a mgh file, which can be used as an overlay for the inflated surface (or any other surface) of the fsaverage subject. The resulting p-map, visualized with FreeSurfers

---

[6] Actually, not the p-value itself, but the common logarithm of the p-value was written to the mgh file, as FreeSurfer handles the p-values in this way. When using the FreeSurfer tools, this kind of file is usually named with the ending *.sig.mgh*.

Freeview tool, was thresholded with $p = 0.05$ depicted by colored areas, the brightly colored areas with a threshold of $p = 0.01$. Colored areas consequently mark areas with significant cortical thickness changes, i.e. significantly reduced cortical thickness.

3.2 Results

When comparing group M+S+ vs. group M+S-, widespread reduction in cortical thickness can be observed in frontal and temporal areas of the cortex. The corrected group M+S+ vs. group M+S- results for the $\theta_{1(j)}$ parameter are summarized in Figure 4 for the left and for the right hemisphere.

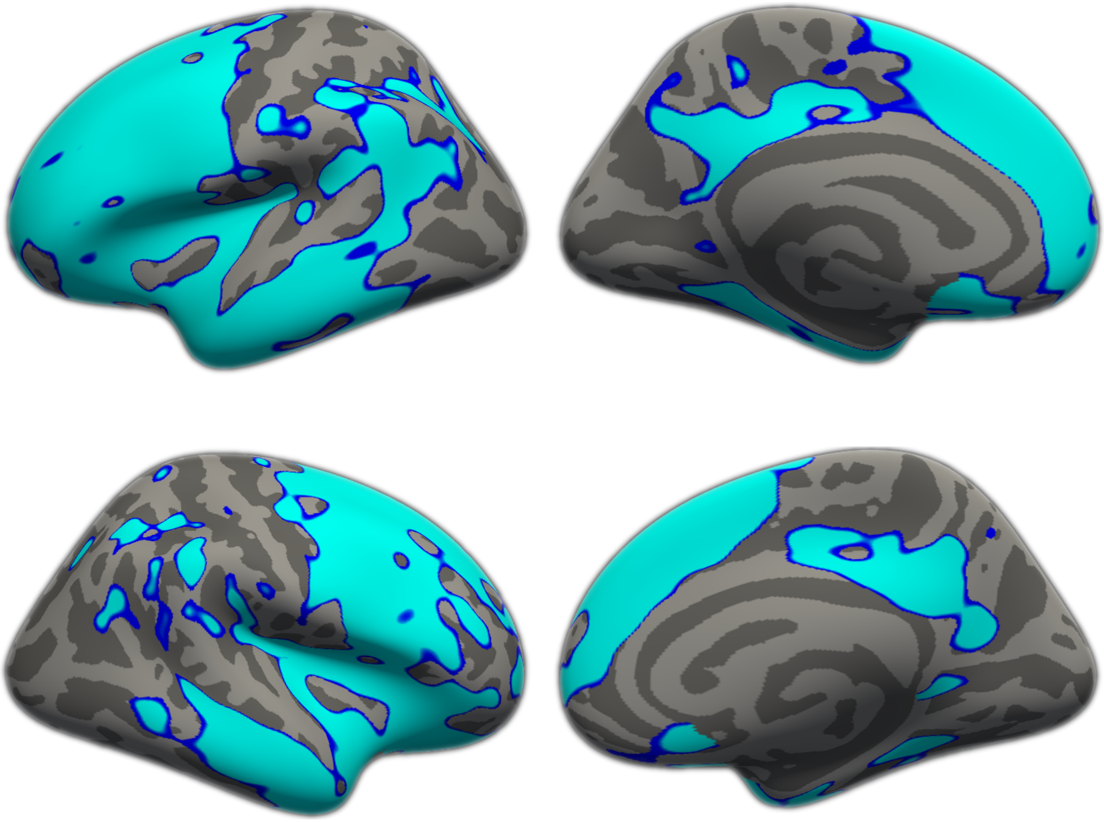

Figure 4: Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+S+ vs. M+S-

Figure 5 summarizes the uncorrected results: it shows the areas with a significant reduction in cortical thickness for all parameters and for the left and right hemisphere respectively.

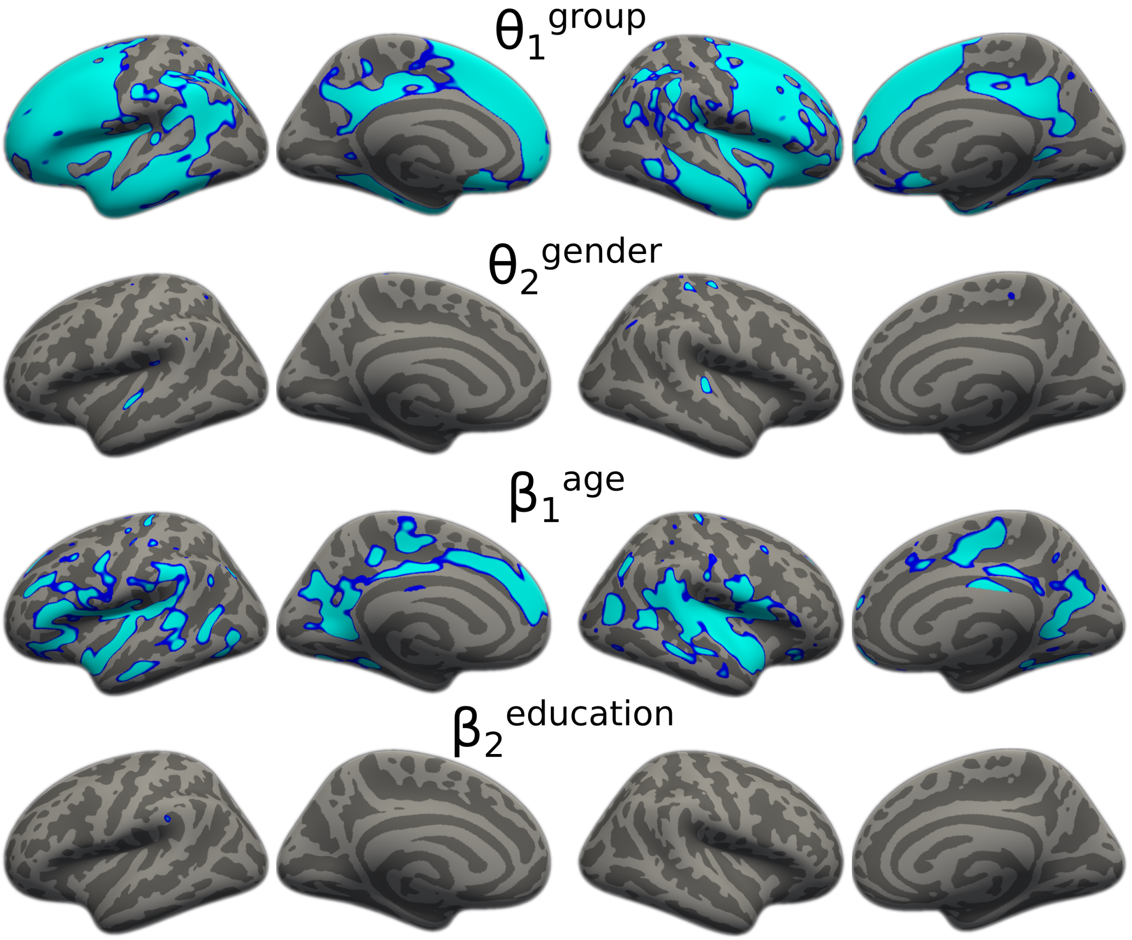

Figure 5: Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$ used on the GENFI dataset with $\theta_{1(j)}$ modelling the difference value between M+S+ vs. M+S-.

3.2.1 HCP Parcellation for the FTLD results

I visualized the data with the aforementioned HCP Parcellation: Figure 6 shows the inflated left and right hemisphere of the HCP 32k fs_LR subject with a p-map overlay from the permutation test of the $\theta_{1(j)}$ parameter for previously described model $\mathcal{M}_{A1}$, which shows the difference in cortical thickness between the group M+S+ and the group M+S-. As another overlay, the areas of the parcellation by Glasser et al. (2016) are shown.

Almost all frontal and temporal areas and some parietal areas of the left hemisphere show significantly reduced thickness. The right hemisphere shows the same pattern less pronounced compared to the left hemisphere.

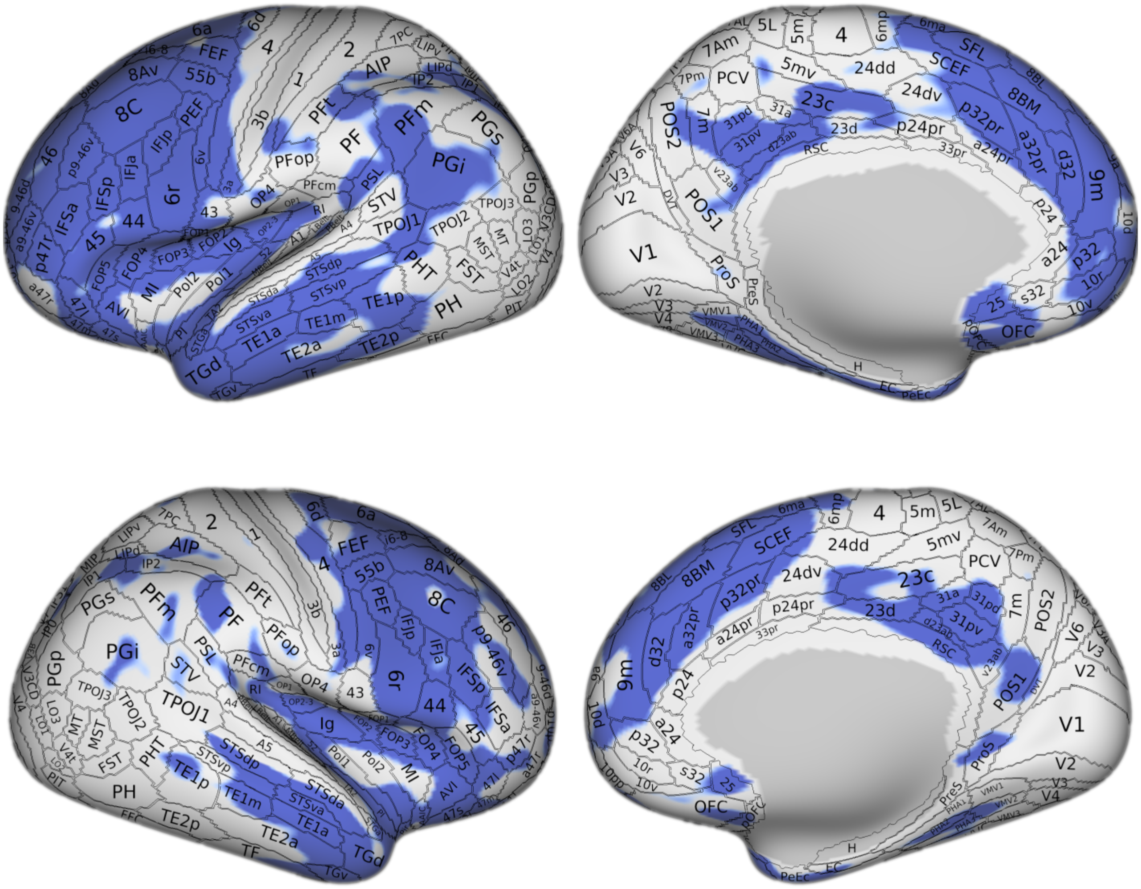
6a
6d
FEF
8Av
55b
8C
4
3b
1
2
AIP
PFt
PF
PFm
PGs
PGi
PFop
PFcm
PSL
STV
TPOJ1
TPOJ2
TPOJ3
PGp
46
p9-46v
IFJa
IFJp
PEF
6v
6r
IFSp
IFSa
44
45
43
OP4
RI
FOP1
FOP2
FOP3
FOP4
FOP5
Ig
MI
PoI1
PoI2
AVI
p47r
a47r
47l
STSdp
STSvp
STSda
STSva
TE1a
TE1m
TE1p
TE2a
TE2p
TGd
PHT
PH
FST
MST
MT
V4t
LO1
LO2
LO3
V3CD
V4
PIT
FFC
TF
TGv
7Am
5L
5m
4
6mp
24dd
5mv
7Pm
PCV
SCEF
SFL
24dv
8BM
8BL
23c
31a
31pd
31pv
23d
d23ab
p24pr
33pr
a24pr
p32pr
a32pr
d32
9m
7m
POS2
POS1
RSC
v23ab
V6
V3
V2
V1
ProS
PreS
p24
a24
p32
10r
s32
25
10v
10d
OFC
H
EC
PeEc
VMV1
VMV2
VMV3
PHA1
PHA2
PHA3

Figure 6: Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between M+S+ and M+S-. In addition, an overlay of the human cortex parcellation is shown.

3.3 Discussion

By comparing symptomatic with presymptomatic genetic FTLD subjects, I could replicate the expected FTLD atrophy pattern with my presented methods (Neary et al., 2005; Hodges et al., 2004). As expected, I found significant generalized atrophy in the temporal and frontal lobe. This atrophy pattern is not new, but it still serves as a confirmation, that the genetic FTLD cohort does indeed show the expected atrophy in the cortex and that my presented methods are able to capture them. Furthermore, they also serve as a reference, for the other results presented. Even more, they provide general FTLD atrophy masks for the presymptomatic analysis, presented later.

My results did not show any considerable gender or education effects, but there was an effect of age. The mean age of the M+S- group is much lower, than the mean age of the M+S+ group, which is one reason why a strong age effect is present in the data and can be seen in the results. Also, age is an important factor for neurodegenerative diseases, as the atrophy due to an underlying disease seems to get enhanced by increasing age. The observed age difference between the groups is thus naturally inherent to a degenerative dataset, that got collected in the presented fashion. Still, by including an age factor in the model, it was possible to at least differentiate between pure age effects, and the effect caused by belonging to the symptomatic group. Future studies could try to include an interaction of age and group, to pin down possible interaction effects.

The further below upcoming analysis about presymptomatic cortical degeneration is of special interest for this group, as presymptomatic changes in the brain are a major point for discussion in recent FTLD research (Bertrand et al., 2018; Papma et al., 2017; Rohrer et al., 2015).

Important to mention is, that the presented approach and the accompanying results additionally served as a usage example for an approach that is generalizable across a wide range of hypothesis driven research questions. In this work, I present two more hypothesis-driven usage examples, which will be presented next.

4 Apraxia

The goal of this analysis was the identification of apraxia related areas in the cortex of genetic FTLD subjects. I hope to highlight the relevance of apraxia for FTLD and also link my results to research about apraxia in stroke.

4.1 Material and Methods

For the apraxia analysis, I needed a specific subset of the GENFI dataset: I selected all subjects among the 294 M+ group, that showed noticeable apraxia in the clinical examination. All subjects, that had an apraxia GERS score above zero on either the left or the right or both sides were selected for the apraxia dataset: this resulted in a sample of 31 subjects in the mutation carriers with apraxia (M+A+) group.

For the selected M+A+ group a matching mutation carriers without apraxia (M+A-) group and a matching M- group had to be selected: I did this by matching with age (± 2) and gender and allowing for multiple matches per subject, including all visits. If a subject had two matching visits, the last visit was included, to avoid duplicate subjects. The resulting dataset is summarized in the fourfold table 4, together with some descriptive statistics for the corresponding groups.

The resulting four groups were: the M+A+ group and the M+A- group, as well as the non-mutation carriers with apraxia (M-A+) group and the non-mutation carriers without apraxia (M-A-) group. Surprisingly, there are two subjects in the M-A+ group, which was unexpected. However, as this subject size is very small, this group was not used for any analysis.

To show how these groups differ in disease severity, I created a histogram of the FTLD-CDR-SOB scores for every generated data group: see Figure 7.

Table 4: Fourfould table for the datased used with detailed information. Descriptive statistics are means with standard deviations in brackets. Male, female and symptomatic gives the count of subjects for the respective group. The two columns segregate M+ and M-, the two rows segregate probands with apraxia (A+) and probands without apraxia (A-).

	M+	**M-**
A+	31 $\begin{cases} C9ORF: & 12 \\ GRN: & 17 \\ MAPT: & 2 \end{cases}$ Age: 61.85 (10.86) Education: 12.29 (4.20) Male: 15 Female: 16 Symptomatic: 26	2 $\begin{cases} C9ORF: & 0 \\ GRN: & 1 \\ MAPT: & 1 \end{cases}$ Age: 57.55 (7.57) Education: 15 (7.07) Male: 1 Female: 1 Symptomatic: 0
A-	127 $\begin{cases} C9ORF: & 49 \\ GRN: & 60 \\ MAPT: & 18 \end{cases}$ Age: 56.09 (12.22) Education: 13.93 (3.81) Male: 63 Female: 64 Symptomatic: 45	74 $\begin{cases} C9ORF: & 26 \\ GRN: & 36 \\ MAPT: & 12 \end{cases}$ Age: 54.82 (11.12) Education: 14.19 (3.30) Male: 31 Female: 43 Symptomatic: 0

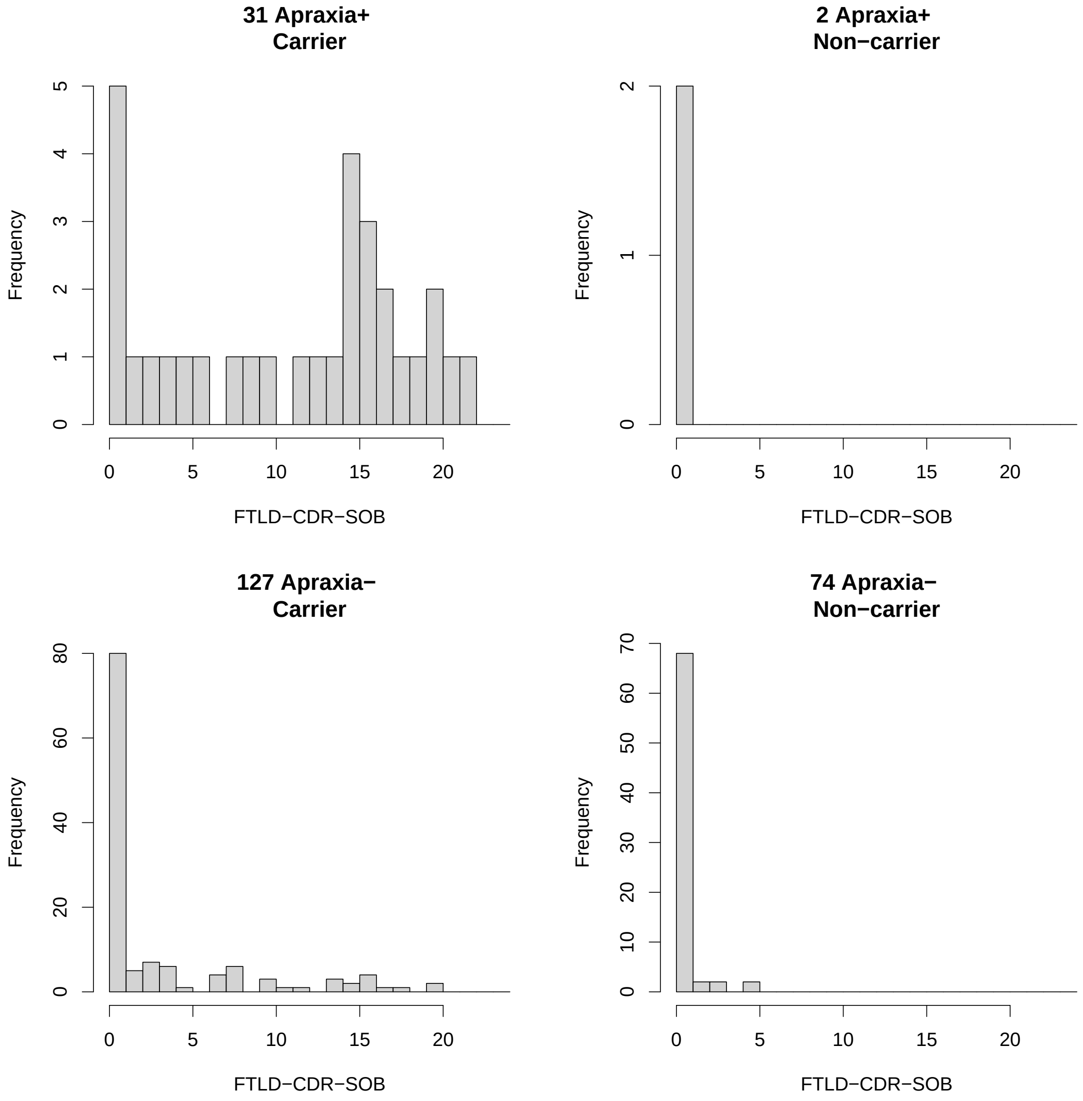

Figure 7: FTLD-CDR-SOB scores histograms for the four generated data groups - from left to right and top to bottom: the M+A+ group, the M-A+ group, the M+A- group and the M-A- group.

The described matching procedure resulted in groups with considerable age differences. In order to have a better control of age effects, I created a second M+A- group, matching each case with apraxia with the two gender matched subjects with the closest age match, resulting in the 62 mutation carriers without apraxia (62M+A-) group.[7] Descriptive statistics for the 62M+A- group are shown in Table 5.

Table 5: Descriptive statistics for group 62M+A-. Descriptive statistics are means with standard deviations in brackets. N is a placeholder for the number of subjects. Male, female and symptomatic gives the count of subjects for the respective group.

N (male/female)	Age	Education	Symptomatic
62 (30/32)	61.41 (10.52)	13.84 (3.73)	29

4.1.1 Specific area detection models and tests

As before, I used cortical thickness as the dependent variable and performed the analysis on both hemispheres. The used model is equivalent to the Model $\mathcal{M}_{A1}$ (shown in Equation 4), but the $\theta_{1(j)}$ parameter this time models the difference value for M+A+ compared to M+A-. All other parameters of the model have the same meaning as previously: $\theta_{2(j)}$ models the difference value for male subjects, $\beta_{1(j)}$ models the value for continuous age changes, $\beta_{2(j)}$ models the value for continuous education changes.

4.1.2 Comparison of the apraxia results to apraxia in stroke

I was further interested in finding out, if the brain substrates of apraxia in FTLD are similar to the brain substrates of apraxia in stroke patients.

Goldenberg and Karnath (2006) report data for the left brain hemisphere from stroke patients with apraxia: they differentiate between finger and hand apraxia tests and provided subtraction masks of patients with distorted finger and hand movements vs. controls respectively. In addition, Goldenberg et al. (2007) report data for the left brain hemisphere from stroke patients with apraxia, that had undergone pantomime apraxia testing and provided subtraction masks of patients with distorted pantomime vs. controls.

Using the actual original data from the two mentioned studies, provided by one of the authors[8], I created one fsaverage surface mask for finger, hand and pantomime, respectively. In addition, I created a combined mask of all three, serving as a kind of unspecific apraxia mask.

---

[7] Using the same, latter mentioned matching procedure, I also selected 62 subjects for the M-A- group. However, as the later presented results did not substantially differ, the report of this group analysis will be spared.

[8] Thanks to Otto Karnath, who provided the original data.

As the original data is only present for the left hemisphere, the later presented results will only show the left hemisphere.

The images of the results were created using the connectome workbench viewer `wb_view` (retrieved from `https://www.humanconnectome.org`).

4.2 Results

4.2.1 Area detection Analysis

For the apraxia area detection analysis, I used the following groups of the previously described apraxia dataset: the M+A+ group, the M+A- group, the 62M+A- group, as well as the M-A- group.

In the following, I present the results of the comparison of the M+A+ group first with the M-A- group, second with the M+A- group and third with the 62M+A- group.

The previously described MLM $\mathcal{M}_{A1}$ has been used for the analysis.

M+A+ vs. M-A-

When comparing the M+A+ group vs. the M-A- group, widespread reduction in cortical thickness can be observed in the frontal, temporal and parietal lobe of the left and right hemisphere. The corrected M+A+ vs. M-A- results for the $\theta_{1(j)}$ parameter are summarized in Figure 8 for the left and for the right hemisphere.

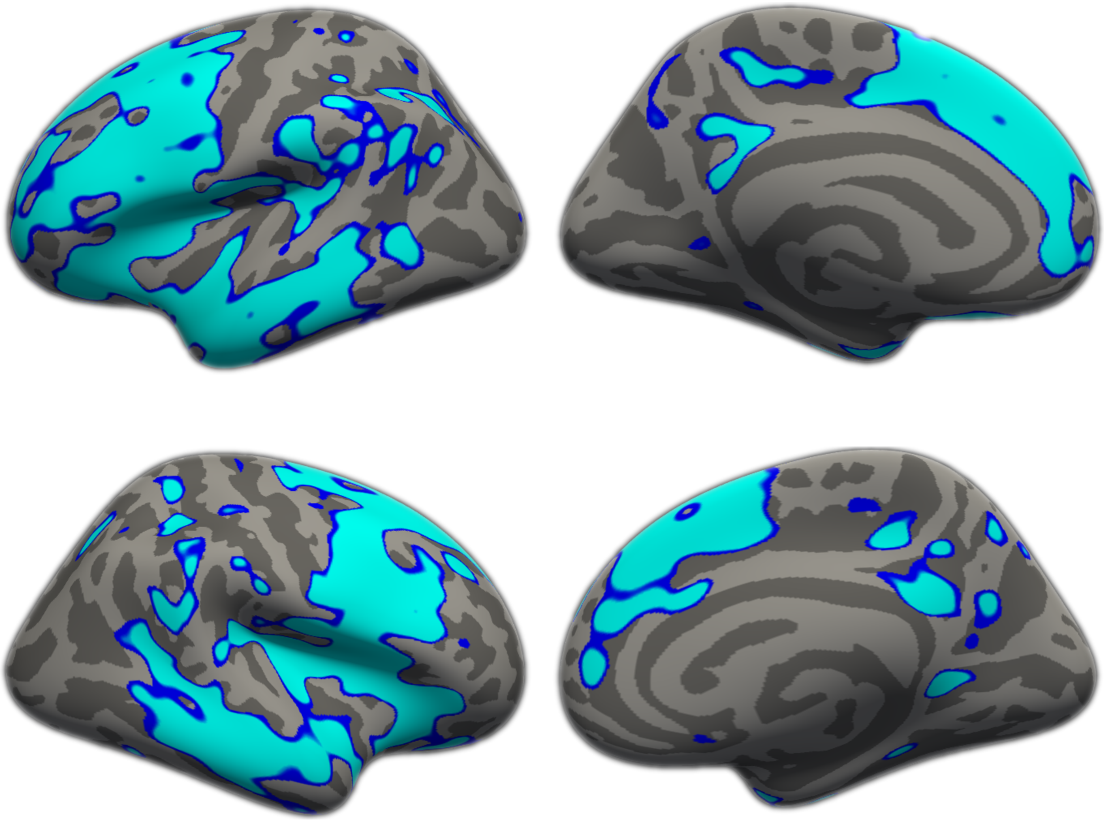

Figure 8: Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M-A-

Figure 9 summarizes the uncorrected results: it depicts the areas with a significant reduction in cortical thickness for all parameters and for the left and right hemisphere respectively.

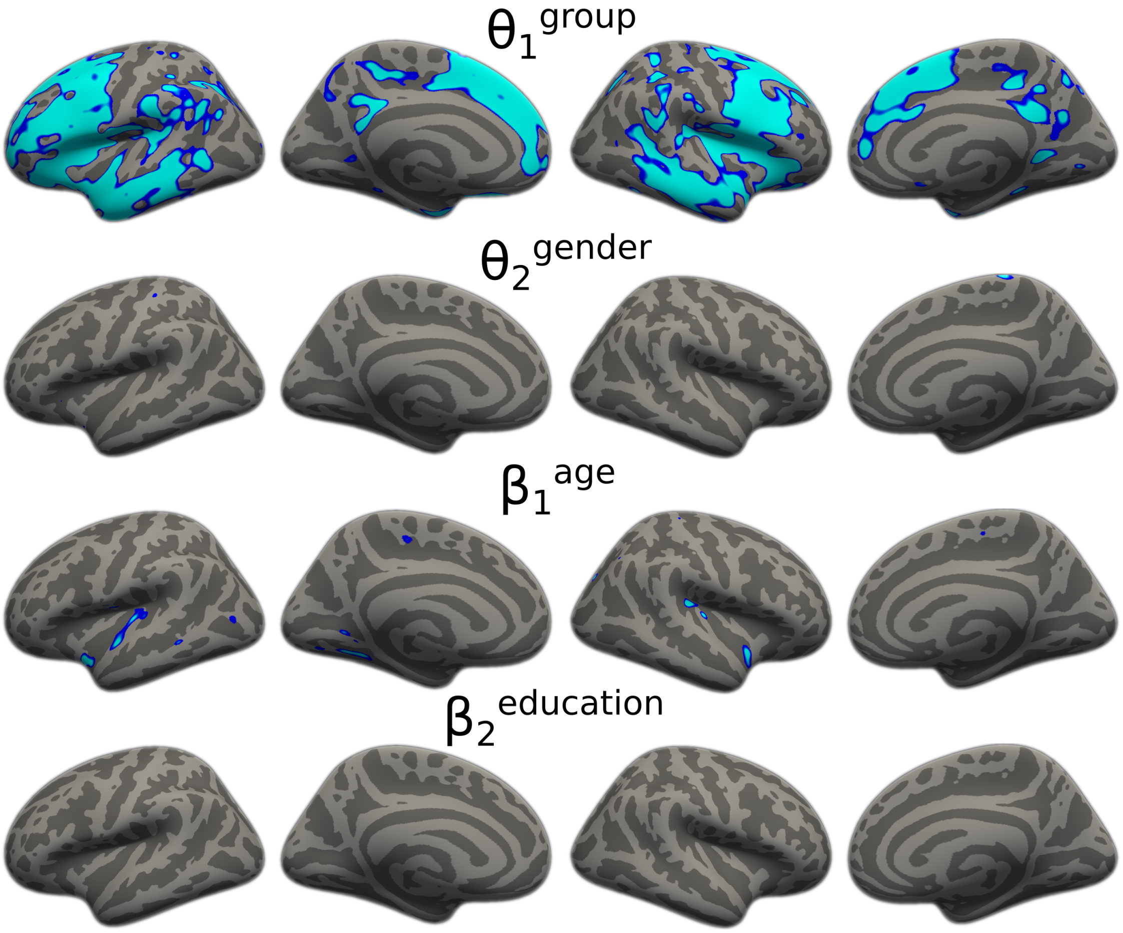

Figure 9: Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M-A-

M+A+ vs. M+A-

The comparison of the M+A+ group vs. the M+A- group, revealed a significant focal reduction of cortical thickness in the premotor cortex, inferior frontal and frontal opercular areas on the left and right hemisphere, as can be seen in Figure 10, which shows the corrected M+A+ vs. M+A- results for the $\theta_{1(j)}$ parameter for the left and for the right hemisphere.

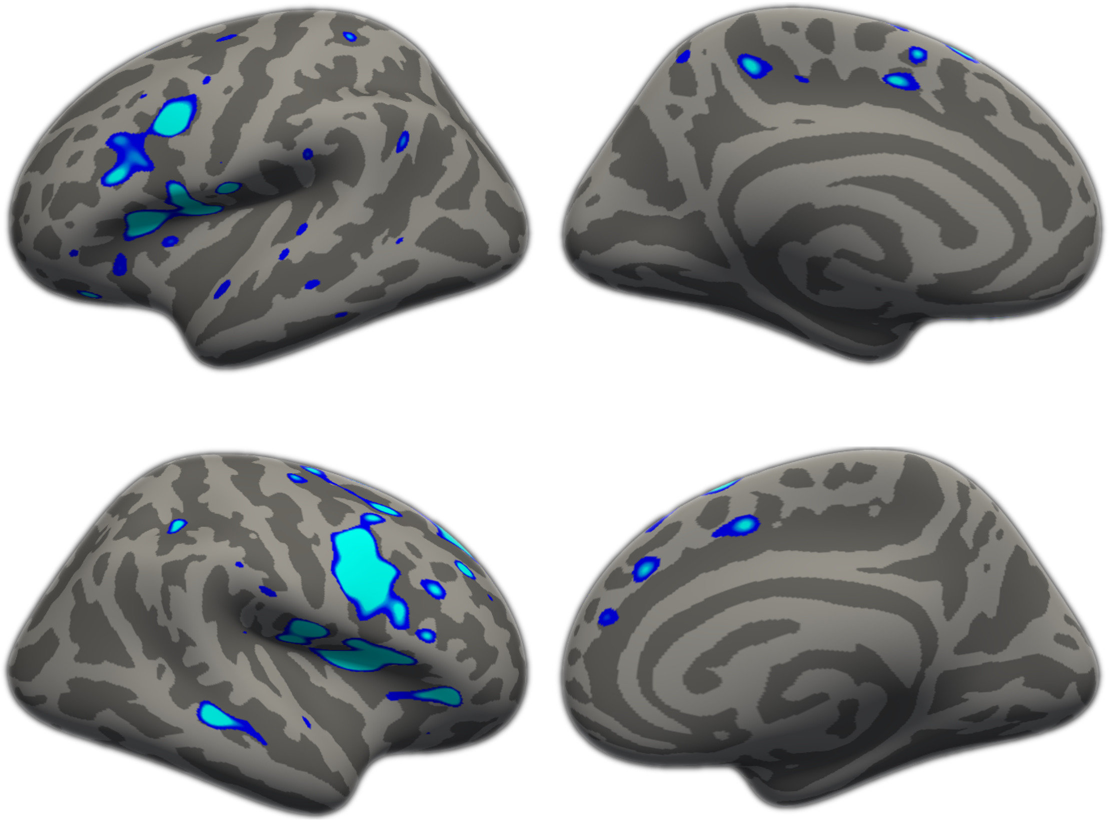

Figure 10: Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M+A-

Figure 11 shows the uncorrected results for all parameters and for the left and right hemisphere respectively. Note that in this analysis, there are noticeable age-related reductions in cortical thickness, captured by the $\beta_{1(j)}$ parameter.

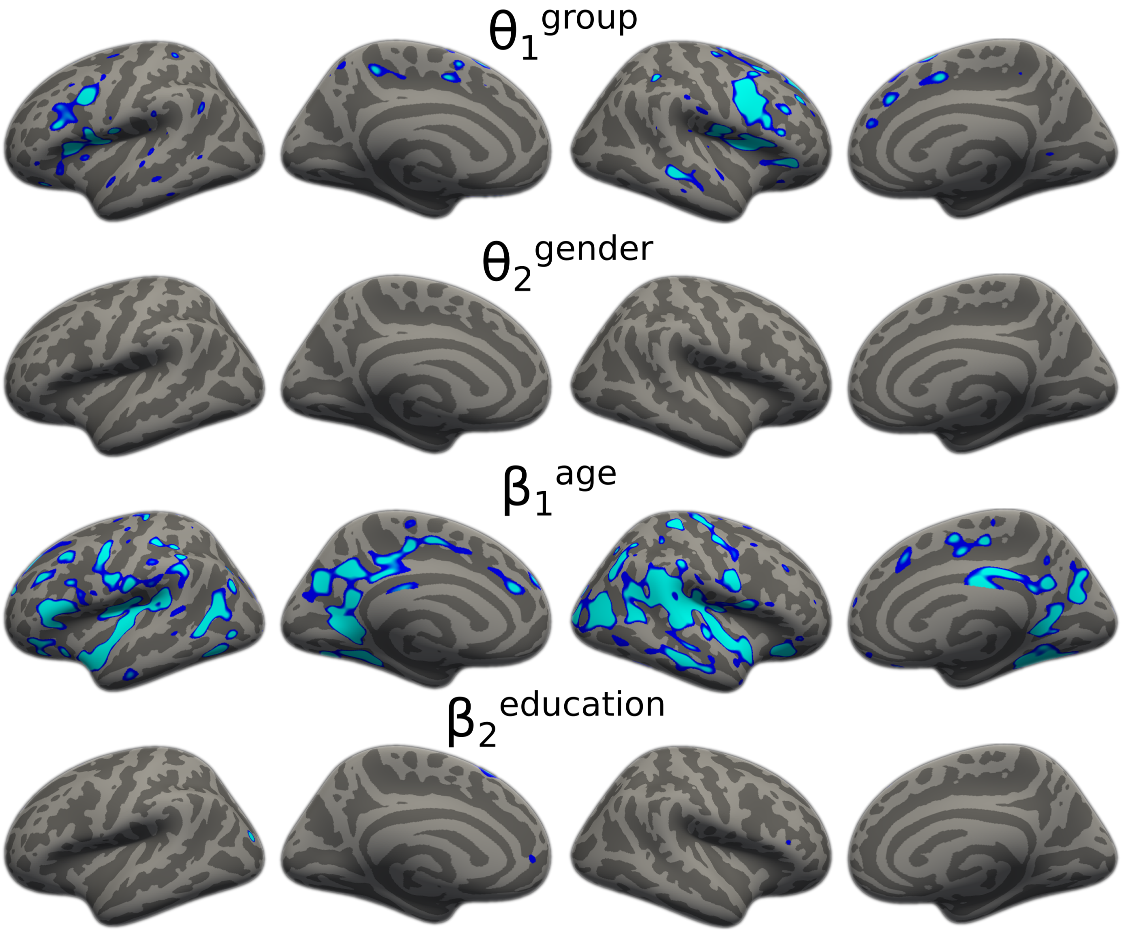
θ_1^{group}
θ_2^{gender}
β_1^{age}
$\beta_2^{\text{education}}$

Figure 11: Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M+A-

M+A+ vs. 62M+A-

Using the same procedure as before to compare the M+A+ group vs. the M+A- group, the comparison of the groups M+A+ vs. 62M+A-, revealed the same areas, but with a smaller area delineated through the significance test: the corrected M+A+ vs. 62M+A- results for the $\theta_{1(j)}$ parameter are summarized in Figure 12 for the left and for the right hemisphere. A lineup of the corrected results for the comparison of M+A+ vs. M+A- and the comparison of M+A+ vs. 62M+A- is shown in Figure 13.

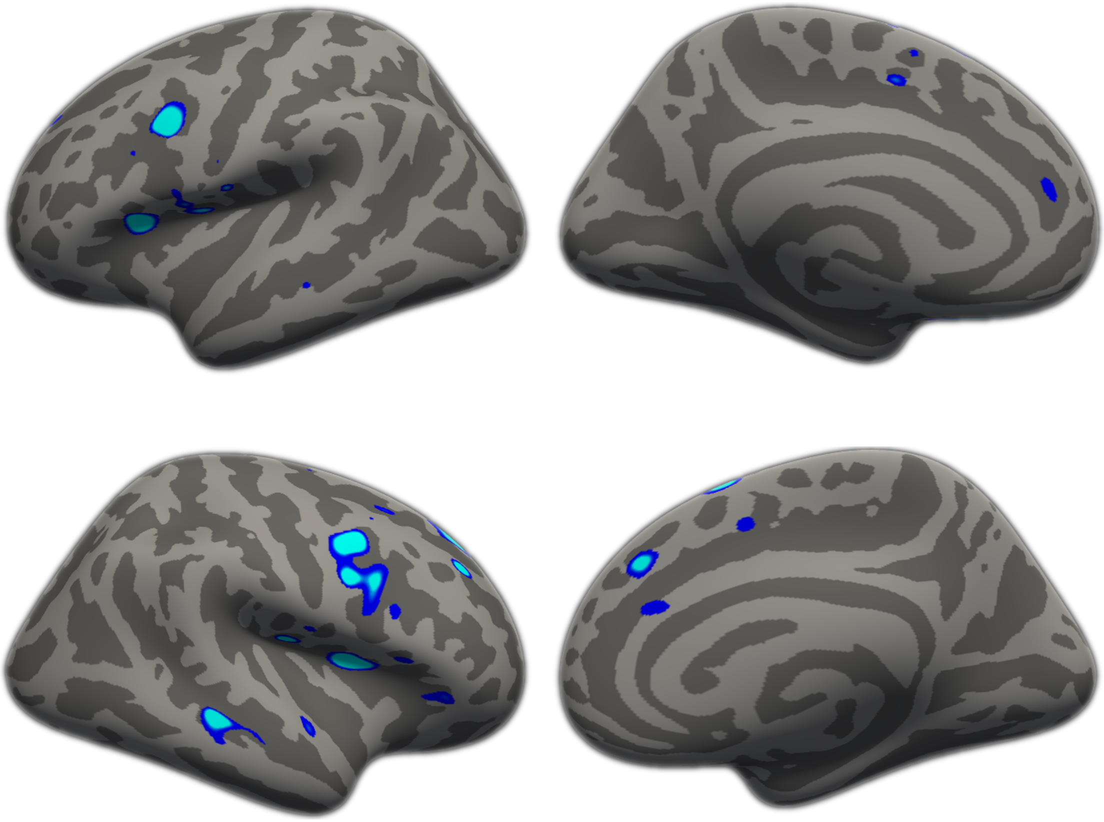

Figure 12: Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. 62M+A-

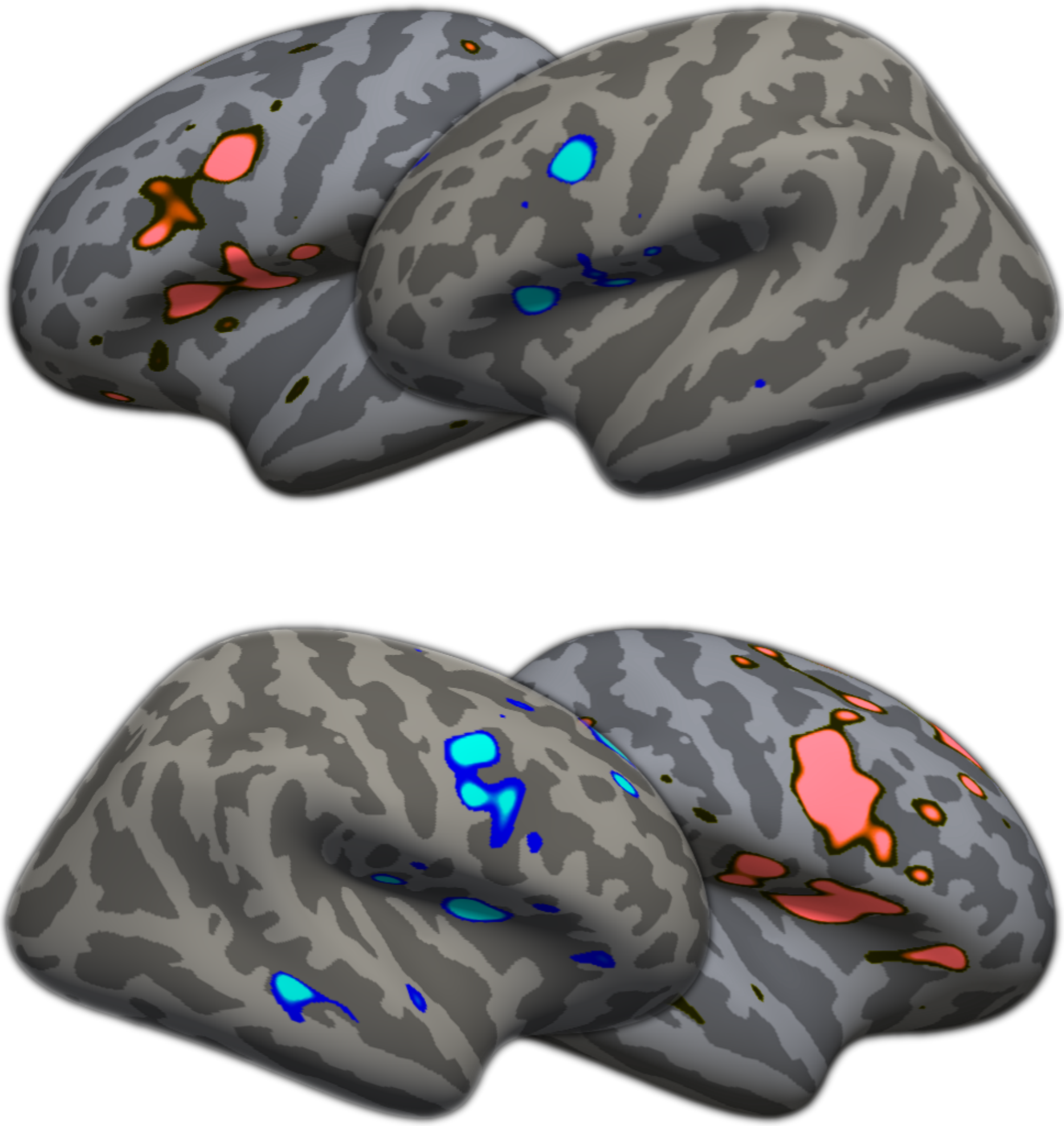

Figure 13: Lineup of the left (top) and right (bottom) hemisphere results for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. Shown are the results for the compared groups: M+A+ vs. M+A- (red) and M+A+ vs. 62M+A- (blue)

Figure 14, shows the uncorrected results for all parameters and for the left and right hemisphere respectively.

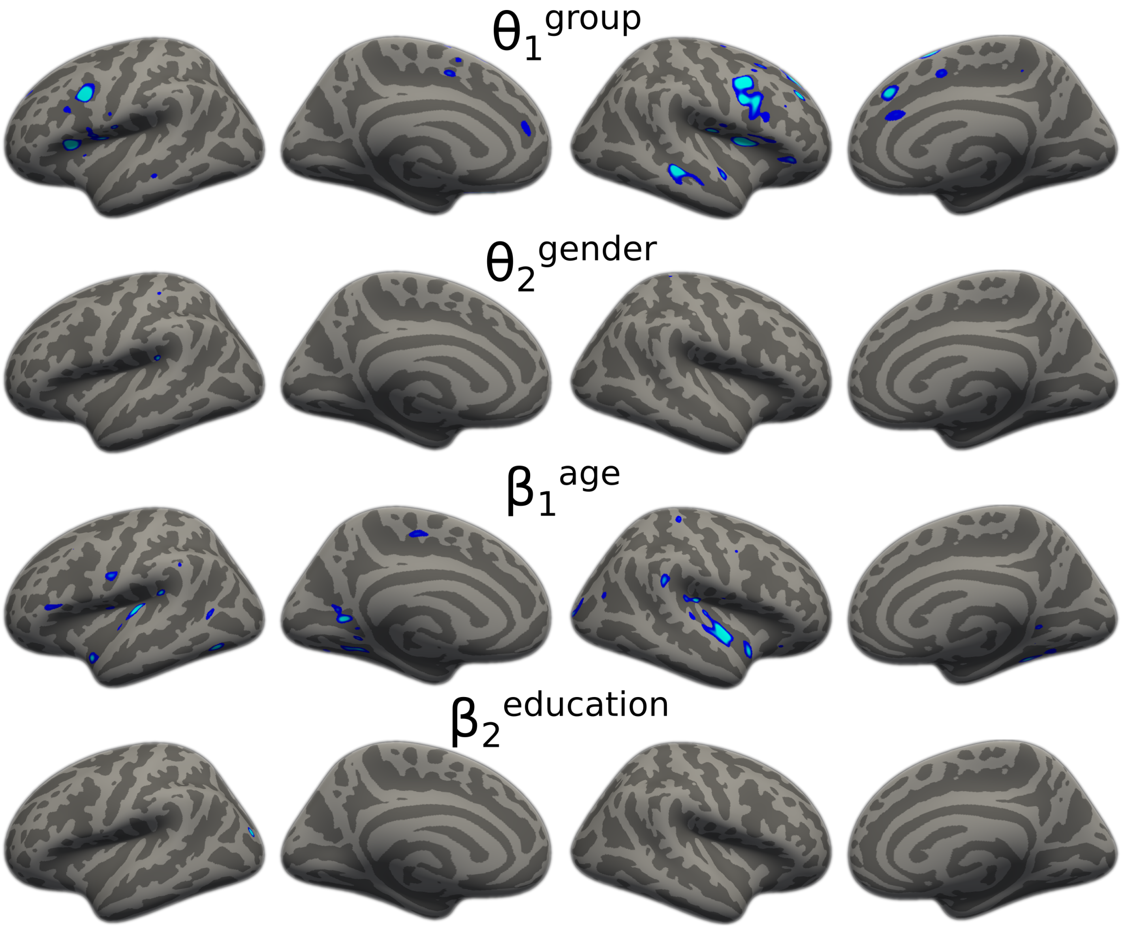
θ_1^{group}
θ_2^{gender}
β_1^{age}
$\beta_2^{education}$

Figure 14: Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$. The compared groups were: M+A+ vs. M+A-

Genetic group differences

Lastly, I also tried to find differences between the genetic groups, however, including a parameter that models the differences between the genetic groups in my models did not yield any significant results.

4.2.2 Comparison of the apraxia results to apraxia in stroke

I compared the results of the $\theta_{1(j)}$ parameter significance test for the M+A+ group vs. the M+A- group with masks coming from apraxia in stroke research, prepared as described in the methods section: I created three specific masks from the results of the study by Goldenberg and Karnath (2006) and Goldenberg et al. (2007) for reported pantomime, finger and hand apraxia; an unspecific mask as a combination of the three masks.

The following figures (Figure 15 and Figure 16) show the inflated left hemisphere of the fsaverage subject with a p-map overlay, which shows the significant difference in cortical thickness between the M+A+ group and the M+A- group. In addition, the figures show the stroke apraxia masks as another overlay.

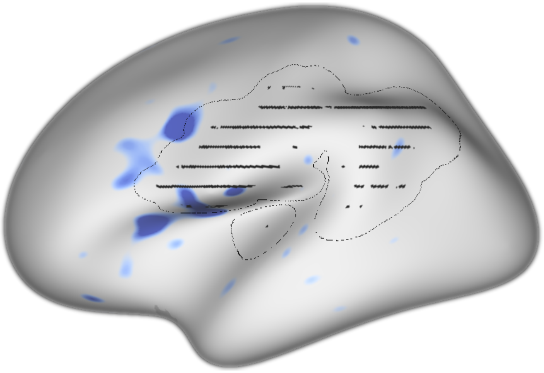

Figure 15: Inflated left hemisphere surface of the fsaverage subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A0}$, capturing the group difference between M+A+ and M+A- (blue areas). In addition, an overlay of the unspecific apraxia stroke mask is shown (black stripes).

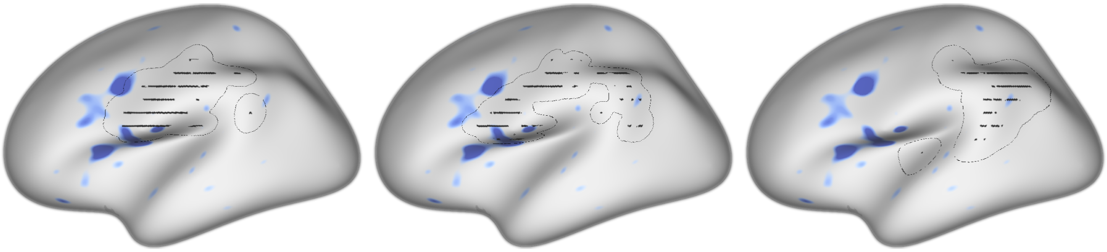

Figure 16: Inflated left hemisphere surface of the fsaverage subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A0}$, capturing the group difference between M+A+ and M+A- (blue areas). In addition, an overlay of the specific apraxia stroke masks are shown (black stripes): from left to right: pantomime, finger, hand.

4.2.3 HCP Parcellation for the apraxia results

I visualized the data with the HCP Parcellation: Figure 17 shows the inflated left and right hemisphere of the HCP 32k fs_LR subject with a p-map overlay from the permutation test of the $\theta_{1(j)}$ parameter for previously described model $\mathcal{M}_{A1}$, which shows the difference in cortical thickness between the group M+A+ and the group M+A-. As another overlay, the areas of the parcellation by Glasser et al. (2016) are shown.

Significantly reduced thickness was found in the following areas on the left hemisphere: PEF, IFJp, Area 8C, IFSp, 6r, 44, FOP1, FOP2, FOP3, FOP4, FOP5; And in the following areas on the right hemisphere: 55b, PEF, IFJp, IFJa, 6r, 44, FOP1, FOP2, FOP3, FOP4, in addition to some small areas between FOP5, 45, 47l, and between STSdp, STSvp, and between OP4, 43.

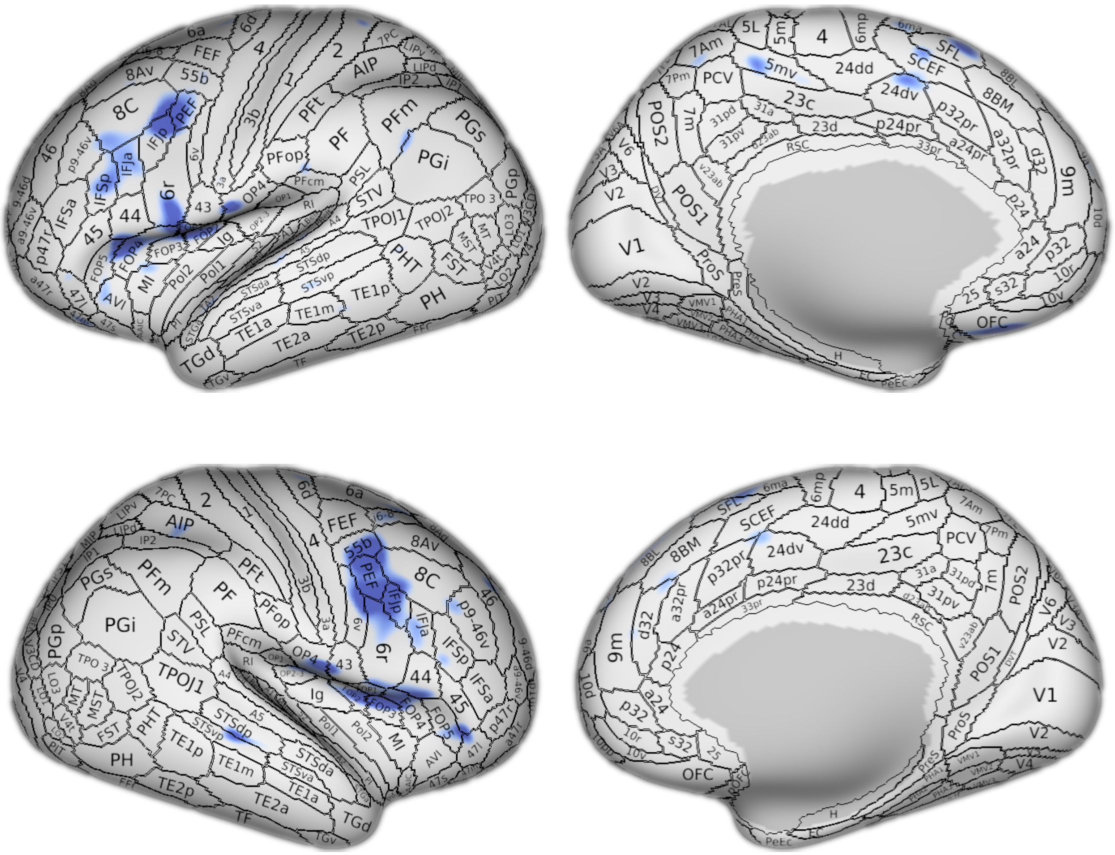
6a
6d
FEF
4
2
7PC
8Av
55b
1
AIP
IP2
8C
PEF
3b
PFt
PFm
46
p9-46v
IFJa
IFJp
PF
PGs
IFSp
6r
PFop
PGi
44
43
OP4
PFcm
PSL
IFSa
45
STV
TPOJ1
TPOJ2
PGp
p47r
FOP5
FOP4
FOP3
Ig
STSdp
STSvp
MT
MST
FST
PHT
PH
MI
PoI2
PoI1
TE1p
TE1m
STSda
STSva
TE1a
TE2a
TE2p
AVI
47l
TGd
5L
5m
6mp
7Am
5mv
24dd
SCEF
SFL
PCV
7Pm
23c
24dv
8BM
8BL
7m
POS2
31pd
31a
31pv
23d
p24pr
p32pr
a24pr
a32pr
d32
9m
V6
V3
V2
POS1
RSC
33pr
p24
V1
ProS
a24
p32
10r
10v
25
s32
OFC
V4
H
PeEc
10d

Figure 17: Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between M+A+ and M+A-. In addition, an overlay of the human cortex parcellation is shown.

4.3 Discussion

In the GENFI data, I identified 31 cases of apraxia among the 294 mutation carriers: this is remarkable, considering the loose definition of apraxia as gesture imitation deficits. A more detailed examination following a more specific apraxia definition and using predefined and unambiguous tests would possibly lead to a higher sensitivity for capturing apraxia, resulting in an even higher identification of apraxia among genetic FTLD subjects. This shows two things: the importance of apraxia for genetic FTLD subjects and the awareness of apraxia as a common symptom for genetic FTLD. Previous studies, as done by Johnen et al. (2016), also show the importance of apraxia for FTLD. A more thorough discussion of this study follows further below in the general discussion.

In order to find brain areas with reduced cortical thickness among genetic FTLD subjects with apraxia compared to supposedly healthy controls, I compared the M+A+ group with the M-A- group and found significant widespread reduction in cortical thickness in frontal and temporal areas of the left and right hemisphere. This comparison also revealed some very small parietal areas, that show significant reduction in cortical thickness, however, these are much less pronounced compared to the frontal and temporal regions. This pattern strongly reminds of the typical FTLD brain atrophy (Neary et al., 2005), and serves as another validation of the methods applied in the current project. It can be interpreted as general FTLD atrophy areas, which also include areas relevant for apraxia.

To identify areas that are specifically related to apraxia, I compared the M+A+ group with the M+A- group. The result was significant focal reduction of cortical thickness in the premotor cortex, inferior frontal and frontal opercular areas on the left and right hemisphere. My interpretation of these results is, that these areas do play a key role for apraxia in genetic FTLD.

Furthermore, this group comparison showed age effects, as the groups did differ in age, even though I tried to control this with a matching procedure. To account for this, the age parameter was included in the model and thus the age effects should be caught by the age parameter and therefore not influence the group effect substantially. To reduce doubts about the effectiveness of the age parameter, I reanalyzed the data with a group that matched the age of the M+A+ group better: the 62M+A- group. There was, however, a trade-off for this better age matching, as this group is only about half as big as the first M+A- group. Thus, I suspected to find the same results, less pronounced, and with much less age effects captured by the age parameter, as the groups do not differ in age as much. This is exactly, what the results showed: the core of the previously delineated areas is still there. However, the delineated areas are considerably smaller. This is, at least partly, due to the loss of statistical power through using a smaller sample size. Also, the impacted brain areas could be more severely affected with increasing age, which could boost the results with stronger age effects. Thus, the results with a better control for age effects could show the more specific beginning of the atrophy process that develops and spreads more over time.

Lastly, I wanted to find different brain areas within the different genetic groups of FTLD. When I differentiated between the different genetic groups, the separate groups used in the model were accordingly much smaller. In the analysis with these groups, I could not find any significant effects at all, almost certainly due to the small group size. As usual, a null effect in classical frequentist testing theory cannot be interpreted as no effect present, but rather that no effect could be detected. This non-detection can have many reasons, the most probable being that the effect size is too small to be caught with the used group size. I still think that differences in the atrophy pattern of the different genetic groups and accordingly different atrophy patterns for the specific apraxia areas could underlie the different genetic groups. However, a different approach is necessary to account for this claim, probably with either a more sensitive method or bigger groups sizes.

The M-A+ group included two subjects: as the M- basically served as healthy controls, this was unexpected. Possible explanations could be false positives, as the FTLD-CDR-SOB scores of these subjects were not impaired. Another possibility could be a different disease as the cause of this symptom. Due to the small subject size and the questionable basis of the group composition, I did not use the M-A+ group for any analysis.

Stroke research is the typical frame for research about apraxia. Thus, it seemed logical to compare my results to apraxia research with a different background. In theory, one would assume that, even tough different disease pathologies are the underlying cause for the symptom apraxia, similar brain areas should be affected. The comparison showed that there is indeed an overlap visible, but also that there are areas, which do not overlap. This could have several reasons, one being that the relevant brain areas for apraxia in stroke do indeed differ from the brain areas relevant for apraxia in genetic FTLD. It is thinkable that different parts of the brain are relevant for specific neuropsychological functions, and if one or more of them fail, it results in very similar symptoms. However, do also keep in mind that the symptom apraxia was differently operationalized in the GENFI study and in the studies by Goldenberg and Karnath (2006); Goldenberg et al. (2007). It could be possible, that the overlapping areas do increase, if the symptom apraxia follows a more strict definition, in accordance to the work by Goldenberg and Karnath (2006); Goldenberg et al. (2007) in future research.

Lastly, to also draw a line to general research about brain area functionality, I wanted to visualize my results within an up-to-date multimodal brain parcellation. I used the parcellation by Glasser et al. (2016). The following areas seemed to be affected: Premotor Eye Field (PEF), Area 8c (8c), Area 55b (55b), Rostral Area 6 (6r), Frontal Opercular Area 1-5 (FOP1-5), Inferior Frontal Sulcus anterior and posterior (IFSp), Inferior Frontal Junction anterior and posterior (IFJa, IFJp), Area OP4 (OP4). These areas are all frontal areas and probably play a key role in action planning and other frontal functions. This does make sense considering that apraxia is a dysfunction for which frontal brain areas theoretically should play an important role. As mentioned before: the question arises, why I did not find any other brain areas in the parietal lobe, as can be found in apraxia in stroke research. It is probable that the way how apraxia was tested in the GENFI study, did only test for frontal brain area dysfunctions and was thus not able to capture probably meaningful areas from other parts of the brain. A much closer look at the testing procedure and accordingly a stricter and more detailed testing procedure considering apraxia could provide a better basis for future research in this respect. This will be further discussed in the general discussion.

5 TMT

The TMT (Bowie & Harvey, 2006) performance is a measure often used for executive dysfunctions, which are typical for FTLD (Rohrer et al., 2015; Demakis, 2004; Reitan & Wolfson, 1995). Therefore, detecting brain correlates of TMT performance was another chosen use case, presented next.

5.1 Material and Methods

For the TMT deficiency Analysis, I used the TMT standardized test data. Remember, as described earlier, that the test scores have been standardized by using the M- in the GENFI cohort as standard population. I identified 49 M+ with a deficiency in TMTB (TMTB|D+), i.e. TMT Subtest B (TMTB) time score equal to or worse than two standard deviations from the standard group, i.e. M- mean. Afterwards, I selected 49 gender and age matched M+ with no deficiency in TMTB (TMTB|D-), i.e. TMTB time score not worse than two standard deviations from the standard group mean. Age-matching was done by selecting a matching participant with the smallest age difference. Additionally, I performed the same procedure on the test time score of the TMT Subtest A (TMTA), identifying 42 M+ with a deficiency in TMTA (TMTA|D+), i.e. TMTA time score equal to or worse than two standard deviations from the standard group mean. Afterwards, I selected 42 gender and age matched M+ with no deficiency in TMTA (TMTA|D-), i.e. TMTA time score not worse than two standard deviations from the standard group mean. Table 6 provides descriptive statistics about the two datasets.

Table 6: Descriptive statistics for the TMTA analysis: the TMTA|D+ group and the TMTA|D- group; as well as for the TMTB analysis: the TMTB|D+ group and the TMTB|D- group. Descriptive statistics are means with standard deviations in brackets. N is a placeholder for the number of subjects. Male, female and symptomatic gives the count of subjects for the respective group.

	N (male/female)	Age	Education	Symptomatic
TMTA|D+	42 (26/16)	63.25 (9.25)	12.00 (3.99)	39
TMTA|D-	42 (26/16)	63.04 (9.44)	14.02 (3.28)	16
TMTB|D+	49 (28/21)	61.99 (10.13)	12.53 (3.61)	45
TMTB|D-	49 (28/21)	60.76 (9.62)	14.57 (3.18)	10

5.1.1 Specific area detection model and tests

As before, I used cortical thickness as the dependent variable and performed the analysis on both hemispheres. The model was equivalent to the Model $\mathcal{M}_{A1}$ (shown in Equation 4), but the $\theta_{1(j)}$ parameter this time models the difference value for TMTA|D+ compared to TMTA|D- and the difference value for TMTB|D+ compared to TMTB|D-, for the TMTA and TMTB analysis respectively. Also note, that M- were excluded for this analysis. All other parameters of the model have the same meaning as previously: $\theta_{2(j)}$ models the difference value for male subjects, $\beta_{1(j)}$ models the value for continuous age changes, $\beta_{2(j)}$ models the value for continuous education changes.

5.2 Results

5.2.1 TMTA analysis

When comparing the TMTA|D+ group vs. the TMTA|D- group, little reduction in cortical thickness can be observed in the frontal cortex of the left hemisphere. The corrected TMTA|D+ vs. TMTA|D- results for the $\theta_{1(j)}$ parameter are summarized in Figure 18 for the left and for the right hemisphere.

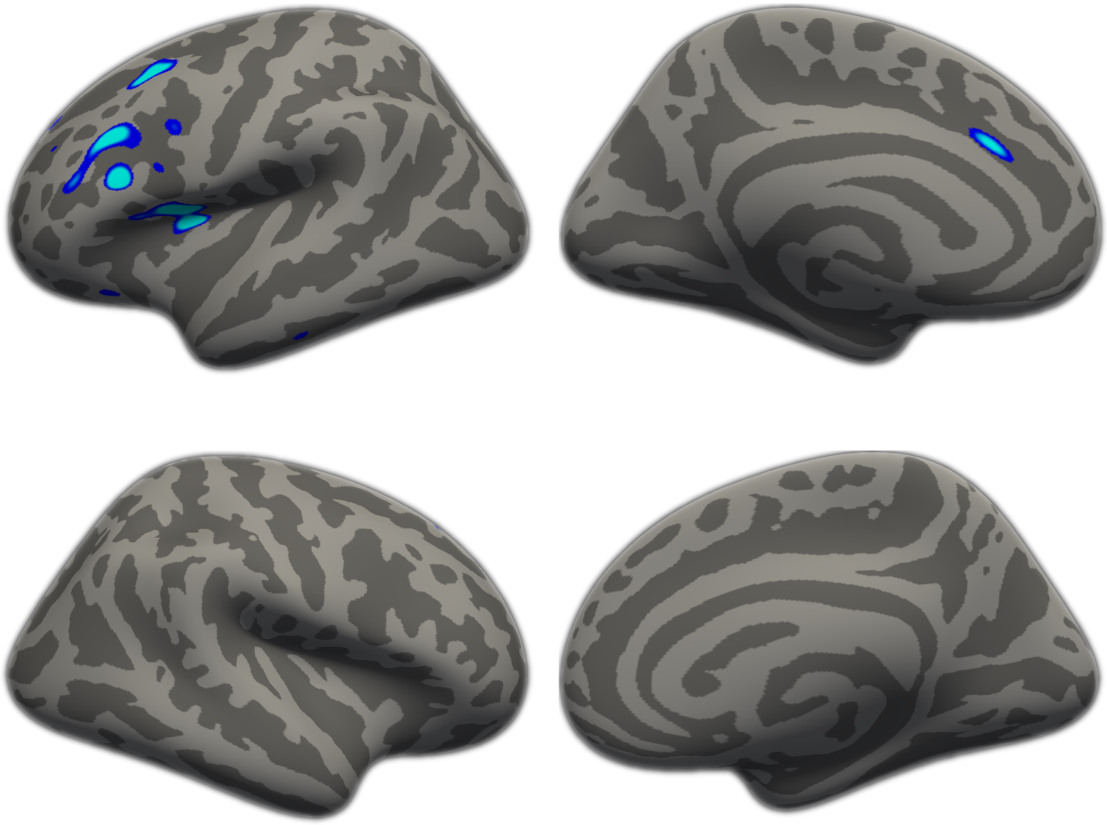

Figure 18: Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: TMTA|D+ vs. TMTA|D-

Figure 19 summarizes the uncorrected results: it shows the areas, that showed a significant reduction in cortical thickness for all parameters and for the left and right hemisphere respectively.

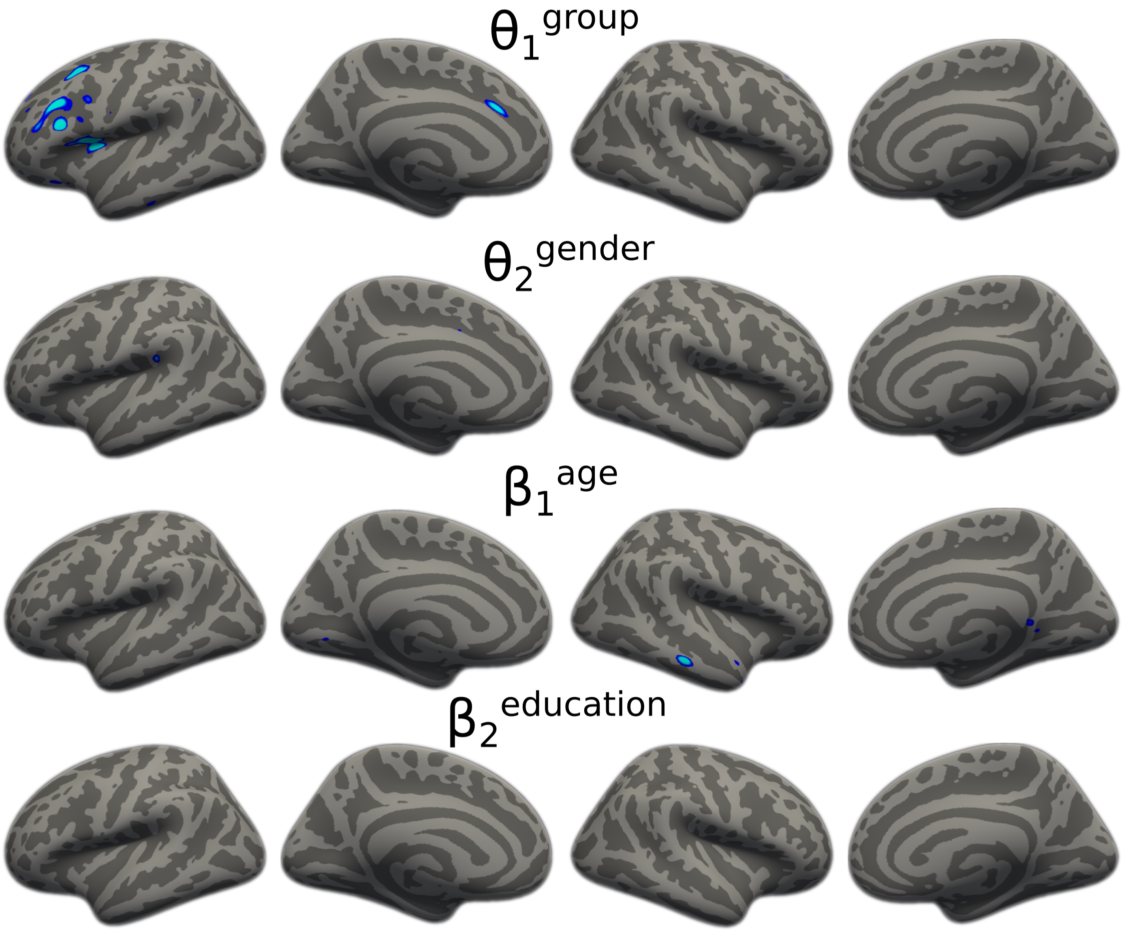

Figure 19: Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$ used on the TMTA dataset with $\theta_{1(j)}$ modelling the difference value between TMTA|D+ vs. TMTA|D-.

5.2.2 HCP Parcellation for the TMTA results

I visualized the data with the HCP Parcellation: Figure 20 shows the inflated left and right hemisphere of the HCP 32k fs_LR subject with a p-map overlay from the permutation test of the $\theta_{1(j)}$ parameter for previously described model $\mathcal{M}_{A1}$, which shows the difference in cortical thickness between the group TMTA|D+ and the group TMTA|D-. As another overlay, the areas of the parcellation by Glasser et al. (2016) are shown.

Significantly reduced thickness was found in the following areas on the left hemisphere: IFSp, 8C, p9-46v, IFSa, FOP1, FOP2, FOP3, FOP4, and a small area between p32pr, a32pr; and no areas on the right hemisphere.

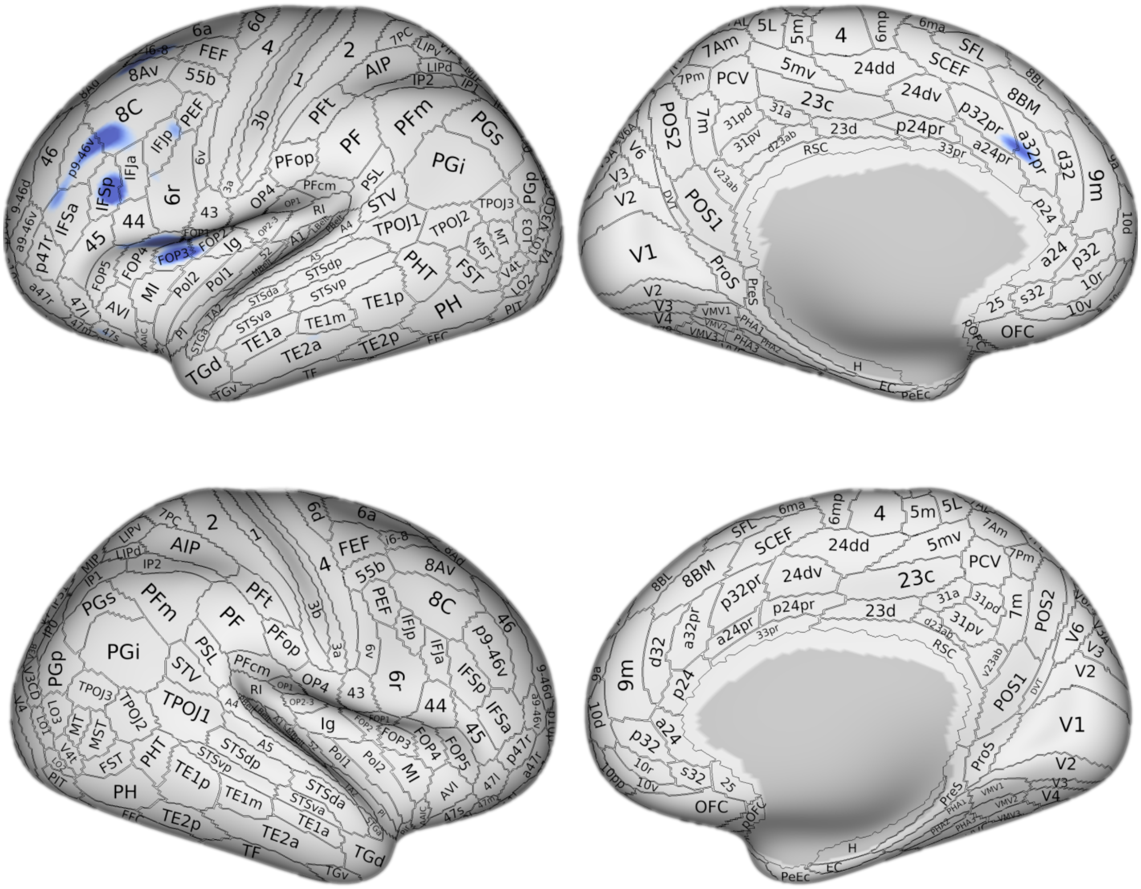
6a
6d
FEF
8Av
55b
8C
PEF
IFJp
IFJa
IFSp
IFSa
p9-46v
46
9-46d
a9-46v
p47r
45
44
6r
6v
4
3b
1
2
AIP
PFt
PF
PFm
PGs
PGi
PFop
PFcm
OP4
43
STV
PSL
TPOJ1
TPOJ2
TPOJ3
MT
MST
FST
PHT
PH
LO3
V4t
PGp
FOP1
FOP2
FOP3
FOP4
FOP5
MI
AVI
47l
a47r
Pol1
Pol2
Ig
RI
A5
STSdp
STSvp
STSda
STSva
TE1a
TE1m
TE1p
TE2a
TE2p
TGd
TF
5L
5m
5mv
7Am
7Pm
PCV
24dd
24dv
6mp
6ma
SFL
SCEF
8BL
8BM
23c
23d
31pd
31pv
31a
d23ab
v23ab
RSC
p24pr
a24pr
33pr
p32pr
a32pr
d32
9m
p24
a24
p32
10r
10v
10d
25
s32
OFC
POS2
POS1
V6
V3
V2
V1
V4
ProS
PreS
VMV1
VMV2
VMV3
PHA1
PHA3
H
EC
PeEc
DVT

Figure 20: Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between TMTA|D+ and TMTA|D-. In addition, an overlay of the human cortex parcellation is shown.

5.2.3 TMTB analysis

When comparing the TMTB|D+ group vs. the TMTB|D- group, a reduction in cortical thickness can be observed in the frontal cortex of the left and right hemisphere. The corrected TMTB|D+ vs. TMTB|D- results for the $\theta_{1(j)}$ parameter are summarized in Figure 21 for the left and for the right hemisphere.

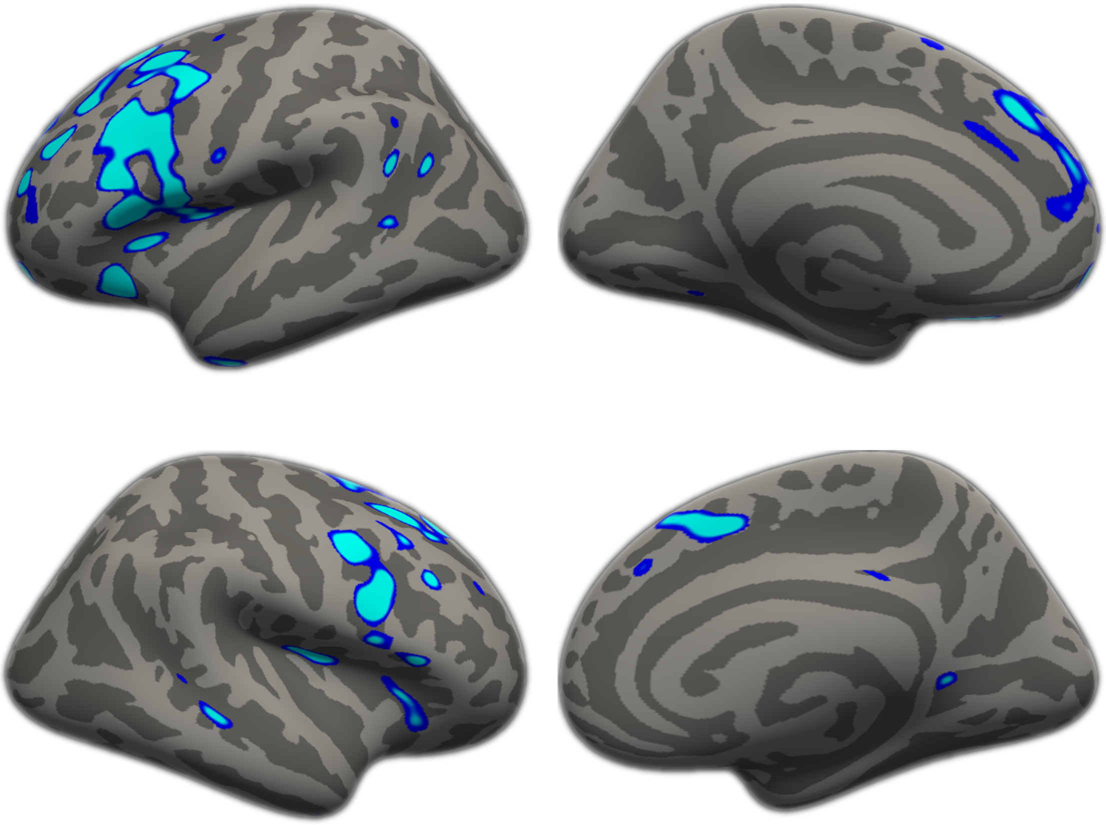

Figure 21: Inflated left (top) and right (bottom) hemisphere surface of the fsaverage subject, with an approximate permutation Wald test significance overlay (threshold $p < 0.05$, brightly colored areas $p < 0.01$) for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_{A1}$. The compared groups were: TMTB|D+ vs. TMTB|D-

Figure 22 summarizes the uncorrected results: it shows the areas, that showed a significant reduction in cortical thickness for all parameters and for the left and right hemisphere respectively.

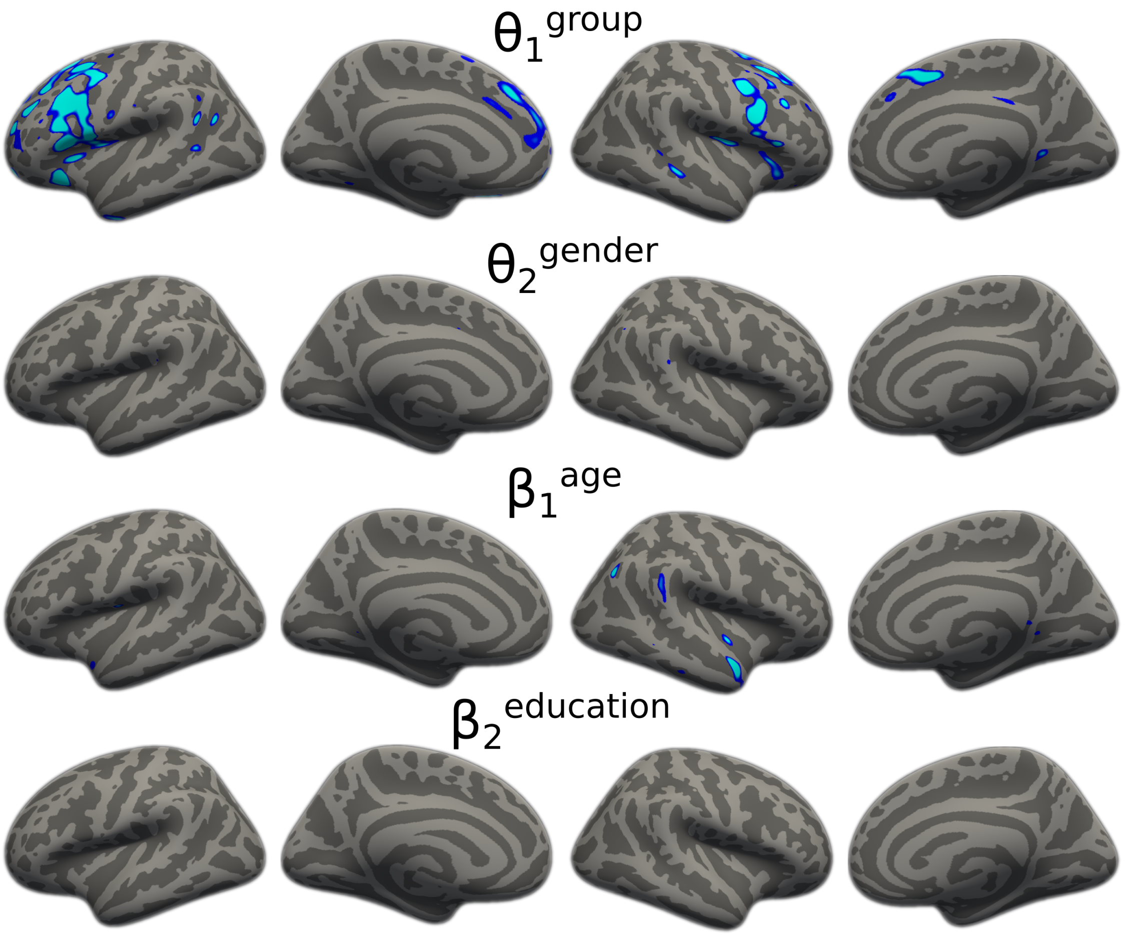
θ_1^{group}
θ_2^{gender}
β_1^{age}
$\beta_2^{\text{education}}$

Figure 22: Inflated left (on the left) and right (on the right) hemisphere surfaces of the fsaverage subject, with a Wald test p-value overlay (threshold $p < 0.0001$, brightly colored areas $p < 0.00001$) for the parameters of Model $\mathcal{M}_{A1}$ used on the TMTB dataset with $\theta_{1(j)}$ modelling the difference value between TMTB|D+ vs. TMTB|D-.

5.2.4 HCP Parcellation for the TMTB results

I visualized the data with the HCP Parcellation: Figure 23 shows the inflated left and right hemisphere of the HCP 32k fs_LR subject with a p-map overlay from the permutation test of the $\theta_{1(j)}$ parameter for previously described model $\mathcal{M}_{A1}$, which shows the difference in cortical thickness between the group TMTB|D+ and the group TMTB|D-. As another overlay, the areas of the parcellation by Glasser et al. (2016) are shown.

Significantly reduced thickness was found in the following areas on the left hemisphere: 44, IFSp, IFJa, 8C, IFJp, 6r, FOP2, FOP3, FOP4, AVI, 47s, 47l, 55b, FEF, PGi, TPOJ1, 9-46d, 46, 8BM, 9M, 9a, TGv; And in the following areas on the right hemisphere: 55b, PEF, IFJp, 8C, p9-46v, 6r, 44, OP2-3, Ig, FOP2, FOP4, AVI, STSdp, 8BM, SCEF, ProS.

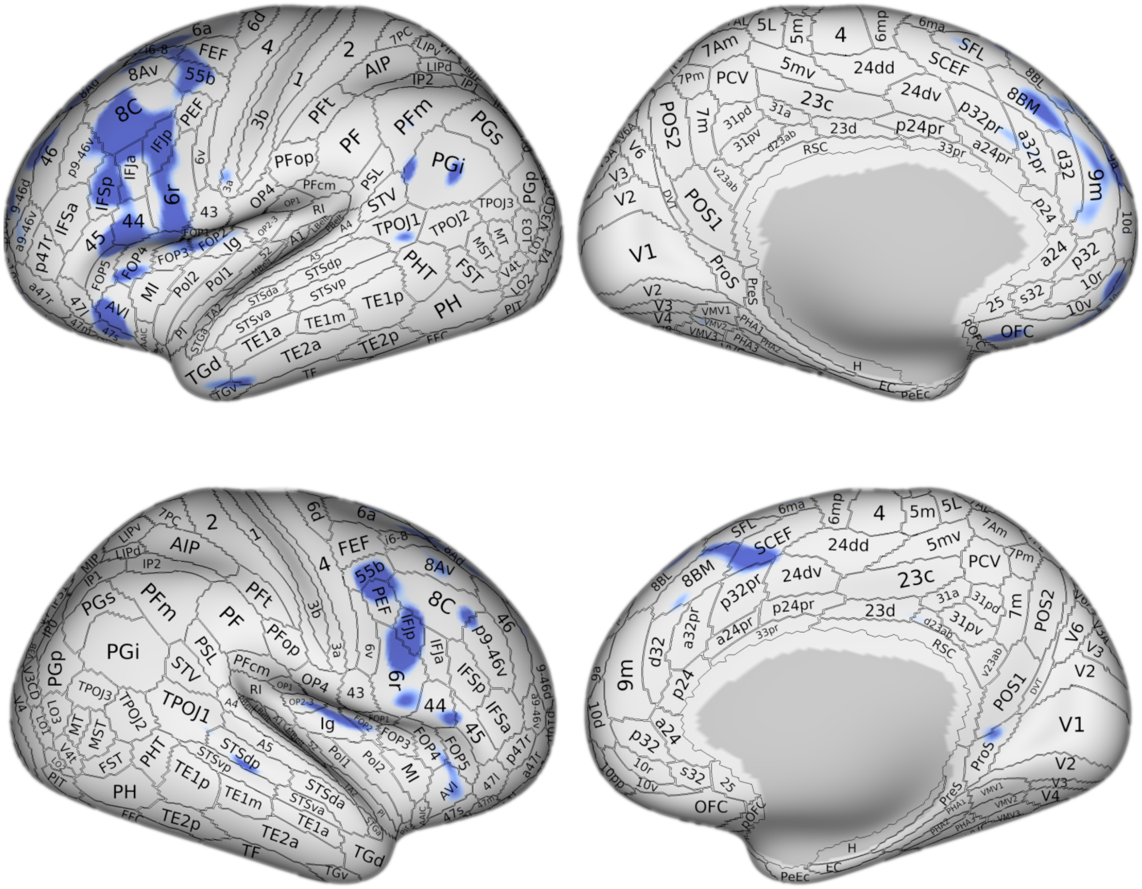
6a
6d
FEF
8Av
55b
4
2
1
3b
AIP
7PC
LIPv
LIPd
IP2
IP1
PFt
PF
PFm
PGs
PGi
PFop
PFcm
8C
IFJp
IFJa
PEF
6v
6r
46
p9-46v
9-46d
a9-46v
IFSp
IFSa
44
45
43
OP4
OP1
RI
A1
A4
A5
PSL
STV
TPOJ1
TPOJ2
TPOJ3
PGp
LO3
MT
MST
FST
PHT
V4t
LO2
PIT
PH
TE1p
TE1m
TE1a
TE2a
TE2p
STSdp
STSvp
STSda
STSva
FOP1
FOP2
FOP3
FOP4
FOP5
MI
Pol1
Pol2
Ig
AVI
47l
47s
47m
a47r
p47r
TGd
TGv
TF
FFC
5L
5m
6mp
6ma
7Am
7Pm
PCV
5mv
24dd
24dv
SCEF
SFL
8BL
8BM
23c
23d
31pd
31pv
31a
d23ab
v23ab
RSC
p24pr
a24pr
33pr
p32pr
a32pr
d32
9m
10d
p24
a24
p32
s32
10r
10v
25
OFC
pOFC
7m
POS2
POS1
V6
V3
V2
V1
V4
ProS
PreS
PHA1
PHA2
PHA3
VMV1
VMV2
VMV3
H
EC
PeEc
DVT

Figure 23: Inflated left (top) and right (bottom) hemisphere surface of the HCP 32k fs_LR subject, with a p-map overlay for the $\theta_{1(j)}$ parameter of Model $\mathcal{M}_A$, capturing the group difference between TMTB|D+ and TMTB|D-. In addition, an overlay of the human cortex parcellation is shown.

5.3 Discussion

A typical measure for frontal lobe function and executive (dys-)function is the TMT (Bowie & Harvey, 2006), although its specific role in this context is still being discussed (Kopp et al., 2015; Demakis, 2004; Stuss et al., 2001; Reitan & Wolfson, 1995). To contribute to this discussion, I used the GENFI data, which provided TMT measures, to present a third use case: I delineated areas relevant for TMT performance within the genetic FTLD cohort. When comparing group TMTA|D+ with group TMTA|D-, I found little but significant reduction in cortical thickness in the frontal cortex of the left hemisphere. When comparing group TMTB|D+ with group TMTB|D-, I found significant reduction in cortical thickness in the frontal cortex of the left and right hemisphere. The resulting pattern seems to include the results of the TMTA analysis, but in addition to that more areas.

The areas relevant for TMTB test performance included the smaller areas, relevant for TMTA test performance. As the TMTA is the easier of the two TMT subtests, there supposedly are less cognitive resources necessary to perform well in the subtest, accordingly the failure of less cognitive resources is sufficient to produce a bad test result. For the TMTB performance, which is the harder of the two tests, more specific functions, including those from the TMTA are necessary. The failure of these specific functions in addition to the more general TMTA functions leads to a bad test result. Relevant brain areas for these more specific functions have been revealed in the analysis. The areas with reduced cortical thickness in the TMTB analysis, that were additionally significant compared to the TMTA analysis, seem to be the areas, associated with the additional test performance that is necessary for the TMTB subtest.

Considering the TMT performance, especially interesting functions are exactly those, capturing the performance difference between the TMTB and TMTA test, as this difference captures the more specific functions, necessary for the TMTB performance, while eliminating the role of the basic functions, necessary for both TMTA and TMTB. The difference score or ratio between TMTA and TMTB reduces influences through functions such as processing speed and visuospatial processing, thus allowing for a cleaner measure of executive function (Varjacic, Mantini, Demeyere, & Gillebert, 2018; Kopp, 2011; Bowie & Harvey, 2006). Alas, due to the unavailability of raw TMT test data I could not conduct an analysis with such a measure: I only had access to normalized test performance scores, therefore I could not calculate a difference score or ratio. Future studies could improve our understanding by delineating even more specific brain areas for the proposed cognitive functions of the TMTB performance compared to the TMTA performance, by using the difference score or ratio.

As I identified problems with the age matching of the first used matching procedure in the previous apraxia analysis, I decided to use the second matching procedure for the TMT analysis. This age matching method did work well again, as the age parameter did capture almost no significant areas, apart from some very small areas in the right hemisphere. Hence, the results show the specific beginning of the atrophy pattern, presumably getting more pronounced when the atrophy develops with increasing age.

A limitation of the TMT results are the standardized test scores: they have been standardized, by using the non-mutation carriers in the dataset as standard population. While this serves as a good reference due to the non-mutation carriers having a similar sociocultural and genetic background, it also limits the external validity.

6 Presymptomatic cortical degeneration

When do brain changes in degenerative disorders present themselves or at least present themselves to a degree that they can be detected is a key question for research about degenerative diseases as FTLD (Bertrand et al., 2018; Rohrer et al., 2015; Bateman et al., 2012). This is why possible presymptomatic effects in the previously detected areas of the human cortex are of special interest. Therefore, I chose to investigate them with another approach that builds upon the previous results, using another statistical model, presented below.

6.1 Methods

I performed the analysis on all 472 GENFI subjects, including 114 second visits with a resulting total sample of 586 data points: I fitted a linear mixed model (LMM) with the mean cortical thickness of masked areas as dependent variable, random effects for site, family and subject, with family being nested within site and subject being nested within family, and fixed effects for age and group, where group is a factor with three levels: M-, M+S- and M+S+. I allowed a different variance estimates per group as the M+S+ group showed a higher variance of mean thickness values in comparison to the other two groups. I calculated Wald F-Tests for the fixed effects age and group with a Bonferroni corrected alpha level of 0.01. Additionally, a Likelihood-Ratio-Test (LRT) was performed to check if a model with an interaction of the two fixed effects would fit the data better. A second LRT was performed to check if a less complex model, which does not allow for difference variance estimates per group, would suffice for a good fit of the data. Equation 5 shows the described model, which will be referred to as Model[9]$\mathcal{M}_{P1}$: y_{pt} denotes the dependent variable, which is the mean value of the cortical thickness in masked areas for person p and time point (i.e. age) t. β variables denote fixed effects, whereas υ variables denote random effects and ε denotes the individual error, which follows a normal distribution with a group wise (index g) variance parameter σ_g.

---

[9] For the actual calculation of the model, I used the `lme` function from the `nlme` R package, with a nested structure of the random effects (subject nested in family nested in site) and a weights argument to define the heteroscedasticity (every group having its own variance).

$$\begin{aligned}
y_{pt} &= \beta_0 + \beta_1 \cdot age_{pt} + \beta_2 \cdot group_p + \upsilon_{0p}^{(site)} + \upsilon_{1p}^{(family)} + \upsilon_{2p}^{(subject)} + \varepsilon_{(g)pt}^{(group)} \\
\upsilon_{0p}^{(site)} &\sim N(0, \sigma_0^{(site)}) \\
\upsilon_{1p}^{(family)} &\sim N(0, \sigma_1^{(family)}) \\
\upsilon_{2p}^{(subject)} &\sim N(0, \sigma_2^{(subject)}) \\
\varepsilon_{(g)pt}^{(group)} &\sim N(0, \sigma_g^{(group)}); \text{ where } \sigma_g^{(group)} = \begin{cases} \sigma_3 & \text{for group M-} \\ \sigma_4 & \text{for group M+S-} \\ \sigma_5 & \text{for group M+S+} \end{cases}
\end{aligned} \tag{5}$$

In a second alternative analysis of presymptomatic cortical degeneration I replaced age with expected years of onset (EYO) as a fixed effect. For the group variable, I did not differentiate between M-, M+S- and M+S+, but only between M- and M+. The neglected differentiation was due to the inclusion of the EYO variable in the model, which claims to inherently contain this information, as it differentiates between presymptomatic and symptomatic. Also, it goes in accordance to the approach of previous studies (Rohrer et al., 2015; Bateman et al., 2012). Other than that, the model was the same as shown in Equation 5 and will be referred to as Model $\mathcal{M}_{P2}$. The EYO variable was calculated, in accordance to Rohrer et al. (2015), as the individual difference between a subject's age and the mean age at onset of symptomatic members of her/his family. This was done for all subjects, even the ones that were already symptomatic. Testing procedure was the same as before, with Wald F-Tests for the fixed effects and an LRT to test the interaction of the fixed effects.

The analysis with Model $\mathcal{M}_{P1}$ will be called main analysis, while the analysis with Model $\mathcal{M}_{P2}$ will be called alternative analysis.

6.2 FTLD related areas

6.2.1 Methods

As I identified FTLD related areas with the FTLD symptomatic dataset, it is of interest to investigate presymptomatic cortical degeneration in those areas. Substantially, this replicates the idea from Rohrer et al. (2015) with a different method and perspective on the data. In order to do so, I created masks from left and right hemispheric areas of reduced cortical thickness identified in the comparison M+S+ vs. M+S- (shown in Figure 4) and calculated the mean cortical thickness in these areas for each of the 586 GENFI subjects (including the 114 second visits). These mean cortical thickness values have then been used as the dependent variable of the Model $\mathcal{M}_{P1}$ (main analysis) and Model $\mathcal{M}_{P2}$ (alternative analysis).

6.2.2 Results

In the main analysis, I found a significant effect for age ($F(1, 109) = 198.4806,\ p < 0.0001$), and a significant effect for the comparison M- vs. M+S+ ($F(1, 109) = 178.1357,\ p < 0.0001$) and M+S- vs. M+S+ ($F(1, 109) = 155.4739,\ p < 0.0001$), but no significant effect for the comparison M- vs. M+S- ($F(1, 109) = 6.0153,\ p = 0.0158$). A more complex model that included interaction coefficients for the interaction of age and the groups did not fit the data significantly better as the main model ($LRT(2) = 0.5735,\ p = 0.7507$). The main model, in comparison to a less complex model, which does not allow for different variance estimates per group, fitted the data better ($LRT(2) = 120.2854,\ p < 0.0001$).

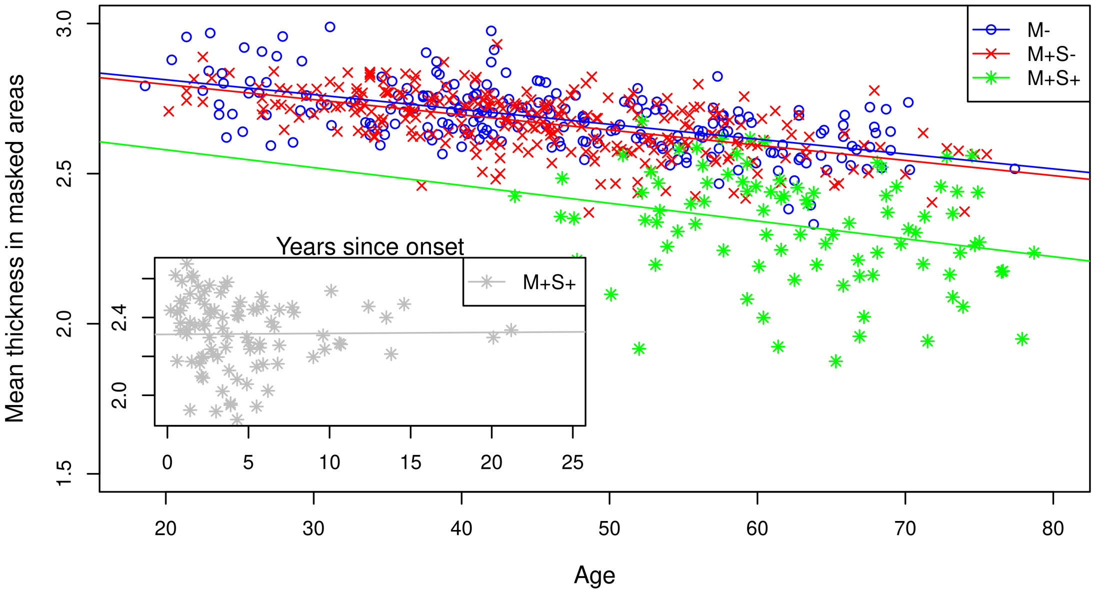

Figure 24: The ordinate shows the mean thickness in masked areas (from the previous results of the M+S+ vs. M+S- analysis); the abscissa shows the age of the subjects. Blue circles denote M-, red crosses denote M+S- and green stars denote M+S+. The lower left inlay shows again the M+S+ with the same ordinate and the years since onset (YSO) on the abscissa. The lines show the group trend of the group with the same color, respectively.

In my alternative analysis, the comparison of the alternative model with a more complex model including an interaction term of the fixed effects was significant ($LRT(1) = 28.2030,\ p < 0.0001$). Thus, the following parameter tests have been conducted with the more complex model: I found a significant effect for the EYO variable ($F(1, 110) = 50.9963,\ p < 0.0001$) and a significant effect for the group comparison M+ vs. M- ($F(1, 265) = 60.2788,\ p < 0.0001$). Figure 25 depicts the data dependent on the EYO variable (allowing for it to be positive, where expected disease onset lay in the past), Figure 26 is very similar to Figure 25, but shows the actual YSO for M+S+ subjects.

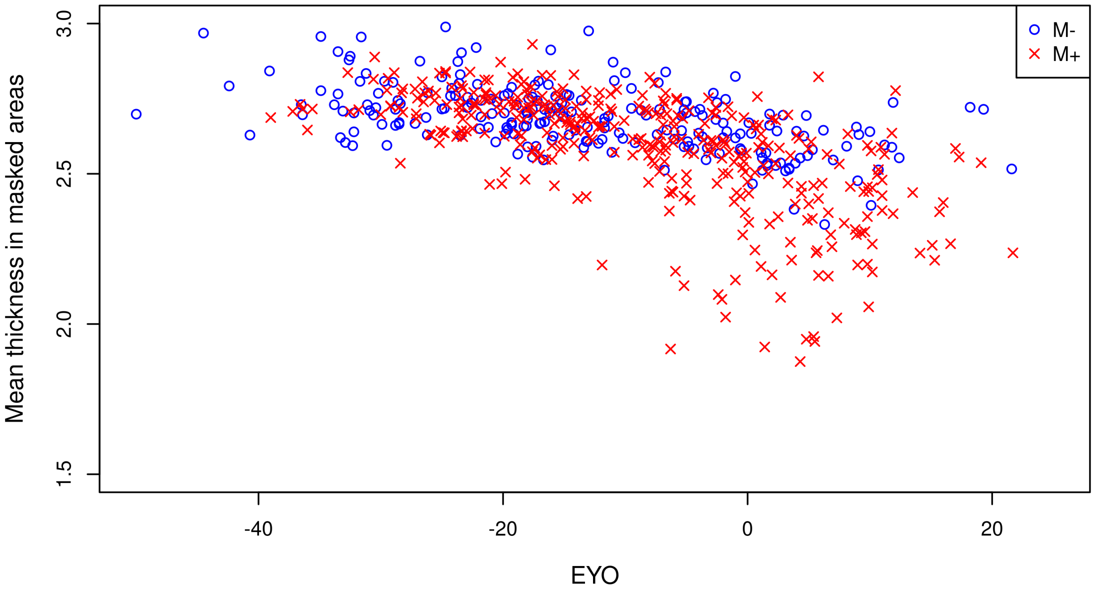

Figure 25: Scatterplot with the mean thickness of the masked areas (by the previous results from the M+S+ vs. M+S- comparison) as ordinate and the EYO as abscissa. Red crosses delineate M+, whereas blue circles delineate M-.

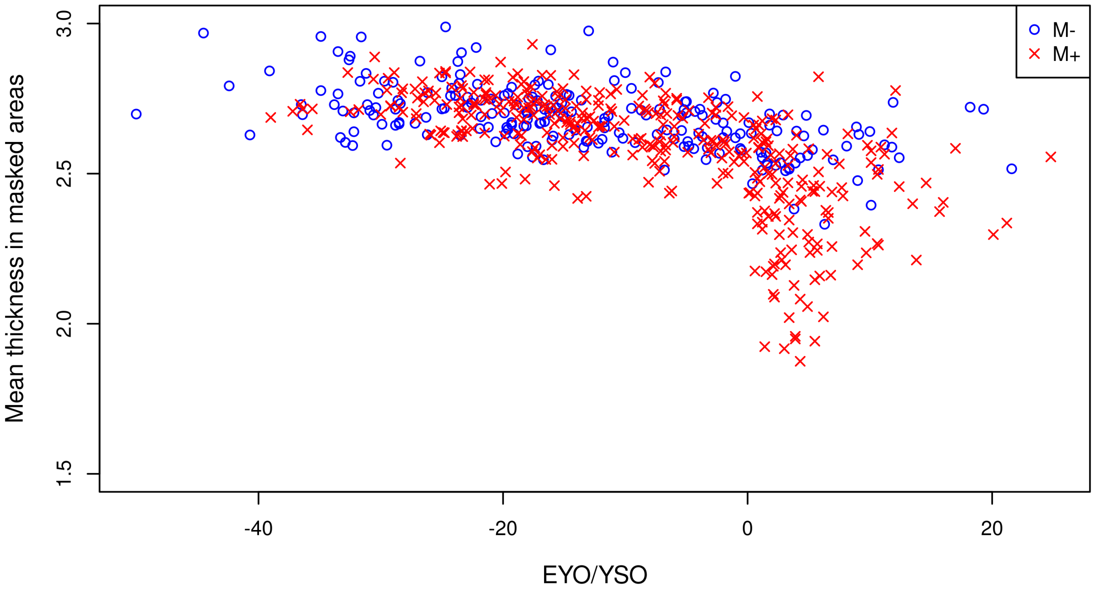

Figure 26: Similar scatterplot as shown in Figure 25, however, for M+S+ datapoints, the EYO variable has been corrected to show the actual YSO.

6.3 Apraxia related areas

6.3.1 Methods

I created masks from left and right hemispheric areas of reduced cortical thickness identified in the comparison M+A+ vs. M+A- (shown in Figure 10) and calculated the mean cortical thickness in these areas for each of the 586 GENFI subjects (including the 114 second visits). These mean cortical thickness values have then been used as the dependent variable of the Model $\mathcal{M}_{P1}$ (main analysis) and Model $\mathcal{M}_{P2}$ (alternative analysis).

6.3.2 Results

In the main analysis, I found a significant effect for age ($F(1, 109) = 189.6241,\ p < 0.0001$), and a significant effect for the comparison M- vs. M+S+ ($F(1, 109) = 117.5038,\ p < 0.0001$) and M+S- vs. M+S+ ($F(1, 109) = 99.2009,\ p < 0.0001$), but no significant effect for the comparison M- vs. M+S- ($F(1, 263) = 6.6259,\ p = 0.0114$). A more complex model that included interaction coefficients for the interaction of age and the groups did not fit the data significantly better as the main model ($LRT(2) = 0.0214,\ p = 0.9894$). The main model, in comparison to a less complex model, which does not allow for different variance estimates per group, fitted the data better ($LRT(2) = 114.3929,\ p < 0.0001$). The predictions of the fixed effects are shown in Table 7. For a visualisation of the data and the described corresponding results, see Figure 27.

Table 7: Predictions for cortical thickness of fixed effects in main analysis. Columns depict age given in first row, rows depict group given in first column. The thickness values are given in mm.

Group ↓ | Age →	**20**	**30**	**40**	**50**	**60**	**70**	**80**
M-	2.7322	2.6780	2.6237	2.5695	2.5153	2.4610	2.4068
M+S-	2.7090	2.6548	2.6005	2.5463	2.4920	2.4378	2.3835
M+S+	2.4791	2.4249	2.3706	2.3164	2.2621	2.2079	2.1537

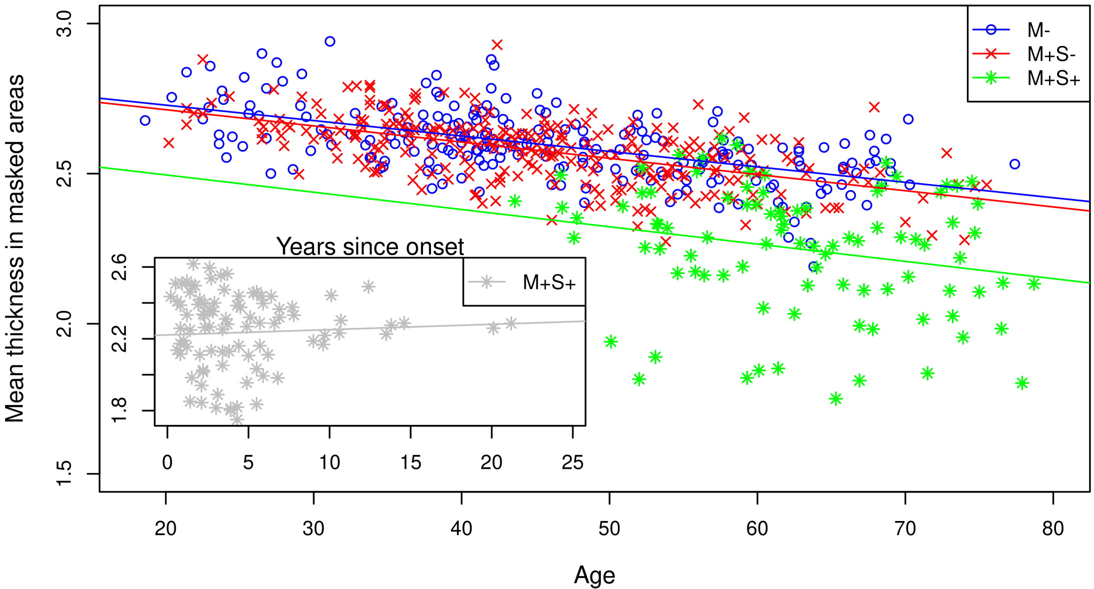

Figure 27: The ordinate shows the mean thickness in masked areas (from the previous results of the M+A+ vs. M+A- analysis); the abscissa shows the age of the subjects. Blue circles denote M-, red crosses denote M+S- and green stars denote M+S+. The lower left inlay shows again the M+S+ with the same ordinate and the YSO on the abscissa. The lines show the group trend of the group with the same color, respectively.

In my alternative analysis, the comparison of the alternative model with a more complex model including an interaction term of the fixed effects was significant ($LRT(1) = 19.4641,\ p < 0.0001$). Thus, the following parameter tests have been conducted with the more complex model: I found a significant effect for the EYO variable ($F(1, 110) = 45.6413,\ p < 0.0001$) and a significant effect for the group comparison M+ vs. M- ($F(1, 265) = 44.73451,\ p < 0.0001$). The predictions of the more complex model are shown in Table 8. Figure 28 depicts the data dependent on the EYO variable (allowing for it to be positive, where expected disease onset lay in the past), Figure 29 is very similar to Figure 28, but shows the actual YSO for M+S+ subjects.

Table 8: Predictions for cortical thickness of fixed effects in alternative analysis. Columns depict age given in first row, rows depict group given in first column. The thickness values are given in mm.

Group ↓ | Age →	**20**	**30**	**40**	**50**	**60**	**70**	**80**
M-	2.6655	2.6381	2.6106	2.5832	2.5557	2.5283	2.5008
M+	2.6810	2.6310	2.5810	2.5310	2.4810	2.4310	2.3810

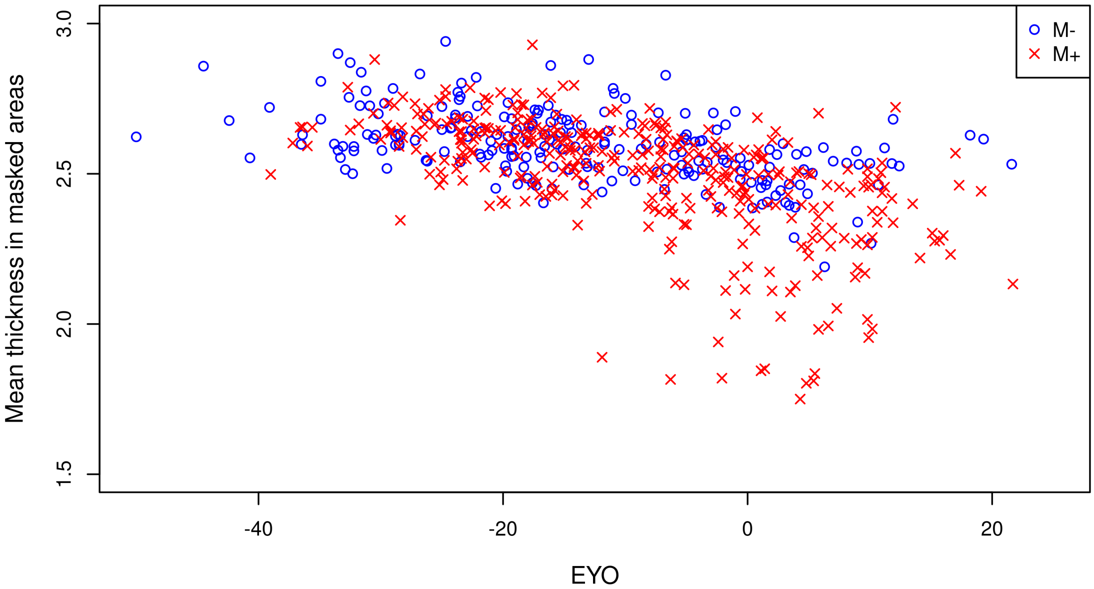

Figure 28: Scatterplot with the mean thickness of the masked areas (by the previous results from the M+A+ vs. M+A- comparison) as ordinate and the EYO as abscissa. Red crosses delineate M+, whereas blue circles delineate M-.

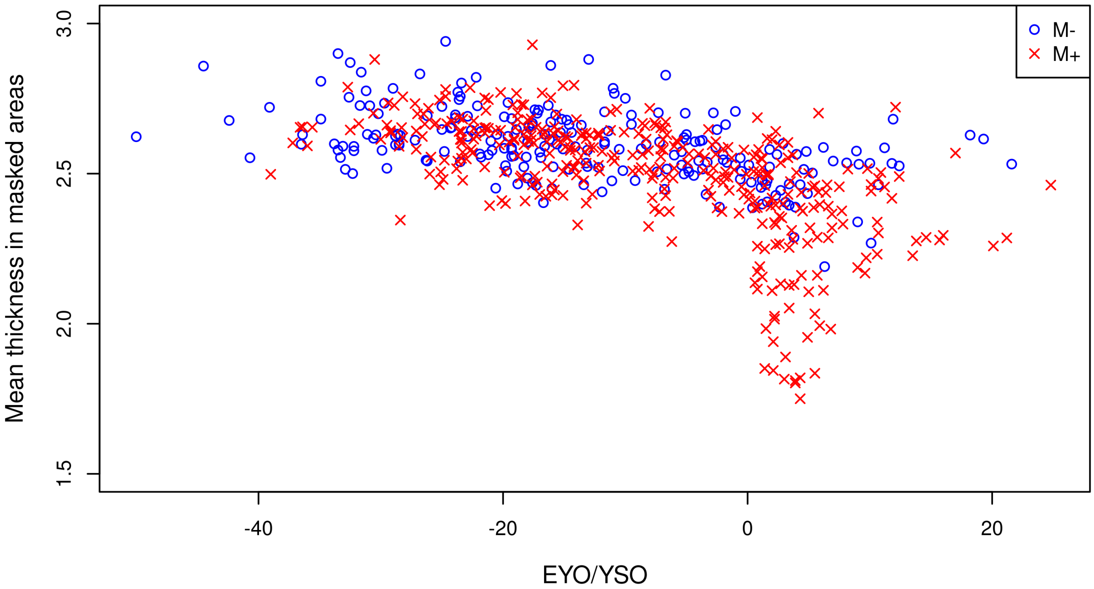

Figure 29: Similar scatterplot as shown in Figure 28, however, for M+S+ datapoints, the EYO variable has been corrected to show the actual YSO.

6.4 TMT related areas

6.4.1 Methods

This analysis followed the same procedure as described previously in Section 6.3.1, with the only difference being the used masks: I created masks from left and right hemispheric areas of reduced cortical thickness identified in the comparison TMTB|D+ vs. TMTB|D- (shown in Figure 21) and calculated the mean cortical thickness in these areas for each of the 586 GENFI subjects (including the 114 second visits). These mean cortical thickness values have then been used as the dependent variable of the Model $\mathcal{M}_{P1}$ (main analysis) and Model $\mathcal{M}_{P2}$ (alternative analysis).

6.4.2 Results

In the main analysis, I found a significant effect for age ($F(1, 109) = 167.5471,\ p < 0.0001$), and a significant effect for the comparison M- vs. M+S+ ($F(1, 109) = 136.1934,\ p < 0.0001$) and M+S- vs. M+S+ ($F(1, 109) = 120.7403,\ p < 0.0001$), but no significant effect for the comparison M- vs. M+S- ($F(1, 109) = 4.336733,\ p = 0.0396$). A more complex model that included interaction coefficients for the interaction of age and the groups did not fit the data significantly better as the main model ($LRT(2) = 0.1985,\ p = 0.9055$). The main model, in comparison to a less complex model, which does not allow for different variance estimates per group, fitted the data better ($LRT(2) = 135.8287,\ p < 0.0001$).

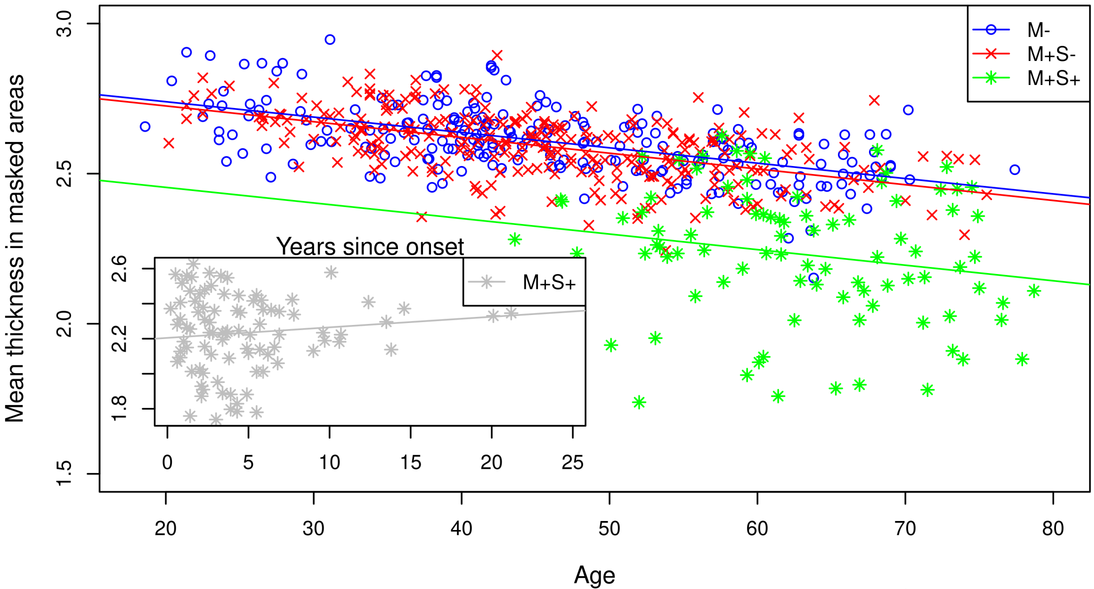

Figure 30: The ordinate shows the mean thickness in masked areas (from the previous results of the TMTB|D+ vs. TMTB|D- analysis); the abscissa shows the age of the subjects. Blue circles denote non-mutation carriers, red crosses denote presymptomatic mutation carriers and green stars denote symptomatic mutation carriers. The lower left inlay shows again the symptomatic mutation carriers with the same ordinate and the YSO on the abscissa. The lines show the group trend of the group with the same color, respectively.

In my alternative analysis, the comparison of the alternative model with a more complex model including an interaction term of the fixed effects was significant ($LRT(1) = 22.8384,\ p < 0.0001$). Thus, the following parameter tests have been conducted with the more complex model: I found a significant effect for the EYO variable ($F(1, 110) = 42.2496,\ p < 0.0001$) and a significant effect for the group comparison M+ vs. M- ($F(1, 265) = 49.3672,\ p < 0.0001$). Figure 31 depicts the data dependent on the EYO variable (allowing for it to be positive, where expected disease onset lay in the past), Figure 32 is very similar to Figure 31, but shows the actual YSO for M+S+ subjects.

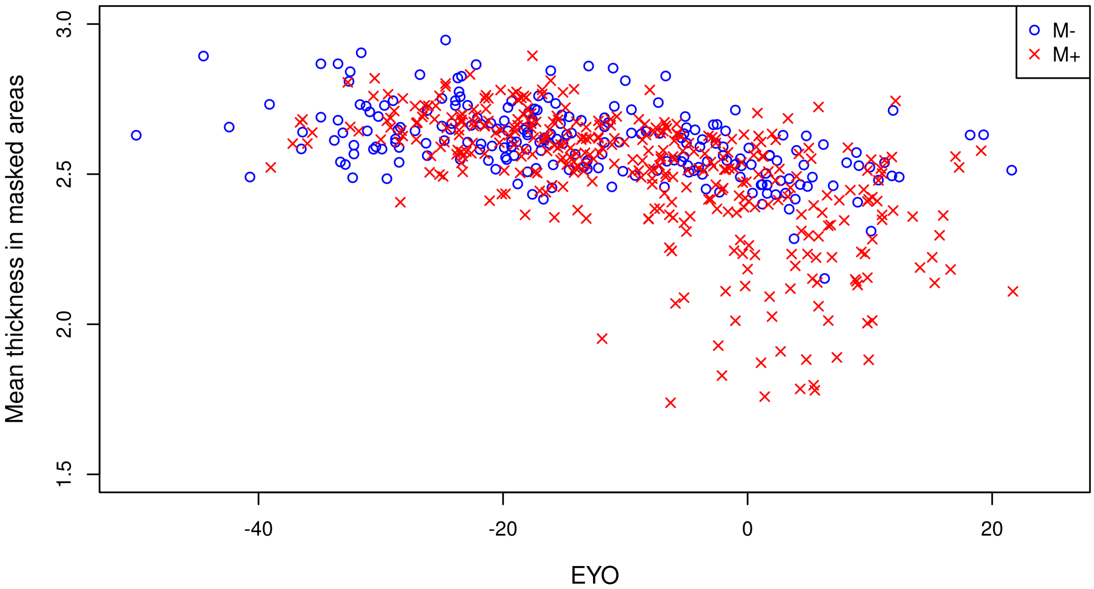

Figure 31: Scatterplot with the mean thickness of the masked areas (by the previous results from the TMTB|D+ vs. TMTB|D- comparison) as ordinate and the EYO as abscissa. Red crosses delineate M+, whereas blue circles delineate M-.

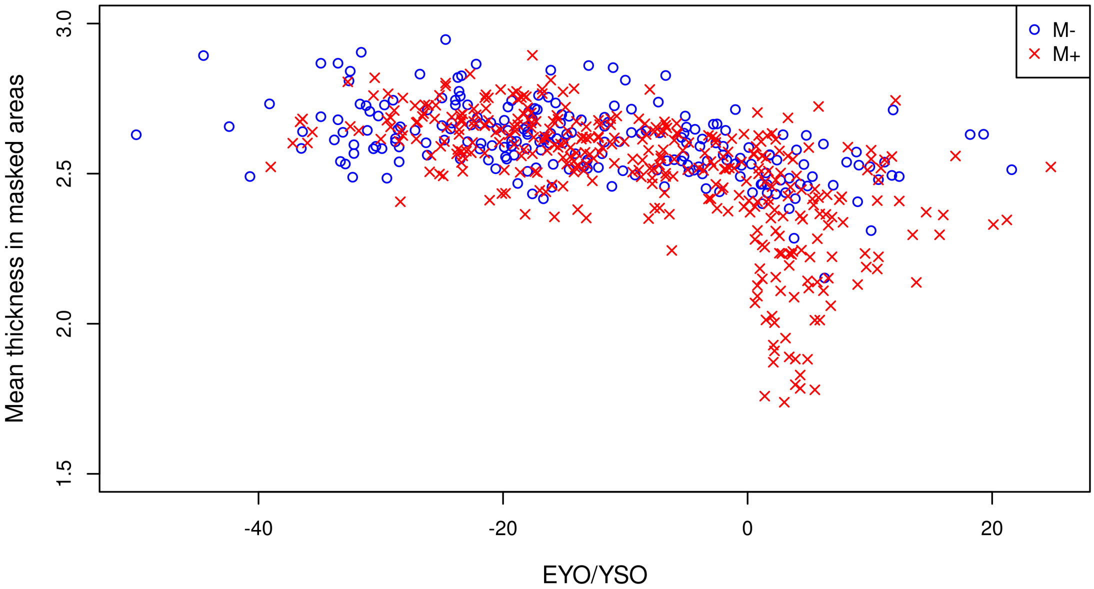

Figure 32: Similar scatterplot as shown in Figure 31, however, for M+S+ datapoints, the EYO variable has been corrected to show the actual YSO.

6.5 Discussion

The preceding analyses aimed at the detection of reduced cortical thickness prior to the symptomatic onset of the disease, in the previously detected brain areas for genetic FTLD subjects. The results for the three subanalyses all showed equal trends: there was a significant effect of age, which showed a decrease in cortical thickness with an increase in age: this is in accordance to previous research about age effects in the cortex (see, e.g. Salat et al., 2004). Also, there were significant group effects, showing that symptomatic FTLD subjects tend to have lower cortical thickness, compared to presymptomatic FTLD subjects and compared to healthy controls. The main analysis did, against my expectations, not show any significant presymptomatic decrease in cortical thickness: I observed no significant interaction of age and group in any of the three subanalyses.

The alternative analysis, which instead of using the real age, used the previously described EYO variable, detected significant presymptomatic changes. However, even though this methodology follows a more typical approach for this kind of analysis (as in Rohrer et al., 2015), I claim that the assumptions for the used variables and the model have certain flaws, which lead to false results and conclusions. Especially the use of the EYO variable is questionable, as it tends to convert symptomatic subjects to presymptomatic subjects and thereby blends symptomatic values into a presumed presymptomatic phase. A close examination of Figure 25 and Figure 26 (or Figure 28 and Figure 29 or Figure 31 and Figure 32 for the respective analysis) visualizes this problem, showing how actually symptomatic data points move from the presymptomatic into the symptomatic phase, if the estimated EYO variable is replaced by the actual YSO variable for subjects who actually are symptomatic. Secondly, beyond the questionable use of the EYO variable, the respective alternative model itself is not accurate for the data, as the group of M+ is not divided into subgroups of M+S+ and M+S-. This is, however, a crucial step as the subgroups showed a different variance, as can be seen in Figure 24 (or Figure 27 and Figure 30 for the respective analysis). The observation of a significantly better data fit of the main model from the main analysis, including the differentiation between symptomatic and presymptomatic subjects, compared to a simpler model without different variance estimates per group, further supported this point.

In the presented univariate mixed effects models, I used the mean values of cortical thickness of the masked areas of the left and right hemisphere as the dependent variable, serving as a global measure for the masked areas on both hemispheres. This means that

I am giving away the information about where exactly in the masked areas specific effects might be located. Also, there exists the possibility of overseeing local effects of cortical thickness, that could cancel each other out by using a mean across the respective regions of interest. The results thus have to be taken exactly the way they were constructed in the model: as a mean effect of the masked areas. On the other hand, as the thickness values in those areas of a subject do follow a normal distribution, taking the mean is, on pure methodological grounds, entirely sufficient and reduces complexity.

A word about the random factors (family, site and subject) is also necessary, as their unbalanced structure does not prevent the models from being fitted, but could influence it: the random factors often do only have one member per group, which reduces the variability that could be bigger in reality, and thus could negatively influence the validity of the model.

As a last limitation of the main analysis, note that the presented model is not perfect: Looking at the plots of the main model for the respective analyses (see Figure 27, Figure 24 and Figure 30), there always seem to be few symptomatic mutation carriers, that are not well depicted by the model (the ones with the lowest mean thickness in masked areas). Maybe there are rare cases with an even worse disease progression. Future studies could try to differentiate between symptomatic subjects especially considering disease severity and try to find out what distinguishes those subjects with a worse disease progression from the others.

7 General Discussion

In this work I investigated the correlation between cortical atrophy patterns of genetic FTLD subjects with specific neuropsychological dysfunctions and if they can be detected at the presymptomatic stage. At the beginning, the general atrophy pattern of symptomatic FTLD subjects has been replicated. Then, in a more specific approach, apraxia deficits operationalized by gesture imitation deficits and deficits in TMT performance have been investigated. In a surface-based analysis of cortical thickness and with the help of a newly established methodology for vertex-based multiple comparison correction, I delineated cortical atrophy patterns for the respective neuropsychological dysfunctions.

First, in an attempt to establish my newly presented methods and to enhance the differences to the subsequent hypothesis driven approach about specific neuropsychological dysfunctions compared to a general approach in search of brain areas altered by disease progression, I replicated expected findings about genetic FTLD subjects: I found the typical atrophy pattern of FTLD within symptomatic genetic FTLD subjects when compared to presymptomatic genetic FTLD subjects.

Second, I looked at apraxia within FTLD subjects. Part of my methods already included a result: the identification of 31 subjects showing apraxia, i.e. gesture imitation deficits, among 294 mutation carriers. I found brain areas with reduced cortical thickness among genetic FTLD subjects compared to non-mutation carriers, who served as supposedly healthy controls. In order to delineate the brain areas relevant for apraxia in genetic FTLD subjects I showed brain areas with reduced cortical thickness specific to genetic FTLD subjects with apraxia compared to genetic FTLD subjects without apraxia. Lastly, I could not delineate mutation specific atrophy patterns for apraxia.

In a third analysis with a similar approach, I presented another use case by investigating TMT deficits among genetic FTLD subjects: this analysis revealed brain areas with reduced cortical thickness specific to genetic FTLD subjects showing TMT performance deficits compared to genetic FTLD subjects, who did not show TMT performance deficits.

The subsequent analysis about presymptomatic brain changes, i.e. changes in cortical thickness before the disease reaches a clinically visible degree, was not able to capture any significant presymptomatic effects: I could not find reduced cortical thickness prior to the symptomatic onset of the disease, in the previously detected brain areas for genetic FTLD subjects. However, this analysis did show an effect for symptomatic subjects, as they commonly have decreased cortical thickness in the respective areas, and an age effect, with cortical thickness decreasing by increasing age in the respective areas.

The broad definition of apraxia as gesture imitation deficits has its advantages, as it enables to capture subjects with apraxia deficits on a simple and practicable basis. Lots of data is present in the GENFI dataset and I used only a small amount of it. To collect such a huge amount of clinical data, it was necessary to not go into to much detail about all specific subfunctions. Moreover, this loose definition of apraxia already enabled the identification of several individuals showing these deficits, as mentioned earlier. It also allowed to conduct the presented analysis emphasizing the importance of neuropsychological deficits for neurodegenerative disorders, as apraxia for FTLD spectrum disorders. Albeit, the definition of apraxia as gesture imitation deficits has its limits. The validity of the presented results for apraxia research is restricted, as a clearer definition of apraxia, e.g. following the definition presented by Goldenberg (2014), is necessary to arrange the results within the existing literature about apraxia. Goldenberg (2014) differentiates between apraxia of gesture imitation of finger and hand movements as supramodality of body part coding; single tool/object use as a supramodality of mechanical problem solving; production of communicative gestures on command as high level cognitive disturbance.[10] Defining apraxia as a unitary disorder of a praxis system is highly problematic, due to the complex and different mechanisms, that underlie the specific apraxia subcategories. In accordance, these specific apraxia subcategories show different correlates of brain regions (see, e.g. Buxbaum et al., 2014; Goldenberg, 2014; Goldenberg et al., 2007; Goldenberg & Karnath, 2006). One of the questions that could arise when looking at the presented results is, why there were no parietal regions delineated, that seem so common to specific subfunctions of apraxia: the answer might underlie the definition of apraxia as gesture imitation of mainly finger movements, which can result from left frontal areas (Goldenberg & Karnath, 2006) and areas of the right hemisphere (Goldenberg & Strauss, 2002).

Another possible explanation for not finding parietal areas could be due to the development of the atrophy pattern. Rohrer et al. (2015) claims, that frontal atrophy occurs prior to parietal atrophy. As the GENFI data includes many subjects at an early stage of the disease process, it is possible that relevant apraxia areas in the frontal lobe get detected, while areas in the parietal lobe, also relevant for apraxia, might show up later, when the disease process progressed further. This could be a bias, due to the specific

---

[10] Goldenberg (2014) provides additional and alternative views together with a historic review of theories, that can not all be mentioned here.

FTLD disease progression. Although, as most subjects showing apraxia deficits are at a later stage of the disease, the strength of influence through this bias is questionable. Also, keep in mind that other research about apraxia suffers from disease specific biases as well.

The typical context for apraxia research is apraxia in stroke. Most cohorts for apraxia in stroke studies rely on data from patients that suffered middle cerebral artery (MCA) strokes. I supply a completely different context for apraxia research without this limitation: not only by using a genetic FTLD cohort, but also by looking at all areas of the cortex. This could, at least partly, explain the overlap of my results with the presented apraxia in stroke (Goldenberg et al., 2007; Goldenberg & Karnath, 2006) results, in the anterior regions of those results. Stroke data could bias study results, by showing inferior parietal areas more often, as they are located in areas supplied mainly by the MCA. However, it is possible that loss of function in other areas of the brain can result in the same or a very similar deficit. Possibly, the actual cortex areas relevant for apraxia in a degenerative disease go beyond the area supplied by the MCA, into more anterior regions. One could even argue, that apraxia in a degenerative disease is due to a very different disturbance of basic brain functions. Moreover, regions on the other half of the brain could also play a different role for apraxia in a degenerative disease. The failure of one relevant area, parts of a relevant area or several (parts of) relevant areas for a neuropsychological function, could result in the same or similar outcome: clinically observable apraxia. A good distinction between different forms of apraxia could improve our understanding. Key areas for apraxia in FTLD could, in comparison to apraxia in stroke, not include parietal regions.

Until now, there exist only few studies investigating apraxia within neurodegenerative diseases and their brain correlates. The work by Johnen et al. (2016) is one, which points towards the importance of apraxia in FTLD and their brain correlates. Johnen et al. (2016) used a mixed dataset, consisting of 18 clinically diagnosed bvFTD and 18 clinically diagnosed subjects with Alzheimer's and 34 healthy control subjects. As apraxia measures, they used two different scores: one for pantomime of object-use and one for limb imitation of meaningful and meaningless gestures. As brain measure, they used grey matter volume originating from a volume-based approach. When using only the bvFTD subset of the data, they could not delineate any apraxia related clusters in the brain. Only when using a mixed dataset of both disease subgroups together, they could delineate shared neural correlates for both their measures in the parietal cortex, which they advocate as shared apraxia correlates, across the two different disease entities.

Concluding, there are many key differences between my results and the work by Johnen et al. (2016): a volume-based vs. surface-based approach, different apraxia measures and a very different datasets. Both results do provide different perspectives for relevant brain substrates of apraxia and show the necessity to investigate shared and unshared correlates of neuropsychological deficits across different disease entities. Degenerative diseases in general could provide a huge impact on research about neuropsychological (dys-)functions and their correlates in the brain, as the different patterns of atrophy do not rely on artery supply and thus are not biased by it.

Apraxia as a symptom is mainly present in subjects already affected by the disease, but those symptomatic subjects showed all degrees of disease severity. The distribution of FTLD-CDR-SOB scores, which have been measured for every subject and are good predictors for the actual disease severity (Borroni et al., 2010), showed, that apart from five subjects with apraxia, all had FTLD-CDR-SOB scores above 0, but also that the scores were distributed across a wide range. Note, that the proportion of subjects with a score of 0 was considerably lower, compared to the group of genetic FTLD subjects without apraxia. This means, that the proportion of symptomatic subjects was high among subjects that showed apraxia as a symptom. However, a more sensitive test for apraxia might also identify more subjects with apraxia within the group of subjects, who were not symptomatic. The advanced disease stage of many subjects with apraxia could probably alter the results, by including general FTLD atrophy, as the proportion of symptomatic and presymptomatic subjects is unbalanced among the compared groups. Choosing only matching symptomatic subjects was not an option for the present dataset, as this would have resulted in a very small matching group. Yet still, the comparison of mutation carriers showing apraxia vs. mutation carriers not showing apraxia is very specific and its differences to the other analyses presented make this point more clear: the results from the comparison of mutation carriers showing apraxia with non-mutation carriers (healthy controls) revealed a more general FTLD atrophy pattern. The whole pattern is not very specific for apraxia, but it includes the areas relevant for it. Another analysis that serves as a good reference is the discussed comparison of symptomatic vs presymptomatic genetic FTLD subjects, which captured the general FTLD atrophy pattern. To further lighten up and enhance the delineation of the specific relevant areas for apraxia within genetic FTLD, future research could take a closer look at the comparison of only symptomatic or presymptomatic subjects with and without apraxia.

To further enhance the discussion about relevant areas of the cortex for specific neuropsychological dysfunctions, it could also help to look at literature about general cortex parcellation. This is why I visualized my results with the multimodal parcellation by Glasser et al. (2016). The basis for the parcellation by Glasser et al. (2016) is the HCP data: a huge and promising investigation about the human connectome. For an overview about the HCP, see Van Essen et al. (2013) and technical details are summarized by Van Essen et al. (2012). The HCP data provides and combines several MRI modalities utilized on a young and representative cohort. It takes into account the latest technological advances and enables studies, like the one by Glasser et al. (2016), to use this data together with existing knowledge about the brain, resulting in a informative and validated parcellation of the brain. The details and scientific basis of this parcellation support closing the gap between disease pathology and healthy brain functioning. This is why visualizing my results in light of this parcellation also enables an easy and up to date comparison of my results to other studies in the past and in the future. Seeing my results within the parcellation by Glasser et al. (2016) could serve as a starting point for future studies, investigating specific subfunctions, hypothetically important due to relevant brain areas being affected, that might be necessary for specific neuropsychological functions, such as apraxia.

Comparing the apraxia related areas from my results with the myelin maps given by Glasser et al. (2016), the areas IFSp, IFJa, IFJp and PEF show a stronger relative myelination than their neighbouring areas and most of the frontal cortex. Thus, the delineated patches in the named areas are probably highly interconnected to other areas, whereas the delineated patches in the areas FOP1-5, 6r, 44, OP4 are less interconnected. More interconnected areas could also have stronger effects on the brain network or be more affected by network effects. With the rising importance of human connectomics (Sporns, 2013; Behrens & Sporns, 2012), the interconnections of the human brain, this could serve as another starting point for future studies: looking for the location of the interconnections of the strong myelinated areas. The areas with stronger myelination could operate through a underlying praxis network. Additionally, the role of the strength of interconnections for areas with strong atrophy in degenerative diseases is exciting. Another comparison with data from Glasser et al. (2016) allows to look at the associations of the apraxia related areas. The affected area 6r is lightly associated with sensorimotor tasks: could this area thus represent the part of apraxia, that is responsible for the apraxia related sensorimotor information? The areas PEF and IFJp seem to have less specific associations. Area 44, IFSp and IFJa also seem less specific and additionally tend to be more task negative, meaning they are less active when specific tasks are conducted. Do these differences in the unspecific frontal areas represent different parts of apraxia related activity? Questions arising from comparisons like these could drive future studies.

In order to find out more about when the degeneration process of the affected brain areas starts, I conducted the presymptomatic analysis. In the presymptomatic analysis, the group effect of symptomatic vs. presymptomatic carriers and of symptomatic vs. non-mutation carriers was significant in all of the three presented cases. These results were expected and confirmed, as we do know that symptomatic subjects suffer from atrophy in cortical areas of the brain. The effect of actual interest was the interaction effect of age and group. This interaction effect would model the hypothetically faster degeneration in presymptomatic carriers compared to non-mutation carriers. However, I found no significant interaction effect in my model. Additionally, there was no significant group difference between presymptomatic carriers and non-mutation carriers. Hence, no presymptomatic effects could be detected. Albeit, the absence of evidence is not evidence of absence: it is possible, that presymptomatic effects are there, but could no be detected with the here presented approach and the used dataset. Presymptomatic effects seem to be subtle and hard to pin down, especially considering the high variability of disease onset. Be aware that the presented presymptomatic analysis seems to suggest, that the data for the model stems from longitudinal data, nevertheless keep in mind that the basis for the analysis was mainly cross-sectional data, with only few one year followup data points. GENFI strives towards supplying a good longitudinal structure within their dataset, by including more and more followup data of the same subjects. This could in the future provide data for a good longitudinal study, that should be better in capturing the subtle differences between presymptomatic and non-mutation carriers.

In addition to the main analysis of the presymptomatic analysis, I also presented a second alternative analysis. It follows a more typical approach inspired by the existing literature in the field, especially the work by Rohrer et al. (2015) and Bateman et al. (2012). The model in this analysis differentiates only between mutation carriers and non-mutation carriers and does not contain an observed age variable, but a variable called EYO, that provides the estimated years until symptomatic onset of the disease. The estimation is based on a familial mean of symptomatic onset, including anamnestic information about disease onset of family members not included in the GENFI dataset themselves. This alternative analysis, especially in comparison with the methodically more correct main analysis, highlights the problems and limitations of inadequate modelling and problematic use of seemingly observed variables. The first problem is the grouping of presymptomatic and symptomatic mutation carriers in one general mutation carriers group: the two groups differ too much to simply treat them as one group, most notably in their variability. Secondly, the construction of the EYO variable is highly problematic, as it actually alters the supposedly presymptomatic area of the model with the false classification of symptomatic subjects as presymptomatic, as their expected onset is at a later time compared to the real onset time of the disease. In accordance with these theoretical problems, the model is also methodically inadequate, due to violations of general model assumptions, like the errors being independent and normally distributed. As a consequence, all substantial interpretation derived from this model are to be heavily questioned. The purpose of the alternative analysis was to sensitize the reader for these kinds of modelling problems and enhance the awareness of critical conclusions on the basis of such models. If the basis of our scientific reasoning relies on the adequacy of a model, all assumptions that are put into the model have to be checked.

Previously, I have talked about age effects in my results. The existence of a common age effect on cortical thickness has been shown in the literature (Salat et al., 2004). With the presymptomatic analysis, I hoped to shed more light on the effects of age on the GENFI data, apart from its main goal, to search for presymptomatic effects. However, especially considering this main goal, it is of utmost importance to make age effects visible and part of a model, claiming to show presymptomatic effects, as they can clearly alter the results. Indeed, the age effect of the main presymptomatic analysis was significant in all of the three presented analyses.

Furthermore, there are relevant age differences between the different groups that were used for the presented area detection analyses. To address them, I included respective age parameters for all used models in my analyses, a strategy often used in similar studies (Cash et al., 2018; Bertrand et al., 2018; Salat et al., 2004). Beyond such a control over age effects through the inclusion of a covariate, I presented a comparison of a group with a rather exact age matching but smaller sample size and a group with less accurate age matching and larger size. As mentioned earlier, the age parameter seems to be capable of catching the age effects pretty well. However, the results with the rather exact age matching, while basically showing the same area locations, do also show show a smaller delineation of these areas. This could be due to enhancements through an age effect, but also due to an effect of the smaller sample size, or both.

Another aspect of this work was the genuine presentation of a surface-based analysis with a vertex-wise correction framework, instead of the cluster-based correction offered by the FreeSurfer software. Blair and Karniski (1993) presented the permutation-based method in a different context and I took the general idea and applied it to the context of vertex-based testing. The presented results supplied the first use-cases of this approach and show their utility. It is especially useful for cases, where small brain areas need to be detected, as a cluster-based approach neglects small areas, but the vertex-wise approach does not. For the presented results, the TMTA subanalysis is a good example for a use-case with small but significant areas, that got detected with this approach.

For the Surface-based analyses presented, the processing pipeline included using a FWHM value of 10 mm. This is basically a smoothing algorithm for the calculated cortical surface. Assuming that there should not be big differences between neighbouring vertices, this smoothing algorithm should improve the cortical surface data. Concurrently it makes the detection of coherent and continuous clusters feasible, while possibly neglecting very small areas of atrophy.

Another necessary step to make vertex-wise group comparisons within a surface-based approach is, that all single subjects cortical surface needs to be registered in one average cortical surface. This is a necessary requirement, to make the vertex-wise comparison meaningful, i.e. compare the areas of one individual with the supposedly same areas of another individual. As an average subject, I used the `fsaverage` subject supplied by the FreeSurfer software. I did this to increase the external validity of my results. As an alternative, I could have used a self created average of the GENFI subjects, which might have provided a better mapping of the subjects, but reduced the external validity. Future research should work out positive and negative sides of using respective average subjects.

Bertrand et al. (2018) present data from a C9ORF72 relatives cohort with 41 presymptomatic subjects carrying the mutation and 39 non-mutation carriers. They did find presymptomatic atrophy in C9ORF72 carriers compared to non-mutation carriers in different and diffuse regions of the cortex: one frontal, three inferior temporal and four parietal regions. Also, they found reduced praxis scores in C9ORF72 carriers compared to non-mutation carriers, using the mean praxis score of a shortened version of a french testing battery for five subtests of: imitation of finger configuration (manual dexterity), motor programming and alternate gestures (melokinetic apraxia), imitation of non-representational gestures, pantomime of intransitive gestures, pantomime of transitive gestures. The different mean (SD) praxis scores were 165.2 (3.4) vs. 167.6 (0.6) in a subcohort of subjects younger than 40 years. The mean (SD) age of presymptomatic C9ORF72 mutation carriers was $39.8(11.1)$ years; they estimate the mean (SD) age of onset at 58.9 (4.9) years, taken from a correlation of age and EYO, with EYO being calculated as I did earlier. It is very unfortunate, that they actually used a test battery with seemingly more distinct definitions of apraxia, but did not report other tests, than the comparison of the mean praxis scores. The same critique as discussed earlier applies here. For the assessment of structural anatomical differences in the cortex, they used a univariate model several times for different brain regions and corrected their results using the Benjamini-Hochberg method (Benjamini & Hochberg, 1995). This model is very similar to the one I used for the presymptomatic analysis of masked areas, but there are two major differences: they used a volume measure as dependent variable and brain regions defined by the Desikan-Killiany atlas, supplied by the FreeSurfer software. Note, that they did not use the EYO variable in their model, but instead used the actual subject's age, as I did and as I promote doing, due to the flaws reported about using the EYO variable. As a side note, they found an unsurprising correlation between age and EYO, but do point out that the estimation of disease onset is very individual in C9ORF72 mutation carriers. Important to highlight is, that in comparison to my results, they found different areas of the brain, used a volume based measure, and could not find a correlation of the reduced praxis scores with the detected brain areas underlying presymptomatic atrophy. Thus, my results do not contradict those results, but give a very different perspective and approach to identifying specific substrates of a neuropsychological dysfunction, and looking for atrophy in exactly and only those regions. Bertrand et al. (2018) did, in accordance with my results, not find presymptomatic atrophy in areas, which where used in my presymptomatic analysis. Further, they could not delineate dysfunction related brain areas, but delineate general presymptomatic atrophy areas, explicitly without finding a relationship to their raised praxis scores. Be aware that the definition of apraxia is very different from the one underlying the GENFI data. Future studies could combine the positive aspects of both, the study and especially the apraxia testing by Bertrand et al. (2018), with explicit use of the subtest results, and my presented approach, to further delineate specific dysfunction related areas.

Stamenova et al. (2015) report about the progression of apraxia in a case study of seven patients, each clinically diagnosed with CBDS. They suggest, that there could be different disease progression speeds within patients, as around half of their cases performed worse on their clinical scores used and progressed quicker, with strong worsening of their scores over time compared to the other half. However, all their patients did show apraxia related difficulties over time. Using single-photon emission computed tomography (SPECT), they found uni- to bilateral hypoperfusion in frontal and parietal to generalized regions, from mildly affected to severely affected patients. As this was a case study with, on scientific grounds, low inclusion criteria (patients needed to be clinically diagnosed with CBDS) and hence many unclear influences from confounding variables, the interpretation of results is very limited. However, the study hints at the importance of apraxia within CBDS, a FTLD spectrum disease, and it emphasizes the progression of cortical change from specific to generalized areas. Also, by using an extensive test battery on apraxia deficits, it shows the diversity of apraxia, and why a global measure of apraxia is insufficient. A connection of the affected brain areas and specific apraxia deficits cannot be drawn based on the presented qualitative results.

Rohrer, Rossor, and Warren (2010) investigated apraxia in 16 patients, clinically diagnosed with Progressive Nonfluent Aphasia (PNFA). They measured limb apraxia in a extensive battery and used the global overall score as a global limb apraxia measure. This measure was then used as explanatory variable in a volume based analysis with voxel intensity as dependent variable. This analysis revealed the following result, which was statistically not controlled for multiple comparisons: atrophy in a small area in the left inferior parietal lobe was associated with limb apraxia. In contrast to my results, they hence found a different area associated with apraxia. However, there are several key differences between the studies: they used a volume based method as opposed to the surface based method I used; the cohort was much smaller and also very different as it consisted of 16 clinically diagnosed PNFA subjects, as opposed to genetic FTLD subjects; the apraxia measures differ entirely as they used a global score composed of several subscores opposed to a single score of gesture imitation deficits. A volume based approach could capture subcortical differences, which hence might influence the results, whereas a surface based approach focuses solely on the cortical structure and is thus only influenced by cortical changes. The usage of a global apraxia measure makes it hard to pin down the driving force behind the results, as it is possible that specific apraxia subdomains are the driving force behind the results. The presented region by Rohrer, Rossor, and Warren (2010) is very small and based on results uncorrected for multiple comparisons, yet it claims to be a global limb apraxia correlate. I suspect, also considering the small sample size coupled with small cortical changes, that many other areas, also important for a global apraxia brain correlate, got neglected in the analysis. Due to all those differences between the studies, they both supply different perspectives for our understanding of apraxia in FTLD and leave room for future studies, to clarify the link between parietal and frontal atrophy and specific apraxia subdomains.

The presented approach has its advantages in concentrating only on the cortical thickness, but at the same time, this had its disadvantages. Most importantly, there is a huge amount of interconnections between cortical and subcortical areas, creating brain networks. Tetreault et al. (2020) have used knowledge about average brain networks in healthy subjects, the normative connectome, to interconnect individual atrophy areas of clinical Alzheimer's subjects using a new method they called atrophy network mapping. With this method, they claim individually varying atrophy to be connected to disease specific brain networks. Tetreault et al. (2020) extended their results by correlating them vertex-wise with memory scores and delineating the significantly correlated areas. Further, they compared delusional with non-delusional subjects by performing a vertex-wise t-test and delineating the significant areas. The areas they delineated for the described functions did show a good overlap with brain lesions from subjects with related deficits. However, there are several limitations to their study. They do claim to identify disease specific brain networks, based on averaged cortical areas they identified with their method: using a cortical average as I did in the presymptomatic analysis. The crucial difference is, that they used the normative connectome to delineate all cortical areas connected to atrophied areas in the cortex. So the final result they present are averaged cortical areas, based on the normative connectome identifying atrophy connected areas, based on their underlying definition of atrophy as cortical thickness of below two standard deviations from the average cortical thickness, compared vertex-wise. Still, it is very interesting to see how the incorporation of a normative connectome can enhance results of small cortical atrophy areas. Going further, it would be interesting to see how my results get altered, when incorporating interconnections of a connectome. As mentioned earlier, some of the areas I delineated show high myelination when compared with the results by Glasser et al. (2016) and could thus reveal a general underlying praxis network, if combined with the approach by Tetreault et al. (2020).

The way of recruiting participants for the GENFI study has many advantages, as a control group with a similar sociocultural background, demographics and genetics got naturally selected through the data acquisition process. However, there were also negative side effects of this recruitment, as there are family relations in the dataset with unknown effects: for example, there is one family in the dataset, that provides more than 30 subjects for the study, on the other hand 123 families provide only one member. The research sites also have a similar unbalanced structure: There is one that provides more than 100 subjects, while four provide less than five subjects. To account for this, I included random factors for family and site in the model for the presymptomatic analysis. However, this inclusion of random factors cannot fully account for potentially systematic biases. Moreover, estimations of independent variables based on family relationships like EYO must be severely affected by the unbalanced representation of familiy members in the whole sample. Especially the calculation of the EYO variable for subjects with only one family member (themselves) is obscure, as it uses anamnestic data not included in the GENFI dataset.

In summary, after establishing my methods by delineating the general atrophy pattern of FTLD, I supplied a new perspective on apraxia, investigating its cortical correlates in a genetic FTLD cohort. I was able to delineate specific cortical areas, related to gesture imitation deficits of genetic FTLD subjects. In addition, I could delineate TMT performance related areas. Further, the provided methodology to correct for multiple comparisons allows for the detection of significant yet small cortical areas. In my presymptomatic analysis, I was able to highlight difficulties and limitations of important statistical methods capturing early structural brain changes, while showing the effect of age and the effect of disease progression in symptomatic subjects within specific areas of the cortex. I aimed to supply innovative views and techniques, by combining established methods in a novel approach. In the end, I hope to contribute with my work not only to apraxia research and FTLD research, but also to research of structural brain changes and its neuropsychological correlates in general.

8 Summary

8.1 Abstract

Neuropsychological deficits, as apraxia, play an important role for neurodegenerative dementias, including frontotemporal dementia (FTD). In this work I delineate cortical atrophy areas relevant for apraxia and Trail Making Test (TMT) performance within a genetic frontotemporal lobar degeneration (FTLD) cohort. Using 472 subjects from the GENFI study, I demonstrate a hypothesis-driven approach for the identification of cortical atrophy areas related to specific neuropsychological dysfunctions: through this I identified atrophy in the premotor cortex, inferior frontal and frontal opercular areas to be specifically related to apraxia; and areas in the frontal cortex to be related to TMT performance. In addition, I present an alternative way of correcting my results within a surface-based approach for multiple comparisons using a permutation based maximum statistic, facilitating the identification of small areas. Lastly, I present an analysis about presymptomatic effects and discuss its limitations and difficulties: I could not delineate presymptomatic cortical degeneration in the previously identified areas. As a result, this work contributes to research about apraxia, genetic FTLD and general relations of neuropsychological functions with cortical areas.

8.2 Zusammenfassung

Neuropsychologische Defizite, wie Apraxie, spielen für neurodegenerative Erkrankungen, speziell auch für Frontotemporale Demenzen (FTD), eine wichtige Rolle. In dieser Arbeit zeige ich abgegrenzte atrophische Areale der Hirnrinde, die für Apraxie und die Leistung im Trail Making Test (TMT) innerhalb einer Kohorte von familiärer FTD relevant sind. Mit einer Stichprobe von 472 Probanden aus der GENFI Studie, demonstriere ich einen Hypothesen getriebenen Ansatz um kortikale atrophische Areale zu identifizieren, die mit spezifischen neuropsychologischen Defiziten zusammenhängen: dadurch konnte ich atrophische Areale im prämotorischen Cortex, Inferioren Frontalcortex und der Pars opercularis des Inferioren Frontalcortex identifizieren, die speziell mit Apraxie zusammenhängen; sowie Areale im Frontalcortex, die mit der Leistung im TMT zusammenhängen. Zusätzlich präsentiere ich einen alternativen Weg, um das Problem multipler Vergleiche von kortikalen Strukturen zu lösen, indem eine permutationsbasierte Maximum Statistik berechnet wird, welche die Identifikation kleiner Areale ermöglicht. Zuletzt zeige ich eine Analyse über präsymptomatische Effekte und diskutiere deren Schwierigkeiten und Limitationen: in den zuvor identifizierten Arealen konnte ich keine präsymptomatische Degeneration feststellen. Schließlich trägt diese Arbeit zur wissenschaftlichen Forschung im Bereich der Apraxie, familiärer FTD sowie genereller Zusammenhänge von neuropsychologischen Defiziten und Hirnrindenbereichen bei.

9 References

Bateman, R. J., Xiong, C., Benzinger, T. L., Fagan, A. M., Goate, A., Fox, N. C., ... others (2012). Clinical and biomarker changes in dominantly inherited alzheimer's disease. *N Engl J Med*, *367*, 795–804.

Behrens, T. E., & Sporns, O. (2012). Human connectomics. *Current opinion in neurobiology*, *22*(1), 144–153.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Bertrand, A., Wen, J., Rinaldi, D., Houot, M., Sayah, S., Camuzat, A., ... others (2018). Early cognitive, structural, and microstructural changes in presymptomatic c9orf72 carriers younger than 40 years. *JAMA neurology*, *75*(2), 236–245.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518–524.

Boeve, B. F., Boylan, K. B., Graff-Radford, N. R., DeJesus-Hernandez, M., Knopman, D. S., Pedraza, O., ... others (2012). Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the ggggcc repeat expansion in C9ORF72. *Brain*, *135*(3), 765–783.

Borroni, B., Agosti, C., Premi, E., Cerini, C., Cosseddu, M., Paghera, B., ... Padovani, A. (2010). The FTLD-modified clinical dementia rating scale is a reliable tool for defining disease severity in frontotemporal lobar degeneration: evidence from a brain spect study. *European journal of neurology*, *17*(5), 703–707.

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the trail making test. *Nature protocols*, *1*(5), 2277.

Buxbaum, L. J., Shapiro, A. D., & Coslett, H. B. (2014). Critical brain regions for tool-related and imitative actions: a componential analysis. *Brain*, *137*(7), 1971–1985.

Cash, D. M., Bocchetta, M., Thomas, D. L., Dick, K. M., van Swieten, J. C., Borroni, B., ... others (2018). Patterns of gray matter atrophy in genetic frontotemporal dementia: results from the genfi study. *Neurobiology of aging*, *62*, 191–196.

Chow, T. W., Miller, B. L., Hayashi, V. N., & Geschwind, D. H. (1999). Inheritance of frontotemporal dementia. *Archives of neurology*, *56*(7), 817–822.

Dale, A., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179 - 194.

Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *Journal of cognitive neuroscience*, *5*(2), 162–176.

Demakis, G. J. (2004). Frontal lobe damage and tests of executive processing: a meta-analysis of the category test, stroop test, and trail-making test. *Journal of clinical and experimental neuropsychology*, *26*(3), 441–450.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968 - 980.

Englund, B., Brun, A., Gustafson, L., Passant, U., Mann, D., Neary, D., & Snowden, J. (1994). Clinical and neuropathological criteria for frontotemporal dementia. *J Neurol Neurosurg Psychiatry*, *57*(4), 416–8.

Fischl, B. (2012). Freesurfer. *Neuroimage*, *62*(2), 774–781.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(20), 11050-11055.

Fischl, B., Liu, A., & Dale, A. M. (2001, Jan). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Medical Imaging*, *20*(1), 70-80.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341-355.

Fischl, B., Salat, D. H., van der Kouwe, A. J., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23*(Supplement 1), S69 - S84. (Mathematics in Brain Imaging)

Fischl, B., Sereno, M. I., & Dale, A. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195 - 207.

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*(4), 272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., … Dale, A. M. (2004). Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, *14*(1), 11-22.

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Müller, K., …

Weishaupt, J. H. (2015). Haploinsufficiency of TBK1 causes familial als and fronto-temporal dementia. *Nature neuroscience*, *18*(5), 631.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … others (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178.

Goldenberg, G. (2014). Apraxia – the cognitive side of motor control. *Cortex*, *57*, 270 - 274.

Goldenberg, G., Hermsdörfer, J., Glindemann, R., Rorden, C., & Karnath, H.-O. (2007). Pantomime of tool use depends on integrity of left inferior frontal cortex. *Cerebral Cortex*, *17*(12), 2769–2776.

Goldenberg, G., & Karnath, H.-O. (2006). The neural basis of imitation is body part specific. *Journal of Neuroscience*, *26*(23), 6282–6287.

Goldenberg, G., & Strauss, S. (2002). Hemisphere asymmetries for imitation of novel gestures. *Neurology*, *59*(6), 893–897.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., … Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180-194.

Hodges, J. R., Davies, R. R., Xuereb, J. H., Casey, B., Broe, M., Bak, T. H., … Halliday, G. M. (2004). Clinicopathological correlates in frontotemporal dementia. *Annals of neurology*, *56*(3), 399–406.

Hyatt, C. J., Haney-Caron, E., & Stevens, M. C. (2012). Cortical thickness and folding deficits in conduct-disordered adolescents. *Biological Psychiatry*, *72*(3), 207–214.

Johnen, A., Brandstetter, L., Kärgel, C., Wiendl, H., Lohmann, H., & Duning, T. (2016). Shared neural correlates of limb apraxia in early stages of alzheimer's dementia and behavioural variant frontotemporal dementia. *Cortex*, *84*, 1–14.

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., … Dale, A. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*(2), 436 - 443.

Kertesz, A., Hillis, A., & Munoz, D. G. (2003). Frontotemporal degeneration, pick's disease, pick complex, and ravel. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *54*(S5 5), S1–S2.

Kertesz, A., McMonagle, P., Blair, M., Davidson, W., & Munoz, D. G. (2005). The evolution and pathology of frontotemporal dementia. *Brain*, *128*(9), 1996–2005.

Kopp, B. (2011). Neuropsychologists must keep their eyes on the reliability of difference measures. *Journal of the International Neuropsychological Society*, *17*(3), 562–563.

Kopp, B., Rösser, N., Tabeling, S., Stürenburg, H. J., de Haan, B., Karnath, H.-O., & Wessel, K. (2015). Errors on the trail making test are associated with right hemispheric frontal lobe damage in stroke patients. *Behavioural Neurology*, *2015*.

Kuperberg, G. R., Broome, M., McGuire, P. K., David, A. S., Eddy, M., Ozawa, F., ... Fischl, B. (2003). Regionally localized thinning of the cerebral cortex in Schizophrenia. *Archives of General Psychiatry*, *60*, 878-888.

Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: Dicom to nifti conversion. *Journal of neuroscience methods*, *264*, 47–56.

Lindroth, H., Nair, V. A., Stanfield, C., Casey, C., Mohanty, R., Wayer, D., ... Sanders, R. D. (2019). Examining the identification of age-related atrophy between t1 and t1+ t2-flair cortical thickness measurements. *Scientific reports*, *9*(1), 1–11.

Mahoney, C. J., Beck, J., Rohrer, J. D., Lashley, T., Mok, K., Shakespeare, T., ... others (2012). Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: clinical, neuroanatomical and neuropathological features. *Brain*, *135*(3), 736–750.

Neary, D., Snowden, J., & Mann, D. (2005). Frontotemporal dementia. *The Lancet Neurology*, *4*(11), 771–780.

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., ... others (1998). Frontotemporal lobar degeneration a consensus on clinical diagnostic criteria. *Neurology*, *51*(6), 1546–1554.

Papma, J. M., Jiskoot, L. C., Panman, J. L., Dopper, E. G., den Heijer, T., Kaat, L. D., ... others (2017). Cognition and gray and white matter characteristics of presymptomatic c9orf72 repeat expansion. *Neurology*, 10–1212.

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Ratnavalli, E., Brayne, C., Dawson, K., & Hodges, J. R. (2002). The prevalence of frontotemporal dementia. *Neurology*, *58*(11), 1615–1621.

Reitan, R. M., & Wolfson, D. (1995). Category test and trail making test as measures of frontal lobe functions. *The Clinical Neuropsychologist*, *9*(1), 50–56.

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, *53*(4), 1181–1196.

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*(4), 1402–1418.

Rohrer, J., Guerreiro, R., Vandrovcova, J., Uphill, J., Reiman, D., Beck, J., … others (2009). The heritability and genetics of frontotemporal lobar degeneration. *Neurology*, *73*(18), 1451–1456.

Rohrer, J., Warren, J., Fox, N., & Rossor, M. (2013). Presymptomatic studies in genetic frontotemporal dementia. *Revue neurologique*, *169*(10), 820–824.

Rohrer, J. D., Nicholas, J. M., Cash, D. M., van Swieten, J., Dopper, E., Jiskoot, L., … Rossor, M. N. (2015). Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the genetic frontotemporal dementia initiative (GENFI) study: a cross-sectional analysis. *The Lancet Neurology*, *14*(3), 253–262.

Rohrer, J. D., Ridgway, G. R., Modat, M., Ourselin, S., Mead, S., Fox, N. C., … Warren, J. D. (2010). Distinct profiles of brain atrophy in frontotemporal lobar degeneration caused by progranulin and tau mutations. *Neuroimage*, *53*(3), 1070–1076.

Rohrer, J. D., Rossor, M. N., & Warren, J. D. (2010). Apraxia in progressive nonfluent aphasia. *Journal of neurology*, *257*(4), 569–574.

Rohrer, J. D., & Warren, J. D. (2011). Phenotypic signatures of genetic frontotemporal dementia. *Current opinion in neurology*, *24*(6), 542–549.

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., … Fischl, B. (2002). Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology*, *58*(5), 695-701.

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., … Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral cortex*, *14*(7), 721–730.

Seelaar, H., Rohrer, J. D., Pijnenburg, Y. A., Fox, N. C., & Van Swieten, J. C. (2010). Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp–2010.

Segonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060 - 1075.

Segonne, F., Pacheco, J., & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Trans Med Imaging*, *26*, 518-529.

Sled, J., Zijdenbos, A., & Evans, A. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*, *17*, 87-97.

Sporns, O. (2013). The human connectome: origins and challenges. *Neuroimage*, *80*, 53–61.

Stamenova, V., Roy, E. A., Szilagyi, G., Honjo, K., Black, S. E., & Masellis, M. (2015). Progression of limb apraxia in corticobasal syndrome: Neuropychological and functional neuroimaging report of a case series. *Neurocase*, *21*(5), 642–659.

Stuss, D. T., Bisschop, S. M., Alexander, M. P., Levine, B., Katz, D., & Izukawa, D. (2001). The trail making test: a study in focal lesion patients. *Psychological assessment*, *13*(2), 230.

Tetreault, A. M., Phan, T., Orlando, D., Lyu, I., Kang, H., Landman, B., . . . Initiative, A. D. N. (2020). Network localization of clinical, cognitive, and neuropsychiatric symptoms in alzheimer's disease. *Brain*, *143*(4), 1249–1260.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., . . . others (2013). The wu-minn human connectome project: an overview. *Neuroimage*, *80*, 62–79.

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., . . . others (2012). The human connectome project: a data acquisition perspective. *Neuroimage*, *62*(4), 2222–2231.

Varjacic, A., Mantini, D., Demeyere, N., & Gillebert, C. R. (2018). Neural signatures of trail making test performance: Evidence from lesion-mapping and neuroimaging studies. *Neuropsychologia*, *115*, 78–87.

Whitwell, J., Jack, C., Boeve, B., Senjem, M., Baker, M., Rademakers, R., . . . others (2009). Voxel-based morphometry patterns of atrophy in FTLD with mutations in MAPT or PGRN. *Neurology*, *72*(9), 813–820.

Whitwell, J. L., Weigand, S. D., Boeve, B. F., Senjem, M. L., Gunter, J. L., DeJesus-Hernandez, M., . . . Josephs, K. A. (2012). Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. *Brain*, *135*(3), 794-806.

10 Authorship – Eigenanteil

10.1 Statement of Authorship

This work has been conducted at the Hertie Institute for Clinical Brain Research under the supervision of Matthis Synofzik and Marc Himmelbach. The GENFI data has been provided to me by David M. Cash from the University College London, through contact by Matthis Synofzik and after an application process and approval through GENFI. The described processing of the data as well as the statistical analysis has been performed by me, with guidance by Marc Himmelbach. I, the Author, confirm that the work presented here has been performed and written solely by myself except where explicitly identified as the contrary. External thoughts and ideas as well as sources and other aids are made recognizable as such without exception.

This work has been continuously composed in first-person singular, note however that large parts of the work originated from team work under the above mentioned supervision; and that the data has been provided by the GENFI consortium.

10.2 Erklärung zum Eigenanteil

Die Arbeit wurde im Hertie-Institut für klinische Hirnforschung unter Betreuung von Matthis Synofzik und Marc Himmelbach durchgeführt. Der GENFI Datensatz wurde mir von David M. Cash vom University College London zur Verfügung gestellt, über Kontakherstellung durch Matthis Synofzik und nach Antragstellung und dessen Genehmigung von GENFI. Die Konzeption der Studie erfolgte in Zusammenarbeit mit Marc Himmelbach und Matthis Synofzik. Die in der Arbeit beschriebene Datensatzaufbereitung sowie die statistische Auswertung erfolgte nach Beratung mit Marc Himmelbach durch mich. Ich versichere, das Manuskript selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht.

Die vorliegende Arbeit ist kontinuierlich in der 1. Person Singular verfasst, dennoch bitte Ich zu berücksichten, dass weite Teile der Arbeit im Team unter obig genannter Supervision enstanden sind; sowie der Datensatz vom GENFI Konsortium bereitgestellt wurde.

Tübingen, October 26, 2021

Appendix A FreeSurfer citation

The details on how FreeSurfer works and how scientific methods are incorporated are given in the following citation, taken from the FreeSurfer Website[11]:

> Cortical reconstruction and volumetric segmentation was performed with the FreeSurfer image analysis suite, which is documented and freely available for download online (`http://surfer.nmr.mgh.harvard.edu/`). The technical details of these procedures are described in prior publications (Dale, Fischl, & Sereno, 1999; Dale & Sereno, 1993; Fischl & Dale, 2000; Fischl, Liu, & Dale, 2001; Fischl et al., 2002; Fischl, Salat, et al., 2004; Fischl, Sereno, & Dale, 1999; Fischl, Sereno, Tootell, & Dale, 1999; Fischl, van der Kouwe, et al., 2004; Han et al., 2006; Jovicich et al., 2006; Segonne et al., 2004; Reuter, Rosas, & Fischl, 2010; Reuter, Schmansky, Rosas, & Fischl, 2012). Briefly, this processing includes motion correction and averaging (Reuter et al., 2010) of multiple volumetric T1 weighted images (when more than one is available), removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Segonne et al., 2004), automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures (including hippocampus, amygdala, caudate, putamen, ventricles) (Fischl et al., 2002; Fischl, Salat, et al., 2004) intensity normalization (Sled, Zijdenbos, & Evans, 1998), tessellation of the gray matter white matter boundary, automated topology correction (Fischl et al., 2001; Segonne, Pacheco, & Fischl, 2007), and surface deformation following intensity gradients to optimally place the gray/white and gray/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale et al., 1999; Dale & Sereno, 1993; Fischl & Dale, 2000). Once the cortical models are complete, a number of deformable procedures can be performed for further data processing and analysis including surface inflation (Fischl, Sereno, & Dale, 1999), registration to a spherical atlas which is based on individual cortical folding patterns to match cortical geometry across subjects (Fischl, Sereno, Tootell, & Dale, 1999), parcellation of the cerebral cortex into units with respect to gyral and sulcal structure (Desikan et al., 2006; Fischl, van der Kouwe, et al., 2004), and creation of a variety of surface based data including maps of

---

[11] `https://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferMethodsCitation`

curvature and sulcal depth. This method uses both intensity and continuity information from the entire three dimensional MR volume in segmentation and deformation procedures to produce representations of cortical thickness, calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface (Fischl & Dale, 2000). The maps are created using spatial intensity gradients across tissue classes and are therefore not simply reliant on absolute signal intensity. The maps produced are not restricted to the voxel resolution of the original data thus are capable of detecting submillimeter differences between groups. Procedures for the measurement of cortical thickness have been validated against histological analysis (Rosas et al., 2002) and manual measurements (Kuperberg et al., 2003; Salat et al., 2004). FreeSurfer morphometric procedures have been demonstrated to show good test-retest reliability across scanner manufacturers and across field strengths (Han et al., 2006; Reuter et al., 2012).

Appendix B R code for Wald Test with permutation based test distribution

Here I present the actual R code, that I used for the calculation of the permutation based maximum statistic that I used to extend the Wald test with a correction for multiple comparisons. Box 1 shows the code for the Wald Test with a permutation based test distribution. A usage example with simulated data is given in Box 2.

```
library(gtools)
library(doMC) # for parallelization
registerDoMC(floor(detectCores()/1.2))

# Returns a list with as many matrices as parameters of the model.
# Every matrix contains the dependent variable as one dimension
# and the permutation number in the other dimension: nY X nPar.
# Every value is a t-value of the parameter
# for the permutation and the dependen variable.
mlmPermDist <- function(formula, data, permcol, n=10000)
{
  mlmobj1 <- lm(formula, data)
  coefnames <- dimnames(coef(mlmobj1))[[1]]
  empdistpar <- foreach (i=1:n) %dopar% {
    data[,permcol] <- permute(data[,permcol]) # requires gtools

    mlmobj <- lm(formula, data)

    sum.mlmobj <- summary(mlmobj)

    mlmpar <- list()
    for (coe in coefnames) {
      mlmpar[[coe]] <- sapply(coef(sum.mlmobj), "[[", coe,3)
    }
    return(mlmpar)
  }
  retval <- list()
  for(coe in coefnames) {
    retval[[coe]] <- sapply(empdistpar, "[[", coe)
  }
  retval
}

# If yvar needs to be a matrix, use this function
# and feed the formula and data only the regressor variables.
mlmPermDist2 <- function(yvar, formula, data, permcol, n=10000)
{
  x <- model.matrix(formula,data)
  mlmobj1 <- lm(yvar ~ 0 + x)
  coefnames <- dimnames(coef(mlmobj1))[[1]]
  empdistpar <- foreach (i=1:n) %dopar% {
    data[,permcol] <- permute(data[,permcol]) # requires gtools

    # appends an x to the coefficients
    x <- model.matrix(formula,data)
    mlmobj <- lm(yvar ~ 0 + x)

    sum.mlmobj <- summary(mlmobj)

    mlmpar <- list()
    for (coe in coefnames) {
      mlmpar[[coe]] <- sapply(coef(sum.mlmobj), "[[", coe,3)
    }
    return(mlmpar)
  }
  retval <- list()
  for(coe in coefnames) {
    retval[[coe]] <- sapply(empdistpar, "[[", coe)
  }
  names(retval) <- dimnames(model.matrix(formula, data))[[2]]
  retval
}

extr <- function(x)
{
x.max <- max(x)
x.min <- min(x)
if (abs(x.max) > abs(x.min))
  { return(x.max) }
else if (abs(x.max) < abs(x.min))
  { return(x.min) }
else # if abs values are equal: randomize
  { return(sample(c(-1,1),1) * x.max) }
}

absmax <- function(x)
{
x.max <- max(x)
x.min <- min(x)
if (abs(x.max) >= abs(x.min))
  { return(x.max) }
else
  { return(abs(x.min)) }
}

absMaxEcdf <- function(permres, parname)
{
  extTpar <- apply(permres[[parname]], 2, absmax)
  ecdf(extTpar)
}

extrEcdf <- function(permres, parname)
{
  extTpar <- apply(permres[[parname]], 2, extr)
  ecdf(extTpar)
}

minEcdf <- function(permres, parname)
{
  minTpar <- apply(permres[[parname]], 2, min)
  ecdf(minTpar)
}

maxEcdf <- function(permres, parname)
{
  maxTpar <- apply(permres[[parname]], 2, max)
  ecdf(maxTpar)
}
```

Box 1: R code for Wald Test with permutation based test distribution.

```
# mlm permtest functions
source("mlm-permtest.R")

# Simulation variables
nsubj = 60 # number of subjects
nvertex = 100 # number of vertices
nperm = 5000 # number of permutations for the test

# Assume a small brain area, consisting of ’nvertex’ vertices:
# for ’nsubj’ subjects
y <- matrix(rnorm(nvertex*nsubj,2.7,0.3),
            nrow=nsubj,
            byrow=TRUE)

# 30% of vertices are atrophied in the second half of subjects
y[(nsubj/2):nsubj,(nvertex/2):(nvertex/2+(nvertex*0.3-1))] <- rnorm((
   nvertex*0.1),2.0,0.3)

# other model variables
# theta1
affected <- factor(c(rep(0,(nsubj/2)), rep(1,(nsubj/2))), levels=c
   (0,1))
# theta2
gender <- factor(sample(c(0,1),nsubj,TRUE), levels=c(0,1), labels=c("
   male", "female"))
# beta1
age <- round(rnorm((nsubj),60,20),0)
# beta2
edu <- round(runif(nsubj,9,20),0)

# multivatiate linear model
m1 <- lm(y~affected+gender+age+edu)
sum.m1 <- summary(m1)

# call the permutation function that does the actual work
m1.x <- mlmPermDist2(y, ~affected+gender+age+edu,
  data.frame(affected=affected,gender=gender,age=age,edu=edu),
  "affected", nperm)

# take most extreme value of all vertices for theta1 parameter
# for all permutations
# --> results in a list of 'nperm' values
# --> the maximum T statistic
theta1ecdf <- extrEcdf(m1.x, "affected1")

# take the t-values for all vertices of the acutal model with the
   original data
theta1m1 <- sapply(coef(sum.m1), "[[", "affected1","t value")

# extract the p value: value of cumulative distribution function
# for the acutal t-values
# under the distribution of maximum T statistic distribution
theta1m1pval <- theta1ecdf(theta1m1)
theta1m1pval
log(theta1m1pval, base=10)
```

```
# note: for a one-tailed test, use the min or max version of this
   function
theta1ecdfmin <- minEcdf(m1.x, "affected1")
theta1m1pvalmin <- theta1ecdfmin(theta1m1)
```

Box 2: Simple simulated example code.