Universität Regensburg

# Mapping abnormality: Towards evaluating MRI scans for individualized diagnosis and validation of clinical course/progression in motor neuron disease

Inaugural-Dissertation zur Erlangung der Doktorwürde der Fakultät für Humanwissenschaften der Universität Regensburg

vorgelegt von Marlene Tahedl aus
Pfarrkirchen
2021

Die Arbeit entstand in gemeinsamer Betreuung durch die Fakultät für Psychologie, Pädagogik und Sportwissenschaft der Universität Regensburg und der Fakultät für Medizin am Universitätsklinikum Regensburg

Regensburg 2021

# 1 Preface

**Gutachter/in (Betreuer/in):** Prof. Dr. Mark W. Greenlee
**Gutachter/in:** Prof. Dr. Ulrich Bogdahn

# 1 Preface

The aim of the thesis entitled „*Mapping abnormality: Towards evaluating MRI scans for individualized diagnosis and validation of clinical course in motor neuron disease*" was to develop and validate a diagnostic tool to evaluate an individual patients' structural MRI scans, to infer on individual patterns of systemic or local atrophy. Of particular interest was the validation of this tool in real-world clinical populations, which were chosen for the different sub-projects composing this thesis – encompassing amyotrophic lateral sclerosis, primary lateral sclerosis, poliomyelitis survivors as well as the frontotemporal dementia spectrum. The four sub-projects of the thesis are summarized in the Abstract (page 8). A list of abbreviations of this thesis is found on page 11). The references of all four subprojects are merged in one bibliography at the end (starting on page 83).

All four studies were accepted at peer-reviewed journals (impact factor > 1.5). Two of the four studies were successfully published (subprojects 1 and 2), two are currently in press (subprojects 3 and 49 as of July 2021). Chapter 2 includes the pre-print versions of published projects. The projects which are "in press" are represented in their most recently revised and accepted versions. The manuscripts of all studies were consistently formatted according to the guidelines of American Psychological Association 7th Edition. Numbers of figures, tables, and equations were adapted. No other changes were made to the manuscripts. Note that some journals required British English spelling, while the remainder of this work is written in American English spelling.

Large parts of this thesis (encompassing subprojects 2–4) were conducted in cooperation with the Computational Neuroimaging Group of Professor Peter Bede at the Biomedical Sciences Institute at Trinity College Dublin. Co-author contributions are listed at the end of each manuscript in chapter 2.

For subproject 3 (in press at the Journal *Neurobiology of Aging*), a spin-off data article was published at *Mendeley Data* entitled "*The evolution of cortical integrity in motor neuron diseases: longitudinal cortical thickness profiles in ALS, PLS and poliomyelitis*" (doi *10.17632/3pnp4hkhnw.1*, see also section Contributions for further details) but is not explicitly discussed in this thesis since it primarily serves as a data source for researchers aiming to replicate the results of subproject 3, or to add these data into a meta-analyses.

In the following, the references for all four subprojects are listed:

(1) **Tahedl, M.** (2020). Towards individualized cortical thickness assessment for clinical routine. *Journal of translational medicine*, *18*(1), 1-12.

(2) **Tahedl, M.**, Chipika, R. H., Lope, J., Shing, S. L. H., Hardiman, O., & Bede, P. (2021). Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use. *Journal of Neurology*, *268*(5), 1913-1926.

*In press:*

*(3)* **Tahedl, M.***, Li Hi Shing, S.*, Finegan, E.*, Chipika, R. H., Lope, J., Hardiman, O., & Bede, P. (2021). Propagation patterns in motor neuron diseases: individual and phenotype-associated disease-burden trajectories across the UMN-LMN spectrum of MNDs. *Neurobiology of Aging, in press.*

(4) McKenna, M. C.*, **Tahedl, M.***, Lope, J., Chipika, R. H., Li Hi Shing, S., Doherty, M. A., Hengeveld, J. C., Vajda, A., McLaughlin, R. L., Hardiman, O., Hutchinson, S., & Bede, P. (2021). Mapping cortical disease-burden at individual-level in frontotemporal dementia: implications for clinical care and pharmacological trials. *Brain Imaging and Behaviour*, *in press*.

** Shared first co-authors*

# 2 Credit where credit is due.

A 3.5 year-long journey is coming to an end. Although I prefer saying, it is coming to a milestone, since I do not whatsoever intend to stop this work here but to continue improving it, expanding it, and carrying it into the future to make it eventually accessible to clinical translation. But the scientific contents of this work will be discussed in every detail in the remainder of this thesis; this is the place where to thank all those who have made this work possible. It is my genuine belief that anybody I've met in the past 3.5 years (and even all those before) contributed in some way to what I can offer herein, given even the smallest encounter shaped my mind, which eventually produced this work. However, there are those whom I owe especial gratitude, and I'm honored for having been able to work with them.

Above all, I'd like to thank my main supervisor, Prof. Mark W. Greenlee, for accepting me as a PhD student. He was the one who first raised my interest in neuroscience in during my Master's, and he encouraged and supported me to initiate a project, although I had very little methodological knowledge back then, but was chasing an idea. Thank you, Prof. Greenlee, for believing in me. I hope I have not disappointed you.

Another very important contributor to my early development was Prof. Robert Weissert, who took a chance on me as a young student and offered his clinical expertise to initiate my first clinical neuroscience project. Although again, I did not have much technological background to convince you with, you did let me convince you with my motivation. You crucially shaped my confidence by the confidence you put in me, Prof. Weissert. Thank you for trusting and appreciating in these values. Aspiring researchers could go nowhere without established scientists like you who believe in them.

It wouldn't have been without the great methodological training of Prof. Jens V. Schwarzbach and his group, especially Dr. Seth M. Levine, that I could pull through with a neuroimaging project. Luckily, I had no idea at the beginning that doing neuroscience would require me to know programming :D I think I would have fled right off. That was five years ago. But you, Jens, trapped me so "softly" that I was writing programs before I knew it. You awakened the child in me and made me love play around with code and to make it do things for me automatically. Surely, this took longer than manual editing in the beginning (and it was a long beginning) but eventually it is now buying me time – time to fill with a little life, time to get inspired and for creativity. Thank you, Jens, for giving me *life time*, while the computer is doing my job. This is probably the greatest of all gifts.

I also owe great gratitude to Prof. Schwarzbach and his lab for hosting me these past years, for training me in scientific writing, presenting, but most of all *thinking*. Thank you for all the inspiring and critical discussions we had, thank you for never wrongly praising my work but continuously asking me to improve it (but most importantly: showing me ways to achieve that). Thank you for teaching me collaboration and autonomy. I know I have not always made it easy for you. Thank you for your patience.

I would also like to express my great appreciation to my secondary supervisor Prof. Ulrich Bogdahn, who crucially shaped the content of this thesis. For many years he has been observing my work and recognized (or shall I say anticipated?) clinical applications for them which I could never even conceive of. It was him who encouraged me to develop the herein presented method further, and who provided his much-valued and constant clinical and scientific support for doing so. Most importantly, he enabled me to carry on with validation by establishing a collaboration with one of his colleagues at Trinity College Dublin, who will be the next to thank. But before that, I'd like to say thank you, Uli, from the bottom of my heart. You were there when I needed you most, not only as a scientific advisor but also as a friend. You reminded me that there was a life besides work, which I was close to forgetting...

Now my thanks turn to Prof. Peter Bede from Trinity College Dublin, who offered me the invaluable opportunity to work with his resources to validate the herein proposed tool. But he did more

than that. He never got tired of discussing with me – even at uncivilized hours –, of shaping and improving the manuscripts and of broadening my horizons for new applications of the tool (and I believe it was you, Peter, who first conceived of the tool's nickname, "mosaic-approach", or did you?). Thank you for your most generous resources and time, Peter. Without your support I could never have put this tool to test and hence could never have presented this thesis in such a structured manner. I hope this was not the last time I had the honor to work with you, and I'm looking forward to more inspiring collaborations in the future. The scientific output I can offer herein was crucially enabled by you, Peter, and I can only begin to describe how grateful I am for this all.

This work wouldn't have been possible without the financial support by two funding sources: On the one hand, the Universität Bayern e.V., who honored me with a PhD grant for the first two years of this work that allowed me great independence and time to develop and improve my own skills. The rest of my PhD time was funded by the Deutsche Multiple Sklerose Gesellschaft (DMSG). Their grant not only allowed me to continue to autonomously work on the project, but also greatly enhanced my own belief in myself and in the project, which was an invaluable backup during the many struggles I faced. Thank you, DMSG, for taking the chance on me and believing in a young researcher who did not have much previous publishing work to offer, but only a high motivation and the possession for an idea. I hope my work will encourage you to support many more aspiring researchers.

The final thanks are reserved for my dear ones, my family and friends. Mom, Dad, Jojo, Giga, and lately also Nick: You were my safe haven whenever I came home on the weekends, back to Pfarrkirchen. You allowed me to rest, to refreshen my mind after some workful weeks. You made coming home truly "coming home" – something to look forward to during the week whenever things got busy (which was the case quite regularly). You took so much work and stress from me, more than you could possibly know. I know I was not always easy to deal with, but you never let this decline your love and support for me. You were there for me even when I did not deserve it. I'm not good to say this with words, so let me put it here in black & white: I love you, more than you know.

Similarly, to all my wonderful friends, who – to reasons I cannot always fully understand in retrospect – remained friends during all these years. I know that there were times when I was not a good friend, either didn't reply or replied late, sometimes stressed. Nevertheless, Markus, Babs, Diana, Micha's Supertruppe, Ellie, Feli, Anis, Flo, Elmar, Philipp, Sepp, Vladi, Martin, Christian, Verena, Laura, Maxi[6] & Nicola[9]: You made and make life worth living; it was you whom I had the best times with – and it was you who made those even better. You listened, you cared, and you provided your support during the bad times – and you know we've had them – and you laughed with me and rejoiced with me during the good times – and you know we've had them. You taught me probably the most important lesson of these past years, you showed me what truly counts in life: caring for your loved ones. And you cared for me. I can't begin to say how deeply, deeply grateful I am and how humble I am to have you as friends. So let me just start by saying "Thank you".

This work wouldn't have been possible without all the people I listed here (and also not without the backing of so many more I do not have the space to thank here – supporters from the university, clinic, from sports, … – I hope you know how much you I owe you, and I'll be happy to buy you a beer if you feel neglected). In light of this, it did not feel right for me to use the pronoun "I" a lot throughout this work, such that I chose to say "we" in most cases. This is another fundamental lesson I've learned during the past years: you can only win together. And you are my together.

Thank you!

Yours, Marlene.

2 Credit where credit is due.

*Table of contents*

# 3 Abstract

The present dissertation proposes a standardized assessment tool for meaningful interpretation of individual patients' MRI data which may be used to infer on atrophy (or hypertrophy). Various MRI-based measures have been subject to a wide range of neuroscience studies and numerous associations between such measures and the diagnosis / prognosis with various pathological conditions have been identified. However, surprisingly, accurate individualized MRI-based assessment is still difficult, and most studies utilize group comparisons in order to describe group-based brain structural or functional differences. One of the reasons for this might be the lack of a standardized system which allows to evaluate individual MRI data. With this dissertation, it was aimed at passing this limit by describing a method to rate single subjects' T1-weighted MRI data. The idea is straightforward, in that it rates single patient data with respect to a matched control population and uses nonparametric statistics to identify regions of unexpected thin (thick) cortical thickness. In this way, signs of atrophy (hypertrophy) can be localized for the individual.

This thesis encompasses four original research papers which taken together describe and validate an individualized atrophy-assessment tool: first, the general procedure of rating an individual's MRI data, specifically cortical thickness data, to a control population is investigated for sensitivity and specificity using simulations. The selected strategy was based on rating topographically distinct cortical regions ("mosaics"/"patches") and is therefore referred to as "mosaic-approach". Given the reference groups were age- and gender matched, we investigated in a second study whether variance associated with these demographic variables was successfully eliminated, while maintaining information on clinical disability, studying a longitudinal data set of amyotrophic lateral sclerosis (ALS) patients (study 2). Finally, we explored the method for "external validity" in a dual approach: First, we tested whether the degree of cortical involvement is mirrored by our tool (which we hypothesized given it exclusively targets supratentorial gray matter regions) by contrasting different motor neuron diseases (MND) against each other (study 3). Second, we also explored if topographically distinct cortical pathology can be correctly localized with our tool, by including different patient subgroups from the frontotemporal dementia (FTD) spectrum. For FTD, as well as for MND, the "ground truth" localization of pathology is well-characterized by histological and previous (group-based) imaging findings, such that we could directly compare the individual results from the mosaic-approach to that knowledge. Another focus of this work was to provide an accessible visualization to easily identify regions of supposed atrophy (or hypertrophy) for the clinical user.

Our results suggest that the here-proposed mosaic-approach which compares single patient's cortical thickness data to matched control data is a viable approach to meaningfully detect signs of atrophy at the individual level; it is furthermore objective, reliable and valid. Despite clinical and methodological limitations, including small sample sizes of reference groups and varying acquisition parameters, which we are currently improving, we have high hopes that this tool can help clinical practice and can ultimately also be used as a novel endpoint for clinical trials.

# 4 Contributions

| Article Title (doi) | Authors (⁺ [shared] first author, * correspondence) | Journal (Impact Factor as most recently available in April 2021) | Publication date | Key points |
|---|---|---|---|---|
| **Towards individualized cortical thickness assessment for clinical routine** (10.1186/s12967-020-02317-9) | Marlene Tahedl ⁺* | Journal of Translational Medicine (4.197) | April 3rd, 2020 (online first) | Development of the method: selecting a strategy with high sensitivity and specificity to rate an individual's T1w MRI data for signs of atrophy with respect to a matched control population |
| **Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use** (10.1007/s00415-020-10368-7) | Marlene Tahedl⁺, Rangariroyashe H. Chipika, Jasmin Lope, Stacey Li Hi Shing, Orla Hardiman, Peter Bede* | Journal of Neurology (3.783) | January 5th, 2021 (online first) | "Internal validation": Demonstration of the method's successful elimination of variance associated with "clinically uninteresting" factors (age, gender) while maintaining information on disease burden (studying a longitudinal ALS data set) |
| **Propagation patterns in motor neuron diseases: individual and phenotype-associated disease-burden trajectories across the UMN-LMN spectrum of MNDs** | Marlene Tahedl⁺, Stacey Li Hi Shing⁺, Eoin Finegan⁺, Rangariroyashe H. Chipika, Jasmin Lope, Orla Hardiman, Peter Bede* | Neurobiology of Aging (4.398) | *Accepted on April 14th 2021 (in press)* | "External validation" (part 1): Demonstration that the method reflects the degree of cortical involvement by contrasting patients of different subgroups of the MND spectrum |

| | | | | |
|---|---|---|---|---|
| *The evolution of cortical integrity in motor neuron diseases: longitudinal cortical thickness profiles in ALS, PLS and poliomyelitis (10.17632/3pnp4hkhnw.1)* | *Marlene Tahedl[+], Stacey Li Hi Shing[+], Eoin Finegan[+], Rangariroyashe H. Chipika, Jasmin Lope, Orla Hardiman, Peter Bede** | *Mendeley Data* | *April 2nd 2021 (online)* | *Complementary data-in-brief article corresponding to the above-listed publication in "Neurobiology of Aging"* |
| Mapping cortical disease-burden at individual-level in frontotemporal dementia: implications for clinical care and pharmacological trials | Mary Clare McKenna[+], Marlene Tahedl[+], Jasmin Lope, Rangariroyashe H. Chipika, Stacey Li Hi Shing, Mark A. Doherty, Jennifer C. Hengeveld, Alice Vajda, Russell L. McLaughlin, Orla Hardiman, Siobhan Hutchinson, Peter Bede* | Brain Imaging and Behavior (3.391) | *Accepted on July 20th 2021 (in press)* | "External validation" (part 2): Demonstration that the method can detect topographically distinct patterns of atrophy (which are in line with histological findings) by contrasting patients of the FTD-/MND spectrum |

# 5 Abbreviations

| | |
|---|---|
| ACC | Anterior cingulate cortex |
| AD | Alzheimer's Disease |
| ALS | Amyotrophic lateral sclerosis |
| ALSFRS-r | Amyotrophic lateral sclerosis functional rating scale – revised |
| ANOVA | Analysis of variance |
| ANCOVA | Analysis of covariance |
| bvFTD | Behavioral variant frontotemporal dementia |
| Cam-CAN | Cambridge Centre for Ageing and Neuroscience |
| CC | Corpus callossum |
| CGM | Conditional growth model |
| CIFTI | Connectivity Informatics Technology Initiative |
| CNS | Central nervous system |
| CSD | Constrained spherical deconvolution |
| CST | Corticospinal tract |
| CT | Cortical thickness |
| DLPFC | Dorsolateral prefrontal cortex |
| DTI | Diffusion tensor imaging |
| DV | Dependent variable |
| EMG | Electromyography |
| EP | Electrophysiology |
| FA | Flip angle |
| FDR | False discovery rate |
| FLAIR | Fluid-attenuated inversion recovery |
| FOV | Field of view |
| FTD | Frontotemporal dementia |
| FWER | Family-wise error rate |
| G-CSF | Granulocyte-colony stimulating factor |
| GM | Gray matter |
| GRAPPA | GeneRalised Autocalibrating Partial Parallel Acquisition |
| HC | Healthy control(s) |
| HCP | Human Connectome Project |
| HSD | Tukey's honestly significant difference |
| HSP | Hereditary spastic paraplegia |
| ICD | International Classification of Diseases |
| ICF | International Classification of Functioning, Disability and Health |
| IR-SPGR | Inversion recovery-prepared spoiled gradient-recalled echo |
| IR-TSE | Inversion recovery turbo spin echo |
| LL | Log-likelihoods |
| LMN | Lower motor neuron |
| Lt | Left |
| M | Arithmetic mean |
| MICCAI | Medical Image Computing and Computer Assisted Intervention Society |
| ML | Machine learning |
| MND | Motor neuron disease |
| MNI152 | Montreal Neurological Institute 152 standard space |

## 5 Abbreviations

| | |
|---|---|
| MPRAGE | Magnetisation-prepared rapid gradient echo |
| Mo | Months |
| MRI | Magnetic resonance imaging |
| MS | Multiple sclerosis |
| NAKO | Nationale Kohorte |
| nfvPPA | Non-fluent variant primary progressive aphasia |
| PCC | Posterior cingulate cortex, |
| PD | Parkinson's Disease |
| PLS | Primary lateral sclerosis |
| PM(S) | Poliomyelitis (survivors) |
| QC | Quality control |
| ROI | Region of interest |
| Rt | Right |
| SBMA | X-linked Spinal and Bulbar Atrophy, |
| SD | Standard deviation |
| SE-EPI | Spin-echo echo planar imaging |
| SENSE | SENSitivity Encoding |
| SMA | Spinal muscular atrophy |
| SPIR | Spectral presaturation with inversion recovery |
| SR | Spatial resolution |
| svPPA | Semantic variant primary progressive aphasia |
| T | Tesla |
| T1w | T1-weighted |
| TCD | Trinity College Dublin |
| TE | Echo time |
| TFCE | Threshold-free cluster enhancement, |
| TI | Inversion time |
| TIV | Total intracranial volume |
| TMS | Transcranial magnetic stimulation |
| TR | Repetition time |
| UGM | Unconditional growth model |
| UK | United Kingdom |
| UL | Upper limb |
| UMN | Upper motor neuron |
| UMM | Unconditional means model |
| V | Version |
| VS | Versus |
| WHO | World Health Organization |
| WM | White matter |
| Y | Year |

# 6 Chapter I: Introduction

## 6.1 The diagnostic algorithm: When a patient seeks treatment

When a person suffers from physical problems and seeks medical treatment, her/his journey usually starts with seeing a primary care physician. The physician will first listen to the patient's reports on their symptoms and proceed with a clinical examination to detect the manifestation and ideally the cause of the problem. Often, the doctor's perceptional skills and clinical experience are sufficient to derive the correct diagnosis – which is imperative to start any treatment – but quite commonly, the diagnosis is not that obvious. Depending on which organ is the suspected cause of the problem, the general practitioner will then refer the patient to a qualified specialist. What virtually all medical disciplines have in common is the employment of diagnostic tools tailored to visually examine the organ they treat – or to put it in a nutshell with the words of best-selling author and psychiatrist Daniel Amen, "How do you know unless you look?" (Amen, 2005): gastroenterologists look at the gastrointestinal tract using endoscopy, cardiologists look at the heart with the help of ultrasound, ophthalmologists look at the eyes with an ophthalmoscope, etc. And neurologists? "Of course," the reader might think, "they look at the brain". While we take the possibility to look at the brain *in vivo* for granted today, the mere thought of that was unconceivable for most neurologists throughout history. However, with Röntgen's discovery of X-rays in 1895 and the aid of contrast agents a few decades later (or even via intrathecal application of air, often a painful procedure, see (Pneumoencephalography, Vogel, 1972)), it has become possible to roughly detect and localize prominent "masses" bearing pathological potential (such as abscesses, tumors or lesions) in the central nervous system (CNS), which formed the basis of an emerging medical discipline called "neuroradiology". Often, the removal of additional masses by the hands of a neurosurgeon would ameliorate the patient's problems (Bull, 1970). The sensitivity of detecting such masses greatly improved first with the introduction of computer tomography but especially magnetic resonance imaging (MRI), since latter technique offers excellent tissue-contrast and importantly is unique in not relying on potentially toxic radiation – unlike X-rays, computer tomography, angiography or even scintigraphy (Lauterbur, 1973). As such, MRI yielded great potential as a diagnostic tool which could be used for regular assessments of neurological patients.

But what about the many cases when a neurological patient's problems are not caused by *additional* masses in the CNS, but rather by the reverse – i.e. the *loss* of masses? Neurodegenerative disorders account for about a third of all deaths related to neurological disorders as well as to a significant proportion of disability-adjusted life years (Feigin et al., 2020): Especially Alzheimer's Disease (AD), but also less commonly observed but equally serious disorders such as Parkinson's Disease (PD), Motor Neuron Diseases (MND) and also Multiple Sclerosis (MS, which bears considerable characteristics of a neurodegenerative disorder besides its immunological components, (Stadelmann, 2011)) contribute to this statistic. By definition, a common pathological mechanism of any neurodegenerative disease is the loss of brain parenchyma, which is also referred to as "atrophy". But how do we detect, measure and monitor such atrophy? In very advanced cases, signs of atrophy can be well-detectable on MRI scans: Figure 6.1.1a illustrates this, showing a 57-year-old female subject who shows deep sulci in the temporal and parietal lobes (compare that to the dense packing of these areas in the healthy young 23-year-old subject displayed in Figure 6.1.1b*)*. But often, the time when atrophy becomes visually apparent on MRI scans is too late to take preventative care (which is hard enough even in early diagnoses as it is, but see (Perneczky, 2019)).

Since focally distinct patterns of gray matter pathology convey important information on differential diagnosis and disability burden, the need for quantification of its extent has early been recognized, e.g. by Koedam and colleagues (2011), who proposed to employ visual analog scales to

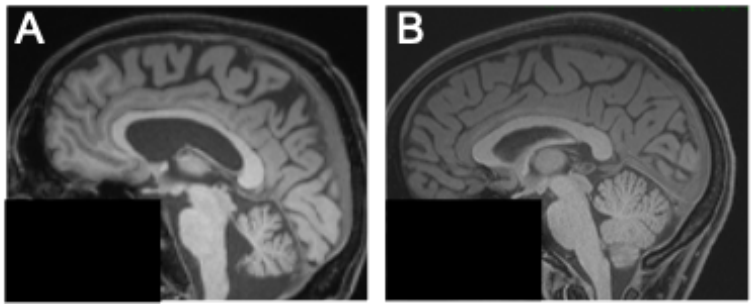



Figure 6.1.1. Atrophic or not? In advanced cases of gray matter loss (as evident by the deep sulci of the 57-year-old female subject shown in A), atrophy can be easily identified by visual inspection (compare e.g. to the relative "dense" packing" of the healthy 23-year-old subject in B with no evident atrophy). However, when atrophy is more subtle, it is not so straightforward to determine and highly subjective. In this dissertation, we propose and validate an objective and reliable tool to detect signs of atrophy for individual patients in a computational and automatized procedure, which uses merely a T1-weighted MRI image along with demographic information on age and gender. The procedure yields high-resolution maps of supposed atrophy which we hope will aid diagnosis / prognosis of neurodegenerative diseases in a inter-rater-independent fashion, which could also serve as a new endpoint in clinical trials.

sensitize raters to more subtle differences and enhance differential diagnostics of dementia (Koedam et al., 2011). However, this subjective method yielded only moderate inter-rater reliability (as low as 0.65), such that an inter-rater independent, objective, measure of atrophy is desirable. Only an objective measure can be reliable and eventually valid to adequately monitor a patient's progression or serve as an appropriate tool to evaluate therapeutic effects of a new drug in clinical trials. Due to its excellent tissue contrast and radiation-free acquisition scheme, MRI is a natural candidate with which to assess such a measure, or "biomarker". In the next section, we will explore how MRI has contributed to the development of such a biomarker, such that we can then discuss which requirements remain to be met in order to provide an assessment tool for gray matter pathology for clinical routine, which is the ambition of this thesis.

## 6.2 Towards a biomarker for neurodegenerative diseases: from estimating neuronal density to inferencing on atrophy

### 6.2.1 Assessing neuronal density: Gray matter volume and cortical thickness

Brain atrophy refers to the degeneration of neuronal cells; therefore, in order to quantify atrophy, the first step is to measure neuronal density. In the human brain, most neuron bodies accumulate at the rims of the brain – the "cortex" – while their connections (or "axons") run within this shell. Since the axons are generously covered with a fatty sheath (which helps to isolate the structures and therefore accelerates information transmission), the physical and chemical constitution between the outer and inner brain tissues differ. T1-weighted (T1w) MRI – due to its emphasis of the spin-lattice relaxation times of protons which rely on these physical and chemical constituents – is therefore prone to localize the border between the two tissue types, providing excellent contrasts of the inner, connecting layer (or "white matter", WM) and the outer, neuronal soma layer ("gray matter", GM). Thus, it makes sense to use T1w MRI to make inferences on neuronal density and thus atrophy.

Generally, there are two major strategies to calculate these metrics: One can either estimate GM volume (by virtually counting the number of voxels which make up the GM), or else one estimates the distance between the gray-/white matter boundary and the dura mater. Latter, when evaluated at each point of the cortex with a perpendicular line, gives an estimate of cortical thickness (CT) at each evaluated point (or "vertex"). However, that also requires a 2D-surface representation of the cortex, which nowadays however can be conceived easily due to automated, highly reliable and also importantly publicly available algorithms, e.g. under the open-source software package FreeSurfer (Dale et al., 1999; B Fischl et al., 1999). Both GM volume and CT have proven as useful tools to make inferences on atrophy in various conditions (Last et al., 2017), and both yield their own special pitfalls which

researchers need to be aware of; however, one great advantage of CT over GM volume is that CT is mostly independent from intracranial volume, which reduces the number of sources of variation (especially since the algorithms to correct for head size vary, (Schwarz et al., 2016)), which is potentially an advantage to develop an individualized biomarker. Also, CT is easy to derive and its variations indicate changes of neuronal density (however one needs to be aware that CT is not only proportional to neuronal density but also influenced by synaptic density, synaptic pruning and intracranial myelination (Fjell & Walhovd, 2010; Huttenlocher, 1979; Huttenlocher et al., 1982; Huttenlocher & Dabholkar, 1997). For these reasons and the fact that CT is nowadays generally discussed to yield less sources of variation as compared to GM volume (Winkler et al., 2010), the main focus of this work will be CT as a measure of neuronal density and eventually atrophy, but not without emphasizing once more that measuring GM volume can yield valuable additional and differential information on neuronal density and should be thoroughly assessed for its validity as a personalized biomarker for neurodegenerative diseases (the interested reader is also referred to (Kong et al., 2015) for an experimental investigation and juxtaposition of the two measures in schizophrenia).

### 6.2.2 Studying cortical thickness at the group level: What have we gained?

Now that we have embarked on CT as a measure for neuronal density, the next challenge is to leverage it for inferencing on atrophy. Since atrophy refers to the loss of neuronal cells, a natural candidate for "biomarking" neurodegeneration is to search for signs of reduced CT. A typical MRI experiment that investigates CT differences compares a group of patients (in our case suffering from some neurodegenerative disorder) against a group of matched healthy controls (HC). Experiments like that have greatly enhanced our understanding of the underlying pathology and progression of a number of neurological and also psychiatric conditions, including AD (Shaw et al., 2016), PD (Zarei et al., 2013), MS (Steenwijk et al., 2016), depression (Liu et al., 2019) and schizophrenia (Assunção Leme et al., 2013). Interestingly, we have learned that these pathological conditions present with topographically distinct patterns of cortical thinning, which are only partly accounted for by factors such as age and genetic components (Chouinard-Decorte et al., 2014; Fjell et al., 2015). These specific focal aspects indicate the potential of CT as a biomarker for differential monitoring of neurodegenerative diseases.

### 6.2.3 Towards development of an individualized biomarker of atrophy

#### *6.2.3.1 The role of machine learning for categorical diagnostics*

In spite of the rich diagnostic information of CT, it is rarely assessed or even considered in clinical practice. One of the reasons for this might be the lack of a standardized system based on which an individual's CT can be rated. As outlined above, most previous studies have compared groups of patients against groups of controls, describing mere statistical effects. But such results are of little practical value to the daily practitioner since they do not treat groups of people but rather individual patients. In recent years, however, the introduction of machine learning (ML) algorithms to neuroscientific studies has allowed to meaningfully appraise a single patient's MRI scan. Typically, such studies incorporate training classifiers to distinguish between patients and controls (thereby aiding diagnosis, e.g. (Bi et al., 2020; Lian et al., 2020; S. Wang et al., 2016)), patients and patients (aiding differential diagnosis, e.g. (O. Ben Ahmed et al., 2017; Basaia et al., 2019; Kim et al., 2019; Singh et al., 2018)), conversion to an advanced disease state (e.g. (Gómez-Sancho et al., 2018; Hojjati et al., 2017; Spasov et al., 2019) or even prognosis (e.g. (Salvatore & Castiglioni, 2018)). Therefore, implementation of ML methods into clinical practice holds potential to greatly support phenotyping.

However, ML is also inherently limited to this kind of "top-down" processing: classifiers need to be "trained" first – they first need to "see" examples of the one class or the other, which means that

accurately labeled images are fundamental for obtaining accurate classifiers. However, provided such accurate labeling (along with an adequate neural network architecture), computers can be incredibly successful in deriving their own conclusions about which class an unseen case belongs to (Young et al., 2020). The successful demonstrations of the power of ML for phenotyping suggests that ML will be a fundamental part of future diagnostics in neurology and especially neurodegeneration.

#### *6.2.3.2 From treating codes to treating patients*

Currently, *International Classification of Diseases* (ICD-10-CM, tenth revision, clinical modification, (WHO, 2021)) lists 241 codes for neurological disorders. As discussed in the previous paragraph, ML bears great potential to help clinicians to disentangle accurately between all these differentials (note here also the ever-more emerging role of genetic sequencing for phenotyping, see (Robinson et al., 2015)). However, there is more to treatment than deriving the correct diagnosis, or as the Canadian physician Dr. William Osler probably not said[1], "A good physician treats the disease, the great physician treats the patient." – even within one diagnosis group, there remains great variability on the exact pathological manifestation, precipitating very unique strengths and needs for each patient. The *World Health Organization* (WHO), who publishes the ICD, recognized early on that the information on diagnosis provided by the ICD is not sufficient to fully capture the health status of an individual patient, and therefore developed a complementary system in 1980 (not approved until 2001) which accounts for the functioning and disability of the patient: the *International Classification of Functioning, Disability and Health* (ICF, (WHO, 2001)). The ICF requires – among others – the consideration of the current individual body structures and functions to obtain a very personalized view of the patient's exact functioning and disability, which can eminently optimize individual treatment plans.

Such a "snapshot" image of an individual's pathological presentation cannot be satisfied by the "top-down" categorization ML algorithms yield. Rather, this requires an objective "bottom-up" scan of the current constitution of the affected organ, which is typically the brain in neurology and more specifically the brain's cortex in neurodegenerative diseases. *To be scientifically useful*, such a "scan" should be objective, reliable and valid; *to be clinically useful*, it should additionally be easy to derive and applicable to individual patients. This would also grant comparability of follow-up measures for individuals, which would allow individualized monitoring of the disease progression. Furthermore, such a system would offer a new endpoint for evaluating therapeutics in clinical trials, which is based on personalized disease status rather than a diagnostic category.

It was the ambitious goal of this dissertation to develop such an objective and reliable technique on the one hand and examine its utility and validity for clinical application on the other. It will be my privilege to present the research concept in the next section to the reader, which will be followed by the original peer-reviewed articles we published on these topics. At the end, I'll discuss the perspective of this work, (some of) its current limitations and how I plan to overcome those in the future. Or maybe I should rather say "improve" those since George E.P. Box just reminds me that "All models are wrong, but some are useful." (Box, 1976).

[1] *„probably not said"* was formulated as a paradox on purpose, given it remains controversial whether Dr. Osler was the true originator of this quote. The source http://www.lessismoremedicine.com/quotes (last accessed on July 20th 2021) specifies this as follows: << (MIS-ATTRIBUTED DR. WILLIAM OSLER; may be derived from Moses Maimonides (Rambam) in his Treatise on Asthma, arguing for a view of the patient as unique: "Any sick individual presents new problems. One can never say one disease is just like the other. *A therapeutic axiom follows: 'the physician should not treat the disease but the patient who is suffering from it') >>.*

## 6.3 Why Motor Neuron Disease?

There is however one last parameter to discuss to finalize the general set-up: As outlined above, this dissertation did not stop after proposing a method for rating an individual's cortical thickness data – rather, one core ambition was the validation of the method. While early validation on simulated data suggested the method's utility (subproject 1), real-world patient data was required to probe its potential clinical benefit. Of the many neurodegenerative disorders, the Motor Neuron Disease (MND) spectrum yields many features which are useful for validation: First, we know very well from post-mortem studies where the pathology is located in MND. As the name suggests, neuronal degeneration from motor-related pathways are the determining pathological feature of MND (P. Ince et al., 2008). Therefore, any method which localizes atrophy *in vivo* should spot these relevant areas. The known "ground truth" localization of pathology is one argument to study MND for validation of an atrophy-detection tool. Second, MND is associated with a fast disease course and pronounced degeneration of the affected areas (Bromberg, 2014). Therefore, the demands for sensitivity of the tool are manageable at this early developmental stage. Third, MND offers a spectrum of differential diagnoses: What all MNDs share is that neurons associated to motor pathways degenerate. However, the execution of any motor task relies on both upper (UMN) and lower motor neurons (LMN), with the first being located in the cortex, the latter in the spinal cord. Different MND disorders involve the different motor neurons to different degrees: Primary lateral sclerosis (PPS) mainly affects UMN, poliomyelitis (PM) mainly LMN, while amyotrophic lateral sclerosis (ALS) affects both systems. Since the method proposed here exclusively targets the cortex (as of yet) and therefore UMN, contrasting the different MND disorders with that measure should directly reflect the different degrees of involvement of the respective UMN. Finally, clinically, some MND cases often resemble (or co-occur with) another neurological disorder from the frontotemporal dementia spectrum (FTD) – although FTD pathology is located at different areas in the brain, as the name suggests. Contrasting MND to FTD variants therefore also allows to put the utility of the method for differential diagnosis at test.

Taken together, MND offers many features which greatly enhance the development and validation for our method which aims to rate an individual patient's cortical thickness, and it was my great privilege to study MND patients for the validation of the method in a collaboration with Prof. Peter Bede from Trinity College Dublin (TCD), who has been researching MND for many years and who very generously made his patient data – along with his invaluable critical and inspiring thoughts – available for these projects. But now, let's move on with the actual research concept and conclude with Goethe, who promts us to finally deliver: "That's enough words for the moment, / Now let me see some action!" (Goethe, 2015).

## 6.4 Research concept

This dissertation is composed of four peer-reviewed research articles, which taken together introduce a scientifically validated method for meaningful individualized assessment of T1w MRI images with respect to cortical thickness. The articles were composed chronologically along the development of a new biomarker, such that with the first article, an adequate objective and reliable method was investigated; in the second article, we internally validated the tool on longitudinal ALS patient data (testing whether the method corrected for confounding effects of age and gender while information on disease burden remained, as implied); and finally, with articles 3 and 4, we probed external validity of the method, contrasting different manifestations of the MND spectrum (article 3), and opposing MND to an independent but often observed comorbidities from the FTD spectrum (article 4). The method is easy to implement for the end user (provided some installed software), relatively cheap (since only a T1w MRI image is required as an input) and straightforward to interpret. Such a tool cannot only be

invaluable for daily clinical routine but also serve as a new endpoint in clinical trials. The following table gives an overview of the four articles:

**Table 6.4.1**. Overview of the research articles corresponding to this dissertation.

| **Article Title** (doi) | **Authors** (⁺ [shared] first author, * correspondence) | **Journal (Impact Factor as most recently available in April 2021)** | **Publication date** | **Key points** |
|---|---|---|---|---|
| **Towards individualized cortical thickness assessment for clinical routine** (10.1186/s12967-020-02317-9) | Marlene Tahedl ⁺* | Journal of Translational Medicine (4.197) | April 3rd, 2020 (online first) | Development of the method: selecting a strategy with high sensitivity and specificity to rate an individual's T1w MRI data for signs of atrophy with respect to a matched control population |
| **Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use** (10.1007/s00415-020-10368-7) | Marlene Tahedl⁺, Rangariroyashe H. Chipika, Jasmin Lope, Stacey Li Hi Shing, Orla Hardiman, Peter Bede* | Journal of Neurology (3.783) | January 5th, 2021 (online first) | "Internal validation": Demonstration of the method's successful elimination of variance associated with "clinically uninteresting" factors (age, gender) while maintaining information on disease burden (studying a longitudinal ALS data set) |
| **Propagation patterns in motor neuron diseases: individual and phenotype-associated disease-burden trajectories across the UMN-LMN spectrum of MNDs** | Marlene Tahedl⁺, Stacey Li Hi Shing⁺, Eoin Finegan⁺, Rangariroyashe H. Chipika, Jasmin Lope, Orla Hardiman, Peter Bede* | Neurobiology of Aging (4.398) | *Accepted on April 14th 2021 (in press)* | "External validation" (part 1): Demonstration that the method reflects the degree of cortical involvement by contrasting patients of different |

| | | | | |
|---|---|---|---|---|
| | | | | subgroups of the MND spectrum |
| ***The evolution of cortical integrity in motor neuron diseases: longitudinal cortical thickness profiles in ALS, PLS and poliomyelitis*** *(10.17632/3pnp4hkhnw.1)* | *Marlene Tahedl[+], Stacey Li Hi Shing[+], Eoin Finegan[+], Rangariroyashe H. Chipika, Jasmin Lope, Orla Hardiman, Peter Bede** | *Mendeley Data* | *April 2nd 2021 (online)* | *Complementary data-in-brief article corresponding to the above-listed publication in "Neurobiology of Aging"* |
| **Mapping cortical disease-burden at individual-level in frontotemporal dementia: implications for clinical care and pharmacological trials** | Mary Clare McKenna[+], Marlene Tahedl[+], Jasmin Lope, Rangariroyashe H. Chipika, Stacey Li Hi Shing, Mark A. Doherty, Jennifer C. Hengeveld, Alice Vajda, Russell L. McLaughlin, Orla Hardiman, Siobhan Hutchinson, Peter Bede* | Brain Imaging and Behavior (3.391) | *Received positive reviews on May 17th 2021; not yet published (in press)* | "External validation" (part 2): Demonstration that the method can detect topographically distinct patterns of atrophy (which are in line with histological findings) by contrasting patients of the FTD-/MND spectrum |

+ (shared) first author

* corresponding author

### 6.4.1 Development of the method (subproject 1): "Towards individualized cortical thickness assessment for clinical routine"

The first subproject aimed at determining an adequate method to meaningfully interpret an individual patient's T1w MRI data to inference on atrophy (Tahedl, 2020). To achieve this, different methods were compared to assess statistical significance of cortical thinning, i.e. atrophy. All compared methods were nonparametric and encompassed rating an individual subject's data set with respect to a control population in a z-scored based approach. Null distributions were calculated using data from the Human Connectome Project (HCP, $n$ = 1000), and an additional HCP data set ($n$ = 113) was used to calculate sensitivity and specificity to compare the different methods, whereas atrophy was simulated for sensitivity / specificity assessments. Validation measures were calculated for the entire cortex ("cumulative") and distinct brain regions ("regional") where possible. The approach yielding the highest combination of specificity and sensitivity implemented generating null distributions for anatomically distinct brain regions, based on the most extreme values observed in the population. With that method, while regional variations were observed, cumulative specificity of 98.9% and cumulative sensitivity at 80% was achieved for simulated atrophy of 23%. This subproject suggested that valid rating of an individual cortical thickness data is possible. Since the proposed method quantitatively evaluates separate brain regions, it can help clinicians in their daily routine to localize signs of atrophy before they become visually apparent on an unprocessed T1w MRI. This can be helpful especially for differential diagnosis / prognosis, since disparate pathologies typically present with distinct atrophy patterns. Furthermore, such a standardized procedure can help to monitor disease progression, which

additionally can be useful for evaluation of therapeutic effects in clinical trials. This method will be also referred to as “mosaic-based approach” (highlighting the evaluation at distinct cortical regions, which we refer to as “mosaics” and “patches” interchangeably).

### 6.4.2 Validation in ALS, or the role of gender, age and disease burden (subproject 2): “Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use”

In subproject 1, we showed that a z-score based approach can be used to meaningfully rate individual patients’ T1w MRI data to infer on atrophy. However, the therein proposed strategy still yielded several drawbacks which limited its utility for clinical routine, namely: (i) the control data was taken from the HCP data repository, which mainly includes young subjects (between 22 and 40 years of age). However, most patients suffering from neurodegenerative diseases are older, such that comparing these patients to this young control cohort will yield many false positive “atrophic” regions due to physiological cortical thinning in ageing (Fjell et al., 2015; Shaw et al., 2016). On that account, for clinical translation of the method, it is fundamental to use better matched control data in terms of age. (ii) Similarly, CT is greatly confounded by gender effects (Sowell et al., 2007). As such, distinct reference distributions for male and female patients should furthermore enhance sensitivity. (iii) We showed that creating separate null distributions for distinct cortical regions greatly increases the power for detecting cortical atrophy. However, we only used a rather coarse cortical parcellation (distinguishing 68 brain regions as defined by the Desikan-Killiany atlas, (Desikan et al., 2006)), within which we still observed physiologically significant variability of CT distributions (cf. also (Wirth et al., 2018), who showed regional differences in ventral vs. dorsal pre-/postcentral gyri with respect to cortical thinning in ALS patients). As such, a finer cortical parcellation scheme should further improve the biological plausibility of the proposed method and therefore also enhance sensitivity of detecting signs of cortical atrophy. (iv) Finally, validation of the method happened on simulated atrophy data in subproject 1. In order to demonstrate clinical utility, validation on real-world patient data is required.

Subproject 2 was designed to overcome all these flaws (Tahedl et al., 2021): To correct for drawback (iv), we tested the method using longitudinal MRI data from 61 ALS patients acquired at Trinity College Dublin (TCD). ALS is a good candidate disease to validate such a cortex-based structural measure since it is a progressive neurodegenerative disorder, affecting both upper and lower motor neurons (Orla Hardiman et al., 2017). As of yet, clinical trials in ALS still rely on functional rating scales, questionnaires, respiratory measures and survival as primary endpoints (Hiroshi Mitsumoto et al., 2014). Surprisingly, neuroimaging measures are not integrated in clinical trials designs (Rangariroyashe Hannah Chipika et al., 2019; Christina Schuster et al., 2015) although the utility of quantitative imaging in ALS has been repeatedly demonstrated in tracking longitudinal changes (Rangariroyashe H Chipika, Siah, et al., 2020; Christina Schuster et al., 2016). A range of promising biomarkers have been evaluated in ALS (Blasco et al., 2018; Devos et al., 2019; Floeter & Gendron, 2018), but the specific appeal of imaging data is that it can be non-invasively acquired and quantitatively interpreted, and it reflects directly on pathological changes in the nervous system. These considerations coupled with the widespread availability of MRI scanners, open-source software, large data repositories, and the fact that the majority of suspected ALS patients undergo brain imaging as part of their diagnostic work-up make a compelling argument to use imaging as means of identifying and tracking pathological change in ALS. Common gray matter metrics, such as regional CT, are typically evaluated in group comparisons. However, the alternative strategy proposed in subproject 1 aims to interpret cortical thickness values in individual patients with respect to carefully chosen reference groups (Tahedl, 2020). To provide a better match of the healthy control population, we used the publicly available data from the Cambridge

Centre for Ageing and Neuroscience data repository (Cam-CAN, (Shafto et al., 2014)), which provides over 650 T1w MRI scans from healthy subjects spread evenly throughout age (ranging from the low twenties to the old eighties) and gender. In addition, we furthermore boosted this control sample size by adding another 125 healthy controls (HC) obtained from the same source as the ALS patient data. To enhance sensitivity of the method, we a) created separate null distributions for males and females (correcting for drawback (iii)) and b) optimized age-matching as follows (correcting for drawback (ii)): for each year of age $X$, we defined a subject-specific reference group which was comprised of HC subjects from the combined Cam-CAN / TCD data set aged between $[X-2; X+2]$, separately for males and females. For example, the reference group for a 54-year-old male ALS patient at baseline included all men between 52 and 56 years from the Cam-CAN / TCD control data set. With this relatively narrow age range, we found still an acceptable median group size for each reference group ($n = 33$, interquartile range = [28.75; 33]) but which still allowed us to detect statistical significance in a permutation testing scheme with a significant alpha-threshold at $p <= 0.05$ (Nichols & Holmes, 2001; Winkler et al., 2014). Finally, drawback (i) was corrected for by employing a finer cortical parcellation scheme (distinguishing 1000 separate, roughly equally-sized brain areas, which we also refer to as "patches" or "mosaics") (Schaefer et al., 2018; Yeo et al., 2011).

We have interrogated this multi-timepoint longitudinal dataset to demonstrate the utility of estimating cortical disease burden and the expansion of cerebral atrophy over time. For validation purposes, we contrasted our strategy to the gold-standard approach (i.e., comparing patient and control groups with correction for alpha-level inflation, age, and gender) to gauge the advantages and drawbacks of our method. Furthermore, we modelled the evolution of cortical integrity (as measured by the new approach) in a conditional growth model, in which we accounted for age, gender, disability, symptom duration, education and handedness. We hypothesized that the variance associated with demographic variables would be successfully eliminated in our approach. Indeed, we could show that the only covariate which modulated the expansion of atrophy was motor disability as measured by the ALSFRS-r ($t(153) = -2.533$, $p = 0.0123$). Using the standard approach, age also significantly influenced progression of CT change ($t(153) = -2.151$, $p = 0.033$) demonstrating the validity and potential clinical utility of our approach – in other words we could show in this study that the herein proposed method successfully corrects for the confounding effects of the "clinically uninteresting" factors age and gender, while still maintaining information on the "clinically interesting" factor disease burden as assessed by the ALSFRS-r rating scale.

### 6.4.3 Differential diagnosis

With the second study of this dissertation, we could provide some evidence suggesting internal validity of the proposed z-scored approach to assess statistical significance of cortical thinning or atrophy. The next step was to also explore its external validity, i.e. to probe the method for its utility for differential diagnosis. We chose a dual approach to achieve this goal: First, we studied whether the method – which as of yet exclusively incorporates cortical involvement – actually *reflects* the degree of cortical involvement. The MND spectrum provides an excellent opportunity to put this to test since it spans differential diagnoses affecting upper and lower motor neurons to varying degrees (subproject 3). In addition, utility for differential diagnosis to a clinically related but pathologically dissimilar condition, namely the frontotemporal dementia spectrum (FTD), was assessed (subproject 4). Conveniently, similarly to MND, FTD provides an excellent opportunity for validation since also for that condition, the "ground truth" localization of pathology is well characterized, namely frontal and temporal cortices (Bang et al., 2015). If our method was valid, it should directly reflect involvement of these areas distinctly

from the ones observed in MND. The following two sections provide closer details on the objectives and findings of subprojects 3 and 4.

#### *6.4.3.1 Within spectrum: MND spectrum (subproject 3)*

The assessment tool we propose herein is (as of yet) exclusively designed to detect atrophy in the cortex of the brain. In order to demonstrate validity, it needs to be shown that the degree of cortical involvement is reflected with this measure. The MND disease spectrum provides an excellent opportunity to address this question: One of the key aspects of the MND disease spectrum is its clinical and pathological heterogeneity with respect to the involvement of upper (UMN) and lower motor neurons (LMN): For example, Primary lateral sclerosis (PLS) mainly affects UMN, poliomyelitis LMN, and ALS impairs both systems. In subproject 3, we contrasted these differential diagnoses and hypothesized that the quantification of atrophy as assessed with our method would mirror the different degrees of cortical involvement, i.e. UMN involvement: in other words, PLS should be attested most pronounced atrophy, ALS second, poliomyelitis third.

In addition, we studied the "other extreme" of cortical thickness changes in this subproject, addressing *hypertrophy* rather than atrophy: There is a notion that compensatory processes may occur in MNDs to adapt for neurodegeneration, which can be reflected in either lack of atrophy or even hypertrophy (Querin, El Mendili, et al., 2019). Irrespective of the MND phenotype studied, the overwhelming majority of published papers only comment on patterns of atrophy, cortical thinning or density reductions (Bede & Hardiman, 2018). Lack of atrophy or 'resilient' regions are only rarely specifically evaluated, but increased CT is either not explored in statistical models, not reported, or not discussed. This seems like a missed opportunity as cortical reorganization may reveal biologically important processes (O Hardiman et al., 2016). With our z-scored based approach, it is straightforward to detect patterns of "unexpectedly thick" cortical regions, simply by assigning p-values to the recorded CT values which reflect the likelihood of observing *the "thickness"* (rather than *the "thinness",* as we do in order to detect atrophy) by chance. We used a multi-timepoint longitudinal data set with uniform follow-up intervals of four months, including 45 poliomyelitis survivors, 61 patients with ALS and 23 patients with PLS. CT alterations were evaluated in a dual analysis pipeline (the herein proposed individualized z-score based, "mosaic-approach" as well as the "gold standard approach", i.e. contrasting patient and control groups with FWER correction). Our results indicated that PLS patients exhibit rapidly progressive cortical thinning in motor regions; ALS patients showed cortical atrophy in both motor and extra-motor regions while poliomyelitis survivors exhibit cortical thickness gains in a number of brain regions.

Clinically, our findings suggested that dynamic cortical changes in MND may depend on relative UMN/LMN involvement, and increased CT in LMN-predominant conditions may represent compensatory adaptive processes. Methodologically, we could provide evidence with this study that the individualized z-score based mosaic-approach directly reflects the degree of cortical involvement which further validated the method; moreover, we showed that the mosaic-approach is also fit to detect patterns of atrophy which might be useful to detect compensatory mechanisms of neurogenesis.

#### *6.4.3.2 Between: ALS vs. FTD (article 4)*

So far, we have (i) proposed an inter-rater independent method for assessing signs of atrophy (hypertrophy) for individual patients, (ii) provided evidence that it successfully eliminates confounding effects of age and gender (while maintaining information on clinical disability) and (iii) shown that it's a direct reflection of cortical involvement. Finally, to also demonstrate the method's utility for differential diagnosis, we studied its application to a clinically related but pathologically unrelated condition, namely

the frontotemporal dementia spectrum (FTD). One argument that speaks in favor of studying FTD patients for validation is that similarly to MND, we know from histology the focus of the pathology (as the name suggests, frontal and temporal cortices, (Bang et al., 2015)). Therefore, the "ground truth" localization of the pathology is established such that we can compare and validate our results based on that knowledge. We studied 62 FTD patients presenting with different FTD subtypes. As before, we evaluated CT changes in a dual pipeline, using standard cortical thickness analyses and an the herein proposed individualized, z-score based mosaic-approach to characterize subject-level disease burden. Phenotype-specific patterns of cortical atrophy were readily detected with both methodological approaches. Consistent with their clinical and pathological profiles, patients with bvFTD exhibited orbitofrontal, cingulate and dorsolateral prefrontal atrophy, patients with ALS-FTD displayed precentral gyrus involvement, nfvPPA patients showed widespread cortical degeneration including insular and opercular regions and patients with svPPA exhibited relatively focal anterior temporal lobe atrophy. Cortical atrophy patterns were reliably detected in single individuals, and these maps were consistent with the clinical categorization.

With these results, we showed that the mosaic-approach is useful to generate maps of suspected cortical atrophy from single patients of various neurodegenerative disorders. This fourth subproject concludes the framework of this thesis. We hope to have shown that the computational interpretation of single T1w scans can greatly enhance assessment of subtle cortical thickness changes in neurodegenerative diseases – a main modulator of disability – which are easily missed by mere visual inspection. The quantitative evaluation of individual MRI data may aid diagnostic classification, clinical decision making, and monitoring longitudinal changes. In the next section, the original research articles are provided, which will be followed by a short recapitulation and practical discussion of this work – along with considerations on its limitations – and concluded by perspectives on future directions.

# 7 Chapter II: Research articles

## 7.1 Subproject 1 "Towards individualized cortical thickness assessment for clinical routine"

### 7.1.1 Abstract

**Background**: Cortical thickness measures the width of gray matter of the human cortex. It can be calculated from T1-weighted magnetic resonance images (MRI). In group studies, this measure has been shown to correlate with the diagnosis/prognosis of a number of neurologic and psychiatric conditions, but has not been widely adapted for clinical routine. One of the reasons for this might be that there is no reference system which allows to rate individual cortical thickness data with respect to a control population.

**Methods**: To address this problem, this study compared different methods to assess statistical significance of cortical thinning, i.e. atrophy. All compared methods were nonparametric and encompassed rating an individual subject's data set with respect to a control data population. Null distributions were calculated using data from the Human Connectome Project (HCP, n = 1000), and an additional HCP data set (n = 113) was used to calculate sensitivity and specificity to compare the different methods, whereas atrophy was simulated for sensitivity assessment. Validation measures were calculated for the entire cortex ("cumulative") and distinct brain regions ("regional") where possible.

**Results**: The approach yielding the highest combination of specificity and sensitivity implemented generating null distributions for anatomically distinct brain regions, based on the most extreme values observed in the population. With that method, while regional variations were observed, cumulative specificity of 98.9% and cumulative sensitivity at 80% was achieved for simulated atrophy of 23%.

**Conclusions**: This study shows that validated rating of individual cortical thickness measures is possible, which can help clinicians in their daily routine to discover signs of atrophy before they become visually apparent on an unprocessed MRI. Furthermore, given different pathologies present with distinct atrophy patterns, the regional validation proposed here allows to detect distinct patterns of atrophy, which can further enhance differential diagnosis/prognosis.

**Keywords**: Cortical thickness, neuroimaging, magnetic resonance imaging (MRI), individual diagnosis, atrophy, neurological assessment

### 7.1.2 Background

Using magnetic resonance imaging (MRI), images with high-tissue contrast (Brant-Zawadzki et al., 1983) of the brain can be acquired without making use of radioactive contamination of patients. Beyond clinical applications, MRI has been widely used for neuroscientific studies. Constantly, methods are being developed which allow to quantify biologic characteristics of the central nervous system and its constituents more and more differentiated, encompassing blood flow, nerve fiber myelination and properties of the cortex or "gray matter" (GM). The GM is the location of the neuron bodies, whereas the extent of cortical thickness seems to be related to synaptic density, synaptic pruning and intracranial myelination (Fjell & Walhovd, 2010; Huttenlocher, 1979; Huttenlocher et al., 1982; Huttenlocher & Dabholkar, 1997), rather than the number of neurons (Fjell & Walhovd, 2010; Herculano-Houzel et al., 2013). A T1-weighted MRI of the brain is sufficient to compute cortical thickness in an automated procedure and can be further optimized with an additional T2-weighted image (Dale et al., 1999; B Fischl et al., 1999). Common algorithms to calculate cortical thickness are publicly available, e.g. under the open-source software package FreeSurfer (Bruce Fischl, 2012).

Cortical thickness has been subject to a wide range of studies, and cortical thinning (i.e. atrophy) has been associated with diagnosis and progression of a number of neurologic conditions, such as Alzheimer's Disease (Shaw et al., 2016), Parkinson's Disease (Zarei et al., 2013) and Multiple Sclerosis (Steenwijk et al., 2016) as well as psychiatric conditions, such as depression (Q. Li et al., 2019) and schizophrenia (Assunção Leme et al., 2013). Interestingly, such pathological conditions present different patterns of cortical thinning and are modified by age and genetic components (Chouinard-Decorte et al., 2014; Fjell et al., 2015). These specific aspects make cortical thickness a good candidate as a biomarker for differential diagnosis/prognosis. However, assessing cortical thickness is rarely incorporated in clinical practice. One of the reasons for this might be the lack of a standardized system, based on which an individual's cortical thickness data can be rated. To pass this limit, the present study aimed to develop a method to rate an individual's cortical thickness data with respect to a control population which detects cortical atrophy with high sensitivity and specificity. To allow detecting distinct patterns of cortical atrophy, the tested methods allow the evaluation of separate brain regions. Such a standardized procedure can help clinicians detect early signs of distinct atrophy patterns and monitor their progression.

### 7.1.3 Methods

#### *7.1.3.1 Subjects*

In order to rate an individual's data with respect to a control population, a large number of standardized data from a representative population sample is required. The Human Connectome Project (HCP) provides such a resource (Marcus et al., 2011, 2013; Van Essen et al., 2013). For this study, data from the HCP's 1200 Subject Release was used. In total, structural data (T1- and T2-weighted sequences) from 1113 subjects was available at the time of this study (507 males, aged between 22 and plus 36). Of the 1113 subjects, 1000 were randomly selected for generating null distributions of cortical thickness, the rest was spared for subsequent validation (see below).

#### *7.1.3.2 Data acquisition and preprocessing*

The HCP data was acquired on a 3 Tesla Connectome Scanner. Two different types of structural sessions were acquired, encompassing a T1-weighted MPRAGE (repetition time (TR) = 2400ms, echo time (TE) = 2.14ms, inversion time = 1000ms, flip angle (FA) = 8°, field of view (FOV) = 224x224, voxel resolution (VR) = 0.7mm$^3$, bandwidth (BW) = 210 Hz/Px, iPAT factor 2, total acquisition time 7 minutes 40 seconds) and a T2-weighted SPACE (TR = 3200ms, TE = 565 ms, FA variable, FOV = 224x224, VR = 0.7mm$^3$, BW = 744 Hz/Px, iPAT factor 2, total acquisition time 8 minutes 24 seconds). The full imaging protocols can be found online at https://www.humanconnectome.org/hcp-protocols. All study procedures of the HCP protocol were approved by the Institutional Review Board at the Washington University in St. Louis.

The HCP offers data which was preprocessed with standardized and validated procedures. The main preprocessing steps encompassed gradient distortion correction, brain extraction, nonlinear registration, surface registration, and registration onto high-resolution (164k mesh) and low-resolution (32k mesh) templates; more details on the exact preprocessing pipeline can be found in (Bruce Fischl, 2012; Glasser et al., 2013; Jenkinson et al., 2002, 2012). The image format of the mesh images is in CIFTI format (Connectivity Informatics Technology Initiative), a file format which combines surface-based cortical data with volumetric-based subcortical/cerebellar data, which was found to enhance alignment to the geometry of the cortex as well as statistical power (Tucholka et al., 2012). The HCP's minimally preprocessed data include cortical thickness maps (generated based on the standardized FreeSurfer pipeline with combined T1-/T2-reconstruction (Dale et al., 1999; B Fischl et al., 1999)). For this study, the high-resolution cortical thickness maps (164k mesh) were used.

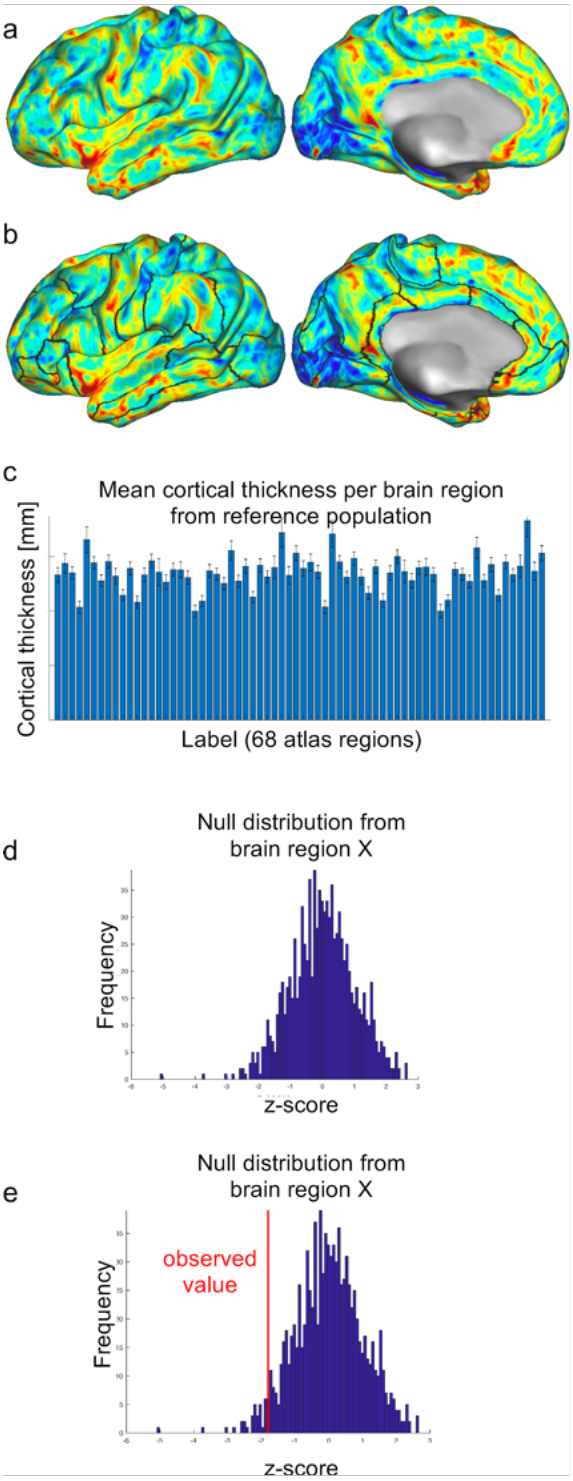



*Figure 7.1.1 Generating a reference system for rating an individual's cortical thickness data with respect to a control population. In methods 3 and 4, each cortical thickness map from a population sample (a) was* divided *into 68 distinct brain regions (borders are indicated as black lines in b). Given that the different brain regions have different means and standard deviations (c), this approach is biologically more plausible than generating one common reference system for all brain regions (as was tested here in methods 1 and 2). Based on these null distributions (see d for an example), the observed values for an individual can be rated within the control population (see red line in e) and statistically significant cortical thinning (i.e. atrophy) can be assessed.*

#### *7.1.3.3 Statistical Analysis*

Statistical analysis of the minimally preprocessed HCP neuroimaging data was carried out with tools from the Connectome Workbench (Marcus et al., 2011, 2013) and MATLAB R2019b (The Mathworks, Natick, USA). First, null distributions were generated using different strategies and subsequently, these methods were validated and compared based on their specificity and sensitivity.

##### *Generating null distributions*

Different strategies to generate null distributions were compared. These can be subdivided into (a) generating one common null distribution for all data points on the cortex (referred to as "vertices" in CIFTI mesh files) and (b) generating separate null distributions for distinct brain regions (Figure 7.1.1a–b)). Note that thickness spreads nonuniformly across the human cortex (Fukutomi et al., 2018; Kennedy, 1998; X. Wang et al., 2010; Zilles et al., 1988) such that different brain regions show different population means (Figure 7.1.1c). Therefore, different null distributions for distinct brain regions might increase sensitivity of detecting atrophy, which is why both approaches were compared in the present study. The two approaches were subdivided further into more and less conservative statistical corrections, such that in total, four methods were compared. Null distributions were computed using nonparametric permutation procedures for all methods (Nichols & Holmes, 2001), since they make less assumptions than parametric models and are therefore considered more robust than parametric tests (Hollander et al., 2014; Sheskin, 2003).

##### *Method 1: Z-min statistic per data point*

The statistically most conservative approach was based on generating one common reference distribution for all 298261 data points of the cortical surface. First, from 1000 HCP data sets, each data set was selected iteratively ("test data set") and standardized with respect to the

remaining 999 data sets ("control data sets"). For that, z-scores were calculated for each vertex using the formula $z_{vertex} = (d_{vertex} - \mu_{vertex}) / \sigma_{vertex}$ , whereas $d_{vertex}$ is the cortical thickness value of one vertex from the test data set, $\mu_{vertex}$ the mean value of that vertex from the control data sets and $\sigma_{vertex}$ the respective standard deviation. From the resulting z-score map, only the minimum value was saved (note that the present research question specifically addresses cortical *thinning*). The result was a reference distribution consisting of 1000 z-scores. Using this distribution, each vertex of an independent validation data set can be rated separately with respect to the reference population, by z-transforming each vertex using the above formula (see section "Validation").

*Method 2: Z-min statistic per data point, averaged across brain regions*

In method 1, a null distribution was calculated based on the most extreme values across the cortex. However, given that cortical thickness is nonuniformly distributed across the cortex physiologically (Fukutomi et al., 2018), potential atrophy will be hard to detect in physiologically thicker brain regions. Method 2 aimed to increase the biological plausibility of the previous method. While the same null distribution was used as in method 1, in method 2, data points were summarized across anatomically distinct brain regions, defined by the Desikan-Killiany atlas (Desikan et al., 2006). This atlas subdivides the cortical surface into 68 regions based on morphologic features ("labels", 34 on each hemisphere). For subsequent validation, statistical significance was determined for the synopsis of all vertices within each of the 68 regions, instead of for each vertex separately (see section "Validation").

*Method 3: Z-min statistic per brain region*

In spite of the increased biological plausibility in method 2, that procedure was still based on one common null distribution from the most extreme values of the cortex. In method 3, this was corrected by calculating distinct null distributions for each of the 68 Desikan-Killiany-labels. For that, the permutation procedure described in method 1 was repeated, however now z-maps were calculated using the formula $z_{vertex} = (d_{vertex} - \mu_{Label}) / \sigma_{Label}$, whereas $z_{vertex}$ was the z-score for a vertex of the test data set, $d_{vertex}$ is the observed cortical thickness value for that vertex from the test data set, $\mu_{Label}$ is the mean value of the respective *label* from the control data sets and $\sigma_{Label}$ its respective standard deviation. On each iteration, the minimum z-score of all vertices composing one common label was saved, such that the result was a 68x1000 matrix, providing a null distribution for each label (Figure 7.1.1d). With these null distributions, each brain region can be rated separately with respect to the reference population, by converting the cortical thickness data into z-scores using the formula $z_{Label} = (d_{Label} - \mu_{Label}) / \sigma_{Label}$ (Figure 7.1.1e).

*Method 4: Z-score per brain region*

Finally, in method 4, null distributions were generated based on *averaging across all vertices* from each brain region instead of using each label's most extreme values, as in method 3. Mean values were calculated for each brain region of the test data set to derive null distributions. These null distributions were generated in analogy to method 3, whereas the formula $z_{Label} = (d_{Label} - \mu_{Label}) / \sigma_{Label}$. Similar to method 3, also in method 4, each brain region can be rated separately with respect to the reference population, by converting the cortical thickness data into z-scores using the formula $z_{Label} = (d_{Label} - \mu_{Label}) / \sigma_{Label}$.

### *Validation*

To validate and compare the proposed methods, specificity and sensitivity were calculated. These measures were calculated for each vertex (method 1) or each label (methods 2–4) separately. For that,

the 113 data sets ("validation data sets") from the 1113 HCP data sets were used which had been spared for the generation of null distributions (see section "Subjects"). Statistical inference tests based on the null hypothesis of no atrophy for a given validation data set were carried out using the above-generated null distributions. For each vertex/label, the number of values of the null distributions that were lower than the observed cortical thickness values in a given validation data set were counted. Dividing this sum by the number of permutations (n = 1000) yielded FWER-corrected p-values ($p_{FWER}$) (Blair & Karniski, 1993; Westfall et al., 1993). Vertices/labels with $p_{FWER} < 0.05$ were considered to indicate lower cortical thickness values than would not be predicted by chance and therefore labeled as "atrophic". In method 2, since data points were summarized within each label, a label was defined as "atrophic" if a certain percentage of its vertices showed $p_{FWER} <= 0.05$. Different percentages were tested (1%, 5%, 10%, 20%, 30%, 40%, 50%). Given that all of these thresholds yielded similarly poor results, hereafter only the results for one threshold (5%, arbitrary choice) are provided. The data for the other thresholds are provided in Supplementary Figure 1 and Supplementary Table 1.

*Specificity*

Specificity defines the rate of true negatives, i.e. the share of patients which are correctly diagnosed as not having the condition of interest (here, "no atrophy"). The validation data set was used to calculate specificity, assuming that – given this data set was a random selection of a data set of healthy young subjects with no history of psychiatric/neurologic disorders – the validation data set can be labeled as non-atrophic. Each of the four methods was applied to all of the 113 validation data sets and specificity was defined as the percentage of vertices (method 1)/labels (methods 2, 3, 4) which were not classified as significantly atrophic. This procedure was repeated for each validation data set separately. Mean and standard deviations of the specificity calculations were determined across all 113 data sets ("cumulative specificity").

To allow evaluation for distinct brain regions, in addition, specificity per atlas region was defined (for methods 2,3 and 4 only, since in method 1, no atlas regions were analyzed). This was done by calculating, per atlas region, the percentage of the 113 validation data sets which were not significantly classified as atrophic in that atlas region ("regional specificity").

*Sensitivity*

Sensitivity defines the rate of true positives, i.e. the share of patients which are correctly diagnosed as having the condition of interest (here, "atrophy"). Given no true atrophy was assumed in the validation data sets, atrophy was simulated: Different degrees of atrophy were simulated as follows (Figure 7.1.2): The original cortical thickness data (each vertex) was multiplied by a number between 0 and 1 (e.g. multiplication by 0.9 represents simulated atrophy of 10%, etc.). For each of the 113 validation data sets, atrophy was simulated from 1% to 100% in steps of 1 percentage points (p.p.). Then, each of the four methods was applied to all of the simulated data sets. For each method and degree of atrophy, sensitivity was calculated separately. Cumulative sensitivity was defined as the percentage of vertices (method 1) or labels (methods 2, 3, 4) which were classified as significantly atrophic, summarized across all 113 data sets ("cumulative sensitivity"). Sensitivity across methods was compared using the degree of atrophy required to achieve cumulative sensitivity of 80% ("cumulative sensitivity threshold"). Note that less sensitive methods will require more pronounced atrophy, therefore higher cumulative sensitivity threshold, in order to detect atrophy.

To allow evaluation for distinct brain regions, additionally, sensitivity per atlas region was defined for each degree of atrophy (for methods 2,3 and 4 only, since in method 1, no atlas regions

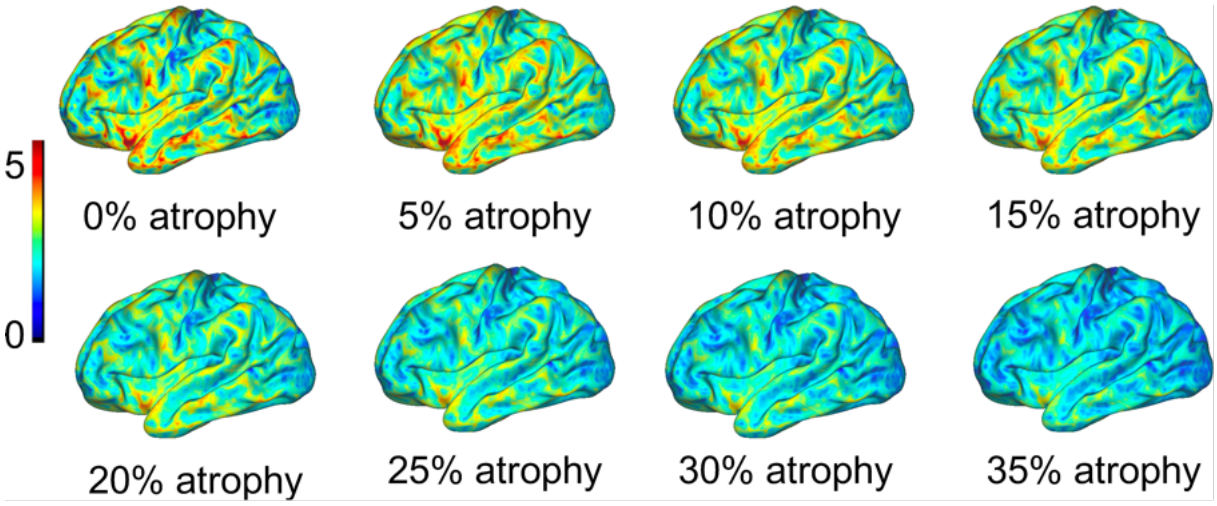



Figure 7.1.2 Atrophy was simulated for sensitivity calculations as follows: The original cortical thickness map from each of the subjects from the control population ("0% atrophy") was multiplied by values ranging between 0 and 1. Multiplication by lower values indicate higher degrees of simulated atrophy. For example, multiplication by 0.9 simulates 10% atrophy, multiplication by 0.8 20% atrophy, etc. In the present study, atrophy was simulated between 1% and 100% in steps of one percentage points. This allows to assess sensitivity by the degree of simulated atrophy. In this Figure, coloring indicates cortical thickness in millimeters.

were analyzed). This was done by calculating, per atlas region, the percentage of the 113 validation data sets which were significantly classified as atrophic in that atlas region ("regional sensitivity").

Note that although cortical thickness was simulated at consistent rates throughout the cortex (which is not how cortical thinning occurs in aging or pathology (Chouinard-Decorte et al., 2014; Fjell et al., 2015; Shaw et al., 2016)), evaluation was performed for each vertex/label independently. Therefore, the proposed methods are fit to analyze also diffuse patterns of cortical thinning.

### 7.1.4 Results

#### *7.1.4.1 Specificity*

Table 7.1.1 summarizes the cumulative specificity calculations for each method. Methods 1 and 2 showed ideal specificity (100%,+/–0 p.p.), such that these methods classified no vertex (method 1)/label (method 2) as significantly atrophic. Method 3 had a mean specificity of 98.9% (+/–1.3 p.p.), and method 4 was less specific with a mean of 93.6% (+/–2.0 p.p.). Figure 7.1.3 shows the regional specificity profiles evaluated across all 68 atlas regions. While the most specific method (method 2, red dashed line) yielded 100% specificity for each label, method 3 showed relatively constant specificity across brain regions except for a slight drop for the right lingual gyrus. Method 3 showed specificity of 100% for almost all labels on the right hemisphere (notice however a slight drop for the right lingual gyrus), while the values were slightly lower for the labels on the left hemisphere. Finally, method 4 (golden dashed line) showed notably lower values throughout all labels as compared to methods 2 and 3.

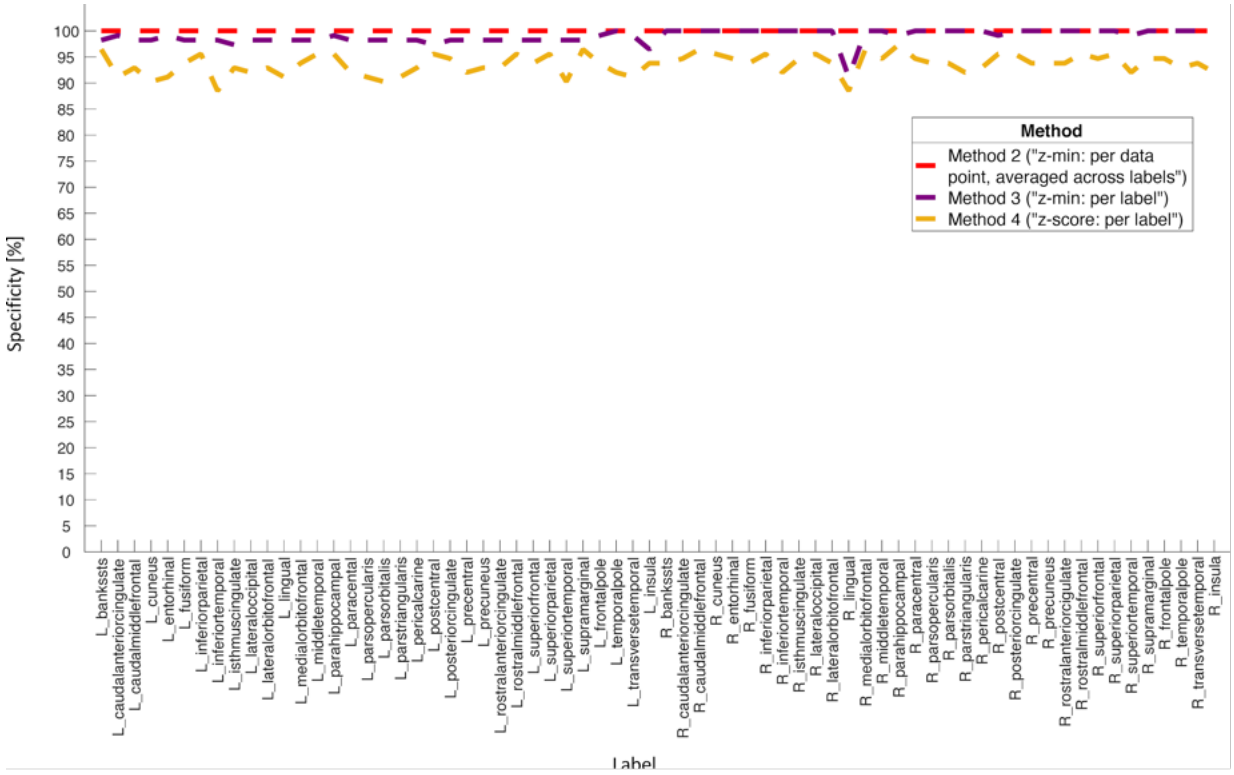



Figure 7.1.3 Comparison of regional specificity profiles between methods 2–4. The statistically most conservative approach (method 2, "z-min: per data point, averaged across labels", red dashed line) yielded ideal specificity for all brain regions, i.e. it correctly assigns "no atrophy" in 100% of cases. The less conservative method 3 ("z-min: per label", purple dashed line) also showed specificity of 100% for many brain regions, but had some drops, e.g. for the right lingual gyrus. The most liberal approach, method 4 ("z-score: per label", golden dashed line) yielded lower specificity for all brain regions. Note that method 1 ("z-min: per data point") is not shown here because it does not allow for labelwise assessment. See also Table 9.1.1 for the cumulative specificity values for each method.

Table 7.1.1 Specificity calculations for the four tested methods.

| | Method 1 ("z-min: per data point") | Method 2 ("z-min: per data point, averaged across labels") | Method 3 ("z-min: per label") | Method 4 ("z-score: per label") |
|---|---|---|---|---|
| Mean specificity (across subjects) | 100% | 100% | 98.9% | 93.6% |
| Standard deviation (across subjects) | 0.0 p.p. | 0.0 p.p. | 1.3 p.p. | 2.0 p.p. |

Abbreviation(s): p.p. = percentage point(s)

### *7.1.4.2 Sensitivity*

Figure 7.1.4 illustrates the cumulative sensitivity profiles for each method relative to the degree of simulated atrophy. The horizontal dashed line denotes sensitivity at 80% (cumulative sensitivity threshold), which was used to compare the different methods. Table 7.1.2 summarizes these results:

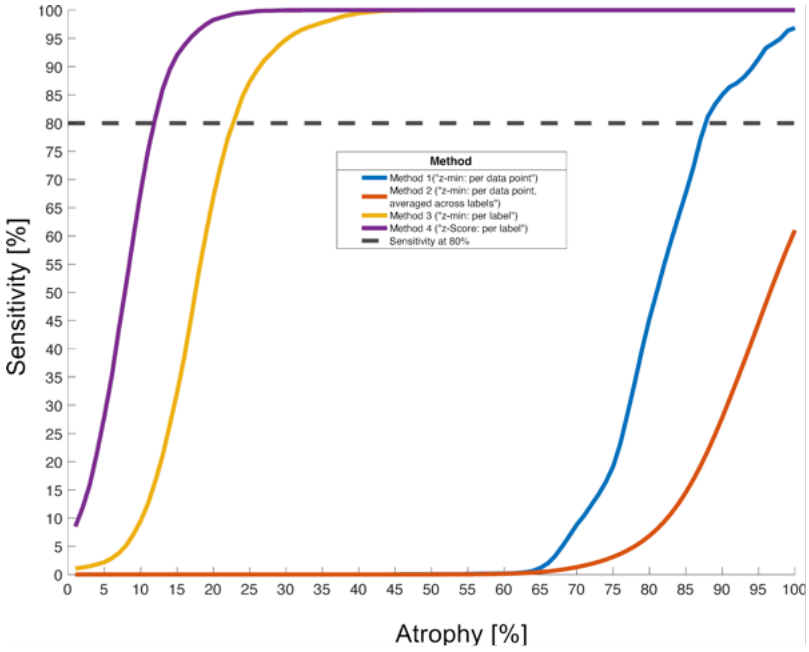



Figure 7.1.4 Cumulative sensitivity relative to the degree of simulated atrophy (across vertices/brain regions), comparison between the four tested methods. All methods detected atrophy more sensitive for more pronounced degrees of atrophy. However, the degree of atrophy the methods required to reach a given level of sensitivity differed. For example, in the current simulation, in order to detect atrophy in 80% of cases (black horizontal dashed line), method 4 ("z-score: per label", purple line) required only 12% atrophy, method 3 ("z-min: per label", golden line) 23%, method 2 ("z-min: per data point, averaged across labels", blue line) 88%, while method 1 ("z-min: per data point", red line) failed to detect atrophy in 80% of cases even for the highest possible degree of atrophy (100%). Compare also Table 9.1.2 for a summary of these results

Method 1 (red line) was extremely unsensitive, such that not even for the highest possible degree of atrophy (literally no brain) did this method detect atrophy in 80% of cases (i.e. cumulative sensitivity threshold not reached). Method 2 (blue line) yielded a cumulative sensitivity threshold for 88% simulated atrophy when a label was considered atrophic if 5% of its vertices had $p_{FWER} < 0.05$ (see methods).

Other tested thresholds for method 2 comprised 1% (cumulative sensitivity threshold for 84% simulated atrophy), 10% (90% atrophy), 20% (94% atrophy), 30% (98% atrophy), 40%/50% (did not reach 80% sensitivity for any degree of simulated atrophy (Supplementary Figure 1 and Supplementary Table 1). Method 3 (yellow line) was clearly superior (cumulative sensitivity threshold 23% simulated atrophy), and for method 4 an even lower value (12% simulated atrophy) was observed.

Table 7.1.2 Cumulative sensitivity thresholds for the four tested methods.

| | Method 1 ("z-min: per data point") | Method 2 ("z-min: per data point, averaged across labels") | Method 3 ("z-min: per label") | Method 4 ("z-score: per label") |
|---|---|---|---|---|
| **Degree of atrophy required for detection of atrophy in 80% of cases (cumulative sensitivity threshold)*** | not available | 88 % | 23% | 12% |

** Note that lower values of atrophy suggest more sensitive methods, since they detect less pronounced atrophy.*

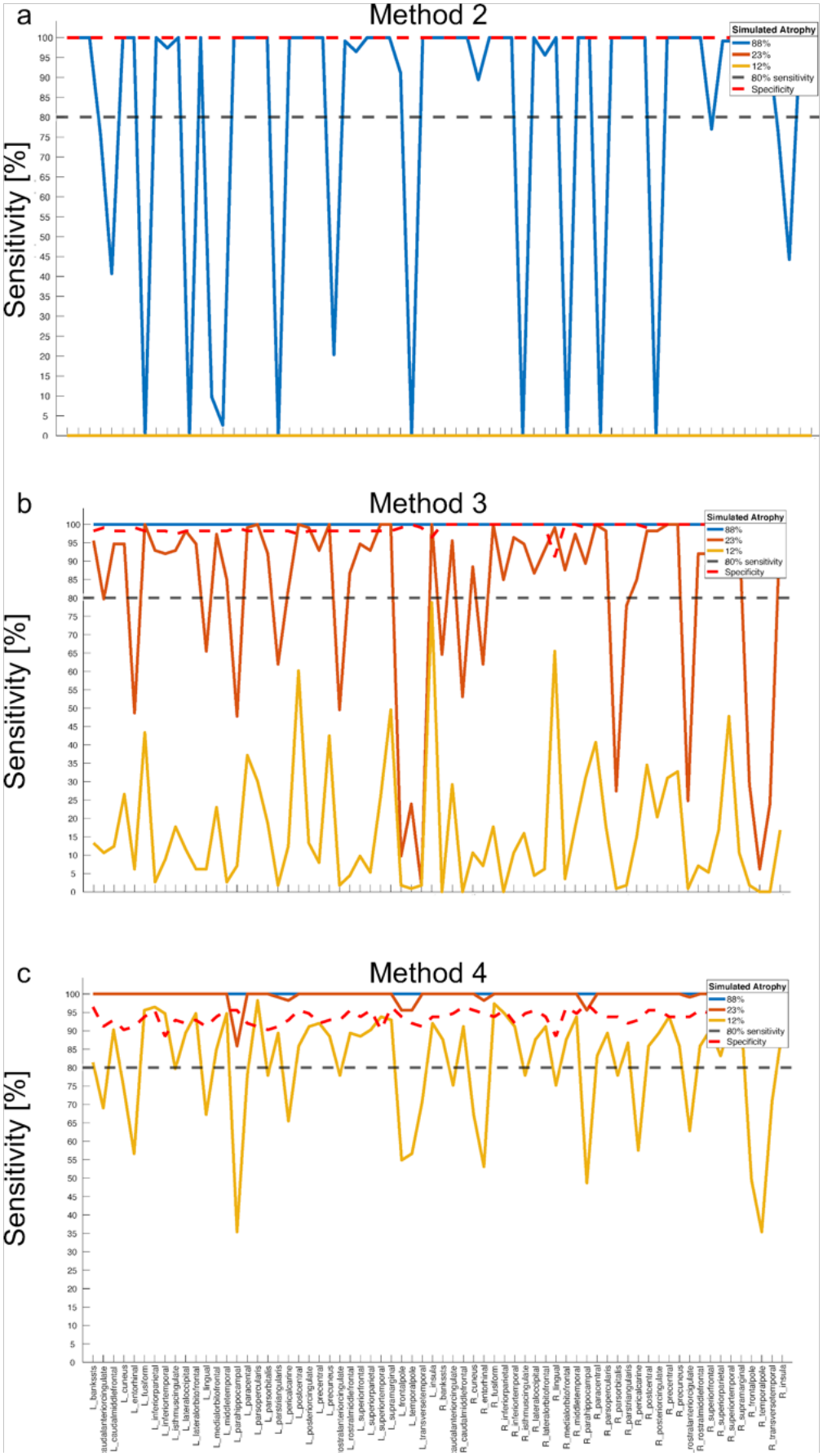



Figure 7.1.5 Regional sensitivity (per brain region) for each region's cumulative sensitivity threshold (i.e. the degree of atrophy each method required to detect atrophy in 80% of cases) for method 2 (a, "z-min: per data point, averaged across labels", method 3 (b, "z-min: per label") and method 4 (c, "z-score: per label"). The cumulative sensitivity threshold for method 2 was 88% atrophy (blue lines), for method 3 23% atrophy (red lines) and for method 4 12% atrophy (golden lines). The 80% sensitivity line is indicated by the gray dashed lines in each panel. In addition, regional specificity is plotted for each method (red dashed lines, compare also Figure 7.1.3). All methods detected atrophy more sensitively for more pronounced degrees of atrophy.

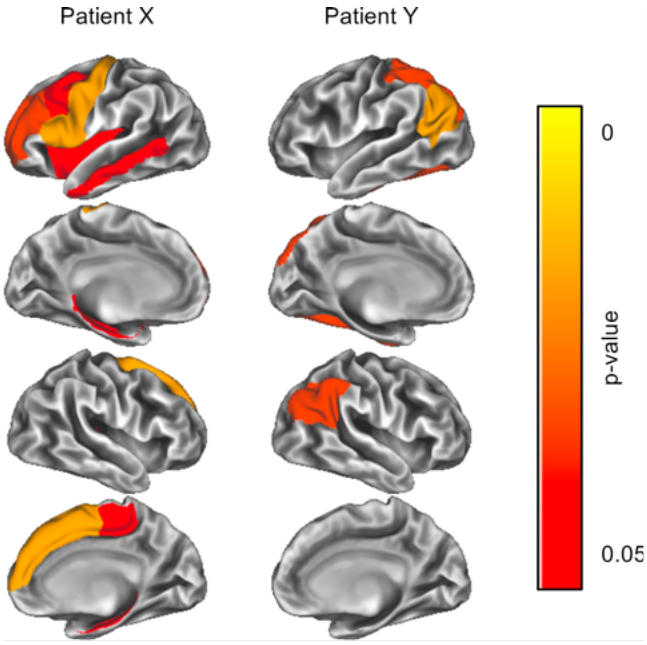



Figure 7.1.6 Exemplary result of analyzing a T1-weighted MRI data set with the current methods. For patient X, cortical thinning was simulated in frontal regions, for patient Y in more posterior regions. Method 3 was used to analyze the data. The emerging p-map indicates where cortical thinning likely occurs in each patient. Using the method proposed in this text, such maps can be created easily and are therefore simple to implement into clinical practice

Figure 7.1.5 shows the results of the regional sensitivity determination for methods 2 (Figure 7.1.5A), 3 (Figure 7.1.5b) and 4 (Figure 7.1.5c). To compare the methods, the regional sensitivity profiles are plotted for each method's cumulative sensitivity threshold (i.e. 88% atrophy for method 2: blue lines, 23% atrophy for method 3: red lines, 12% atrophy for method 4: golden lines). To enhance orientation, 80% sensitivity is indicated with a gray dashed line in Figure 7.1.5a–c. Additionally, regional specificity for each method is plotted (red dashed lines).

Figure 7.1.5a illustrates poor sensitivity of method 2, given it reaches sensitivity of > 0% for none of the cumulative sensitivity thresholds of the other methods. Additionally, the regional sensitivity profile for its own cumulative sensitivity threshold (88% simulated atrophy) shows strong variations across labels. Method 3 (Figure 7.1.5b) is clearly superior: while the variations for its own cumulative sensitivity threshold (23% simulated atrophy) are less pronounced as compared to method 2, it yields perfect (i.e. 100%) sensitivity for the cumulative sensitivity threshold of method 2. However, no region

reaches 80% sensitivity for the cumulative sensitivity threshold of method 3. Finally, method 4 (Figure 7.1.5c) is the most sensitive of the tested methods. It yields almost perfect regional sensitivity for the cumulative sensitivity thresholds of methods 2 and 3, and the regional sensitivity profile for its own cumulative sensitivity threshold (12% simulated atrophy) shows less variations than the other methods. Note however the relatively low specificity (red dashed line) of this method as compared to the others.

Nevertheless, it is evident from Figure 7.1.6 that there are regional variations for the cumulative sensitivity thresholds for each method. Supplementary Table 2 lists the labels which show less regional sensitivity than 80% for each method and its respective cumulative sensitivity threshold. For example, for method 3, among the brain regions that yielded least sensitivity for that method's cumulative sensitivity threshold (23% atrophy) are, on the left hemisphere, parahippocampal gyrus (49.56% sensitivity), temporal pole (23.89% sensitivity), frontalpole (9.73% sensitivity), temporalpole (23.89% sensitivity) and transversetemporal gyrus (1.77% sensitivity), and on the right hemisphere, parsorbitalis (27.43% sensitivity), rostralanteriorcingulate (24.78% sensitivity), frontalpole (29.20% sensitivity), temporalpole (6.19% sensitivity) and transversetemporal gyrus (23.89% sensitivity).

### 7.1.5 Discussion

The goal of this study was to develop a method which allows to rate a single patient's cortical thickness data and identify atrophy sensitively and specifically with respect to a control population. This study was motivated by the many previous reports which have found pronounced associations of cortical thinning with the diagnosis/progression of diverse neurological and psychiatric conditions. In addition, given that

different pathologies present with different patterns of cortical thinning, another goal was to allow the evaluation of cortical thinning for distinct brain regions. To provide such a resource, a reference system was developed by generating population-based distributions of expected cortical thickness data, both for the entire cortex as well as for distinct brain regions. 1000 data sets from young and healthy participants were used to generate expected population null distributions using a permutation procedure. To assess statistically significant cortical thinning (i.e. atrophy), different methods were tested and compared using sensitivity and specificity calculations for the entire cortex ("cumulative") as well as for distinct brain regions ("regional"), calculated from 113 additional subjects. The statistically most stringent methods were based on one common null distribution for all brain regions, which showed ideal specificity but poor sensitivity. Other methods were based on distinct null distributions for different brain regions, which increased sensitivity but decreased specificity. However, when generating distinct null distributions for different brain regions based on the most extreme values within each label (method 3), the drop in cumulative specificity was only very subtle (98.9%), while cumulative sensitivity could still be detected at 80% for 23% simulated atrophy. Variations of regional differences were observed for some brain regions, but decreased for more pronounced degrees of atrophy.

These results emphasize that in order to sensitively detect cortical atrophy for individual patients, it is reasonable to create different null distributions for distinct brain regions. Cortical thickness is not spread uniformly across the cortex,(He et al., 2007) such that for example neurite density is higher for motor regions as compared to regions associated with higher cognitive functions (Fukutomi et al., 2018). Therefore, a single reference distribution to rate any cortex region is biologically implausible and will result in decreases of sensitivity, which was shown here in methods 1 and 2. Furthermore, with this approach, sensitivity is relatively constant for different brain regions, although regional variations are observed (Figure 7.1.5b).

One drawback of working with several null distributions for different brain regions as opposed to a common one is that specificity decreases, which was shown in methods 3 and 4. In method 3, a strategy was suggested to minimize this loss in specificity while maintaining a high level of sensitivity: The idea of method 3 was to generate null distributions for different brain regions based on the (minimally) most extreme values within each brain region across a control population, instead of working with averages across brain regions. With this strategy, atrophy could be detected in 80% of cases when the cortex was roughly three quarters of its original thickness. However, in cases where the clinician wishes to detect atrophy more sensitively, method 4 might be preferred – there, null distributions were generated from population *averages* (rather than from their most extreme values). In this study, that method could detect atrophy in 80% of cases already when the cortex was thinned by a factor of only 12% (also here, regional variations were observed, see Figure 7.1.5c). However, that approach would imply risking to detect false positives, given its lower specificity. Depending on the situation, the clinician can flexibly choose between more sensitivity or more specificity.

One limitation of the suggested reference system is that it was generated from a relatively homogenous control population of healthy young adults. However, cortical thickness declines even in physiological aging, such that the comparison of an elderly individual to that reference group will result in more pronounced atrophy detection, which would not necessarily have to be pathologic (Fjell et al., 2015; Shaw et al., 2016). Nevertheless, given that the regions that exhibit cortical thinning differ in physiological and pathological aging (for example, atrophy of brain regions such as the precuneus and the inferior temporal region can be indicative of early signs of dementia (Lee et al., 2019)), it is still possible to detect such potential pathologic signatures using the method proposed here. This is possible because the reference system suggested herein was generated and evaluated for different brain regions separately. This allows to rate different brain regions independently, such that different atrophy

patterns can be identified. Figure 7.1.6 illustrates this: For patient X, atrophy was simulated in frontal areas, for patient Y in more posterior regions. Using method 3, the resulting p-map indicates where cortical thinning occurred for that patient. Such maps can be generated easily with a given patient's T1 MR image using the procedure proposed here, and are therefore easy to implement into clinical practice.

The atlas used in this work was the Desikan-Killiany atlas, a brain atlas defined by morphologic features of the cortex and is therefore surface-based. This is an important feature because cortical thinning is modified by genetic components (Chouinard-Decorte et al., 2014; Fjell et al., 2015), and such genetic patterns yield high resemblance to surface-based features (Chen et al., 2013). Additionally, patterns of genetic overlap seem to be coarse-grained across the human cortex (current optimal solutions suggest between 9 and 12 labels per hemisphere (Chen et al., 2013; Chouinard-Decorte et al., 2014)), such that the Desikan-Killiany atlas (34 labels per hemisphere) allows a more fine-grained resolution than proposed by genetic commonalities. However, especially in early pathology, cortical thinning may be more localized, such that future work should investigate the benefit of using a more fine-grained atlas for such cases. Furthermore, a more fine-grained atlas might also help to enhance regional sensitivity of those brain regions which showed poor sensitivity with the Desikan-Killiany altas (such as the left frontalpole as well as the left and right transversetemporal gyri). The evaluation of these regions with the current method and atlas should be evaluated with caution given their lower sensitivity.

Finally, the current reference system allows to progress-monitor an individual's condition: given the composition of the reference standard does not change, any potential changes between two measurement time points can be more likely attributed to changes in the individual. Finally, it should be emphasized that atrophy was only simulated in this study, and it is subject to future work to validate the present simulations with real data. It will also be necessary to show that the system is applicable to data acquired from different types of MR scanners and sequence parameters (here, data from a 3 Tesla MR scanner with optimized parameters for T1 imaging were analyzed).

### 7.1.6 Conclusions

Taken together, the here suggested reference system can be used for sensitive and specific detection of cortical atrophy for distinct brain regions (defined by the Desikan-Killiany atlas) for age groups comparable to the reference population (22-40 years), which allows to detect differential patterns of cortical thinning. However, some brain regions are detected less sensitively such that those regions should be evaluated with care. The method should therefore be further validated with data from different pathologies and using different atlases. Although distinct reference systems for different age groups will further help to establish this method in clinical practice, the current method already allows to rate elderly individuals, however these cases should be treated with caution given the risk of detecting false positives due to effects of physiological aging. However, progress-monitoring of elderly individuals is possible with the current system if the individual is compared to its own ranking within the control population for each measurement time point. Therefore, the tool proposed in this work represents a first step of the translation of cortical thickness measures into clinical practice.

### 7.1.7 Declarations

#### *7.1.7.1 Ethics approval and consent to participate*

All study procedures of the Human Connectome Project protocol were approved by the Institutional Review Board at the Washington University in St. Louis.

#### *7.1.7.2 Consent for publication*

The (sole) author declares her consent for publication.

#### *7.1.7.3 Availability of data and material*

All data used in this study are freely and openly available for scientific interrogations from the Human Connectome Project. Researchers can access them online at https://db.humanconnectome.org/app/template/Login.vm;jsessionid=A3E03522D3DEC91B2D2A09FB80CCE6CF.

#### *7.1.7.4 Competing interests*

The author declares no competing or conflicts of interests.

#### *7.1.7.5 Funding*

M. T. is supported by a grant from the German Multiple Sclerosis Society (Deutsche Multiple Sklerose Gesellschaft, DMSG).

#### *7.1.7.6 Authors' contributions*

M. T. is the sole author of this work. She developed the idea for this work and is responsible for data analysis and statistics, conception of the manuscript, generating figures, tables and any other work related to this manuscript.

#### *7.1.7.7 Acknowledgements*

This study was supported by the German Multiple Sclerosis Society (Deutsche Multiple Sklerose Gesellschaft, DMSG) (see "Funding"). Materials and support for the analysis were made available by the *Biomedical Imaging Group at Department of Psychiatry and Psychotherapy, University of Regensburg, Germany, led by Jens V. Schwarzbach, which the author would like to thank.*

## 7.2 Subproject 2: "Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use"

### 7.2.1 Abstract

**Introduction**: The majority of imaging studies in ALS infer group-level imaging signatures from group comparisons, as opposed to estimating disease burden in individual patients. In a condition with considerable clinical heterogeneity, the characterisation of individual patterns of pathology is hugely relevant. In this study, we evaluate a strategy to track progressive cortical involvement in single patients by using subject-specific reference cohorts.

**Methods**: We have interrogated a multi-timepoint longitudinal dataset of 61 ALS patients to demonstrate the utility of estimating cortical disease burden and the expansion of cerebral atrophy over time. We contrast our strategy to the gold- standard approach to gauge the advantages and drawbacks of our method. We modelled the evolution of cortical integrity in a conditional growth model, in which we accounted for age, gender, disability, symptom duration, education and handedness. We hypothesised that the variance associated with demographic variables will be successfully eliminated in our approach.

**Results:** In our model, the only covariate which modulated the expansion of atrophy was motor disability as measured by the ALSFRS-r ($t(153) = -2.533$, $p = 0.0123$). Using the standard approach, age also significantly influenced progression of CT change ($t(153) = -2.151$, $p = 0.033$) demonstrating the validity and potential clinical utility of our approach.

**Conclusion:** Our strategy of estimating the extent of cortical atrophy in individual patients with ALS successfully corrects for demographic factors and captures relevant cortical changes associated with clinical disability. Our approach provides a framework to interpret single T1-weighted images in ALS and offers an opportunity to track cortical propagation patterns both at individual subject level and at cohort level.

**Keywords:** Amyotrophic lateral sclerosis, Biomarkers, Neuroimaging, Cortical thickness, Longitudinal study, MRI

### 7.2.2 Introduction

Amyotrophic lateral sclerosis is a progressive neurodegenerative condition with considerable clinical heterogeneity and a long presymptomatic phase (Eisen et al., 2014; Lulé et al., 2020; Querin, Bede, et al., 2019). A record number of clinical trials have been conducted in 2020 including ground-breaking therapies aimed at specific genotypes (Miller et al., 2020). Despite the considerable advances in therapy development, current clinical trials in ALS continue to rely on functional rating scales, questionnaires, respiratory measures and survival as primary endpoints (Hiroshi Mitsumoto et al., 2014). Perplexingly, quantitative biomarker panels such as electrophysiology measures, bio- fluid assays and neuroimaging are not integrated in clinical trials designs (Rangariroyashe Hannah Chipika et al., 2019; Proudfoot et al., 2019; Christina Schuster et al., 2015). While the utility of quantitative imaging in ALS has been repeatedly demonstrated in tracking longitudinal changes, aiding diagnostic classification, ascertaining presymptomatic changes, describing genotype- associated signatures, the transition from academic use to clinical applications is yet to materialize (Rangariroyashe H Chipika, Siah, et al., 2020; Christina Schuster et al., 2016). A range of promising wet and dry biomarkers have been evaluated in ALS (Blasco et al., 2018; Devos et al., 2019; Floeter & Gendron, 2018), but the specific appeal of imaging data is that

it can be non-invasively acquired and quantitatively interpreted, and it reflects directly on pathological changes in the nervous system. These considerations coupled with the widespread availability of MRI scanners, open-source soft- ware, large data repositories, and the fact that the majority of suspected ALS patients undergo brain imaging as part of their diagnostic work-up make a compelling argument to use imaging as means of identifying and tracking pathological change in ALS. Over 100 structural MRI studies have been published in ALS to date, many of them longitudinal, using a variety of methods from morphometry to cortical thickness analyses through various volumetric, subcortical and atlas-based parcellation approaches (Bede, Querin, et al., 2018; Menke et al., 2018; Omer et al., 2017; Trojsi et al., 2018). Early ALS imaging studies made group-level, cross-sectional observations by contrasting clinically and/or genetically characterised symptomatic patient groups to healthy controls (Bede et al., 2016; Christidi, Karavasilis, Rentzos, et al., 2018; Trojsi et al., 2020). These studies were gradually superseded by large longitudinal studies (Agosta et al., 2009; Bede & Hardiman, 2018; Floeter & Gendron, 2018; Meier et al., 2020; Menke et al., 2018), presymptomatic studies (Lulé et al., 2020; Christina Schuster et al., 2015) and studies interpreting individual imaging data sets (Foerster et al., 2014; Grollemund et al., 2019a; Querin, El Mendili, et al., 2018a; Christina Schuster et al., 2017; Welsh et al., 2013). Common grey matter metrics, such as regional cortical thickness (CT), is typically evaluated in group comparisons. However, alternative strategies have been proposed to interpret cortical thickness values in individual patients with respect to carefully chosen reference groups (Tahedl, 2020). The objective of this study is to evaluate the utility of cortical measures in tracking longitudinal changes in individual patients, using demographically matched control data while minimising variance associated with age and gender. Instead of implementing commonly used longitudinal models (Reuter et al., 2010) which have been repeatedly applied to ALS cohorts (Bede & Hardiman, 2018; van der Burgh et al., 2020), we explore a longitudinal approach which tracks the expansion of cortical atrophy over time. To demonstrate the utility of this strategy, we contrast our method with the gold- standard model. Our hypothesis is that with that longitudinal cortical pathology can be meaningfully tracked in individual patients over multiple timepoints with reference to external datasets, thereby simulating a clinical trial scenario.

### 7.2.3 Methods

#### *7.2.3.1 Study participants*

Sixty-one patients with amyotrophic lateral sclerosis (ALS) and 125 healthy controls (HC) were included in this study in Trinity College Dublin. All participants provided informed consent in accordance with the ethics approval of the Ethics (Medical Research) Committee—Beaumont Hospital, Dublin, Ireland. Participating ALS patients had 'probable' or 'definite' ALS according to the El Escorial criteria (Brooks et al., 2000). Each patient had at least three scans; 27 patients had three and 34 patients had a total of four scans with a uniform inter-scan interval of four months. The 125 healthy controls (HC, 62 males) were scanned with the same scanning parameters but were not invited for follow-up imaging. Core demographic data, including age, gender, handedness, symptom onset (for patients) and years of education, were care- fully recorded for each participant. Additional information with regard to comorbidities, previous surgeries, medications, alcohol intake, smoking, occupation and family history was also carefully recorded. Inclusion criteria included the ability to lie supine in the scanner for the duration of data acquisition. Exclusion criteria for participating patients included implanted non-MRI-compatible devices, comorbid neurological diagnoses, psychiatric conditions, previous head injuries or established vascular risk factors. Exclusion criteria for healthy controls included any neurological or psychiatric diagnosis, previous head injuries, prior neuro- surgery, claustrophobia or established vascular risk factors.

#### *7.2.3.2 Neuroimaging*

*Magnetic resonance imaging*

T1-weighted images of 61 patients with ALS and 125 HCs were acquired on a 3 T Philips Achieva system with an eight-channel receive-only head coil using a 3D inversion recovery-prepared spoiled gradient-recalled echo (IR-SPGR) pulse sequence with following parameters: TR/TE = 8.5/3.9 ms, TI = 1060 ms, field of view (FOV): 256 × 256 × 160 mm, spatial resolution: 1 mm3, flip angle (FA) = 8° and SENSE factor = 1.5. FLAIR images from each participant were individually assessed for vascular white matter lesion load. FLAIR images were acquired in axial orientation using an inversion recovery turbo spin echo (IR-TSE) sequence: FOV = 230 × 183 × 150 mm, spatial resolution = 0.65 × 0.87 × 4 mm, 30 slices with 1-mm gap, TR/TE = 11,000/125 ms, TI = 2800 ms, 120° refocusing pulse, with flow compensation and motion smoothing and a saturation slab covering the neck region. A total of 651 external reference T1w data sets (321 males) were used from the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) repository. T1w MPRAGE data from the Cam- CAN data base were acquired on a 3 T Siemens Magnetom TrioTrim system at the University of Cambridge with the following sequence parameters: FOV = 256 × 240 × 192 mm, spatial resolution 1 mm3, TR/TE 2.25/2.99 ms, TI 900 ms, FA 9° and GRAPPA factor 2.

*Pre-processing*

For each subject, image segmentation and surface reconstruction were performed using FreeSurfer v.6.0 (Dale et al., 1999; B Fischl et al., 1999; Bruce Fischl, 2012). Data were subsequently converted into the grayordinate- based CIFTI format using the *ciftify* package (Dickie et al., 2019) for further analyses. Subject-specific cortical thickness (CT) maps at a resolution of 32 k meshes per hemisphere were used for subsequent analyses. To reduce the number of statistical comparisons and to enhance interpretability, each cortical thickness map underwent atlas-based parcellation. As our objective was to track cortical alterations over time with

high spatial accuracy, we opted for the 1000-region atlas of the Schafer-Yeo-7-network solution (Schaefer et al., 2018; Yeo et al., 2011) which is a high-resolution anatomical map with equally sized cortical labels. We refer to the atlas-defined small cortical regions as ‘mosaics’.

#### 7.2.3.3 Quality control (QC): cross-sectional group comparison

In an initial quality control step, we performed standard cross-sectional analyses by contrasting the 61 ALS patients to the 125 controls from our centre merged with the 651 HC from the Cam-CAN to demonstrate the validity of using external control groups. Accordingly, we initially performed a one-way analysis of variance (ANOVA), contrasting CT maps smoothed with a Gaussian kernel of 5 mm between the two groups. For the ALS patients, we only considered baseline data. We used false discovery rate (FDR) corrections with a voxelwise thresholding at 0.05 and 5000 per- mutations to account for alpha inflation and controlled for age and gender. The statistical comparisons were run within the MATLAB-based (The MathWorks, Natick, MA, USA) toolbox MRM (McFarquhar et al., 2016), which is specifically designed for handling multivariate neuroimaging data.

#### 7.2.3.4 Patch-wise interpretation of cortical atrophy

In this work, we aimed to evaluate an alternative strategy for assessing cortical thickness (CT) over time. In brief, we propose to quantify how many cortical regions are significantly atrophic with respect to age-/sex-matched controls. We refer to this approach as ‘mosaic approach’. The current gold-standard way to interpret cortical thickness is to assess CT differences based on raw CT values in millimetres. We refer to this approach as the ‘standard approach’. In order to appraise the advantages and drawbacks of our strategy, we contrast the two methods for each step of the statistical modelling. Values for the standard

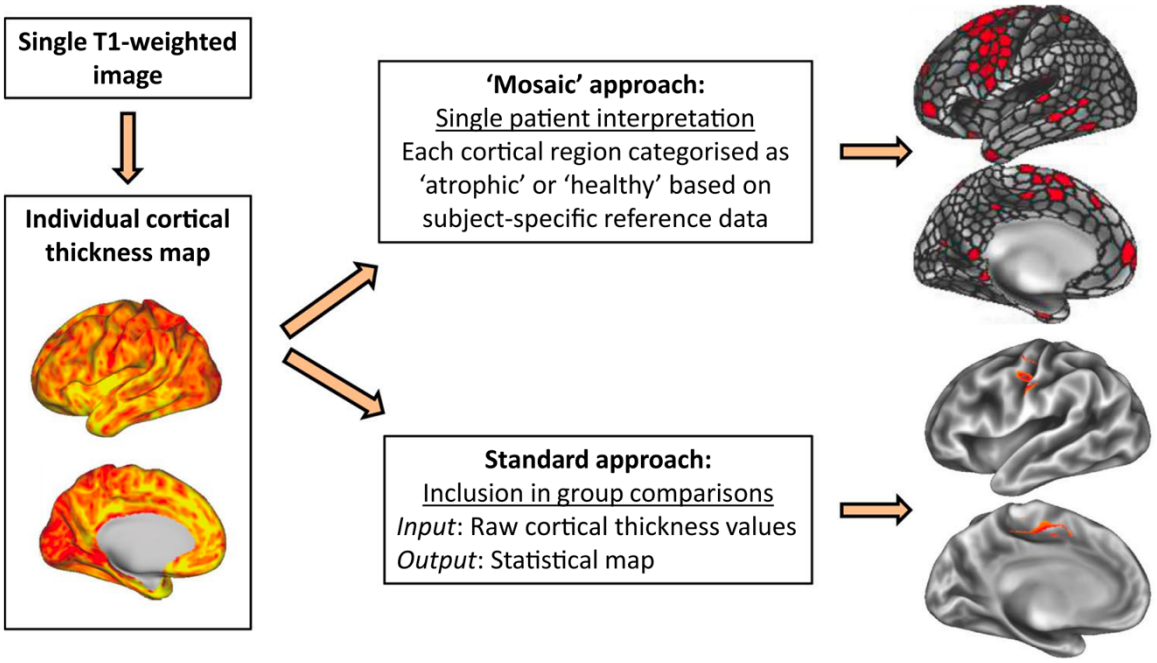



Figure 7.2.1 Method overview. The standard versus the 'mosaic' approach for interpreting corti- cal thickness (CT). The standard approach interprets cortical thickness values across the entire supratentorial cortex in statistical models such as group comparisons. In the 'mosaic' approach, cortical thickness is first averaged across 1000 atlas-defined tiles, and each region is rated with respect to an age-/sex-matched control group. p-values are calculated using nonparametric permuta- tion testing, and each cortical region is subsequently classified as 'healthy' or 'atrophic'.

approach are easily derived, since one standard output from the FreeSurfer pre- processing pipeline is CT maps, providing estimates of CT in millimetres.

*Definition of reference groups*

For each year of age X, we defined a subject-specific reference group which was comprised of HC subjects from the combined Dublin/Cam-CAN data set aged between [X–2; X + 2], separately for males and females. For example, the reference group for a 54-year-old male ALS patient at base- line included all men between 52 and 56 years from our own and the Cam-CAN data set. For that same patient, the reference group at 1-year follow-up slightly changed to include all men between 53 and 57 years. With this 'sliding-window' approach, we aimed to meticulously account for the effect of physiological ageing.

*Calculation of cortical thickness distributions for each reference group*

First, the cortical thickness profile of each control subgroup was established by calculating the raw CT values in millimetres for each of the 1000 cortical patches, z-scoring, and saving the values in a $n_{subgroup}$ × 1000 matrix. This allows the rating of a single patient's data set with respect to a demo- graphically matched reference group. Each patient's patch thickness was then z-scored by subtracting the subgroup's mean and dividing by the subgroup's standard deviation from that same cortical patch. We subsequently converted the resulting z-values to p-values, by exhaustive permutation testing to obtain family-wise error rate (FWER) corrected p-values. We considered $p_{FWER}$-values < 0.05 as significantly thin, i.e. atrophic with reference to demographically matched controls. Our analysis strategy is illustrated in Figure 7.2.1.

### *7.2.3.5 Statistics*

The central idea of this study was to test which predictors modified the progression of cortical thickness using the standard approach and the 'mosaic' approach. We hypothesised that randomness associated with the variables 'age' and 'sex' will be eliminated with the new approach, since correction for age and sex is inherent in our method. (Reference groups are age and sex matched.) In order to test this

assumption and accommodate the complex covariance structure associated with longitudinal data, we adopted a multilevel modelling approach. This was carried out within R (R Core Team, R Foundation for Statistical Computing, Vienna, Austria), using the 'nmle' package (Pinheiro et al., 2020). In multilevel modelling, the ultimate model which includes the factor time and all predictors is built step-wise, starting from an unconditional means model (with no repeated measure, testing only the grand mean), to a unconditional growth model (including the factor time—either as fixed or random effect—but no other predictors). These intermediate models are built so that the addition of more complexity can be directly assessed by comparing the model fit to the less complex precursor model. The predictors we included in the final model ('conditional growth') were age, gender, handedness, months since symptom onset, years of education and ALSFRS-r score. We note that we computed each model separately for the standard and the 'mosaic' approach.

*Unconditional means model (null model)*

The first model we considered within our multilevel approach was the unconditional means model (UMM). The UMM serves as a null model to compare subsequent, more complex models against. It does not consider any predictors or the repeated-measures factor time (i.e. all data points from all subject's follow-ups are treated as if they were cross- sectional data). As such, the UMM tests the grand mean of CT (standard approach)/number of thin patches (mosaic approach) against zero to reject the null hypothesis. We set up the UMM using a linear mixed-effects model within R's nmle package (Laird & Ware, 1982):

(7.2-1) UMM ~ lme(DV ~ 1, random ~ 1 | ID, data = X, method = 'ML'),

where DV is the dependent variable (either CT for the standard or number of thin patches for the new approach), which is modelled with no slope (note the ' ~ 1' following DV). Additionally, no random effects are being accounted for ('random ~ 1'). Note that ID, which is an identifier for the timepoint/session (i.e. between 1 and 4 for each patient), serves as nested variable, since in longitudinal analyses, data are per definition nested by individuals. X denotes the data tabular containing information on all the specified variables. The method we used for estimating the model was maximum likelihood, 'ML'. Our primary interest was the p-value of the fixed effect, i.e. the DV, which tested the hypothesis of the grand mean's intercept being different from 0. Additionally, for later analyses, we were interested in the log-likelihood (LL) of the model, since this should increase for better fits, i.e. subsequent models considering additional variables.

*Unconditional growth model (time fixed/time random)*

Since with the UMM, we could reject the null hypotheses of the DVs (for both the standard and the mosaic approach) being equal to zero, we could then proceed to add a slope to our data, i.e. consider time as a repeated measures fac- tor, which is referred to as the unconditional growth model (UGM). More specifically, we sought to test whether time could be more accurately modelled as a fixed or random effect for the two approaches. A better slope fit implies that each member of the group follows a similar pattern over time, e.g. patients experience the same degree of CT thinning over time, revealing limited individual variability associated with other random variables. Our hypothesis was that for the new approach, modelling time as a random variable versus a fixed one should not improve the model fit, since much of the variance (i.e. variance associated with age and gender) is already accounted for in calculating the number of thin patches. In contrast, modelling the standard approach with a random effect of time vs. a fixed effect of time should improve that model fit, since in the standard approach, randomness associated with age and gender is not inherently corrected for. Note that no other

predictor than time was included in the UGM analyses; rather, we assumed that age and gender were the most significant predictors of cortical thickness, such that accounting for the variance associated with these variables would be reflected in the hypotheses suggested above. The modelling of our UGMs with time fixed (TF, Eq. (7.2-2 and time random (TR, (7.2-3 Eq. 3 can be summarised as follows:

(7.2-2) Time fixed: UGM_TF ~ lme(DV ~ time, random ~ 1 | ID, data = X, method = 'ML'),

(7.2-3) Time random: UGM_TR ~ lme(DV ~ time, random ~ time | ID, data = X, method = 'ML').

For both approaches, we compared the UGM with the UMM, both for time fixed and time random. Additionally, we compared the two UGMs against each other, representing time fixed vs. random, using deviance statistics, which tests whether one version of the model is a better fit to the data than the other version. Within R's *nmle* package, this is simply realised with the following notation:

(7.2-4) model_comparison ~ anova(model1, model2).

*Conditional growth model*

Modelling time as both fixed and random variable improved the model fit in both the standard and the new approach compared to the UGM. However, when comparing the UGM with time as a random vs. time as a fixed effect, we only found improved fit in the standard approach. This suggested that in the mosaic approach, randomness is reduced com- pared to the standard approach, but is not fully accounted for. Accordingly, we next sought to test which additional variables would account for the remaining randomness of the two approaches. Therefore, we proceeded to test the 'full' model, i.e. the model with all predictors (i.e. age, sex, handedness, months since symptom onset, years of education and ALSFRS-r), which is referred to as conditional growth model (CGM). We set up that model as follows:

(7.2-5) CGM ~ lme(DV ~ time + age + sex + handedness + SymptomDuration + yearsOfEducation + ALS- FRS-r, random ~ time | ID, data = X, method = 'ML').

To assess potential improvements of model fit by adding these predictors, we ran deviance statistics comparing the CGM against the UGM_TR.

### 7.2.4 Results

#### *7.2.4.1 Demographics and reference groups*

The demographic profile of the ALS cohort is summarised in Table 7.2.1. The combined Dublin/Cam-CAN healthy control group comprised 776 subjects in total, 125 from Dublin and 651 from Cam-CAN; 383/393 males/females; and mean age in years (SD) = 55.08 (17.63). The youngest ALS patient was 41 years old at baseline and the oldest 83 years old (at follow-up scan 3). Therefore, we created 86 individualised reference groups (43 reference groups for males and 43 for females, the youngest ranging between years [43;47] and the oldest between years [81;85]. The smallest reference group was composed of 16 subjects (females between 81 and 85), the largest of 49 (males between 64 and 68) (median group size = 33, interquartile range [28.75;33]. Cross-sectional comparison of cortical thickness maps from ALS patients at baseline vs. the combined Dublin/Cam-CAN HC group confirmed cortical thinning in the pre- and paracentral gyri. This analysis was merely performed as a quality control

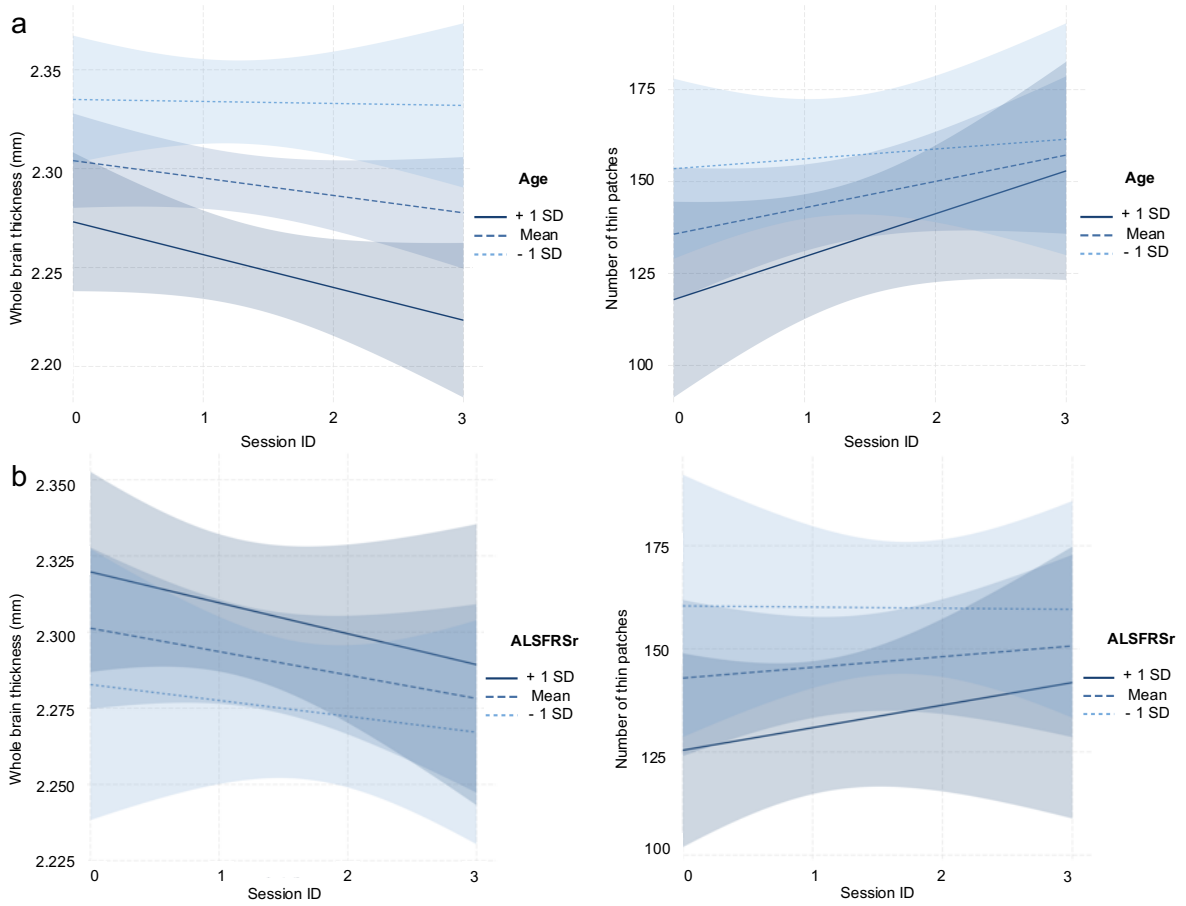



Figure 7.2.2 The effect of confound- ers (age, ALSFRS-r). Effects of the co-variates age (a) and ALSFRS-r (b) on the predic- tion of cortical thickness (CT). Age is a significant modulator of CT in the standard approach (left column in a; t = − 2.140, p = 0.034), but not in the new approach (right column in a; t = − 1.379, p = 0.170). ALSFRS-r is a significant modulator for both standard (left column in b; t = 2.543, p = 0.012) and new approach (right column in b; t = − 2.538, p = 0.012).

procedure, and since these results are consistent with the imaging signature of ALS centred on motor cortex atrophy, we proceeded to more complex modelling.

Table 7.2.1 The demographic profile of participating patients

| Timepoints 3/4 | Male/female | Age (base-line), y, mean (SD) | Symptom duration, mo, mean (SD) | Handedness Right/left | Years of education, y, mean (SD) | ALSFRS-r mean (SD) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Baseline | Month 4 | Month 8 | Month 12 |
| 27/34 | 43/18 | 60.20 (9.62) | 24.89 (13.67) | 52/9 | 12.84 (3.23) | 39.46 (6.38) | 36.92 (7.82) | 34.48 (7.95) | 33.44 (9.16) |

Abbreviations: ALSFRS-r = revised amyotrophic lateral sclerosis functional rating scale, SD = standard deviation

### *7.2.4.2 Unconditional means model*

In the UMM (see Eq. (7.2-1, we made a total of 217 observations across 61 groups. The session ID for each subject (i.e. the indicator of the scan timepoint for each patient, i.e. 'time') was modelled as a nested variable, such that each patient was its own group (having either three or four scans in total). The UMM showed that both approaches were significantly different from 0 (standard approach: t(156) = 159.643, p < 0.001; new approach t(156) = 14.421, p < 0.001). The log-likelihoods (LL) for both

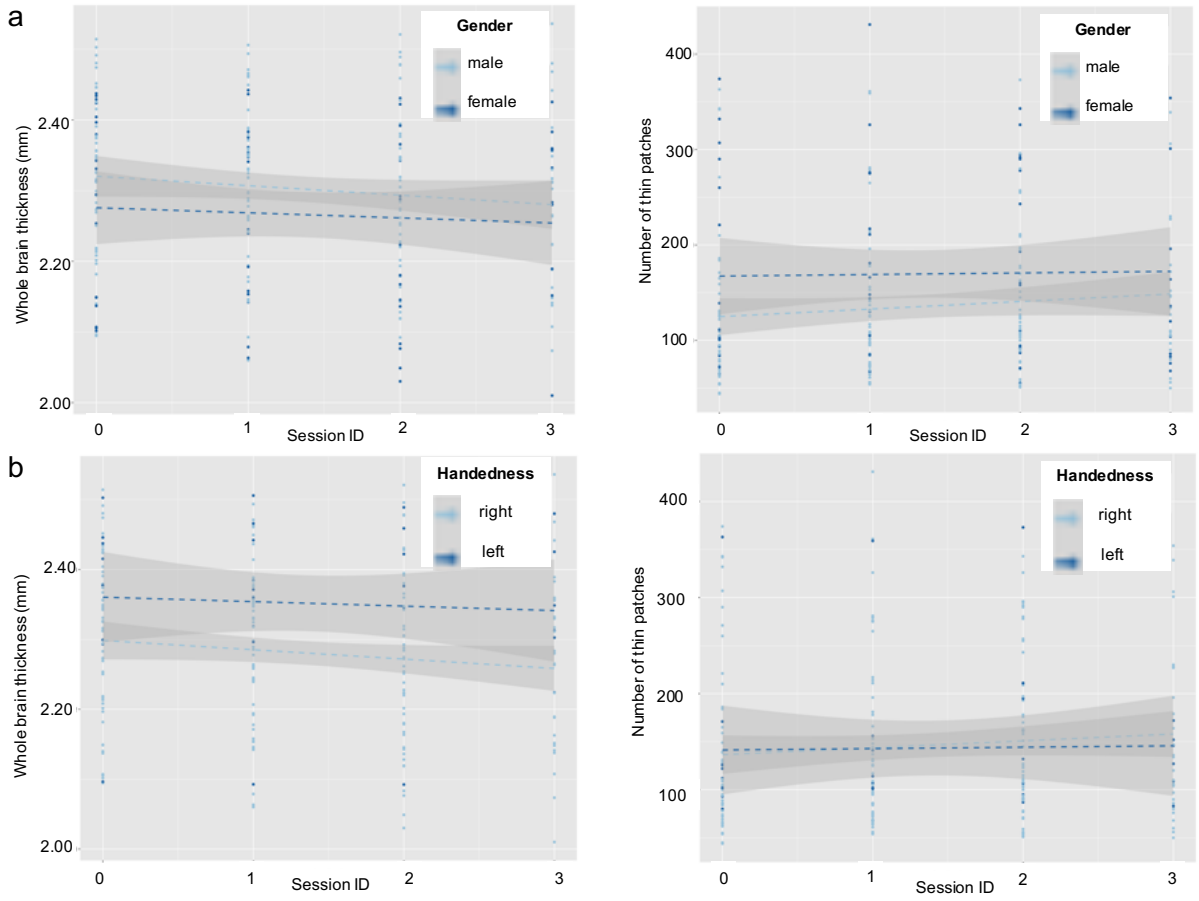



Figure 7.2.3 The effect of confounders (gender, handedness). Effects of categorical co-variates gender (a) and handedness (b) on the prediction of cortical thickness (CT). Neither gender nor handedness is significant modulators of CT progression for both approaches (gender—standard approach (left column in a): t = − 1.470, p = 0.147; gender—new approach (right column in a): t = 1.578, p = 0.120; handedness—standard approach (left column in b): t = 0.767, p = 0.447; handedness—new approach (right column in b): t = − 0.122, p = 0.904).

approaches were used to test any potential improvements of subsequent, more complex models: for the standard approach, the LL was 276.442 and for the new approach − 1173.996.

### *7.2.4.3 Unconditional growth model*

We set up an unconditional growth model (UGM_TF) by including the variable 'time' as fixed effect into the model (see Eq. (7.2-2. We found that for both approaches, time can be treated as a fixed variable (standard approach: t(155) = − 5.2103, p < 0.001; new approach: t(155) = 2.954, p = 0.0036). Note that the t-value for the standard approach was negative, indicating a decline in cortical thickness over time (Figure 7.2.2a, Figure 7.2.3a, Figure 7.2.4a). As expected, the t-value for the new approach was positive, suggesting an increase in the number of pathologically thin patches over time (Figure 7.2.2bFigure 7.2.3b, Figure 7.2.4b). The model fit assessed with the LL for both approaches was significantly higher as compared to the UMM (standard approach: LLUMM_TF = 276.442 vs LLUGM_TF = 289.040, p < 0.001; new approach: LLUMM_TF = − 1173.996 vs LLUGM_TF = − 1169.716, p = 0.0034). Next, we set up an UGM where we treated time as a random variable (UGM_TR, see Eq. (7.2-3, i.e. we allowed the slopes of CT/number of thin patches changes to vary among individuals, and tested whether that improved the model fit as compared to the UMM and the UGM_TF. For both models, we found evidence that modelling time as a random variable is statistically justified (standard approach: t(155) = − 4.792, p < 0.001; new approach: t(155) = 2.886, p = 0.0045). This was further confirmed by deviance statistics which suggested a superior model fit of the UGM_TR (LL = 291.516, p < 0.001) compared to the UMM in the standard approach. Similarly, the LL for the UGM_TR model in the new approach was higher than

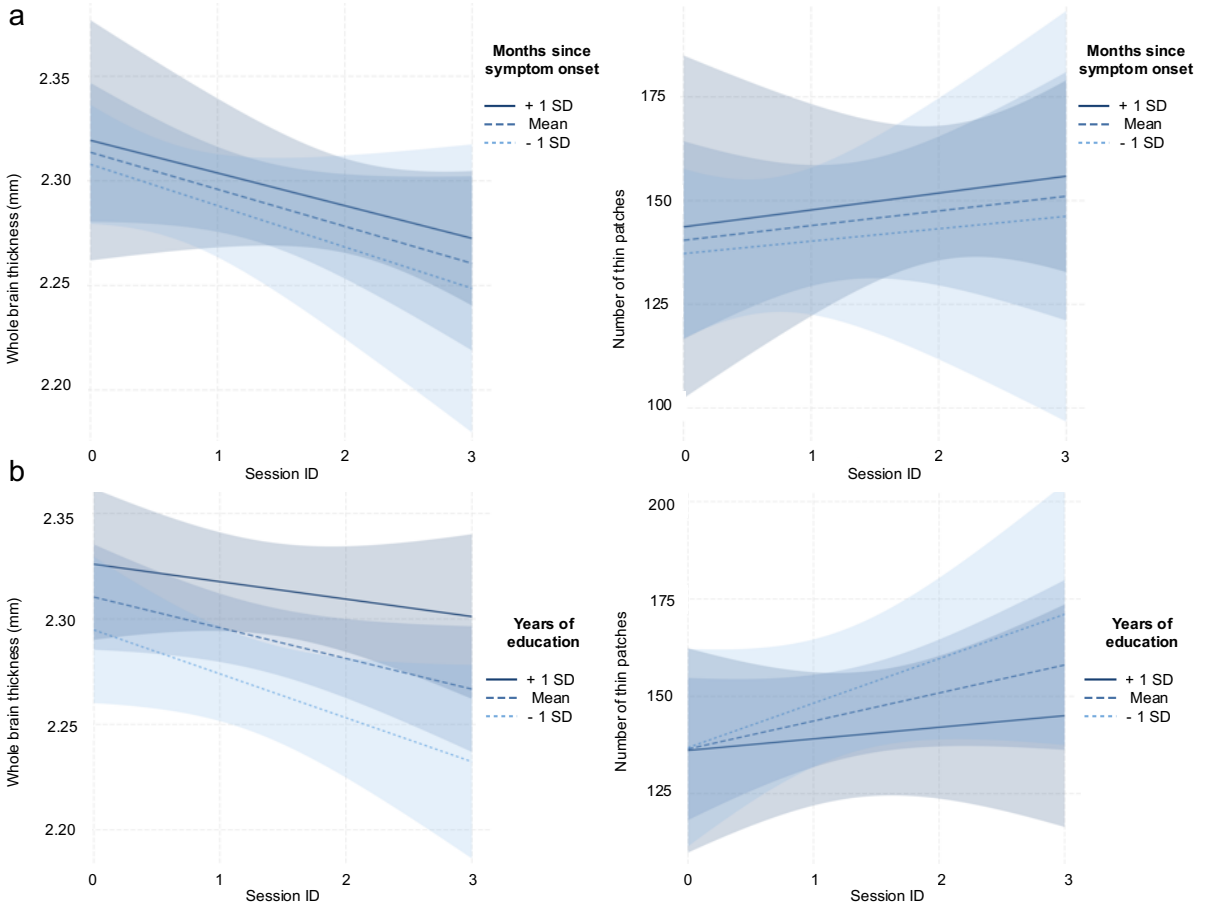



Figure 7.2.4 The effect of confounders (months since symptom onset, years of education). Effects of the two co-variates months since symptom onset (a) and years of education (b) on the prediction of cortical thickness (CT). Neither months since symptom onset nor years of education were significant modulators of CT progression for both approaches (months since symptom onset – standard approach (left column in a): t = 0.998, p = 0.320; months since symptom onset – new approach (right column in a): t = – 0.281, p = 0.779; years of education – standard approach (left column in b): t = 0.929, p = 0.357; years of education – new approach (right column in b): t = – 0.997, p = 0.323).

for the UMM (LL = – 1169.315, p = 0.0248). When comparing the fit of the UGM_TF vs. UGM_TR, there was no evidence for improved model fit for treating time as a random as opposed to modelling time as a fixed effect, although we observed a tendency for a better fit of UGM_TR for the standard approach (p = 0.0841; new approach: p = 0.6691).

### *7.2.4.4 Conditional growth model*

The above results suggested that although both CT decline and the number of pathologically thin patch behave similarly over time, there was still some randomness left in the data, with a tendency towards more randomness within the standard approach. Therefore, we included any observed predictors (age, gender, handedness, years of education, symptom duration and ALSFRS-r) into the full conditional growth model (see Eq. (7.2-5, to test which predictor impacts on the progression of the dependent variable. For the standard approach, we found evidence that both age (t(153) = – 2.151, p = 0.033) and ALSFRS-r (t(153) = 2.401, p = 0.0175) accounted for variability over time in terms of CT decline, while for the mosaic approach only the predictor ALSFRS-r (t(153) = – 2.533, p = 0.0123) was significant (Figure 7.2.3). Neither of the other categorical co-variates (i.e. gender and handedness, Figure 7.2.4) nor continuous co-variates (i.e. symptom duration and years of education, Figure 7.2.5) were significant modulators for the standard or the mosaic approach. Table 7.2.2 and Table 7.2.3 summarise the statistical details of the CGM. Deviance statistics revealed a better fit of CGM as compared to the UGM with time treated as a random variable (standard approach: LL = 299.873, p = 0.0051; new approach: LL = – 1163.273, p = 0.0337).

#### 7.2.4.5 *Propagation patterns in individual patients and at group level*

To illustrate the utility of our approach to track anatomical propagation, we have mapped the expansion of pathological cortical regions over time in individual patients. Our examples demonstrate the spread of cortical pathology from the initial timepoint where pathology is centred on pre- and paracentral gyri and gradually expand to include prefrontal and frontotemporal regions (Figure 7.2.5). While our intention was to develop a method to interpret single patient scans, our method can also be utilised to decipher disease-, phenotype- or genotype-specific propagation patterns at a cohort level. In Figure 7.2.6, we show the ability of our method to depict group-level traits by plotting the percentage of patients exhibiting cortical atrophy in specific 'patches' at each timepoint. This visualisation approach allows the illustration of group-level propagation patterns.

### 7.2.5 Discussion

In this study, we aimed to demonstrate the utility of a novel strategy for appraising cortical thickness in individual ALS patients using z-score derived subject-specific reference data sets. We propose to first parcellate the cortex into small but equally sized cortical regions, or 'patches', using high-resolution atlas-based segmentation. For each cortical patch, we calculate a normal distribution of cortical thick- ness (CT) values from age- and sex-matched control group, based on which individual CT profiles can be interpreted. In accordance with a previous simulation study (Eisen et al., 2014), we then converted that z-score rating into p-values using nonparametric permutation testing. As a result, a given cortical patch with $p < 0.05$ would be considered pathologically thin, i.e. atrophic. This strategy offers a high-resolution assessment of the cortex, categorising cortical regions of individual

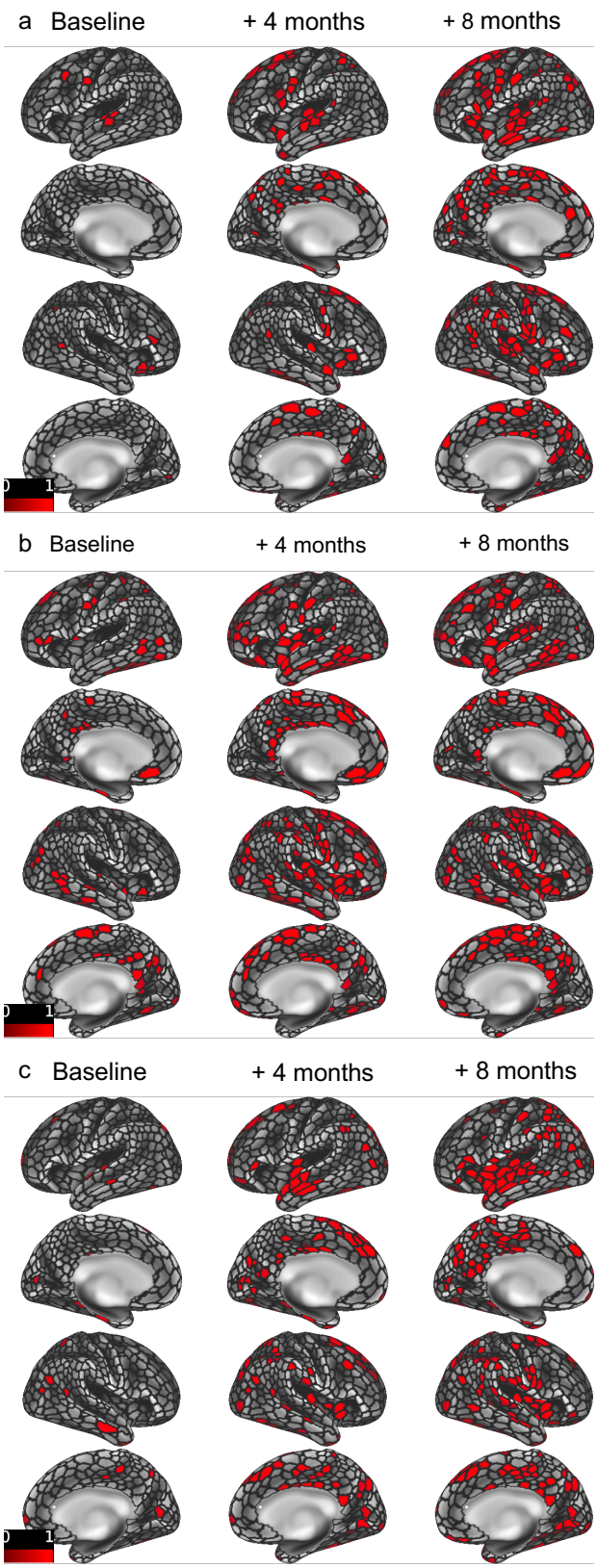



Figure 7.2.5 Expansion of pathological cortical regions over time in individual patients. Expansion of cortical degeneration over time in individual patients. Atrophic regions are defined according to the 'mosaic' method with references to demographically matched healthy controls and marked in red. For patients a) and b), disease burden is particularly evident in the motor cortex, while in patient c) atrophy is most evident in the insula. For all patients, atrophy gradually spreads anteriorly over time.

patients as 'atrophic' or 'healthy' compared to normative values in demographically matched controls (Figure 7.2.1). Given that variance associated with age and sex should be accounted for in our approach, our validation centred on the demonstration of the successful removal of that variance as opposed to the standard approach, i.e. measuring CT in millimetres and computing group comparisons based on these values. To achieve that, we used existing longitudinal data from 61 ALS patients, each having either three or four scans. We contrasted the progression of CT change (raw values, 'standard approach') versus the change in the number of pathologically thin patches over time ('mosaic approach'). We gradually increased the complexity of our model, accounting for potential modulators of the two dependent variables. We found that in both approaches, DV change can both be modelled as a fixed or random effect, which suggests that although in general, the progression pattern of both

Table 7.2.2 Conditional growth model for standard approach (whole-brain thickness).

| | Value | Standard deviation | Degrees of freedom | t-value | p-value |
|---|---|---|---|---|---|
| Intercept | 2.326 | 0.128 | 152 | 18.106 | < 0.001* |
| Session ID | − 0.018 | 0.012 | 152 | − 1.425 | 0.156 |
| Age | − 0.003 | 0.002 | 152 | − 2.140 | 0.034* |
| Gender | − 0.044 | 0.030 | 57 | − 1.470 | 0.147 |
| ALSFRS-r | 0.003 | 0.001 | 152 | 2.543 | 0.012* |
| Symptom duration (months) | 0.001 | 0.001 | 152 | 0.998 | 0.320 |
| Years of education | 0.004 | 0.004 | 57 | 0.929 | 0.357 |
| Handedness | 0.031 | 0.041 | 57 | 0.767 | 0.447 |

**Significant at an alpha level of 0.05*

Table 7.2.3 Conditional growth model for new approach (number of thin patches).

| | Value | Standard deviation | Degrees of freedom | t-value | p-value |
|---|---|---|---|---|---|
| Intercept | 354.137 | 96.725 | 152 | 3.661 | < 0.001* |
| Session ID | 6.411 | 9.335 | 152 | 0.687 | 0.493 |
| Age | − 1.549 | 1.124 | 152 | − 1.379 | 0.170 |
| Gender | 34.943 | 22.144 | 57 | 1.578 | 0.120 |
| ALSFRS-r | − 2.230 | 0.879 | 152 | − 2.538 | 0.012* |
| Symptom duration (months) | − 0.211 | 0.751 | 152 | − 0.281 | 0.779 |
| Years of education | − 3.214 | 3.225 | 57 | − 0.997 | 0.323 |
| Handedness | − 3.676 | 30.219 | 57 | − 0.122 | 0.904 |

**Significant at an alpha level of 0.05*

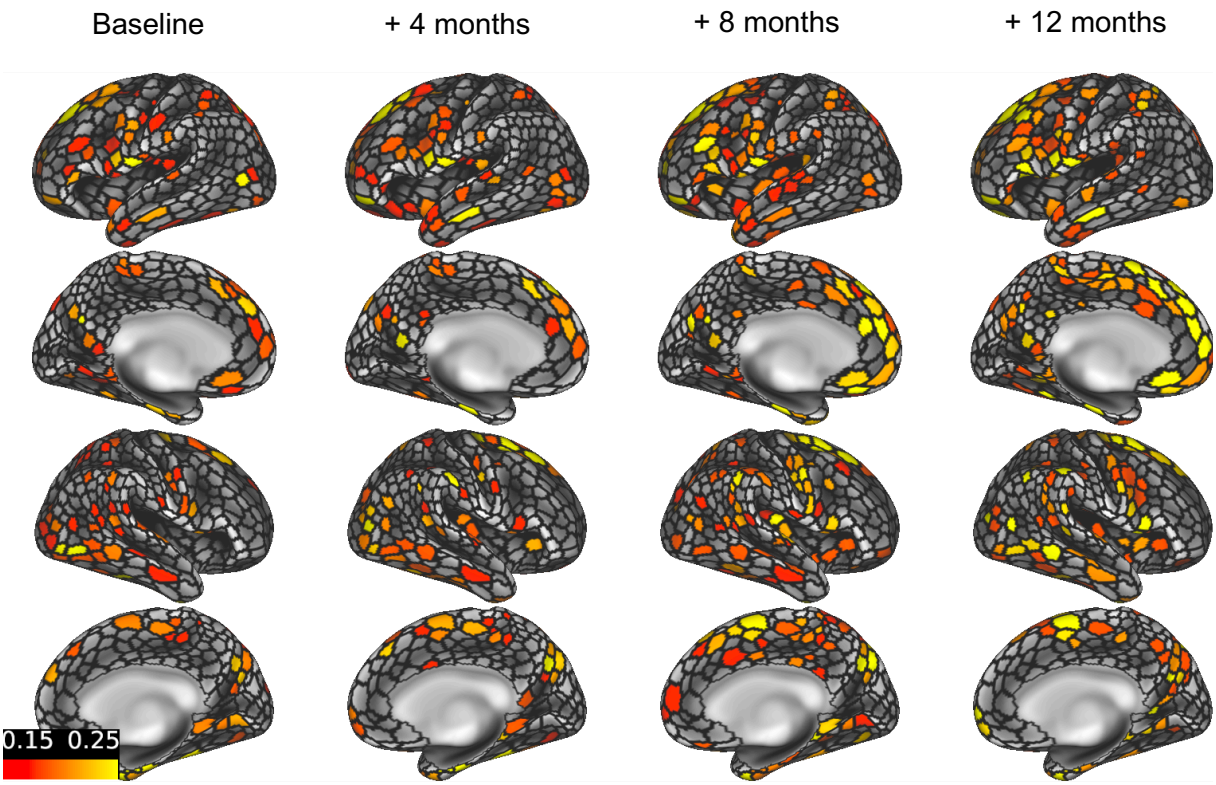



Figure 7.2.6 Cohort-specific cortical propagation pattern. Accumulation of cortical atrophy over time in the ALS cohort. The parcellation is colour-coded by the percentage of patients exhibiting atrophy in a specific cortical region, e.g. a value of 0.20 in a specific cortical patch means that 20% of patients have reduced cortical thickness in that area compared to their demographically matched controls. Initial disease burden is evident in the pre- and paracentral gyrus and the medial orbitofrontal cortex with gradual expansion to frontotemporal regions over time.

approaches is relatively similar across the ALS population, there is some randomness left. We observed a tendency towards more randomness in the standard approach. In the full model, we then considered the modulating effects of the co-variates age, gender, ALSFRS-r, symptom duration, years of education and handedness. Using the standard approach, both age and ALSFRS-r significantly confounded CT decline. With the implementation of the 'mosaic' approach, the variance associated with age was successfully corrected for, so that only disability (i.e. ALSFRS-r score) explained variance in thin-patch-load over time. We note that none of the other co-variates significantly modulated the progression of the DVs over time. Therefore, we have demonstrated that the new strategy successfully corrects for demographic factors. Given the sexual dimorphism associated with physiological and ALS-associated neuroradiological signatures, the meticulous adjustments for demographic variables are particularly important, especially in small studies and in the interpretation of single data sets (Bede et al., 2014; Luders et al., 2014; Menzler et al., 2011). An additional strength of our method is that disability as estimated by the ALSFRS-r score, modulates individual slopes of 'thin-patch' expansion over time. Therefore, our T1w-image-derived metric is resilient against demographic factors, but sensitive to clinical progression. More importantly, our proposed method enables the interpretation of individual patients. This is a considerable advantage over existing methods, which mainly rely on group comparisons. Grey matter atrophy in ALS is typically evaluated by voxel-based morphometry, cortical thickness analyses or various region-of-interest (ROI) approaches (Agosta et al., 2018; Bede et al., 2019, 2020; Rangariroyashe H Chipika, Finegan, et al., 2020; Christidi, Karavasilis, Velonakis, et al., 2018). ALS exhibits considerable clinical heterogeneity with respect to initial disability profiles, progression rates, upper vs lower motor neuron predominance and neuropsychological deficits (M. Abidi et al., 2020;

Burke, Pinto-Grau, et al., 2016; Christidi et al., 2020; Feron et al., 2018; Finegan, Chipika, Shing, et al., 2019; O Hardiman et al., 2016; Lebouteux et al., 2014; Yunusova et al., 2019). Single data sets of individual patients cannot be readily interpreted, which is a major barrier to routine clinical use (Foerster et al., 2013). Existing imaging studies of ALS typically stratify patients into groups of genetically (Bede, Omer, et al., 2018; Westeneng et al., 2016) or phenotypically defined (Agosta et al., 2016; Finegan, Chipika, Li Hi Shing, et al., 2019; Finegan, Li Hi Shing, et al., 2019) cohorts to reduce to effect of disease heterogeneity before running comparative statistics to ascertain group-level signatures. The drawback of this approach is that the resulting statistical maps reflect on the specific cohort included in the study, are 'over-fitted' by definition, and reveal very little about individual patients. From a clinical perspective, the quantification of disease burden in individual patients is hugely important to help diagnostic classification, prognostic categorisation, tracking the course of the disease and potentially gauging response to therapy (Rangariroyashe Hannah Chipika et al., 2019). Existing approaches for individual patient data interpretation in ALS implement a variety of complex machine learning (ML) models and require large patient samples for adequate training, yet, despite these efforts, model overfitting is a common problem (Agosta et al., 2010; Bede et al., 2017; Marwa Elamin et al., 2015; Grollemund et al., 2019a, 2020; Querin, El Mendili, et al., 2018a; Christina Schuster et al., 2017; Welsh et al., 2013; Westeneng et al., 2018). While the concept disease burden quantification and monitoring is firmly integrated in the management and clinical trials of other neurological conditions, such as lesion load/ lesion volume measurements in multiple sclerosis (MS), no comparable strategies have been implemented in ALS trials to date (Filippi et al., 1998). ALS patients often face a circuitous diagnostic journey, and the interval between symptom onset and diagnosis may be over 12 months (Household & Swash, 2000). Diagnostic uncertainty is not only a source of considerable distress for patients and caregivers but also delays recruitment into clinical trials (Kraemer et al., 2010; Zoccolella et al., 2006). Compared to other neurological conditions, clinical trials of ALS are small; therefore, a proxy of cerebral disease burden, such as the extent cortical atrophy, may have practical utility obviating the need for large patient cohorts. Our approach also demonstrates that relying on large, high- quality external control data is a viable approach and enables the designation of demographically defined individualised control groups, which is the best match for an individual patient. Our method offers an objective longitudinal marker which is capable of estimating progressive cortical disease burden in single subjects without requiring data from other patients. Disease progression in ALS is currently evaluated by a range of complementary measures, many of which are not objective. In a clinical setting, upper motor neuron and lower motor neuron degeneration is evaluated by a meticulous neurological examination, which, despite best efforts to document, is suboptimal to capture gradual and subtle changes. Disease-specific instruments, composite scores of disability and questionnaires are often confounded by reporting bias, fatigue, mood, practice effects, motivation to be included into clinical trial, etc. (Burke et al., 2017; Cedarbaum et al., 1999; Marwa Elamin et al., 2017). The multi-timepoint mapping of progression patterns in the 61 patients confirmed disease-associated vulnerability patterns centred on the precentral and paracentral gyri at baseline, and the gradual involvement in the pre- and supplementary motor cortices, insula, medial orbitofrontal and dorsolateral prefrontal cortex (DLPFC) over time (Figure 7.2.6). While these observations are reminiscent of the TDP-43 burden described based on post-mortem observations and also consistent with neuropsychology studies, these maps merely represent the evolution of cortical pathology in this specific cohort (Brettschneider et al., 2013, 2014; Burke et al., 2017; M. Elamin et al., 2013). This study is not without limitations. The pipelines implemented only segment supratentorial cortical regions: cerebellar and subcortical grey matter regions have not been evaluated despite being established foci of ALS pathology (Rangariroyashe H Chipika, Finegan, et al., 2020; Pradat et al., 2009; Pradat & El Mendili, 2014; Westeneng et al., 2015). The analyses presented

herein centre of cortical grey matter alone, even though white mater degeneration, connectivity alterations and functional changes are important early feature of ALS (Agosta, Canu, et al., 2014; Agosta, Galantucci, et al., 2014; Christidi et al., 2014; Lule et al., 2007, 2010; Nasseroleslami et al., 2019; C. Schuster et al., 2016; Trojsi et al., 2018, 2020). All of the patients included in this study fulfilled the diagnostic criteria for ALS. The inclusion of presymptomatic mutation carriers would have helped to evaluate the detection sensitivity of our method further. Notwithstanding these limitations, we have demonstrated that cortical atrophy can be reliably appraised in individual patients in ALS irrespective of their demographic profile and accumulation of disease burden can be captured over relatively short follow-up intervals. Our method is based on T1-weighted input data, which is routinely acquired as part of any clinical and research protocol, and can be reliably and non-invasively attained at any field strength on any MRI platform. The presented approach offers a cheap, non-invasive and objective way to estimate disease burden and track cerebral progression across multiple timepoints in symptomatic patients, with potential utility in clinical management and pharmaceutical trials.

### 7.2.6 Conclusions

Disease burden in ALS may be appraised at an individual level using external reference imaging data. The transition from group-level comparisons to individual patient characterisation is a prerequisite for viable clinical applications, individualised therapies and precision monitoring in pharmaceutical trials.

Acknowledgements We are grateful for all patients with ALS for agreeing to participate in this research study, and we are also indebted to their caregivers and families for their support. Without their generosity, this study would have not been possible. We also thank all patients who expressed interest in this study, but were unable to participate for medical or logistical reasons. We also thank the Irish Motor Neuron Disease Association for facilitating recruitment and providing unrelenting support to all patients with ALS and PLS.

### 7.2.7 Author contributions

MT and PB were involved in the conceptualisation of the study, neuroimaging analyses and drafting the manuscript. RC, JL, SLHS and OH contributed to clinical profiling and data acquisition. MT, RC, JL, SLHS, OH and PB were all involved in the revision of the manuscript for intellectual content.

### 7.2.8 Funding

Marlene Tahedl is funded by the Deutsche Multiple Sklerose Gesellschaft (DMSG), Grant number 2018_DMSG_08. Professor Peter Bede is supported by the Health Research Board (HRB EIA-2017-019), Spastic Paraplegia Foundation, Inc. (SPF), the EU Joint Programme—Neurodegenerative Disease Research (JPND), the Andrew Lydon scholarship, the Irish Institute of Clinical Neuroscience (IICN) and the Iris O'Brien Foundation. The sponsors of the authors had no bearing on the opinions expressed herein.

### 7.2.9 Compliance with ethical standards

#### *7.2.9.1 Conflicts of interest*

The authors have no conflicts of interest to declare.

#### *7.2.9.2 Ethics approval*

All participants provided informed consent in accordance with the approval of the Ethics (Medical Research) Committee— Beaumont Hospital, Dublin, Ireland.

## 7.3 Subproject 3: "Propagation patterns in motor neuron diseases: individual and phenotype-associated disease-burden trajectories across the UMN-LMN spectrum of MNDs"

(accepted at *Neurobiology of Aging* on April 14th 2021 – unpublished version)

### 7.3.1 Abstract

Motor neuron diseases encompass a divergent group of conditions with considerable differences in clinical manifestations, survival, and genetic vulnerability. One of the key aspects of clinical heterogeneity is the preferential involvement of upper (UMN) and lower motor neurons (LMN). While longitudinal imaging patters are relatively well characterised in ALS, progressive cortical changes in UMN- and LMN-predominant conditions are seldom evaluated. Accordingly, the objective of this study is the juxtaposition of longitudinal trajectories in three motor neuron phenotypes; a UMN-predominant syndrome (PLS), a mixed UMN-LMN condition (ALS), and a lower motor neuron condition (poliomyelitis survivors). A standardised imaging protocol was implemented in a prospective, multi-timepoint longitudinal study with a uniform follow-up interval of four months. Forty-five poliomyelitis survivors, 61 patients with amyotrophic lateral sclerosis (ALS) and 23 patients with primary lateral sclerosis (PLS) were included. Cortical thickness alterations were evaluated in a dual analysis pipeline, using standard cortical thickness analyses, and a z-score-based individualised approach. Our results indicate that PLS patients exhibit rapidly progressive cortical thinning primarily in motor regions; ALS patients show cortical atrophy in both motor and extra-motor regions, while poliomyelitis survivors exhibit cortical thickness gains in a number of cerebral regions. Our findings suggest that dynamic cortical changes in motor neuron diseases may depend on relative UMN/LMN involvement, and increased cortical thickness in LMN-predominant conditions may represent compensatory, adaptive processes.

Keywords: Motor neuron disease, Primary lateral sclerosis, Poliomyelitis, Neuroimaging, Clinical trials.

### 7.3.2 Introduction

Motor neuron disease (MND) is an umbrella term for a wide spectrum of clinically dissimilar conditions. While similarities exist in disability profiles among some of the MNDs, their progression trajectories and the overall prognosis is markedly different. Compared to ALS, other MNDs are strikingly understudied with respect to their phenotype-specific neuroimaging signatures. PLS and ALS groups are sometimes contrasted to evaluate disease-specific traits (Pioro et al., 2020), but lower motor neuron predominant MNDs, such as X-linked spinal and bulbar atrophy (SBMA), poliomyelitis or spinal muscular atrophy (SMA) are often only studied in contrast to healthy controls (Querin, Bede, et al., 2018; Shing et al., 2019). The imaging literature of MNDs is disproportionately dominated by cross-sectional studies which often include mixed cohorts in various stages in their individual disease course. Existing longitudinal MND studies, suffer from considerable attrition rates and an inclusion bias to patients with limited disability and slower progression rates (Rangariroyashe Hannah Chipika et al., 2019). The meaningful interpretation of longitudinal imaging data requires careful adjustments for physiological aging, sexual-dimorphism, and ideally, the inclusion of disease-controls (Christina Schuster et al., 2015). From a clinical perspective, there is a pressing and unmet need to monitor single, individual patients in a transparent and observer-independent fashion (Tahedl et al., 2021; Verstraete et al., 2015). From a monitoring viewpoint, a stereotyped question asked by patients and their caregivers is whether their condition have progressed since their previous clinic visit irrespective of other patients, other phenotypes, or age-matched healthy populations. From a clinical trial standpoint, the objective tracking of individual patients is a key requirement, preferably with the use of quantitative outcome measures

(Rangariroyashe Hannah Chipika et al., 2019). Accordingly, the development of methods to track pathology in vivo, at an individual-patient level, is a clinically relevant quest. The characterisation of individual disease trajectories is also hugely relevant academically. There are a number of emerging concepts in MND research, which are either derived from post mortem insights, inferred from clinical observations or based on animal models and have not been reassuringly validated by in vivo human data. These theories include cognitive reserve, motor reserve, stage-wise propagation, compensatory processes, prion-like propagation and selective network vulnerability (Costello et al., 2021; Dukic et al., 2019; Meier et al., 2020). There is also a notion that compensatory processes may occur in MNDs to adapt for relentless neurodegeneration (Querin, El Mendili, et al., 2019). Activation studies in ALS suggest that the pre-, supplementary- and contralateral motor cortex, the cerebellum, and subcortical grey matter structures increasingly contribute to the execution of motor tasks with the gradual degeneration of primary motor areas (M. Abidi et al., 2020; Bede et al., 2021; Nasseroleslami et al., 2019; Proudfoot et al., 2019). Irrespective of the MND phenotype studied, the overwhelming majority of published papers only comment on patterns of atrophy, cortical thinning, or density reductions (Bede & Hardiman, 2018). Lack of atrophy or 'resilient' regions are sometimes specifically evaluated, but increased cortical thickness is either not explored in statistical models, not reported, or not discussed. This seems like a missed opportunity as cortical reorganisation may reveal biologically important processes (O Hardiman et al., 2016). Based on these considerations, we have embarked on a comparative neuroimaging project to contrast the longitudinal course of three MNDs; a lower-motor predominant condition (poliomyelitis survivors), a mixed UMN-LMN syndrome (ALS), and a UMN-predominant condition (PLS). Our main objectives were (a) the characterisation of phenotype-specific propagation patterns (b) the juxtaposition of the rate of decline in LMN, UMN and mixed MNDs and (c) the targeted evaluation of increased thickness.

### 7.3.3 Methods

#### *7.3.3.1 Participants*

45 poliomyelitis survivors (PMS), 23 patients with PLS and 61 patients with ALS were included in a prospective neuroimaging study. Twenty-seven ALS patients had two, 34 had three follow-up scans; 14 PMS had one, 6 had two follow-up scans; 7 PLS patients had two, 16 had three follow-up scans. A uniform inter-scan interval of four months was implemented. All participants provided informed consent in accordance with the ethics approval of the study by the Medical Research Committee of Beaumont Hospital, Dublin, Ireland. ALS patients had 'probable' or 'definite' ALS according to the El Escorial criteria (Brooks et al., 2000) and PLS patients were diagnosed based on the Gordon criteria (Gordon et al., 2006). Reference data from 776 healthy control subjects were included (383 males) in our analyses (125 from Dublin, 651 from the Cam-CAN database (Shafto et al., 2014), with a mean age of 55.08 years and a standard deviation (SD) of 17.63 years.

#### *7.3.3.2 Neuroimaging*

T1-weighted (T1w) MRI data from all patients and controls was acquired on a 3 Tesla Philips Achieva scanner with an 8-channel receiver head coil, using a 3D Inversion Recovery Prepared Spoiled Gradient Recalled Echo (IR-SPGR) pulse sequence with following imaging parameters; repetition time (TR) / echo time (TE) = 8.5/3.9 ms, inversion time (TI) = 1060 ms, field-of-view (FOV): 256 x 256 x 160 mm, spatial resolution: 1 mm3. A total of 651 external reference T1-weighted MPRAGE images were also used in this study from the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) repository (Shafto et al., 2014).

#### *7.3.3.3 Standard cortical thickness analysis*

T1w datasets were segmented and the surface reconstructed with FreeSurfer's *recon-all* tool (Bruce Fischl, 2012). Data were then converted to the Connecitivty Informatics Technology Initiative (CIFTI) file format (Van Essen et al., 2013) using the CIFTIFY toolbox at a resolution of 32,000 vertices per hemisphere. As part of 'standard' cortical thickness analyses, three group comparisons were performed (PMS vs. HC, ALS vs. HC, PLS vs. HC) using vertexwise, non-parametric permutation testing using FSL's *randomise* tool with 5000 iterations, controlling for age and gender (Winkler et al., 2014). To correct for alpha-level inflation, we considered threshold-free cluster enhancement (TFCE) corrected p-maps only (Salimi-Khorshidi et al., 2011). Since we investigated two contrasts for each comparison (patients > controls, patients < controls), the alpha threshold was set to 0.05/2 = 0.025.

#### *7.3.3.4 Normalised, 'mosaic-based' interpretation*

'Standard' cortical thickness (CT) analyses were complemented by a z-score-based approach (Tahedl, 2020; Tahedl et al., 2021) where regional cortical thickness in single patients is interpreted with respect to demographically matched controls, categorising each cortical region ('mosaic') into 'thin', 'thick' or 'comparable' (Figure 7.3.2). Following pre-processing, CT maps were parcellated into 1000 equally sized 'mosaics' using an existing parcellation scheme (Schaefer et al., 2018) and implementing a 7-network approach (Yeo et al., 2011). A subject-specific reference group was generated for each patient based on age (+/– 2 years) and gender matching. To account for the effects of physiological ageing a 'sliding-window' reference matching was implemented. For example, for a 56-year old male patient, his control group comprised male controls aged between 54 and 58 years. At one-year follow-up, his reference group shifted to male controls between 55 and 59 years. Regional cortical thickness in single patients was interpreted based on the CT distribution of demographically matched controls. Individual patient's regional thickness was first z-scored by subtracting the control's mean and dividing by the control's standard deviation. Subsequently, z-values are converted to p-values by exhaustive permutation testing to obtain FWER corrected p-values. A threshold of $p \leq 0.05$ was used to define 'thick' or 'thin'. The main output variables from the 'mosaic approach' are the number of significantly thin and thick patches across the entire cortex and in the motor cortex defined by the pre- and paracentral labels of the Desikan-Killiany atlas (Desikan et al., 2006). To infer which patches were significantly thin or thick in the groups at baseline, Monte-Carlo permutation testing was used to correct for family-wise error rates (FWER), separately for the three study groups as well as for thin and thick patches. The alpha-threshold was set to 0.025 to account for two-sided testing. The analyses were conducted in MATLAB 2019b (The Mathworks, Natrick, MA, USA).

#### *7.3.3.5 Cross-sectional group comparisons*

As demographic matching is inherent in the 'mosaic' method, one-way, three-level, between-subjects analyses of variance (ANOVAs) was utilised to investigate group differences (PMS, ALS and PLS) of the means of the following four dependent variables (DVs): (i) number of whole-brain thin patches, (ii) number of motor-cortex thin patches, (iii) number of whole-brain thick patches, (iv) number of motor-cortex thick patches. For (v) whole-brain cortical thickness (CT) in millimetres and (vi) motor-cortex CT in millimetres, analyses of covariance (ANCOVA) were utilised to correct for age and gender. As the ANOVAs and ANCOVAs reached significance for six DVs, post-hoc pairwise comparisons were conducted using Tukey's honestly significant difference (HSD) testing. Statistical analyses were carried out with RStudio version 1.3.1093 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria).

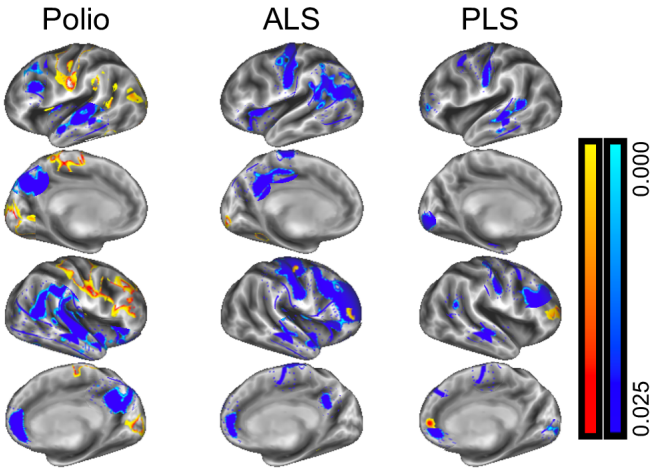



Figure 7.3.1 Cortical thickness patterns based on vertexlwise permutation testing and corrections for age and gender in poliomyelitis survivors (left column), ALS (middle column) and PLS patients (right column).

#### *7.3.3.6 Longitudinal analyses using linear mixed effects models*

To evaluate progressive cortical thickness alterations with the 'mosaic' approach, linear mixed effects (LME) models were implemented using R's *nmle* package (Pinheiro et al., 2020):

(7.3-1) mod <- lme(DV ~ Time*Diagnosis, random =~ Time | ID, data = X , method = 'ML').

DV denotes the dependent variable, which is modelled as dependent upon the factor 'Time' and the categorical variable 'Diagnosis' (PMS / ALS / PLS). ID indicates the timepoint, which serves as a nested variable, since in longitudinal analyses data is per definition nested by individuals. Finally, X is the data containing information on all the specified variables. The method we used for estimating the model was maximum likelihood, 'ML'.

Given that in the standard approach (cortical thickness in millimetres), age and gender are not inherently corrected for, we set up these models as follows:

(7.3-2) mod2 <- lme(DV ~ Time*Diagnosis + age + gender, random =~ Time | ID, data = X , method = 'ML').

Although the LMEs captured no significant 'Time x Diagnosis' interaction, visual representation of the data (Figure 7.3.5a–b) suggested that PLS patients accumulated thin patches more rapidly than the other patient groups. Since the data suggested this effect may be driven by the one-year follow-up of the PLS patients, a supplementary 'Time x Diagnosis' LME analysis was added, restricted only to the baseline and one-year follow-up data.

### 7.3.4 Results

A total of 776 healthy control subjects (383 males) were included in our analyses (125 from Dublin, 651 from the Cam-CAN database, with a mean age of 55.08 years (SD: 17.63y). The mean age of the 45 PMS patients (20 males) at baseline was 66.07 years (+/–6.52), the mean interval between their acute poliomyelitis infection in infancy and their initial brain scan was 62.75 years. The mean age of the 61 ALS patients (43 males) was 60.20 (+/– 9.62 years), and their mean symptom duration was 43 months. The age profile of the 23 PLS patients (12 males) was 58.52 years (+/– 9.79 years) and mean symptom duration 52 months. In accordance with the demographic matching procedure (Tahedl et al., 2021), 82 individualised reference groups were generated. The cross-sectional comparisons of CT maps between

PMS patients and HCs (Figure 7.3.1) revealed significant cortical thinning of the bilateral superior temporal gyri, the bilateral medial posterior cingulate cortices (PCC), the right anterior cingulate cortex (ACC) and parts of the left frontal cortex. Intriguingly, we observed increased thickness in the bilateral sensorimotor cortices, as well as parts of the bilateral medial visual cortices and the right frontal cortex. In ALS patients, cortical thinning was most evident in the bilateral motor cortices, left frontal and bilateral temporal cortices, as well as bilateral PCC and right ACC. Small regions of thicker cortex were also observed in the right pre- and frontal cortices. In PLS, cortical thinning was most evident in the bilateral motor cortices as well as widespread frontotemporal regions. Small regions of thicker cortex were observed in the right rostral middle frontal and the right medial orbitofrontal cortex. The mosaic-based CT analysis resulted in individual subject brain maps, indicating regional cortical changes with respect to subject-matched controls. Representative patient maps are shown in Figure 7.3.2 to illustrate individual cortical patterns across the timepoints.

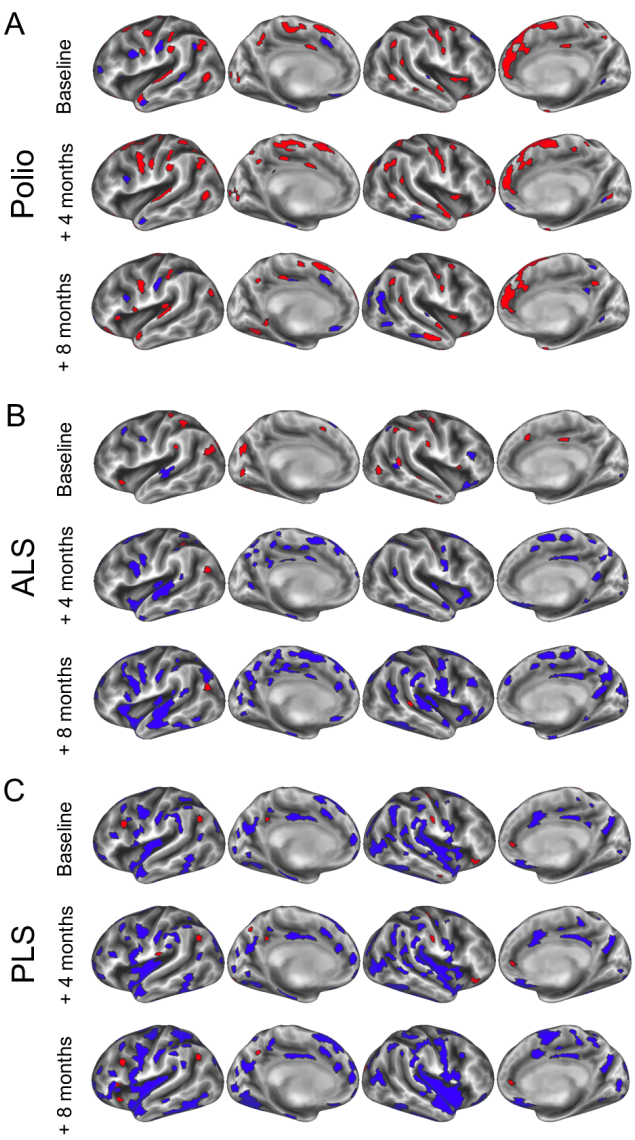



Figure 7.3.2 The 'mosaic-based' approach permits the evaluation of individual patient scans with respect to age- and gender matched controls. 'Thin' regions are colour-coded in blue and 'thick' regions in red. Representative examples of individual patients are shown at baseline and follow-ups in (A) a poliomyelitis survivor (B) ALS patient and (C) a PLS patient.

### *7.3.4.1 Inferential statistics of the mosaic brain maps*

Permutation testing was used to infer which cortical regions were significantly 'thin' or 'thick' in each phenotype. As illustrated in Figure 7.3.3, PMS patients exhibit 'thick' patches (red-yellow) around the bilateral sensorimotor cortices. A cluster of atrophic patches (blue-light blue) was observed in the right temporal cortex, in addition to thin patches scattered across the cortex.

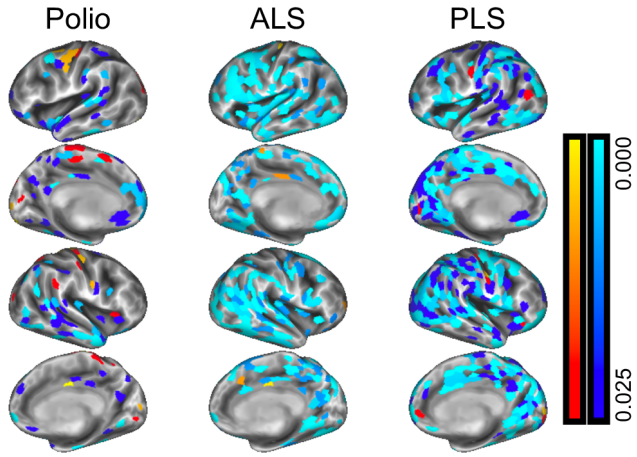



Figure 7.3.3 Baseline cortical characteristics in poliomyelitis survivors (left column), ALS (middle column) and PLS using the 'mosaic approach'.

In ALS and PLS, atrophic patches dominate the entire cortical surface, with only occasional 'thick' patches observed in middle frontal and inferior precentral areas in PLS.

#### *7.3.4.2 Cross-sectional group comparisons*

Significant cross-sectional differences were noted in all six dependent variables (Figure 7.3.4a–f): (i) whole-brain thin patches ($F(2) = 24.51$, $p = 1.02e-09$), (ii) motor-cortex thin patches ($F(2) = 38.16$, $p = 1.1e-13$), (iii) whole-brain thick patches ($F(2) = 6.468$, $p = 0.0021$), (iv) motor-cortex thick patches ($F(2) = 7.717$, $p = 6.9e-04$), (v) whole-brain thickness ($F(2) = 20.20$, $p_{corr} = 2.55e-08$), and (vi) motor-cortex thickness ($F(2) = 32.50$, $p_{corr} = 4.482e-12$). Post-hoc Tukey's HSD confirmed differences in the following pairwise comparisons at $p < 0.05$: (i) whole-brain thin patches: PMS (mean 56.07, SD 46.92) < ALS (98.90 +/– 87.64) ($p = 0.02061$), PMS < PLS (200 +/– 107.77) ($p = 0.0000$), and ALS < PLS ($p = 3.0e-06$); (ii) motor-cortex thin patches: PMS (5.62 +/– 4.69) < ALS (14.43 +/– 14.74) ($p = 0.0022$), PMS < PLS (34.65 +/– 18.12) ($p = 0.0000$) and ALS < PLS ($p = 0.0000$); (iii) whole-brain thick patches: PMS (31.53 +/– 33.97) > PLS (9.7 +/– 9.09) ($p = 0.0016$); (iv) motor-cortex thick patches: PMS (7.27 +/– 6.81) > ALS (4.51 +/– 4.76) ($p = 0.0261$), PMS > PLS (2.09 +/– 3.03) ($p = 0.0007$); (v) whole-brain thickness: PMS (2.34mm +/– 0.09mm) > PLS (2.2mm +/– 0.11mm) ($p_{corr} = 1.8e-06$), ALS (2.31mm +/– 0.11mm) > PLS ($p_{corr} = 0.0001$); (vi) motor-cortex thickness: PMS (2.31mm +/– 0.11mm) > ALS (2.25mm +/– 0.13mm) ($p_{corr} = 0.0293$), PMS > PLS (2.09mm +/– 0.11mm) ($p_{corr} = 0.0000$), ALS > PLS ($p_{corr} = 5.0e-07$).

#### *7.3.4.3 Longitudinal analysis using linear mixed effects models*

Longitudinal progression of the six DVs was evaluated using linear mixed effect models (Figure 7.3.5a–f). We found that for thin patches and raw cortical thickness – but not for thick patches – the main effect "Time" was significant, showing an increase of the thin-patch count and a decrease of CT over time across the diagnoses groups [whole-brain thin patches: $t(235) = 2.2956$, $p = 0.0226$; motor-cortex thin patches: $t(235) = 2.4767$, $p = 0.0140$; whole-brain CT: $t(234) = -4.1241$, $p_{corr} = 0.0000$; motor-cortex CT: $t(234) = -5.1505$, $p_{corr} = 0.0000$). Additionally, we found a main effect "Diagnosis" for the thin-patch count (Fig 5A–5B), separating both PMS and PLS from ALS, both for the whole-brain as well as the restricted motor-cortex count (whole-brain PLS: $t(126) = 4.1607$, $p = 0.0001$; whole-brain PMS: $t(126) = -2.5651$, $p = 0.0115$; motor-cortex PMS: $t(126) = -3.1327$, $p = 0.0022$; motor-cortex PLS: $t(126) = 5.4094$, $p = 0.0000$). In contrast, for the thick-patch count (Figure 7.3.5c–d), we observed no main effect of time,

and in terms of the main effect of diagnosis, only the PMS group differed from the ALS group (whole-brain: t(126) = 2.1188, p = 0.0361; motor-cortex: t(126) = 2.8325, p = 0.0054). For the standard approach (Figure 7.3.5e–f, the main effect "Diagnosis" was evident for PMS vs. ALS patients (t(125) = 2.9811, $p_{corr}$ = 0.0035), as well as PLS vs. ALS patients for the whole-brain analysis (t(125) = -3.8097, $p_{corr}$ = 0.0002); likewise for the restricted motor-cortex analysis, both PMS and PLS patients differed from ALS (PMS vs. ALS: t(125) = 3.8037, $p_{corr}$ = 0.0002; PLS vs. ALS: t(125) = -4.9094, $p_{corr}$ = 0.0000). Given the differences in the available follow-up scans in the different study groups, an supplementary analysis was conducted only taking into account the baseline and one-year follow-up data in PLS and ALS patients. This LME revealed significant interaction effects of Time x Diagnosis, for both whole-brain (t(48) = 2.3004, p = 0.0258) and motor-cortex-only (t(48) = 2.6795, p = 0.0101) thin-patch count. For standard CT values, whole-brain average CT revealed significant differences (t(48) = -2.4035, p = 0.0202), while the restricted motor-cortex analysis only approached significance (t(48) = -1.9350, p =0.0589).

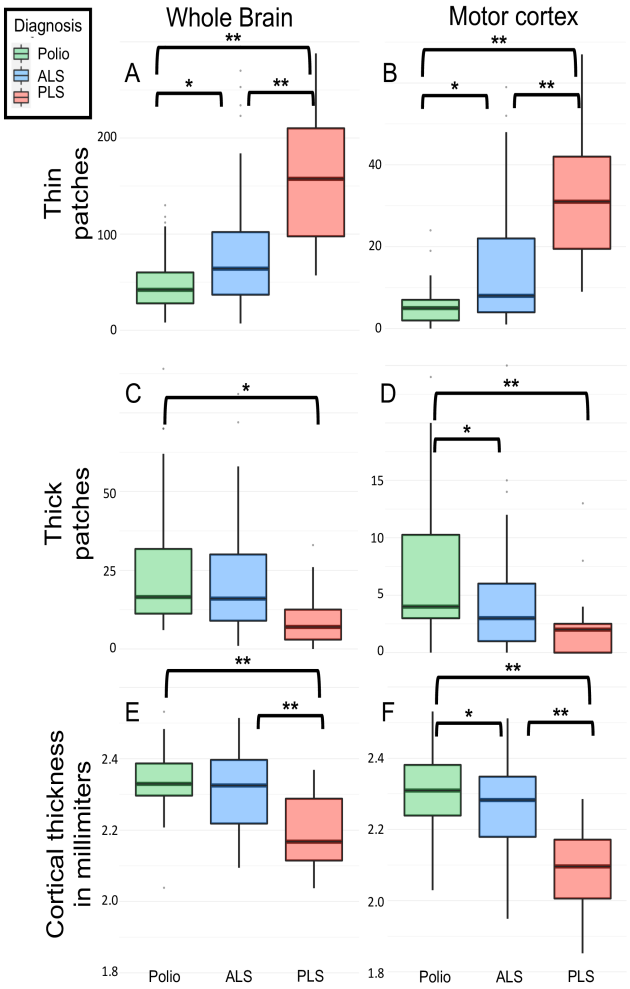



Figure 7.3.4 Cortical thickness over the entire brain and in the motor cortex in poliomyelitis survivors (green), ALS (blue) and PLS (red). Box-plots represent the interquartile range and whiskers indicate IQRx1.5. Tukey's honestly significant difference testing was used for post-hoc pairwise comparisons: (*) indicates p-values < 0.05 (**) indicates p-values < .001.

### 7.3.5 Discussion

Our findings indicate divergent cortical signatures in UMN-predominant, LMN-predominant and mixed UMN-LMN MND phenotypes. Our results also reveal co-existing atrophy and increased cortical thickness in MND, in both LMN- and UMN-predominant phenotypes. Additionally, our results confirm the feasibility of individual patient MRI interpretation using patient-specific normative data sets.

Longitudinal imaging in MND is one of the few ways to characterise progressive pathological changes in vivo to evaluate propagation patterns, verify concepts of disease-biology, and validate staging systems (Rangariroyashe Hannah Chipika et al., 2019; Müller et al., 2016, 2020). Longitudinal imaging studies have been published in symptomatic ALS patients and asymptomatic mutation carriers, but the majority of studies lacked disease controls and relied solely on healthy aging populations for data interpretation (Querin, Bede, et al., 2019). With few exceptions (Clark et al., 2017), no robust longitudinal studies have been published in PLS, despite the long course and relatively good prognosis associated with the condition. Based on clinical observations, there is a notion that PLS may be slowly progressive compared to ALS, but no robust quantitative imaging studies exist to support this notion (Finegan, Chipika, Shing, et al., 2019). Recent post-mortem studies confirmed both motor cortex and extra-motor TDP-43

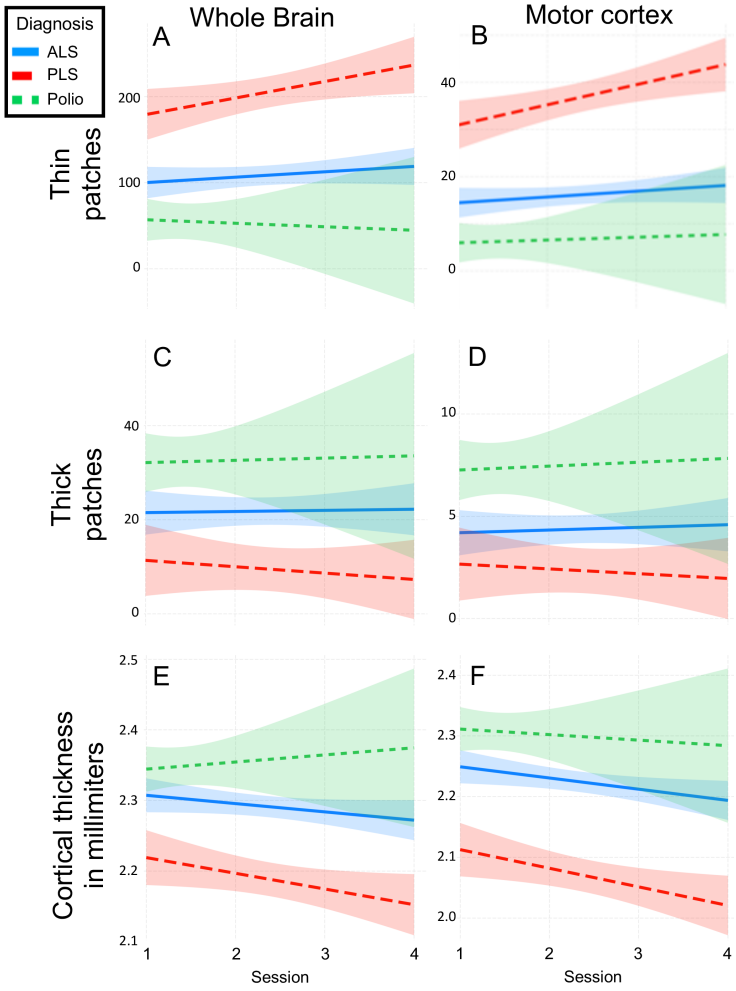



Figure 7.3.5 Linear mixed effects models confirming that for all dependent variables are associated with cortical thinning (A, B, E, F), time was a significant predictor of worsening. In contrast, hypertrophy (C and D) was stable over time, showing only a main effect of "Diagnosis" with respect to the poliomyelitis survivors. Lines denote the best-fit line of the regressions, the shadowed area around the lines indicate 95% confidence intervals.

pathology in PLS (Mackenzie & Briemberg, 2020) and neuropsychology studies also highlighted considerable cognitive deficits (de Vries et al., 2019). It is increasingly recognised that similarly to ALS (Burke, Pinto-Grau, et al., 2016), PLS is also associated with considerable extra-motor manifestations (Finegan et al., 2021). Imaging studies in PLS have largely revealed comparable cerebral signatures in ALS and PLS with the shared involvement of the PMC, brainstem, cerebellum, CSTs and CC (Bede et al., 2019, 2020; Müller et al., 2012; van der Graaff et al., 2011). Very few PLS studies suggested distinguishing radiological changes from ALS; and these proposed that contrary to ALS, the postcentral gyrus may spared (Finegan, Chipika, Li Hi Shing, et al., 2019) and the subcortical signature of PLS may be different from ALS (Rangariroyashe H Chipika, Finegan, et al., 2020; Finegan, Li Hi Shing, et al., 2019; Finegan, Hi Shing, et al., 2020). Some studies suggest greater PMC atrophy (Kiernan & Hudson, 1994; Menke et al., 2018; van der Graaff et al., 2010) and more marked CST degeneration in PLS (Agosta, Galantucci, et al., 2014; van der Graaff et al., 2011) than in ALS, but others have not replicated these findings (Ferraro et al., 2017; Müller et al., 2018). Given the relative lack of convincing cross-sectional differences between ALS and PLS, longitudinal studies may be better suited to capture distinguishing features (Christidi, Karavasilis, Rentzos, et al., 2018). The distinction of PLS and ALS is hugely important in the first years of symptom onset, as there is often a diagnostic dilemma in early PLS, and apprehension that a patient with UMN-predominant symptoms may transition to ALS (Finegan, Li Hi Shing, et al., 2020; Finegan, Siah, et al., 2020; Turner et al., 2020; Yunusova et al., 2019).

MNDs include a range of LMN-predominant conditions (Lebouteux et al., 2014) including progressive muscular atrophy (PMA), Spinal muscular atrophy (SMA), Kennedy's disease (SBMA), progressive bulbar palsy (PBP), monomelic amyotrophy, flail arm / flail leg syndrome etc (O Hardiman et al., 2016). The cerebral imaging literature of these syndromes is particularly scarce as the primary pathology is in the spinal anterior horns and the brainstem nuclei (Li Hi Shing, Lope, Chipika, et al., 2021; Li Hi Shing, Lope, McKenna, et al., 2021). Cerebral involvement PMA is controversial. Recent PMA studies captured cervical spinal cord atrophy (van der Burgh et al., 2019) without overt cerebral connectivity alterations (Basaia et al., 2020). Some PMA studies suggested decreased cerebral white

matter integrity (Prudlo et al., 2012) and abnormal prefrontal activation patterns (Raaphorst et al., 2014), while others did not detect FA reductions (Cosottini et al., 2005; H. Mitsumoto et al., 2007). Tractography in a PMA cohort revealed FA reductions subjacent to the PMC with concomitant FA increase in the rostral internal capsule/corona radiata (van der Graaff et al., 2011). Post mortem studies have also captured corticospinal tract degeneration in patients labelled with PMA on clinical grounds (P. G. Ince et al., 2003) casting doubt on whether PMA can be viewed as a distinct entity. Radiological reports of SBMA are just as conflicting. Cerebellar, pyramidal tract, and limbic white matter degeneration was described by some studies (Kassubek et al., 2007; Pieper et al., 2013; Unrath et al., 2010), but not confirmed by others (Echaniz-Laguna et al., 2005; Nelles et al., 2008; Spinelli et al., 2019). Varying degree of frontal lobe atrophy (Sperfeld et al., 2005) and hypometabolism (Lai et al., 2013) have also been described. The biomarker literature of SMAs is dominated by electrophysiology studies (Querin, Lenglet, et al., 2018), and existing imaging studies are strikingly inconsistent. While cerebral changes have been described in SMA type 0 (Mendonça et al., 2019), type II (Losito et al., 2020), and adult SMA variants (de Borba et al., 2020), other studies specifically highlight the lack of cerebral involvement despite considerable spinal cord degeneration (Querin, El Mendili, et al., 2019).

Our findings reveal distinct cortical signatures in LMN-predominant, UMN-predominant and mixed MND phenotypes. The juxtaposition of the three phenotypes helps to highlight disease-specific traits. Poliomyelitis survivors exhibit increased cortical thickness in both motor and extra-motor regions revealing a unique imaging signature. This pattern is readily captured by both traditional cortical thickness analyses and the individualised 'mosaic-based' approach using subject-specific normative data. The co-existing patterns of cortical atrophy and cortical hypertrophy in polio survivors showcases the importance of two-way analyses in MND, and highlights the pitfalls of implementing hypothesis-aligned, one-way contrasts only. The widespread regions showing increased cortical thickness in poliomyelitis survivors are reminiscent of other LMN conditions such as adult forms of SMA (Querin, El Mendili, et al., 2019). UMN-predominant MND patients were represented by a group of PLS patients in this study. Consistent with previous reports, they exhibit primary motor cortex (PMC) atrophy with additional frontotemporal involvement. It is noteworthy, that both at group-level and at individual level, PLS patients also show regions of increased cortical thickness. The discussion around the interpretation of these findings is stimulating, as PLS is an archetypal neurodegenerative condition with considerable disability which is relatively well-characterised with regards to TDP-43 burden (Mackenzie & Briemberg, 2020). Despite representing the opposing extremes of the LMN-UMN spectrum, a shared feature of PMS and PLS is the long survival which may permit compensatory processes to take place leading to structural reorganisation. Adaptive cortical changes have been observed following spinal cord injuries (Nishimura & Isa, 2009), in response to repetitive tasks, arduous physical training, dexterity associated with musical instruments, and association with cognitive training. Adaptive motor cortex changes have been consistently noted after unimanual training (Sale et al., 2017), aerobic exercise (Colcombe et al., 2006), post stroke (Sterr et al., 2013), and in professional musicians (Bruchhage et al., 2020; Gaser & Schlaug, 2003; Hudziak et al., 2014). Extra-motor changes have also been observed following various exercise regimes (Pereira et al., 2007; Thomas et al., 2016), and cognitive tasks (Lazar et al., 2005). Despite ample examples of cortical volume gains in a range of conditions, cortical thickness increases or volume gains are seldom evaluated or reported in MNDs. In the ALS cohort, we have not observed widespread hypertrophic regions which may be explained the shorter symptom duration of this cohort. Our findings also suggest that longitudinal imaging may be better suited to differentiate MND phenotypes than cross-sectional imaging, as propagation patterns and rate of progression may be more specific to MND phenotypes than a snapshot of cortical disease burden.

The comparable cortical disease burden observed in PLS and ALS stands in stark contrast with the much shorter survival in ALS. Our longitudinal analyses also demonstrate faster cortical thinning in PLS than ALS. These observations suggest that cortical disease burden does not drive survival, and ALS patients are likely to succumb to the sequelae of LMN degeneration. While hypertrophic brain regions are seldom reported in ALS, unaffected regions have been specifically investigated (Bede et al., 2016). Furthermore, some ALS studies have found divergent imaging profiles between UMN-predominant and LMN-predominant ALS patients (Malek Abidi et al., 2021). Even tough increased cortical thickness has not been reported in ALS, functional studies have consistently captured adaptive changes with the increasing involvement of cerebellar, subcortical and contralateral motor regions in the execution of motor tasks (Bede et al., 2021; Proudfoot et al., 2019).

We presented a z-score-based approach to interpret single-patient imaging data, but several other methods have been successfully explored in ALS (Grollemund et al., 2019b). Wet biomarkers (Blasco et al., 2018; Devos et al., 2019), clinical parameters (Marwa Elamin et al., 2015; Westeneng et al., 2018) and MRI data (Bede, Querin, et al., 2018; Welsh et al., 2013) been used in a variety of machine learning applications to categorise single subjects into diagnostic (Bede et al., 2017; Christina Schuster et al., 2016) or prognostic groups (Grollemund et al., 2020, 2021; Christina Schuster et al., 2017).

This study is not without limitations. Despite our dual-methodological approach, only cortical grey matter changes have been evaluated and white matter integrity has not been assessed in this study. We have no supporting post mortem data to interrogate regions of increased cortical thickness and assess the histological underpinnings of our imaging findings. Owing to the low incidence of PLS we have a relatively limited sample size at our disposal. While physiological aging is accounted for in our 'mosaic' model by implementing the 'sliding window' approach, the availability of longitudinal control data would have permitted more fine-grained statistical modelling. The inclusion of additional LMN conditions may have helped to support our interpretation that cerebral reorganisation may occur in anterior horn pathologies. Despite these limitations we have demonstrated different cortical trajectories along the LMN-UMN spectrum of MNDs. We have also shown that cortical disease burden can not only be interrogated at a cohort-level but may also be meaningfully evaluated at an individual-level. Our study draws attention to areas of increased cortical thickness which are notoriously under evaluated in neurodegenerative conditions despite representing an important facet of disease biology.

### 7.3.6 Conclusions

The longitudinal analysis of cortical disease-burden distinguishes MND phenotypes, which exhibit markedly differing trajectories. MND patients don't solely exhibit cortical atrophy, but LMN-phenotypes in particular, show regions of increased cortical thickness as well. These changes may represent adaptive processes, but the post mortem correlates of these findings remain to be elucidated. Cortical disease-burden may be interpreted at an individual-level using demographically matched normative data which may be particularly useful in a clinical or clinical trial setting.

### 7.3.7 Acknowledgements

We are most thankful for the participation of each patient and healthy control, and we also thank all the patients who expressed interest in this research study but were unable to participate for medical or logistical reasons. We also express our gratitude to the caregivers of MND patients for facilitating attendance at our neuroimaging centre. Without their generosity this study would have not been possible.

### 7.3.8 Funding information

This study was sponsored by the Spastic Paraplegia Foundation, Inc. (SPF). Peter Bede and the computational neuroimaging group (CNG) are supported by the Health Research Board (HRB EIA-2017-019), the EU Joint Programme – Neurodegenerative Disease Research (JPND), the Andrew Lydon scholarship, the Irish Institute of Clinical Neuroscience (IICN), and the Iris O'Brien Foundation. Marlene Tahedl is funded by the Deutsche Multiple Sklerose Gesellschaft (DMSG), grant number 2018_DMSG_08.

### 7.3.9 Conflicts of interest

None

### 7.3.10 Authors' contribution

Conceptualization, Methodology, Analyses, drafting the original manuscript: MT, PB
Investigation, Project administration, revision for intellectual content: SLHS, EF, RHC, JL, OH, PB.

## 7.4 Subproject 4: "Mapping cortical disease-burden at individual-level in frontotemporal dementia: implications for clinical care and pharmacological trials"

(accepted at *Brain Imaging and Behavior* on July 20th 2021 – unpublished version)

### 7.4.1 Abstract

Imaging studies of FTD typically present group-level statistics between large cohorts of genetically, molecularly or clinically stratified patients. Group-level statistics are indispensable to appraise unifying radiological traits and describe genotype-associated signatures in academic studies. However, in a clinical setting, the primary objective is the meaningful interpretation of imaging data from individual patients to assist diagnostic classification, inform prognosis, and enable the assessment of progressive changes compared to baseline scans. In an attempt to address the pragmatic demands of clinical imaging, a prospective computational neuroimaging study was undertaken in a cohort of patients across the spectrum of FTD phenotypes. Cortical changes were evaluated in a dual pipeline, using standard cortical thickness analyses and an individualised, z-score based approach to characterise subject-level disease burden. Phenotype-specific patterns of cortical atrophy were readily detected with both methodological approaches. Consistent with their clinical profiles, patients with bvFTD exhibited orbitofrontal, cingulate and dorsolateral prefrontal atrophy. Patients with ALS-FTD displayed precentral gyrus involvement, nfvPPA patients showed widespread cortical degeneration including insular and opercular regions and patients with svPPA exhibited relatively focal anterior temporal lobe atrophy. Cortical atrophy patterns were reliably detected in single individuals, and these maps were consistent with the clinical categorisation. Our preliminary data indicate that standard T1-weighted structural data from single patients may be utilised to generate maps of cortical atrophy. While the computational interpretation of single scans is challenging, it offers unrivalled insights compared to visual inspection. The quantitative evaluation of individual MRI data may aid diagnostic classification, clinical decision making, and assessing longitudinal changes.

Keywords: Frontotemporal Dementia, Cerebellum, PPA, Behaviour, MRI, Cortical thickness

### 7.4.2 Introduction

The majority of imaging studies in FTD stratifies patients based on clinical, molecular or genetic categories and describes group-specific radiological traits (Omer et al., 2017; Rohrer et al., 2011; J. L. Whitwell et al., 2012; Jennifer L. Whitwell et al., 2005, 2011). These data however are difficult to apply to individual patients is everyday clinical practice. The current role of MR imaging in the diagnostic pathway of FTD is limited to 'ruling-out' structural mimics and alternative diagnoses. MR images acquired in a clinical setting are typically only subjectively and qualitatively interpreted with regards to atrophy (Adachi et al., 2004; Aizpurua et al., 2019; Baez et al., 2014; Campanella et al., 2014; De Maindreville et al., 2015; Di Fede et al., 2019; Harper et al., 2014; Kito et al., 2009; Kotagal et al., 2012; McKeon et al., 2007; Mueller et al., 2018; MUQIT, 2001; Nishio et al., 2003; Way et al., 2019; Younes et al., 2018). This is a missed opportunity, as raw MRI datasets contain rich, spatially coded information with regards to cortical thickness, subcortical volumes and white matter integrity that cannot be meaningfully appraised on visual inspection. In contrast, computational imaging offers objective, observer-independent, reference-based quantitative image interpretation (Christidi, Karavasilis, Rentzos, et al., 2018). The potential translation of quantitative MR analysis frameworks to routine clinical practice may offer a number of practical benefits, including the generation of individualised atrophy maps, the objective assessment of longitudinal changes, and the classification of single scans

into likely phenotypic categories. Ultimately, quantitative imaging may enable 'ruling-in' patients into specific groups, as opposed to merely 'ruling-out' differential diagnoses (Bede, Querin, et al., 2018; Grollemund et al., 2019b). From a practical point of view, MR platforms are widely available, MR imaging is non-invasive, relatively cheap, and a multitude of open-source software are available for computational data analyses (Du et al., 2006). Access to 18F-FDG PET imaging on the other hand may be limited and the costs of routine PET imaging may be prohibitive in some health care systems (McMahon et al., 2003; Shivamurthy et al., 2015).

The current diagnostic approach to FTD subtypes – bvFTD, ALS-FTD, nfvPPA, svPPA - requires meeting specific clinical criteria and a definitive diagnosis may only be confirmed in vivo by identifying a pathogenic genetic mutation or typical histopathological findings (Brettschneider et al., 2013; Geser et al., 2009; Gorno-Tempini et al., 2011; Hodges et al., 2010; Perry et al., 2017; Rajagopalan & Pioro, 2015; Rascovsky et al., 2011; Snowden et al., 2007; Strong et al., 2017). The recent development, optimisation and validation of serum and CSF biomarkers panels will not only aid diagnostic classification but help the exclusion of alternative neurodegenerative diagnoses such as Alzheimer's pathology (R. M. Ahmed et al., 2014; Blasco et al., 2018; Devos et al., 2019; Meeter et al., 2019; Paterson et al., 2018; Rascovsky et al., 2011; Steinacker et al., 2017; Swift et al., 2021). As with all diagnostic criteria, there are practical shortcomings with regards to sensitivity and specificity: some symptomatic patients do not meet proposed thresholds for diagnosis, despite subsequent pathological confirmation. In a subset of FTD cases, the diagnosis may never be reached in vivo, or a considerable diagnostic delay is experienced (Harris et al., 2013; Piguet et al., 2009). Diagnostic uncertainty often creates undue stress for the patient and their family. The insidious onset of apathy, lack of interest and social withdrawal may be mistaken for depression, amongst other misdiagnoses (Besser & Galvin, 2020; Rasmussen et al., 2019). Early behavioural symptoms may be difficult to articulate, which is further complicated by the disparity in those perceived by the patients and their caregivers. Early cognitive deficits may also be difficult to identify, particularly due to the masking effect of cognitive reserve and the lack of sensitivity of generic screening instruments (Marwa Elamin et al., 2017; Rasmussen et al., 2019). Primary care physicians may reassure patients and caregivers based on neuropsychological screening tests and a 'grossly' normal MR imaging whilst awaiting lengthy specialist referrals (Rasmussen et al., 2019). Diagnostic delay in neurodegenerative conditions has a number of adverse implications. From a patients' perspective, timely diagnosis is important to inform realistic expectations over coming years (Spreadbury & Kipps, 2019). It helps to guide targeted genetic testing that may be of significance to other family members. Accurate and early diagnostic classification enables prompt multidisciplinary team referrals and appropriate lifestyle adjustments with regards to employment, finances, driving, and childcare (Spreadbury & Kipps, 2019). In those with language impairment, there is a critical time-window to explore alternative communication options e.g. 'voice-banking' to create a digital library for assisted communication devices (Fried-Oken et al., 2015). A timely diagnosis is also important for resource allocation and advanced care planning to ensure that the patients' end-of-life preferences are recognized (Harrison Dening et al., 2019). Early diagnostic categorisation is also indispensable for the timely inclusion of patients in clinical trials, which in turns enables longer follow-up (Finegan, Chipika, Li Hi Shing, et al., 2019). Based on these considerations, we have undertaken a quantitative imaging study across the spectrum of FTD phenotypes to test a framework to interpret cortical atrophy patterns at both individual- and group-level.

### 7.4.3 Methods

#### *7.4.3.1 Recruitment*

A total of 227 participants were included in this study; 12 patients with non-fluent variant primary progressive aphasia ('nfvPPA' 6 females, mean age 61.50±2.97), 3 patients with semantic variant primary progressive aphasia ('svPPA' 1 female, mean age 61.67±6.43), 7 patients with behavioural variant FTD ('bvFTD' 3 females, mean age 60.71±3.30 years, 20 ALS-FTD patients with *C9orf72* hexanucleotide expansions ('C9+ALSFTD' 8 females, mean age 58.65±11.22), 20 ALS-FTD patients without C9orf72 hexanucleotide expansions ('C9–ALSFTD' 7 females, mean age 59.95±7.67), 40 ALS patients with no cognitive impairment ('ALSnci' 21 females, mean age 58.70±11.33) as disease controls and 125 healthy controls (HC). Methods for screening for GGGGCC hexanucleotide repeat expansions in *C9orf72* have been previously described (Byrne et al., 2012; Rangariroyashe H. Chipika, Christidi, et al., 2020). All participants provided written informed consent in accordance with the ethics approval of the Ethics Medical Research Committee of Beaumont Hospital, Dublin, Ireland. 651 additional HCs were also included from the Cambridge Center for Ageing and Neuroscience (Cam-CAN) data base resulting in a total of 776 healthy controls (HC: 393 females, mean age 55.08±17.63 years) (Shafto et al., 2014).

#### *7.4.3.2 Imaging pulse sequences*

All local participants were scanned with uniform scanning parameters on a 3 Tesla Philips Achieva scanner using an 8-channel receiver head coil. As described previously (Bede et al., 2019), a 3D Inversion Recovery Prepared Spoiled Gradient Recalled Echo (IP-SPGR) pulse sequence was utilised to acquire T1-weihted images. Acquisition details: repetition time (TR) / echo time (TE) = 8.5/3.9 ms, inversion time (TI) = 1060 ms, field-of-view (FOV): 256 x 256 x 160 mm, spatial resolution: 1 mm3. To assess vascular white matter lesion load FLAIR images were also acquired from each participant. The Cam-CAN control subjects were scanned with a T1-weighted MPRAGE sequence on a 3T Siemens Magnetom TrioTrim scanner at the University of Cambridge, using the following image acquisition parameters: TR/TE 2.25/2.99 ms, TI 900 ms, FOV= 256 x 240 x 192 mm; spatial resolution 1 $mm^3$ (Shafto et al., 2014).

#### *Pre-processing*

All subjects' T1-weighted data were first pre-processed with FreeSurfer's *recon-all* pipeline to reconstruct and parcellate the cortical surface and generate a cortical thickness (CT) map, which estimates CT at each vertex point of the cortical surface. All CT maps were subsequently transformed to the CIFTI file format at a 32k resolution per hemisphere (Connectivity Informatics Technology Initiative (Marcus et al., 2011; Van Essen et al., 2013), using the *Ciftify* toolbox (Dickie et al., 2019). Finally, each subject's CT map was parcellated into 1000 equally-sized patches, or 'mosaics', using a local-global cortical parcellation scheme proposed by (Schaefer et al., 2018), which further refines a previously published 7-brain-network cortical parcellation framework published by (Yeo et al., 2011).

#### *7.4.3.3 Statistical analysis: The standard approach*

A one-factorial, two-level, between-subjects comparison was first conducted between each patient group and controls controlling for age and gender. To correct for alpha-level inflation, we used a Monte-Carlo permutation procedure to obtain family-wise error-corrected (FWER) p-values (5000 permutations; thresholded at the voxel-level). These analyses were ran within the SPM-based toolbox (http://www.fil.ion.ucl.ac.uk) *Multivariate and repeated measures* (McFarquhar et al., 2016).

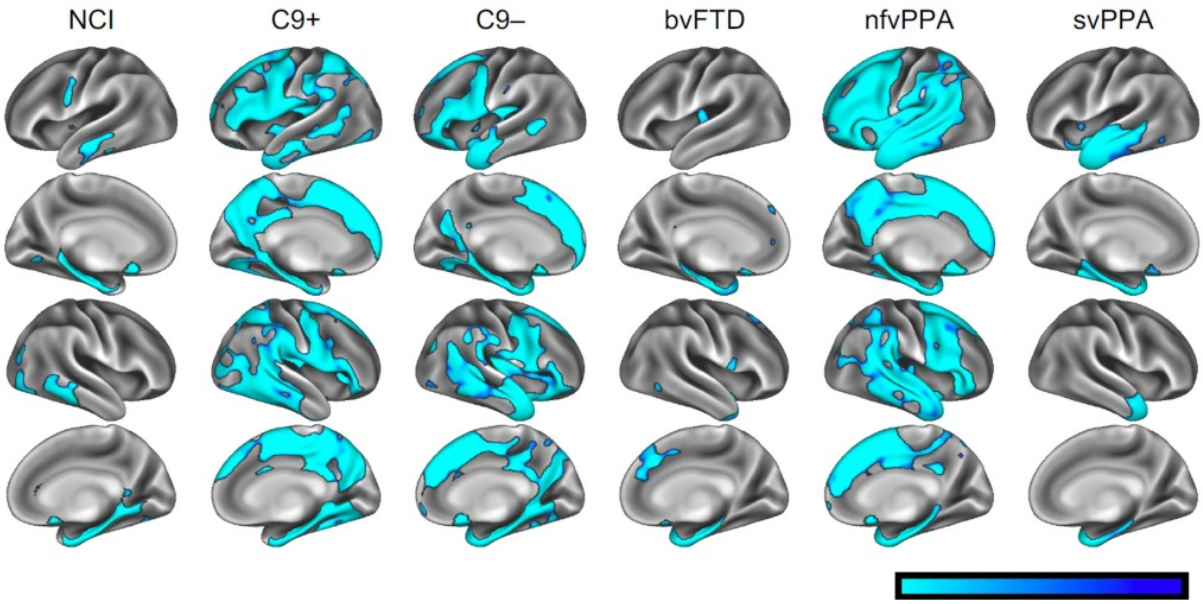



Figure 7.4.1 'Standard' cortical thickness analyses using voxelwise permutation testing, corrected for age and gender; family-wise error corrected p-maps are presented for the six clinical groups with reference to healthy controls. NCI: ALS patients with no cognitive impairment, C9+: ALS-FTD patients with *C9orf72* hexanucleotide expansions, C9-: ALS-FTD patients without *C9orf72* hexanucleotide expansions, bvFTD: behavioural variant FTD, nfvPPA: non-fluent variant primary progressive aphasia, svPPA: semantic variant primary progressive aphasia

#### *7.4.3.4 Statistical analyses: the 'mosaic' approach*

To appraise cortical thinning at an individual level, each CT map was rated with respect to an age- and sex-matched control group. Since neurite density varies significantly across the cortex (Fukutomi et al., 2018), CT was averaged across small 'mosaics', defined by a 1000-patch atlas. For each mosaic, null distributions were built non-parametrically as follows: First, the average CT value of each HC was z-scored with respect to all remaining controls to obtain a distribution at the size of the control group. Likewise, an individual patient's CT was z-scored with respect to all HC. P-values reflecting expected probabilities of cortical thinning were then calculated by counting how many values in the control distribution were smaller than the observed patient's and dividing that count by the number of subjects in the control group. We considered mosaics with p-values ≤ 0.05 as significantly thin or 'atrophic'. To account for confounding effects of age and gender (Trojsi et al., 2020), we customized the reference groups: For each patient, we only included age- and gender-matched controls from the mixed control cohort (in total 776 HC). 'Age-matched' was defined as +/– 2 years from the patients' age. As demonstrated before (Tahedl et al., 2021), this strategy successfully corrects for variance introduced by demographic confounders. This strategy generates a binary atrophic/not-atrophic label to each cortical mosaic with reference to demographically matched controls, enables the calculation of the number of 'significantly thin' mosaics throughout the cortex, as well as its fraction with respect to all evaluated mosaics. To co-validate the output of this method with the 'gold standard' approach we juxtaposed our findings with standard cortical thickness analyses.

#### *7.4.3.5 Inferential statistics of 'mosaic' maps*

The output maps of the mosaic approach can be readily visualized for individual patients indicating whether a cortical region (mosaic) is atrophic ('hit') or not with respect of demographically matched controls. (Figure 7.4.2). However, these outputs can also be at group level; we employed a Monte-Carlo permutation testing scheme to compare each of the clinical groups to HCs. In brief, we first generated a matrix with the dimensions of $n_{Patients}$ x $n_{mosaics}$ for each clinical group, indicating for each element either

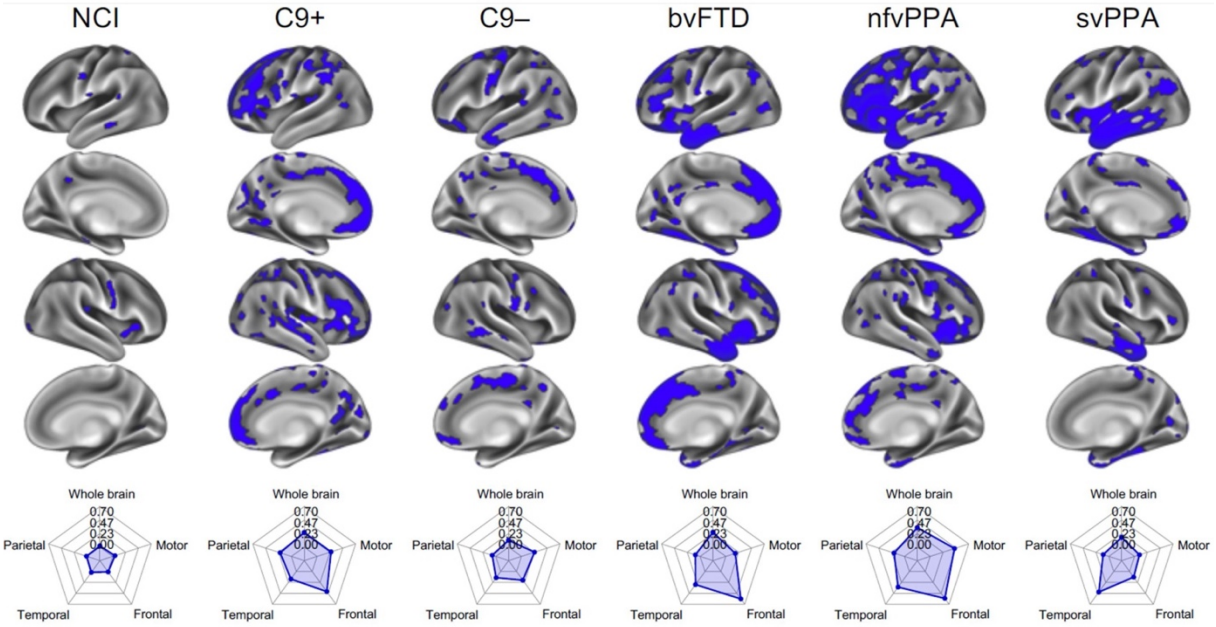



Figure 7.4.2 Individual data interpretation in single patients using the 'mosaic' pipeline; representative examples are shown from each clinical groups. Blue colour indicates cortical thinning with respect to demographically matched controls. Radar charts indicate the fraction of affected 'mosaics' in frontal, parietal, temporal and motor cortices as well as over the entire cortex. NCI: ALS patients with no cognitive impairment, C9+: ALS-FTD patients with *C9orf72* hexanucleotide expansions, C9-: ALS-FTD patients without *C9orf72* hexanucleotide expansions, bvFTD: behavioural variant FTD, nfvPPA: non-fluent variant primary progressive aphasia, svPPA: semantic variant primary progressive aphasia.

the presence ('1') or absence ('0') of regional atrophy. We then shuffled that matrix 100,000 times across mosaics, whereby we saved the count of patients with 1s at each iteration. As a result, we obtained non-parametric distributions, comprised of 100,000 values per mosaic, based on which FWER p-values can be calculated by counting the number of values exceeding the observed number of hits in the data and dividing that count by the number of iterations. We considered p-values ≤ 0.05 as statistically significant. Mathematical analyses were conducted within MATLAB version R2019b (The Mathworks, Natick, MA, USA).

#### *7.4.3.6 Between group contrasts*

Based on the 'mosaic' approach, a one-way, six-level analysis of variance (ANOVA) was conducted to ascertain differences among means of whole-brain thin-patch-fractions between the clinical groups. Based on the 'standard' approach, the means of raw CT values were also compared with the inclusion of age and gender as covariates (ANCOVA), since, as opposed to the mosaic approach, these are not inherently accounted for. As the ANOVA/ANCOVA revealed statistically significant effects, post-hoc testing was conducted. Tukey's honestly significant difference testing (HSD) using type III errors were utilised for pairwise comparisons. For post-hoc testing, age was converted into a categorical variable by assigning each patient to one of six separate age groups, since only categorical confounders can be accounted for in Tukey HSD. All statistical analyses were conducted within RStudio (version 1.3.1093, R Core Team, R Foundation for Statistical Computing, Vienna, Austria).

#### *7.4.3.7 Region-of-interest statistics*

To further characterise regional disease-burden, we calculated fractional thin-patch-counts for four large regions of interest (ROIs): motor cortex (i.e. pre-/paracentral gyri), parietal, temporal and frontal

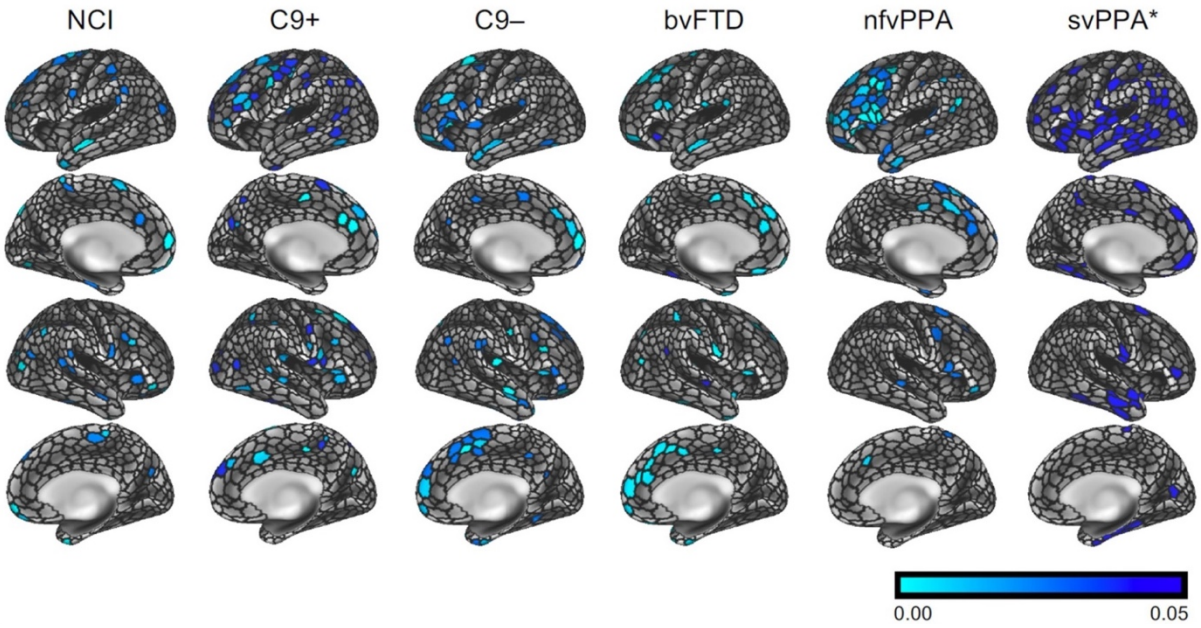



Figure 7.4.3 Inferential statistics; group-level atrophy patterns derived from the 'mosaic' approach. Family-wise error-corrected p-maps are presented at p<.05. For svPPA a threshold of p<.06 is shown. NCI: ALS patients with no cognitive impairment, C9+: ALS-FTD patients with *C9orf72* hexanucleotide expansions, C9-: ALS-FTD patients without *C9orf72* hexanucleotide expansions, bvFTD: behavioural variant FTD, nfvPPA: non-fluent variant primary progressive aphasia, svPPA: semantic variant primary progressive aphasia.

cortices. The 1000-patch mosaic-parcellation was overlaid the anatomically-defined Desikan-Killiany atlas (Figure 7.4.4a) resulting in 122 mosaics in the motor, 185 in the parietal, 150 in the temporal and 200 in the frontal cortices. For each patient, we calculated the fraction of atrophic mosaics, and averaged that fraction across subjects in each clinical subgroup. To highlight the preferential involvement of main brain regions in each phenotype, we generated radar plots (Figure 7.4.4b), in which whole-brain fractional thin-patch-counts were also incorporated. Regional radar plots were also generated to characterise regional involvement in individual patients (Figure 7.4.2).

### 7.4.4 Results

Standard cortical thickness analyses confirmed subgroup-specific patterns of cortical atrophy consistent with the clinical diagnosis (Figure 7.4.1). The 'mosaic-based' approach has successfully generated individual atrophy maps for each patient with reference to controls (Figure 7.4.2). Group-level observations could also be inferred from the 'mosaic-based' approach following permutation testing. (Figure 7.4.3) These results were anatomically consistent with the outputs of the 'standard approach'. (Figure 7.4.1). Group-level traits deduced from the 'mosaic-based' approach produced more focal and better demarcated atrophy maps than those generated by the standard approach. This is best demonstrated by the C9+ALS-FTD group where atrophy is not just more widespread than the C9–ALS-FTD group, but the precentral gyrus is more affected. Cortical atrophy patterns derived from the 'mosaic-approach' are also more focal and less noisy in the nfvPPA group than the in the maps generated by the standard approach.

Both the 'mosaic' and the 'standard' approach indicated intergroup differences (Figure 7.4.4a/c) (mosaic approach: $F(5) = 14.86$, $p = 8.73e\text{-}11$; standard approach: $F(5) = 14.89$, $p = 9.50e\text{-}11$). Post-hoc testing revealed that least affected study group was ALSnci compared to all the other diagnostic categories (Figure 7.4.4b). ALSnci vs. C9–(0.202 +/– 0.132), $p_{adj} = 1.76e\text{-}04$; ALSnci vs. C9+(0.214 +/– 0.100), $p_{adj} = 2.54e\text{-}05$; ALSnci vs. bvFTD (0.208 +/– 0.076), $p_{adj} = 2.01e\text{-}02$; ALSnci vs. nfvPPA (0.321 +/– 0.121), $p_{adj} < 0.0001$. The same pattern was observed for the standard approach (Figure 7.4.4d), where

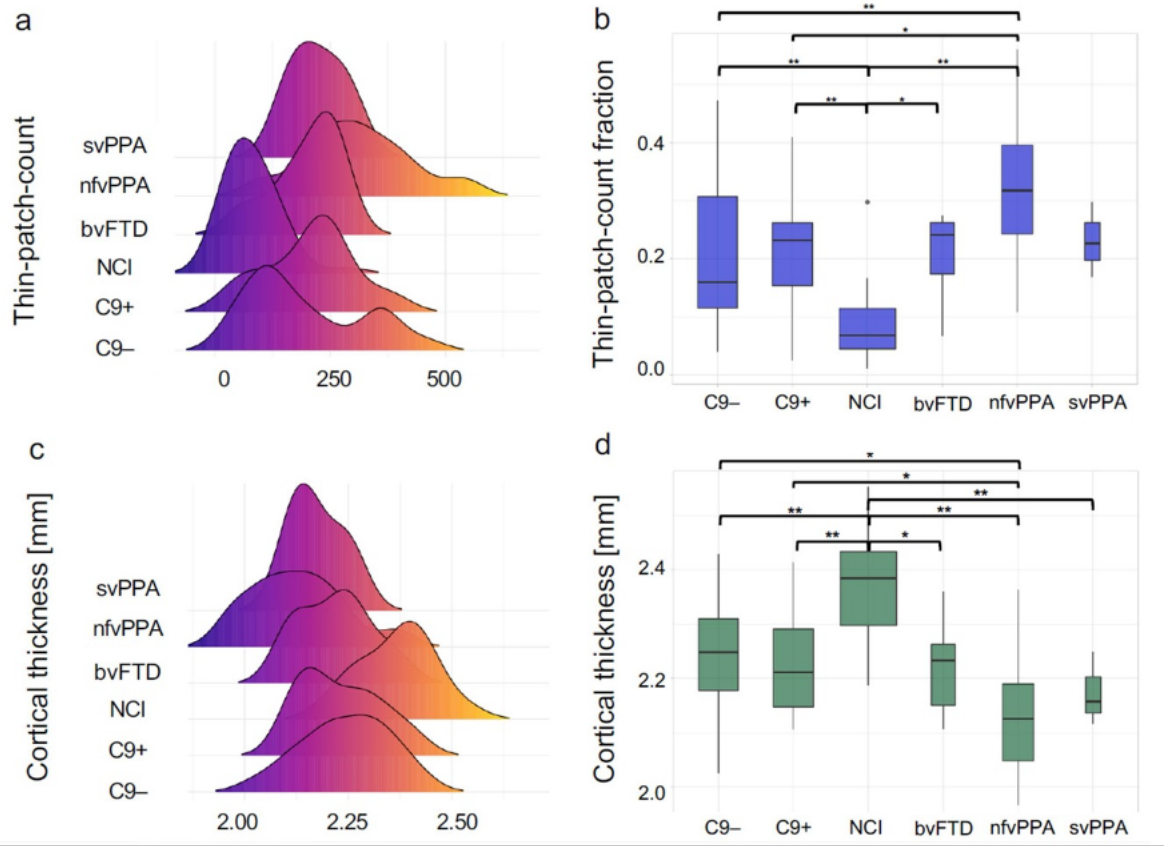



Figure 7.4.4 The comparison of group profiles; distribution of the number of thin patches derived from the 'mosaic approach' (a) and cortical thickness values as calculated by the 'standard approach' (c). Group differences in the number of thin patches (b) and mean cortical thickness (d). * indicates post hoc intergroup difference at $p_{adj} \leq 0.05$, (**) at $p_{adj} \leq 0.001$ following Tukey HSD testing. The widths of box plots indicate sample size and error bars represent 1.5 times the interquartile range. NCI: ALS patients with no cognitive impairment, C9+: ALS-FTD patients with C9orf72 hexanucleotide expansions, C9-: ALS-FTD patients without C9orf72 hexanucleotide expansions, bvFTD: behavioural variant FTD, nfvPPA: non-fluent variant primary progressive aphasia, svPPA: semantic variant primary progressive aphasia.

the ALSnci group exhibited higher CT in the pairwise comparisons than all other groups: ALSnci vs. C9– (2.24 mm +/–0.11 mm), $p_{adj}$ = 1.85e-04; ALSnci vs. C9+ (2.23 mm +/–0.10 mm), $p_{adj}$ = 2.86e-05; ALSnci vs. bvFTD (2.22 mm +/–0.09 mm), $p_{adj}$ = 4.30e-03; ALSnci vs. nfvPPA (2.13 mm +/–0.11 mm), $p_{adj}$ < 0.0001; ALSnci vs. svPPA (2.17 mm +/–0.07 mm), $p_{adj}$ = 1.61e-02. In contrast, the most affected clinical group was nfvPPA, where the mean thin-patch-count fraction was not only higher than that of the ALSnci group, but also the C9–ALSFTD ($p_{adj}$ = 9.45e-03) and the C9+ALSFTD ($p_{adj}$ = 2.80e-02). Again, this pattern was mirrored by the standard approach, where the ALSnci group not only showed higher mean values than the nfvPPA group, but just as in the mosaic approach, also the C9–ALSFTD ($p_{adj}$ = 1.72e-02) and the C9+ALSFTD ($p_{adj}$ = 4.66e-02) groups.

Our region-of-interest statistics evaluated thin-patch-count fraction per 'ROI' (Figure 7.4.5a) and confirmed the preferential involvement of ROIs in the study groups. (Figure 7.4.5b) The most anatomically widespread disease-burden was detected in nfvPPA (largest radius), the least pathology in ALSnci (smallest radius) and the most focal involvement in svPPA (temporal cortex).

### 7.4.5 Discussion

Our findings demonstrate the feasibility of interpreting single T1-weighted images from individual patients to generate maps of atrophy. We have shown that cortical regions can be successfully categorised as atrophic or unaffected in single subjects with respect to a databank of controls. A z-score

based approach not only enables the appraisal of cortical disease-burden in individual-subjects, but group-level patterns may also be inferred. The output maps of the proposed 'mosaic' approach are anatomically concordant with gold standard cortical thickness analyses. The topography of cortical thinning can be reported visually, numerically and in an ROI-based representation at both individual- and group-level. The pipeline is based on quantitative cortical thickness measurements, an atlas-based parcellation and is fully observer independent. In its current form it is computationally demanding, but all the mathematical steps utilised could be integrated into a single computer script and run either as a cloud-based solution or installed locally on the MR platform or data server.

In this paper we have demonstrated the utility of this approach in FTD phenotypes, but this method could potentially also be utilised in neurodegenerative conditions where the ascertainment of cortical atrophy patterns is clinically relevant (Malek Abidi et al., 2021; Christidi et al., 2019; Finegan, Chipika, Li Hi Shing, et al., 2019; Nasseroleslami et al., 2019; Seo et al., 2010). The technique relies on the binary labelling of cortical regions as 'atrophic' or 'normal'. This is fundamentally a reductionist approach, but given the very high number of cortical regions (mosaics), it is a successful strategy as demonstrated by the detection of confluent cortical areas. The generation of putative atrophy maps provides an instant representation of the anatomical expansion, focality and lobar predominance of disease burden. These colour coded maps are potentially useful to illustrate affected regions to patients, caregivers and members of the multidisciplinary team. This starkly contrasts with the current practice of pointing at presumed regions of atrophy on black and white 2D images which are difficult to decipher by laypeople (Harper et al., 2014). The z-score derived, 'mosaic' method may not only be applied to those with an established diagnosis, but also to those with a suspected diagnosis or pre-symptomatic mutation carriers to characterise disease burden distribution.

In a clinical setting, progressive frontotemporal pathology is often monitored by validated neuropsychological tests (Burke et al., 2017; Burke, Elamin, et al., 2016; Marwa Elamin et al., 2017). Cognitive assessment however may be particularly challenging in certain FTD phenotypes, especially in ALS-FTD where motor disability and dysarthria may preclude the use of certain tests (Burke, Elamin, et al., 2016; Verstraete et al., 2015; Yunusova et al., 2019). In other FTD phenotypes, performance on neuropsychological testing may be confounded by mood, apathy, cognitive reserve and practice-effects which highlight the role of neuroimaging in tracking progressive changes (Costello et al., 2021; Radakovic et al., 2016).

Quantitative cortical thickness mapping may also give additional reassurance to those who fear a particular diagnosis despite scoring high on neuropsychological tests (O Hardiman et al., 2016). This is often a significant source of anxiety for patients, particularly for those who have first-hand witnessed a family member or close friend carrying a certain a diagnosis. Immediate answers would provide early reassurance, alleviating the sense of heightened stress and anxiety. The implementation of this method may be relatively straightforward as most patients undergo a routine MRI brain scan as part of the current diagnostic pathway (Harper et al., 2014).

Despite the clinical rationale to devise such frameworks, our study has a number of limitations. The sample size of the various patient groups in this pilot study is relatively small necessitating validation

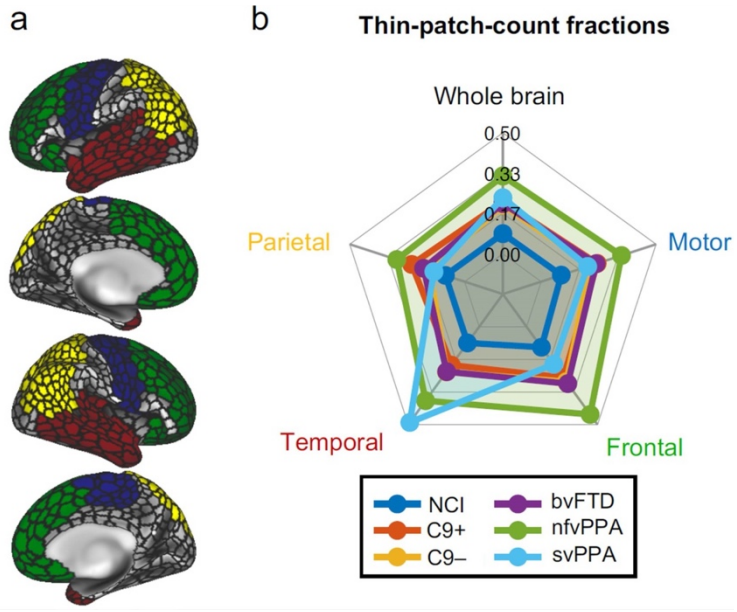



Figure 7.4.5 Regional disease burden; cortical thinning was further evaluated in four atlas-defined regions-of-interest (ROIs) in the motor (blue), parietal (yellow), temporal (red) and frontal (green) cortices and over the entire cerebral cortex (a). The fraction of atrophic 'mosaics' was calculated in each patient within each ROI with respect to the total number of mosaics comprising the given ROI. The distribution of disease burden in the patient groups is presented as radar charts (b). NCI: ALS patients with no cognitive impairment, C9+: ALS-FTD patients with C9orf72 hexanucleotide expansions, C9-: ALS-FTD patients without C9orf72 hexanucleotide expansions, bvFTD: behavioural variant FTD, nfvPPA: non-fluent variant primary progressive aphasia, svPPA: semantic variant primary progressive aphasia.

in larger external datasets. All patients in our study had an established diagnosis; thus, the sensitivity of this method needs to be further evaluated in those with a suspected diagnosis, early-stage disease or in asymptomatic mutation carriers (Rangariroyashe H Chipika, Siah, et al., 2020; Li Hi Shing, McKenna, Siah, et al., 2021; Querin, Bede, et al., 2019). Moreover, only grey matter analyses were conducted, despite the contribution of white matter pathology to the clinical manifestations of these phenotypes (Bede et al., 2016; Bede, Omer, et al., 2018; Qin et al., 2021; C. Schuster et al., 2016; Zhou et al., 2010). Finally, while our approach provides individualised atrophy maps, supervised and unsupervised machine learning approaches offer direct individual patient categorisation into diagnostic and prognostic groups (Bede et al., 2017; Grollemund et al., 2020, 2021; Querin, El Mendili, et al., 2018b; Christina Schuster et al., 2016).

We envisage future applications for this methodological approach in both clinical practice and potentially in clinical trials. Consecutive MR datasets could be compared to the patients' initial scan; allowing for the objective measurement of disease-burden accumulation and the evaluation of progression rates (Bejanin et al., 2020; Rangariroyashe Hannah Chipika et al., 2019; Christina Schuster et al., 2015). Alternative imaging metrics such as spinal cord measures, network integrity indices, white matter diffusivity parameters or subcortical grey matter metrics could also be readily investigated in a similar z-score based framework (Malek Abidi et al., 2021; Dukic et al., 2019; El Mendili et al., 2019; Proudfoot et al., 2019; Querin, El Mendili, et al., 2019). Future applications would require the validation of our findings in large multicentre studies, ideally incorporating diverse patient populations across a variety of neurodegenerative disorders.

### 7.4.6 Conclusions

Our preliminary findings indicate that T1-weighted MRI data from individual patients may be meaningfully interpreted and maps of cortical atrophy can be readily generated. The outputs of our analyses are anatomically analogous with gold standard methods. This is a promising approach to interpret single subject scans with viable clinical and clinical trial utility.

### 7.4.7 Ethical approval

This study was approved by the Ethics (Medical Research) Committee – Beaumont Hospital, Dublin, Ireland.

### 7.4.8 Consent to participate

All subjects provided informed consent to participate in this study in accordance with the approval of the Ethics (Medical Research) Committee – Beaumont Hospital, Dublin, Ireland.

### 7.4.9 Consent to Publish

Study participants consented to the presentation of research findings.

### 7.4.10 Author contributions

Study design (MCMcK, MT, JL and PB), clinical assessments (MCMcK, RHC, SLHS, OH, SH, PB), genetic analyses (MAD, JCH, AV, RLM), MR data interpretation (MCMcK, MT, JL, PB), drafting the manuscript (MCMcK, MT, PB) revision of the manuscript for intellectual content (MCMcK, MT, JL, RHC, SLHS, MAD, JCH, AV, RLM, OH, SH, PB)

### 7.4.11 Funding

Peter Bede and the Computational Neuroimaging Group are supported by the Spastic Paraplegia Foundation (SPF), Health Research Board (HRB EIA-2017-019), the EU Joint Programme – Neurodegenerative Disease Research (JPND), the Andrew Lydon scholarship, the Irish Institute of Clinical Neuroscience (IICN), the Iris O'Brien Foundation. MT is funded by the Deutsche Multiple Sklerose Gesellschaft (DMSG), grant number 2018_DMSG_08.

### 7.4.12 Competing interests

The authors have no conflicts of interest to declare.

### 7.4.13 Availability of data and materials

Raw imaging data cannot be shared due to institutional policies. Additional information on processing pipelines and statistical frameworks may be requested from the corresponding author.

# 8 Chapter III: Concluding remarks

## 8.1 Recapitulation

With this dissertation, we aimed to demonstrate and validate a strategy for appraising cortical thickness (CT) changes in individual patients. Specifically, we suggest to first parcellate the cortex into small but roughly equally sized brain regions, or "patches"/"mosaics", based on existing brain-wide atlases (we therefore referred to our approach as "mosaic-approach"). For each patch, we propose to calculate a normal distribution of CT values from an age- and sex-matched control group, based on which an individual's CT can be z-scored patch-wise. We then propose to convert these z-scores into p-values using exhaustive nonparametric permutation testing. As such, individual CT profiles can be analyzed for signs of atrophy (specifically, we suggested to consider patches with $p <= .05$ as "significantly thin", or atrophic). In subproject 1, we contrasted different strategies to convert z-scores to p-values and selected the one with the best combination of sensitivity and specificity for further validation (subprojects 2–4).

Given our age-/sex-matched approach implied that confounding effects of age and gender should be corrected for, the successful elimination of variance associated with these factors needed to be formally shown, which we demonstrated in subproject 2 (studying a longitudinal data set from ALS patients). Central to clinical translation, we could also provide evidence in that study that information on disease burden, as measured by the ALSFRS-r rating scale, was maintained in the mosaic-approach.

Another implication of our measure – which is exclusively targeting the cortex – is that it should mimic the degree of cortical involvement (as opposed to independent atrophic pathology, e.g. associated with the spinal cord). The motor neuron disease (MND) spectrum offers an excellent opportunity to investigate this issue, since its different clinical subpopulations yield different degrees of upper motor neuron (and therefore cortical) and lower motor neuron (spinal cord) involvement, which we contrasted in subproject 3: PLS patients, representing a mainly upper MND, Polio survivors, representing mainly a lower MND, and finally ALS patients, representing a mixed upper-/lower MND. The results suggested that, as hypothesized, the degree of upper motor neuron involvement is mirrored by the CT involvement as assessed by the mosaic-approach, such that PLS showed the most evident cortical atrophy, Polio only very little, and ALS ranged between the two extremes. Furthermore, the mosaic-approach also allows to look for signs of unexpectedly "thick" cortical regions (by just flipping the logic for detecting atrophy and assigning p-values to each patch which reflect the likelihood of its observed *thickness* rather than its *thinness*). Such hypertrophy might signify compensatory mechanisms and neurogenesis. In our study (subproject 3), only Polio survivors exhibited such hypertrophic clusters, differentiating that clinical population from both ALS and PLS, which is in line with clinical and pathological profiles of these differential diagnoses.

Finally, to further demonstrate clinical utility, we aimed to probe whether our T1w derived metric was capable of detecting signs of atrophy at topographically distinct cortical regions (up until here we had only studied MND patients and therefore mainly motor cortex involvement). To achieve this goal, we considered diagnoses groups from a clinically similar (but pathologically dissimilar) and commonly observed comorbidity of MND-patients, namely patients suffering from different subtypes of frontotemporal dementia (FTD), in subproject 4. Although an umbrella term reflecting vast clinical (and pathological) heterogeneity, one commonality of most FTD patients is frontal and temporal neurodegeneration. Therefore, we knew the "ground truth" localization of pathology. And indeed, we found that evaluating atrophy at an individual level with our mosaic-based approach yielded the topographically distinct patterns of atrophy for the different subgroups as one would expect based on previous histological and (group-based) imaging findings.

In brief, in this dissertation we propose a strategy to assess cortical atrophy at a high-resolution level, which categorizes cortical regions of individual patients as either atrophic or not compared to demographically matched control subjects. We demonstrated high sensitivity and specificity of the method (study 1), provided evidence that it is resilient against confounding demographic factors while maintaining information on clinical disease burden (study 2), shown that it is sensitive to the degree of cortical involvement (study 3) and finally it is capable of correctly detecting atrophy at topographically distinct regions (study 4).

## 8.2 Clinical (and general) discussion

Our findings demonstrate the feasibility to **generate individualized maps of atrophy** using T1w MRI data from single patients. More specifically, we suggest categorizing distinct cortical regions as atrophic or not with respect to a matched control group. Furthermore, we demonstrated that it is viable to rely on external, high-quality control data to derive such matched control groups. As shown in study 2, this strategy is **resilient to demographic confounding effects** induced by age and gender, which is in line with previous reports highlighting the need for controlling for such variables, especially in small studies and in the interpretation of single data sets (Bede et al., 2014; Luders et al., 2014; Menzler et al., 2011). We consider it a particular strength of our mosaic-approach that information on clinical burden is maintained in spite of resilience to demographic variables.

While the proposed binary labeling of cortical regions is a reductionist approach, we provided evidence that it's a successful strategy by the capture of confluent cortical regions. The generation of putative atrophy maps provides an instant representation of the anatomical expansion, focality and lobar predominance of the disease burden. The resulting maps can be color-coded, indicating presence or absence of atrophy at distinct cortical regions (as displayed in the original associated research articles), which is potentially useful to illustrate affected regions to patients, caregivers and members of the multidisciplinary team. This overtly contrasts with the current practice of pointing at presumed regions of atrophy on black and white 2D images which are difficult to decipher by laypeople (Harper et al., 2014). This inter-rater independence is especially important to monitor patients suffering from neurodegenerative diseases such as ALS, since the current gold standard includes a range of subjective complementary measures and neurological examinations, which – despite best clinical efforts – is inadequate to capture gradual and subtle changes over time. Confounding factors including reporting bias, mood, practice effects, training and fatigue, which ultimately compromise **objectivity and reliability** of such assessment tools, may be better respected (Burke et al., 2017; Cedarbaum et al., 1999; Marwa Elamin et al., 2017). Our automated and computational z-score based approach overcomes these limits.

The key to the mosaic-approach, we believe, is however that it allows for **interpretation of individual patients**. This is a considerable advantage over existing methods, which mainly rely on group comparisons. Gray matter atrophy in neurodegenerative disorders such as ALS is typically evaluated by voxel-based morphometry, CT analyses or various region-of-interest (ROI) approaches (Agosta et al., 2018; Bede et al., 2019, 2020; Rangariroyashe H Chipika, Finegan, et al., 2020; Christidi, Karavasilis, Rentzos, et al., 2018). ALS exhibits considerable clinical heterogeneity with respect to initial disability profiles, progression rates, upper vs lower motor neuron predominance and neuropsychological deficits (Burke, Pinto-Grau, et al., 2016; Christidi et al., 2020; Feron et al., 2018; Finegan, Chipika, Shing, et al., 2019; Lebouteux et al., 2014), such that single data sets of individual patients cannot be readily interpreted – a major barrier to routine clinical use (Foerster et al., 2013). Existing imaging studies of ALS typically stratify patients into groups of genetically (Bede, Omer, et al., 2018; Westeneng et al., 2016) or phenotypically (Agosta et al., 2016; Finegan, Chipika, Li Hi Shing, et al., 2019; Finegan, Li Hi

Shing, et al., 2019) defined cohorts to reduce the effect of disease heterogeneity before running comparative statistics to ascertain group-level signatures. The drawback of this approach is that the resulting statistical maps reflect on the specific cohort included in the study, are 'over-fitted' by definition, and reveal very little about individual patients. From a clinical perspective, the quantification of disease burden in individual patients as by our approach is hugely important to help diagnostic classification, prognostic categorization, tracking the course of the disease and potentially gauging response to therapy (Rangariroyashe Hannah Chipika et al., 2019). It also yields several advantages over existing methods to interpret individualized MRI data implementing using machine learning algorithms, given those require large patient samples, accurate labeling and adequate training, and yet, overfitting is a well-known problem (Agosta et al., 2010; Bede et al., 2017; Marwa Elamin et al., 2015; Grollemund et al., 2019a, 2020; Querin, El Mendili, et al., 2018a; Christina Schuster et al., 2017; Welsh et al., 2013; Westeneng et al., 2018).

For **validation** purposes, another advantage of our method is that the mosaic-approach not only enables the appraisal of cortical disease-burden in individual subjects, but group-level patterns may also be inferred by employing permutation testing across patients (studies 2–4). This allowed us to compare the output maps of the proposed mosaic-approach to the gold standard, group-based, imaging CT analyses to demonstrate the anatomical concordance between the two methods. Moreover, the cerebral signatures revealed by the mosaic-approach were also in line with histologic findings of cortical involvement (Brettschneider et al., 2013, 2014; Mackenzie & Briemberg, 2020), independent imaging studies (Bede et al., 2019, 2020; Finegan, Chipika, Li Hi Shing, et al., 2019; Müller et al., 2012; Christina Schuster et al., 2013; van der Graaff et al., 2011) and neuropsychological findings (de Vries et al., 2019; Finegan et al., 2021).

One aspect of our method which we hope will further promote clinical translation is that the topography of **cortical thinning can be reported visually**, numerically and in an ROI-based representation at both individual- and group-level. Although computationally expensive in its current form (requiring on average 6.5 hours per subject on an i9 3.6 GHz processor with 64 GB RAM[2] – however parallel processing is possible), but all the mathematical steps utilized can be integrated into a single computer script and run either on a cloud-based software solution or installed locally on the MR platform or data server.

Our findings emphasize the utility of our approach for **differential diagnosis**: we were able to identify divergent cortical signatures in UMN-predominant, LMN-predominant and mixed UMN-LMN MND phenotypes (study 3). Our results also revealed co-existing atrophy and increased cortical thickness in MND, in both LMN- and UMN-predominant phenotypes. Furthermore, the utility of this approach for related conditions such as FTD phenotypes was demonstrated in study 4, which suggests that this method could potentially also be employed to diagnose or differentiate other neurodegenerative conditions where the ascertainment of patterns of cortical atrophy is clinically relevant (Du et al., 2006; M. Li et al., 2017; Seo et al., 2010). In addition, future work will need to explore whether the z-score derived, mosaic-approach may not only be applied to those with a suspected diagnosis, but also to those with an established diagnosis or pre-symptomatic mutation carriers, offering the opportunity to quantify the extent of disease burden. Quantitative CT mapping may also give reassurance to those who fear a particular diagnosis despite scoring high on neuropsychological tests. This is often a significant source of anxiety for patients, particularly for those who have first-hand witnessed a family member or close friend carrying a certain a diagnosis. Immediate answers could

[2] The provided processing time was observed on the PhD candidate's machine, which is an iMac 3.6 GHz 8-core Intel i9 (2019) with 64 GB RAM and a Radeon Pro Vega 48 8 GB graphics card (note that eight subjects could be processed in parallel, each taking ~6.5 hours) and was not tested on different systems.

provide early reassurance, alleviating the sense of heightened stress and anxiety. The implementation of this method may be relatively straightforward as most patients undergo a routine T1w MRI brain scan as part of the current diagnostic pathway (Harper et al., 2014). Early diagnosis can not only facilitate distress for patients and their caregivers facing diagnostic uncertainty but also accelerate recruitment into clinical trials (Zhou et al., 2010; Zoccolella et al., 2006), which notoriously suffer from small sample sizes in relatively rare neurological conditions such as ALS. A proxy of cerebral disease burden, such as the extent of cortical atrophy we propose to assess herein, may have practical utility obviating the need for large patient cohorts. Interestingly, while the concept of individually-assessed disease burden quantification and monitoring is firmly integrated in the management and clinical trials of other neurological conditions, such as lesion load / volume measurements in MS, no comparable strategies have been implemented in ALS trials to date (Filippi et al., 1998).

Finally, two clinically interesting "by-products" of our methodological development shall be highlighted, namely (i) the utility of longitudinal imaging for differential pathologic characterization of clinical subgroups and (ii) the observation of hypertrophy in selected neurological disorders. Our findings from study 3 suggested that longitudinal studies may be better suited to differentiate MND phenotypes than cross-sectional imaging, as propagation patterns and rate of progression may be more specific to MND phenotypes than a snapshot of cortical disease burden. This is in line with previous reports that failed to provide consistent cross-sectional differences between ALS and PLS (many studies yield contradictory results, e.g. (Ferraro et al., 2017; Müller et al., 2018) and (Agosta, Galantucci, et al., 2014; van der Graaff et al., 2011)), which suggests that longitudinal studies may be better suited to capture distinguishing features (Rangariroyashe Hannah Chipika et al., 2019; Müller et al., 2016, 2020). Clinically, however, the distinction of PLS and ALS is hugely important in the first years following symptom onset, as there is often a diagnostic dilemma in early PLS, and apprehension that a patient with UMN-predominant symptoms may transition to ALS.

Furthermore, we have showcased the need for two-way analyses in MND – investigating both atrophy and hypertrophy – given that polio-survivors exhibited increased CT in both motor and extra-motor regions (study 3), revealing a unique cerebral signature. This also suggests that hypothesis-aligned, one-way contrasts may lead to skewed results. Such biases in analyses can lead to missing signs of adaptive cortical reorganization, which are well-known to occur especially in spinal cord injuries (Nishimura & Isa, 2009), in response to repetitive tasks, arduous physical training, dexterity associated with musical instruments, and association with cognitive training. Adaptive motor cortex changes have been repeatedly noted after unimanual training (Sale et al., 2017), aerobic exercise (Colcombe et al., 2006), post stroke (Sterr et al., 2013), and in professional musicians (Bruchhage et al., 2020; Gaser & Schlaug, 2003; Hudziak et al., 2014). Extra-motor changes have also been observed following various exercise regimes (Pereira et al., 2007; Thomas et al., 2016) and cognitive tasks (Lazar et al., 2005). Despite ample examples of cortical volume gains in a range of conditions, CT increases or volume gains are seldom evaluated or reported in MNDs. Here, the mosaic-based approach may help to increase the awareness of such regenerative processes and potentially serve as a tool to capture and quantify brain repair in clinical trials (but see the final section "Future perspectives" on an extended discussion on this opportunity).

## 8.3 Limitations

We have proposed a framework for meaningful interpretation of single patients' cortical thickness MRI data to infer on atrophy. The main idea is to first z-score an individual's data to a matched control population and then assign p-values to these scores based on nonparametric permutation testing. To note, the idea to rate an individual's MRI data with respect to a control population is not new, and we

do not want to take credit for that. For example, it has been extensively studied and tested in the AD literature, whereas standardized volume changes of various gray matter structures including the medial temporal lobe structures (Ma et al., 2016), the posterior cingulate cortex or the precuneus (Shimoda et al., 2015) as well as the hippocampus (den Heijer et al., 2010) were discussed as candidates for clinical diagnosis / prognosis. However, to our knowledge, we are the first ones to provide (i) carefully selected reference groups (in terms of age and gender), (ii) along with a thorough investigation of sensitivity and specificity of opposing permutation testing procedures for a (iii) a high-resolution parcellation of the cortex. Moreover, our strategy is easy-to-implement such that visualizations of potentially atrophic patches are simple to derive from a T1w scan. Despite the clinical rationale to devise such a method, this work has a number of limitations, both methodologically and clinically. We will discuss the most important limitations in the next sections.

### 8.3.1 Methodological considerations

One limitation of our currently proposed pipeline are the relatively small sample size of the various control groups in our pilot studies. Although we combined data from two sources (namely Cam-CAN and HC data from Trinity College Dublin, TCD), the smallest reference group was $n$ = 16 (females between 81 and 85) and the largest $n$ = 49 (males between 64 and 68; cf. also study 2). Given we assigned p-values to the z-scores in a permutation testing scheme (following a "singleton vs. group" design), the resulting p-values were highly discrete (Winkler et al., 2014). Although this procedure still produces valid tests even for small sample sizes, the smallest possible discrimination value is ($n$+1)-choose-$n$, which is 1/17 = 0.0588 for our smallest and 1/49 = 0.0204 for our largest reference group. Therefore, significance could only be attributed to a patient matched to the smallest reference group if the observed value was *smaller than all values* found in the reference distribution (resulting in a $p$-value of 0.000). Such insufficient numbers of permutations may interfere with the sensitivity and specificity of the method, and it will be necessary for clinical translation to boost the sample size of the control population to produce valid results. Currently, we are making efforts to obtain control data from the UK Biobank and the German NAKO study ("Nationale Kohorte") for further development.

Another methodological issue which needs to be investigated are the effects of the different scanner / acquisition parameters from the control population (in our case HCP [study 1] and Cam-CAN / TCD [studies 2–4], respectively) and the patient data (in our case the TCD-derived data). In our present studies, we have combined HC data from Cam-CAN and TCD (whereas the HC data from TCD yielded the same scanner / acquisition parameters as the patients'). To get a first impression of the extent of the effects, we ran an analysis of covariance to investigate effects of time (i.e. ageing) on CT between the two scanning sites (Cam-CAN vs. TCD, see *Figure 2*). While we found significant main effects of site (TCD data being thinner by a mean value of 0.544 mm, $F(1) = 66.730$, $p$ = 1.27e-15) and age (with CT decreasing over time, $F(1) = 768.391$, $p$ = 2.6e-16), there was no significant interaction effect between the two factors ($F(1) = 2.254$, $p = 0.134$). This indicates that CT is decreasing with age commensurably between Cam-CAN and TCD subjects. However, it is important to note that given the TCD data yielded thinner CT to start with, atrophy might have been assigned to patients more readily. This is especially problematic for cross-sectional interpretation of the data and is another reason why larger and better-matched HC cohorts need to be considered for building reference populations when aiming for clinical translation. Nevertheless, longitudinal CT changes are less problematic to interpret: given the reference data does not change, any longitudinally observed CT changes in a patient are likely attributable to true changes in that individual. Therefore, we believe that the tool is already now useful to progress-monitor an individual patient's disease course.

### 8.3.2 Clinical considerations

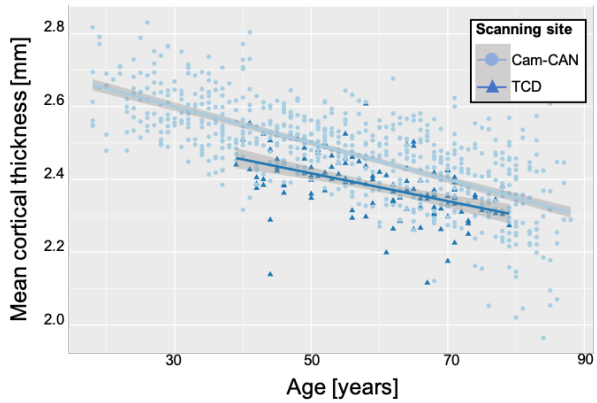



Figure 8.3.1. In studies 2–4, we combined reference data from two healthy control populations, namely the open-source repository from the Cambridge Centre for Ageing and Neuroscience ("Cam-CAN") and local data obtained at Trinity College Dublin ("TCD"; note that our patient data was obtained at the same scanning site as TCD using the identical acquisition parameters). We compared the evolution of cortical thickness (CT) in the two groups over time (i.e. ageing) in an analysis of covariance. We found no significant interaction effect (i.e. CT decreases similarly in both groups over age), however there was a main effect of scanning site, such that atrophy might be assigned more readily to patient data from TCD. This needs to be considered for cross-sectional interpretation of the current pipeline and suggests harmonization of acquisition schemes when aiming for clinical translation of the proposed method.

From a clinical perspective, one limitation is that all included patients (ALS / PLS / Polio survivors / FTD) fulfilled the diagnostic criteria for their respective condition. The inclusion of pre-symptomatic patients, e.g. mutation carriers of the C9orf72 gene, which predisposes towards developing ALS (Bede et al., 2013), would have helped to evaluate the detection sensitivity of our method further.

Such subjects are hard to recruit, however, and hence we are currently working on improving another limitation of our method: Currently, our pipeline only considers supratentorial cortical regions, although also cerebellar and subcortical gray matter regions are well-known foci of pathology of neurodegenerative diseases such as ALS (Rangariroyashe H Chipika, Finegan, et al., 2020; Pradat et al., 2009; Pradat & El Mendili, 2014; Westeneng et al., 2015). In particular, however, *white matter pathology* is an often observed clinical manifestation of these diagnoses (Bede, Omer, et al., 2018; Bede & Hardiman, 2018; Qin et al., 2021; Zhou et al., 2010). To acknowledge the central role of white matter pathology in neurodegenerative disorders, we aim to also provide a tool for meaningful interpretation of white matter microstructure alterations of single patients. The basic idea is very similar to what we presented herein for cortical thickness: First, null distributions for diverse diffusion tensor imaging (DTI)-derived scalars (including axial diffusivity, fractional anisotropy, radial diffusivity) for anatomically distinct tracts are built. Based on these, a single patient's metrics can be rated with respect to his or her matched control population. While this is still work in progress, we show in *Figure 3* how the results can look like: We display the results from this pipeline for a single MND patient (59 years, male) and show the results for radial diffusivity (left panel), fractional anisotropy (middle panel) and axial diffusivity (right panel). The presence of a tract is to be interpreted as "significant unexpected" extent of the observed DTI-metric; the color denotes the direction: red indicates "significantly increased", blue "significantly decreased". In this example, radial and axial diffusivity for this patient are increased for a wide number of tracts, including the tracts histologically found to be affected in MND such as the corticospinal tracts and cerebellar peduncles (this is in line with the current opinion that axial and radial diffusivity increases reflect axonal and myelin damage, (Winklewski et al., 2018), which is often also mirrored by decreases of fractional anisotropy (Baek et al., 2020), as also observed in this patient). We hope that with the release of that work, we will be able to provide another tool for meaningful interpretation of a single patient's MRI scan with respect to white matter pathology (in this case diffusion-weighted imaging) to further help clinical translation of advanced neuroimaging techniques.

Here, I would like to make a brief excursion on the rationale of using DTI-derived metrics: Over the past years, it has emerged that those yield interpretational problems in regions of crossing fibers, such

that e.g. both increases and decreases of one and the same metric can reflect pathology (Alba-Ferrara & de Erausquin, 2013). To overcome this limit, more advanced white matter metrics have been proposed, which often rely on constrained spherical deconvolution (CSD, (Tournier et al., 2008)) rather than the tensor model (Basser et al., 1994). *Apparent fibre density* and *fibre-density cross section* (D. Raffelt et al., 2012; D. A. Raffelt et al., 2017) are just two examples which target distinct fiber populations in a voxel (rather than all fiber population's average, as do tensor-derived metrics) and quantify their morphometry. These measures have already been shown to be developmentally and clinically relevant (Pecheva et al., 2019) but are more sensitive as compared to tensor-derived metrics (Pannek et al., 2018) given they are resilient to confounding by crossing fibers (Gajamange et al., 2018). These promising results – along with the argument that up to 90% of brain voxels contain crossing fibers (Jeurissen et al., 2013) – were compelling enough for the PhD candidate to delve thoroughly into the methods of CSD-based analyses. I have summarized my results in an extensive tutorial (along with complementary tutorial data) which is freely available on OSF (https://osf.io/fkyht/). This tutorial has been widely used in the community (currently counting > 3,200 downloads, as of July 2021) and won second prize at the 2018 tutorial contest of the MICCAI society (Medical Image Computing and Computer Assisted Intervention Society). Efforts to turn this tutorial into a protocol paper are ongoing; however, one intention of this work was to apply these CSD-based metrics for the proposed pipeline for rating single patient's white matter microstructure. The reason why this has not been realized with the above-described work in progress is that CSD-based metrics require certain acquisition standards such that they are straightforward to interpret (*b*-values of > 3,000 s/mm$^2$ have been discussed (D. Raffelt et al., 2012), but even higher *b*-values as well as multi-shell acquisition schemes should further improve interpretability, (Genc et al., 2020)). Unfortunately, the diffusion-weighted data we currently have available for developing the method is single-shell data with *b*-values of 1,000 s/mm$^2$. Accordingly, we decided to stick with the well-characterized tensor-based metrics with these data, and progress development with more advanced white matter quantifications once sophisticated data are disposable.

## 8.4 Future perspectives: Towards detecting brain repair?

In this dissertation, we have repeatedly highlighted the potential utility of the here-proposed atrophy-detection tool for clinical practice, which can aid to quantitatively progress-monitor a patient's disease course. In addition, we have raised the possibility of such a measure serving as a new endpoint in clinical trials. In this section, we would like to discuss the rationale of the latter opportunity further.

We have commenced this thesis with a patient's medical journey: Typically, once they notice physical problems, they go to see a general practitioner who then might refer them to a specialist if an accurate diagnosis cannot be derived. Once the diagnosis is established, there are guidelines which aid the doctor to initiate an appropriate therapy. And this is what the patient ultimately seeks – therapy to stop, ameliorate or ideally reverse their problems. Often, the presence of such medical guidelines is taken for granted – but the journey for any therapy to find a way into generally accepted guidelines is long and requires approval by a range of administrative bodies, such as the Food and Drug Administration (FDA), which can (ideally) only be satisfied by rigorous, exhaustive, scientifically- and medically-thorough clinical investigations. Coarsely speaking, a successful clinical trial requires evidence that efficacy of a new therapy outperforms that of existing treatment strategies. But what defines efficacy? Typically, efficacy demonstration requires comparative evaluation on a measure which indicates the current state of the pathology of a patient, also referred to as "biomarker". Such biomarkers need to be objective, reliable and valid, in order to meaningfully mimic the degree of disease burden (Craig-Schapiro et al., 2009). As we have shown, the atrophy assessment tool we have proposed, which we referred to as mosaic-approach, meets all of these criteria; however, to be useful for revealing

*curing* aspects of a therapy – the ultimate goal of medical treatment – there need to be ways to change it. Therefore, it is crucial to evaluate whether cortical thickness can actually be changed by any drug when attempting to translate this tool as an endpoint for clinical trials where brain repair is aspired.

One obvious way to increase CT is to increase the number of neurons – however, for a long time, the common opinion was that neurons cannot be generated de novo *in vivo*. This notion changed fundamentally with experiments like the London Taxi Driver Study, where hippocampal sizes were shown to correlate with the navigation experiences of taxi drivers (E A Maguire et al., 2000). Many similar studies followed (Hüfner et al., 2010; Eleanor A. Maguire et al., 2006; Martin et al., 2007; Mechelli et al., 2004) – including work from our university where hyperthalamic hypertrophy was demonstrated in narcoleptic patients (Draganski et al., 2002) – such that nowadays it is established that even in adult brains, additional neurons can be generated (Bergmann et al., 2015; Boldrini et al., 2018). Such "neuroplasticity" is typically observed in circumscribed subcortical regions of the brain (Jurkowski et al., 2020), whereby the hippocampus (Augusto-Oliveira et al., 2019; Kozareva et al., 2019; Toda et al., 2019) and the subventricular zone are (Ponti et al., 2013; Shapiro et al., 2009) among the best studied "neurogenic niches". For treating or eventually curing neurodegenerative diseases such as the here-study MND population, neurogenesis in cortical regions – crucially the primarily affected motor cortex – is required. But can such targeted neurogenesis be achieved? Filgrastim, a granulocyte-colony stimulating factor (G-CSF), has been suggested to induce also cortical neurogenesis (Schneider et al., 2005; Wallner et al., 2015), and was recently shown to yield positive effects on survival in ALS by an associate research group (Johannesen et al., 2021). In that study, biomarkers reflecting stem cell mobilization, neuroinflammatory response but also white matter microstructure (as measured by DTI-derived fractional anisotropy) were shown to be the main modulators differentiating G-CSF responders from non-responders.

Nonetheless, it has not been evaluated in that study whether the putative neurogenesis was actually reflected in macroscopic CT changes. The here-proposed biomarker offers to evaluate such macroscopic changes, which is why we initiated a cooperation with the G-CSF study group to investigate CT-changes in the filgrastim-treated ALS population. In *Figure 4*, we demonstrate some preliminary results by an illustrative example from an individual patient (with the kind approval of U. Bogdahn who was in charge of that study) to show how our tool may be useful for clinical trials. The presented patient was male and diagnosed with ALS at the age of 44. He was included in the filgrastim-study and treated with a mean dose of 480 Mio IU/month (range 90 to 2,160 Mio IU/month, injected subcutaneously), mainly as five-day treatment block once or twice a month, or continuously on single days (mostly every second day; but see the original reference for details on the treatment plan). Every three months, he was invited for follow-up clinical assessments, which included a T1w MRI scan. This patient exhibited exceptionally long survival (> 45 months; but compare to the median survival of 15.8 months of ALS patients following diagnosis, (Traxinger et al., 2013)). For each consecutive scan, we z-scored the data as suggested herein, with respect to the patient's age- and sex-matched male control group, with reference ages ranging between years [$X$-2; $X$; $X$+2], whereby $X$ denotes the patient's age at a respective

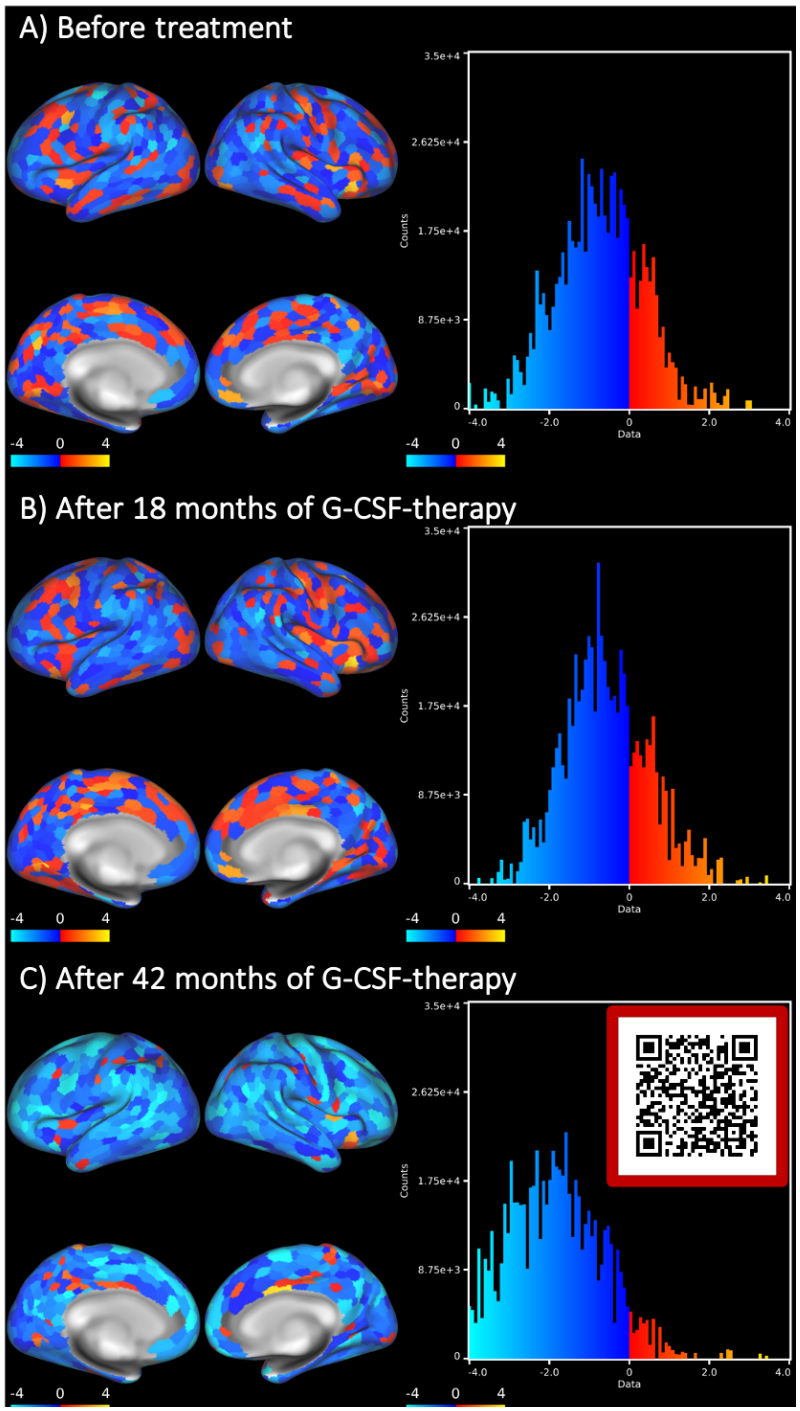



Figure 8.4.1. Can the mosaic-approach indicate neurogenesis? A male patient was diagnosed with amyotrophic lateral sclerosis (ALS) at 44 years of age and included in a clinical evaluation at the University of Regensburg, where he was treated with filgrastim (G-CSF), a growth factor believed to indirectly allow induction of neurogenesis. This patient exhibited exceptionally long survival (> 42 months) and was assessed for clinical follow-ups every three months. We applied the mosaic-approach on each of his follow-up MRI data, and show some exemplary time points here. The left panel shows the z-score maps from the mosaic-approach, the right panel the distribution of z-scores across all brain "patches". Hot colors indicate relative thicker, cool colors relative thinner cortex as compared to his reference group. Of particular interest to the study was the relative increase of cortical thickness at his (right) motor cortex and insula after 18 months of filgrastim treatment (compare respective regions in B vs. the baseline shown in A). However, at the end of his disease (C), atrophy was evident across the entire cortex, such that any potential drug effects might have been outweighed by ALS pathology. Although merely preliminary and descriptive data, they are in line with the supposed molecular mechanism of filgrastim and motivate us to explore the potential utility of the mosaic-approach to mirror neurogenesis further. This could offer a new opportunity quantify in vivo neurogenesis in clinical trials for the emerging pharmacological field of neuroregenerative medicine. Notice also the QR code in C, which links to a video where we show all of the patient's 15 follow-ups consecutively.

scan date. However, the control data for this analysis was only derived from the Cam-CAN control population and did not contain any HC from TCD (due to data protection policies on the one hand and due to different scanner parameters on the other, as outlined in section "Methodological considerations" – note however that if mixing the control data from Cam-CAN and TCD to rate the G-CSF patient on, we would have mixed three scanner acquisition schemes / parameters instead of just two [as we did in the associated studies], which we considered too much without proper investigation of confounding effects, as discussed above). This however yielded even smaller reference groups as previously (also cf. section "Methodological considerations" for the problematic nature of small reference groups). Therefore, conversion of z-scores to p-values with the proposed Monte-Carlo permutation procedure would have resulted in unsatisfactory little (and potentially insufficient) numbers of permutations such that the associated p-values were not meaningfully interpretable, as discussed above. As such, for demonstrative purposes, we decided to only consider the z-scored maps for a first impression of possible CT changes associated with G-CSF treatment. In spite of all these shortcomings, *Figure 8.4.1* suggests a highly dynamic nature of CT: the left panel shows the

patient's z-scored maps at some exemplary time points, the right panel shows the respective distributions of the z-scores across all patches; cool colors indicate relative lower, hot colors relative higher CT with respect to the control group. Of particular interest to this retrospective clinical study is the relative increase of CT in the motor cortex (especially on the right hemisphere) after 21 months of filgrastim treatment (notice the hot colors at the motor cortex in *Fig. 4B vs. 4A*). However, at the end of his disease and after 42 months into his diagnosis (*Fig. 4C*), cool colors dominated his cortex, indicating atrophy occurring throughout his brain. The dynamics of CT changes however really becomes evident when inspecting all consecutive scans, which is why we provide a video showing all 15 scans chronologically and which can be accessed by the QR code in *Figure 4C*. We are however fully aware that the displayed results are noisy such that statistically relevant CT changes cannot be concluded from these images but can at most serve exploratory purposes. However, we also believe that the suggested CT increase in the motor cortex might be real since it is in line with the hypothesized molecular mechanism associated with G-CSF (i.e. neurogenesis); moreover, the "spreading" CT decline towards the end of his disease is in line with the clinical presentation of this devastating disorder, with disability rapidly progressing over time (Bromberg, 2014). These results motivate us to investigate this tool further for demonstration of *in vivo* neurogenesis, which will require larger control populations, ideally obtained with the identical scanner / acquisition parameters as the patient data. We are currently planning to realize this in a prospective clinical trial on G-CSF with Prof. Bogdahn's study group. One methodological goal of this study will be to assess whether the number of significantly "thin patches" or "mosaics" (and eventually also the number of thick patches) correlates with respondence to G-CSF, i.e. survival, as well as clinically observed disease burden. Such a demonstration would highly suggest the utility of the here-proposed method to evaluate slowing down of neurodegeneration and eventually also indicate neurogenesis. This would make our tool highly attractive for clinical trials probing therapy in the ever-more emerging field of CNS-regenerative medicine (Salgado, 2020).

## 8.5 Conclusions

To summarize, this dissertation suggested an inter-rated-independent tool to reliably appraise cortical disease burden in individual patients by rating single patient's cortical thickness at distinct regions with respect to a matched control group (study 1; "mosaic-approach"). We have demonstrated that our method is independent of the demographic profile (but not independent of the clinical profile, study 2). Furthermore, it is a valid tool for quantifying the degree of cortical involvement (as opposed to e.g. spinal cord pathology, study 3) and lastly is powerful to detect topographically distinct atrophy patterns, which are in line with histological findings on the respective conditions (study 4). Sine the mosaic-approach is merely based on T1w MRI data, which is routinely acquired as part of clinical and academic protocols, it offers a cheap and non-invasive way to estimate and quantify cortical disease burden, which can aid to progress-monitor individual patient's disease course which in turn can aid clinical management. We are currently undertaking efforts to improve the tool further (including increasing the sample size of the reference populations and harmonizing acquisition parameters) and compliment it with a related tool to assess white matter pathology individually. We have high hopes that our tool might also be useful for pharmaceutical trials. The utility for demonstration of therapeutic efficacy is currently furthermore motivated by very precautious results suggesting that even neurogenesis might be reflected by our tool (cf. preliminary data from G-CSF study presented in part III), but such a correlation will require further and thorough scientific investigation, which can however be probed soon since the mosaic-approach is planned to be used in a prospective clinical trial on G-CSF in ALS patients. To conclude, we envisage ample future application for our methodological approach, both for daily clinical routine and clinical trials. However, future applications will require further validation of our

findings in larger and ideally multi-center studies, investigating various clinical subgroups across a variety of neurodegenerative disorders.

# 9 References


Abidi, M., Marco, G., Couillandre, A., Feron, M., Mseddi, E., Termoz, N., Querin, G., Pradat, P. -F. & Bede, P. (2020). Adaptive functional reorganization in amyotrophic lateral sclerosis: coexisting degenerative and compensatory changes. *European Journal of Neurology*, *27*(1), 121–128. https://doi.org/10.1111/ene.14042

Abidi, Malek, Marco, G., Grami, F., Termoz, N., Couillandre, A., Querin, G., Bede, P. & Pradat, P. (2021). Neural Correlates of Motor Imagery of Gait in Amyotrophic Lateral Sclerosis. *Journal of Magnetic Resonance Imaging*, *53*(1), 223–233. https://doi.org/10.1002/jmri.27335

Adachi, M., Kawanami, T., Ohshima, H., Sugai, Y. & Hosoya, T. (2004). Morning Glory Sign: A Particular MR Finding in Progressive Supranuclear Palsy. *Magnetic Resonance in Medical Sciences*, *3*(3), 125–132. https://doi.org/10.2463/mrms.3.125

Agosta, F., Canu, E., Inuggi, A., Chiò, A., Riva, N., Silani, V., Calvo, A., Messina, S., Falini, A., Comi, G. & Filippi, M. (2014). Resting state functional connectivity alterations in primary lateral sclerosis. *Neurobiology of Aging*, *35*(4), 916–925. https://doi.org/10.1016/j.neurobiolaging.2013.09.041

Agosta, F., Ferraro, P. M., Riva, N., Spinelli, E. G., Chiò, A., Canu, E., Valsasina, P., Lunetta, C., Iannaccone, S., Copetti, M., Prudente, E., Comi, G., Falini, A. & Filippi, M. (2016). Structural brain correlates of cognitive and behavioral impairment in MND. *Human Brain Mapping*, *37*(4), 1614–1626. https://doi.org/10.1002/hbm.23124

Agosta, F., Galantucci, S., Riva, N., Chiò, A., Messina, S., Iannaccone, S., Calvo, A., Silani, V., Copetti, M., Falini, A., Comi, G. & Filippi, M. (2014). Intrahemispheric and interhemispheric structural network abnormalities in PLS and ALS. *Human Brain Mapping*, *35*(4), 1710–1722. https://doi.org/10.1002/hbm.22286

Agosta, F., Gorno-Tempini, M. L., Pagani, E., Sala, S., Caputo, D., Perini, M., Bartolomei, I., Fruguglietti, M. E. & Filippi, M. (2009). Longitudinal assessment of grey matter contraction in amyotrophic lateral sclerosis: A tensor based morphometry study. *Amyotrophic Lateral Sclerosis : Official Publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, *10*(3), 168–174. https://doi.org/10.1080/17482960802603841

Agosta, F., Pagani, E., Petrolini, M., Sormani, M. P., Caputo, D., Perini, M., Prelle, A., Salvi, F. & Filippi, M. (2010). MRI predictors of long-term evolution in amyotrophic lateral sclerosis. *European Journal of Neuroscience*, *32*(9), 1490–1496. https://doi.org/10.1111/j.1460-9568.2010.07445.x

Agosta, F., Spinelli, E. G. & Filippi, M. (2018). Neuroimaging in amyotrophic lateral sclerosis: current and emerging uses. *Expert Review of Neurotherapeutics*, *18*(5), 395–406. https://doi.org/10.1080/14737175.2018.1463160

Ahmed, O. Ben, Benois-Pineau, J., Allard, M., Catheline, G. & Amar, C. Ben. (2017). Recognition of Alzheimer's disease and Mild Cognitive Impairment with multimodal image-derived biomarkers and Multiple Kernel Learning. *Neurocomputing*, *220*, 98–110. https://doi.org/10.1016/j.neucom.2016.08.041

Ahmed, R. M., Paterson, R. W., Warren, J. D., Zetterberg, H., O'Brien, J. T., Fox, N. C., Halliday, G. M. & Schott, J. M. (2014). Biomarkers in dementia: clinical utility and new directions. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(12), 1426–1434. https://doi.org/10.1136/jnnp-2014-307662

Aizpurua, M., Selvackadunco, S., Yull, H., Kipps, C. M., Ironside, J. W. & Bodi, I. (2019). Variably protease-sensitive prionopathy mimicking frontotemporal dementia. *Neuropathology*, *39*(2), 135–140. https://doi.org/10.1111/neup.12538

Alba-Ferrara, L. M. & de Erausquin, G. A. (2013). What does anisotropy measure? Insights from increased and decreased anisotropy in selective fiber tracts in schizophrenia. *Frontiers in Integrative Neuroscience*, *7*. https://doi.org/10.3389/fnint.2013.00009

Amen, D. G. (2005). *Making a Good Brain Great: The Amen Clinic Program for Achieving and Sustaining Optimal*. Harmony Books.

Assunção Leme, I. B., Gadelha, A., Sato, J. R., Ota, V. K., Mari, J. de J., Melaragno, M. I., Smith, M. de A. C., Belangero, S. I. N., Bressan, R. A. & Jackowski, A. P. (2013). Is there an association between cortical thickness, age of onset, and duration of illness in schizophrenia? *CNS Spectrums*, *18*(6),

315–321. https://doi.org/10.1017/S1092852913000333

Augusto-Oliveira, M., Arrifano, G., Malva, J. & Crespo-Lopez, M. (2019). Adult Hippocampal Neurogenesis in Different Taxonomic Groups: Possible Functional Similarities and Striking Controversies. *Cells*, *8*(2), 125. https://doi.org/10.3390/cells8020125

Baek, S.-H., Park, J., Kim, Y. H., Seok, H. Y., Oh, K.-W., Kim, H.-J., Kwon, Y.-J., Sim, Y., Tae, W.-S., Kim, S. H. & Kim, B.-J. (2020). Usefulness of diffusion tensor imaging findings as biomarkers for amyotrophic lateral sclerosis. *Scientific Reports*, *10*(1), 5199. https://doi.org/10.1038/s41598-020-62049-0

Baez, S., Couto, B., Torralva, T., Sposato, L. A., Huepe, D., Montañes, P., Reyes, P., Matallana, D., Vigliecca, N. S., Slachevsky, A., Manes, F. & Ibanez, A. (2014). Comparing Moral Judgments of Patients With Frontotemporal Dementia and Frontal Stroke. *JAMA Neurology*, *71*(9), 1172. https://doi.org/10.1001/jamaneurol.2014.347

Bang, J., Spina, S. & Miller, B. L. (2015). Frontotemporal dementia. *Lancet (London, England)*, *386*(10004), 1672–1682. https://doi.org/10.1016/S0140-6736(15)00461-4

Basaia, S., Agosta, F., Cividini, C., Trojsi, F., Riva, N., Spinelli, E. G., Moglia, C., Femiano, C., Castelnovo, V., Canu, E., Falzone, Y., Monsurrò, M. R., Falini, A., Chiò, A., Tedeschi, G. & Filippi, M. (2020). Structural and functional brain connectome in motor neuron diseases. *Neurology*, *95*(18), e2552–e2564. https://doi.org/10.1212/WNL.0000000000010731

Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., Filippi, M. & Alzheimer's Disease Neuroimaging Initiative. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage. Clinical*, *21*, 101645. https://doi.org/10.1016/j.nicl.2018.101645

Basser, P. J., Mattiello, J. & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*(1), 259–267. https://doi.org/10.1016/S0006-3495(94)80775-1

Bede, P., Bogdahn, U., Lope, J., Chang, K., Xirou, S. & Christidi, F. (2021). Degenerative and regenerative processes in amyotrophic lateral sclerosis: motor reserve, adaptation and putative compensatory changes. *Neural Regeneration Research*, *16*(6), 1208. https://doi.org/10.4103/1673-5374.300440

Bede, P., Bokde, A. L., Byrne, S., Marwa Elamin, M., Russell McLaughlin, M. L., Kenna, K., Fagan, A. J., Pender, N., Bradley, D. G. & Hardiman, O. (2013). Multiparametric MRI study of ALS stratified for the C9orf72 genotype Results: A congruent pattern of cortical and subcortical involvement was identified in those with the. *Neurology*, *81*(4), 361–369. www.neurology.org

Bede, P., Chipika, R. H., Finegan, E., Li Hi Shing, S., Chang, K. M., Doherty, M. A., Hengeveld, J. C., Vajda, A., Hutchinson, S., Donaghy, C., McLaughlin, R. L. & Hardiman, O. (2020). Progressive brainstem pathology in motor neuron diseases: Imaging data from amyotrophic lateral sclerosis and primary lateral sclerosis. *Data in Brief*, *29*, 105229. https://doi.org/10.1016/j.dib.2020.105229

Bede, P., Chipika, R. H., Finegan, E., Li Hi Shing, S., Doherty, M. A., Hengeveld, J. C., Vajda, A., Hutchinson, S., Donaghy, C., McLaughlin, R. L. & Hardiman, O. (2019). Brainstem pathology in amyotrophic lateral sclerosis and primary lateral sclerosis: A longitudinal neuroimaging study. *NeuroImage: Clinical*, *24*, 102054. https://doi.org/10.1016/j.nicl.2019.102054

Bede, P., Elamin, M., Byrne, S. & Hardiman, O. (2014). Sexual dimorphism in ALS: exploring gender-specific neuroimaging signatures. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *15*(3–4), 235–243. https://doi.org/10.3109/21678421.2013.865749

Bede, P. & Hardiman, O. (2018). Longitudinal structural changes in ALS: a three time-point imaging study of white and gray matter degeneration. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *19*(3–4), 232–241. https://doi.org/10.1080/21678421.2017.1407795

Bede, P., Iyer, P. M., Finegan, E., Omer, T. & Hardiman, O. (2017). Virtual brain biopsies in amyotrophic lateral sclerosis: Diagnostic classification based on in vivo pathological patterns. *NeuroImage: Clinical*, *15*, 653–658. https://doi.org/10.1016/j.nicl.2017.06.010

Bede, P., Iyer, P. M., Schuster, C., Elamin, M., Mclaughlin, R. L., Kenna, K. & Hardiman, O. (2016). The selective anatomical vulnerability of ALS: "disease-defining" and "disease-defying" brain regions. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *17*(7–8), 561–570. https://doi.org/10.3109/21678421.2016.1173702

Bede, P., Omer, T., Finegan, E., Chipika, R. H., Iyer, P. M., Doherty, M. A., Vajda, A., Pender, N., McLaughlin, R. L., Hutchinson, S. & Hardiman, O. (2018). Connectivity-based characterisation of

subcortical grey matter pathology in frontotemporal dementia and ALS: a multimodal neuroimaging study. *Brain Imaging and Behavior*, *12*(6), 1696–1707. https://doi.org/10.1007/s11682-018-9837-9

Bede, P., Querin, G. & Pradat, P.-F. (2018). The changing landscape of motor neuron disease imaging: the transition from descriptive studies to precision clinical tools. *Current Opinion in Neurology*, *31*(4), 431–438. https://doi.org/10.1097/WCO.0000000000000569

Bejanin, A., Tammewar, G., Marx, G., Cobigo, Y., Iaccarino, L., Kornak, J., Staffaroni, A. M., Dickerson, B. C., Boeve, B. F., Knopman, D. S., Gorno-Tempini, M., Miller, B. L., Jagust, W. J., Boxer, A. L., Rosen, H. J. & Rabinovici, G. D. (2020). Longitudinal structural and metabolic changes in frontotemporal dementia. *Neurology*, *95*(2), e140–e154. https://doi.org/10.1212/WNL.0000000000009760

Bergmann, O., Spalding, K. L. & Frisén, J. (2015). Adult Neurogenesis in Humans. *Cold Spring Harbor Perspectives in Biology*, *7*(7), a018994. https://doi.org/10.1101/cshperspect.a018994

Besser, L. M. & Galvin, J. E. (2020). Diagnostic experience reported by caregivers of patients with frontotemporal degeneration. *Neurology: Clinical Practice*, *10*(4), 298–306. https://doi.org/10.1212/CPJ.0000000000000738

Bi, X., Li, S., Xiao, B., Li, Y., Wang, G. & Ma, X. (2020). Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology. *Neurocomputing*, *392*, 296–304. https://doi.org/10.1016/j.neucom.2018.11.111

Blair, R. C. & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518–524. https://doi.org/10.1111/j.1469-8986.1993.tb02075.x

Blasco, H., Patin, F., Descat, A., Garçon, G., Corcia, P., Gelé, P., Lenglet, T., Bede, P., Meininger, V., Devos, D., Gossens, J. F. & Pradat, P.-F. (2018). A pharmaco-metabolomics approach in a clinical trial of ALS: Identification of predictive markers of progression. *PloS One*, *13*(6), e0198116. https://doi.org/10.1371/journal.pone.0198116

Boldrini, M., Fulmore, C. A., Tartt, A. N., Simeon, L. R., Pavlova, I., Poposka, V., Rosoklija, G. B., Stankov, A., Arango, V., Dwork, A. J., Hen, R. & Mann, J. J. (2018). Human Hippocampal Neurogenesis Persists throughout Aging. *Cell Stem Cell*, *22*(4), 589-599.e5. https://doi.org/10.1016/j.stem.2018.03.015

Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, *71*(356), 791–799. https://doi.org/10.1080/01621459.1976.10480949

Brant-Zawadzki, M., Enzmann, D. R., Placone, R. C., Sheldon, P., Britt, R. H., Brasch, R. C. & Crooks, L. A. (1983). NMR imaging of experimental brain abscess: comparison with CT. *American Journal of Neuroradiology*, *4*(3), 250–253. http://www.ncbi.nlm.nih.gov/pubmed/6410714

Brettschneider, J., Arai, K., Del Tredici, K., Toledo, J. B., Robinson, J. L., Lee, E. B., Kuwabara, S., Shibuya, K., Irwin, D. J., Fang, L., Van Deerlin, V. M., Elman, L., McCluskey, L., Ludolph, A. C., Lee, V. M.-Y., Braak, H. & Trojanowski, J. Q. (2014). TDP-43 pathology and neuronal loss in amyotrophic lateral sclerosis spinal cord. *Acta Neuropathologica*, *128*(3), 423–437. https://doi.org/10.1007/s00401-014-1299-6

Brettschneider, J., Del Tredici, K., Toledo, J. B., Robinson, J. L., Irwin, D. J., Grossman, M., Suh, E., Van Deerlin, V. M., Wood, E. M., Baek, Y., Kwong, L., Lee, E. B., Elman, L., McCluskey, L., Fang, L., Feldengut, S., Ludolph, A. C., Lee, V. M.-Y., Braak, H. & Trojanowski, J. Q. (2013). Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Annals of Neurology*, *74*(1), 20–38. https://doi.org/10.1002/ana.23937

Bromberg, M. B. (Ed.). (2014). Clinical Anatomy, Physiology, and Pathology of Motor Neuron Disease. In *Motor Neuron Disease in Adults* (pp. 12–17). Oxford University Press. https://doi.org/10.1093/med/9780199783113.003.0003

Brooks, B. R., Miller, R. G., Swash, M. & Munsat, T. L. (2000). El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*, *1*(5), 293–299. https://doi.org/10.1080/146608200300079536

Bruchhage, M. M. K., Amad, A., Draper, S. B., Seidman, J., Lacerda, L., Laguna, P. L., Lowry, R. G., Wheeler, J., Robertson, A., Dell'Acqua, F., Smith, M. S. & Williams, S. C. R. (2020). Drum training induces long-term plasticity in the cerebellum and connected cortical thickness. *Scientific Reports*, *10*(1),

10116. https://doi.org/10.1038/s41598-020-65877-2

Bull, J. W. (1970). The history of neuroradiology. *Proc R Soc Med.*, *63*(6), 637–643. https://doi.org/10.1136/bmj.2.3474.205

Burke, T., Elamin, M., Bede, P., Pinto-Grau, M., Lonergan, K., Hardiman, O. & Pender, N. (2016). Discordant performance on the 'Reading the Mind in the Eyes' Test, based on disease onset in amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *17*(7–8), 467–472. https://doi.org/10.1080/21678421.2016.1177088

Burke, T., Lonergan, K., Pinto-Grau, M., Elamin, M., Bede, P., Madden, C., Hardiman, O. & Pender, N. (2017). Visual encoding, consolidation, and retrieval in amyotrophic lateral sclerosis: executive function as a mediator, and predictor of performance. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *18*(3–4), 193–201. https://doi.org/10.1080/21678421.2016.1272615

Burke, T., Pinto-Grau, M., Lonergan, K., Elamin, M., Bede, P., Costello, E., Hardiman, O. & Pender, N. (2016). Measurement of Social Cognition in Amyotrophic Lateral Sclerosis: A Population Based Study. *PLOS ONE*, *11*(8), e0160850. https://doi.org/10.1371/journal.pone.0160850

Byrne, S., Elamin, M., Bede, P., Shatunov, A., Walsh, C., Corr, B., Heverin, M., Jordan, N., Kenna, K., Lynch, C., McLaughlin, R. L., Iyer, P. M., O'Brien, C., Phukan, J., Wynne, B., Bokde, A. L., Bradley, D. G., Pender, N., Al-Chalabi, A. & Hardiman, O. (2012). Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: a population-based cohort study. *The Lancet Neurology*, *11*(3), 232–240. https://doi.org/10.1016/S1474-4422(12)70014-5

Campanella, F., Shallice, T., Ius, T., Fabbro, F. & Skrap, M. (2014). Impact of brain tumour location on emotion and personality: a voxel-based lesion–symptom mapping study on mentalization processes. *Brain*, *137*(9), 2532–2545. https://doi.org/10.1093/brain/awu183

Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B. & Nakanishi, A. (1999). The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *Journal of the Neurological Sciences*, *169*(1–2), 13–21. https://doi.org/10.1016/s0022-510x(99)00210-5

Chen, C. H., Fiecas, M., Gutiérrez, E. D., Panizzon, M. S., Eyler, L. T., Vuoksimaa, E., Thompson, W. K., Fennema-Notestine, C., Hagler, D. J., Jernigan, T. L., Neale, M. C., Franz, C. E., Lyons, M. J., Fischl, B., Tsuang, M. T., Dale, A. M. & Kremen, W. S. (2013). Genetic topography of brain morphology. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17089–17094. https://doi.org/10.1073/pnas.1308091110

Chipika, Rangariroyashe H., Christidi, F., Finegan, E., Li Hi Shing, S., McKenna, M. C., Chang, K. M., Karavasilis, E., Doherty, M. A., Hengeveld, J. C., Vajda, A., Pender, N., Hutchinson, S., Donaghy, C., McLaughlin, R. L., Hardiman, O. & Bede, P. (2020). Amygdala pathology in amyotrophic lateral sclerosis and primary lateral sclerosis. *Journal of the Neurological Sciences*, *417*, 117039. https://doi.org/10.1016/j.jns.2020.117039

Chipika, Rangariroyashe H, Finegan, E., Li Hi Shing, S., McKenna, M. C., Christidi, F., Chang, K. M., Doherty, M. A., Hengeveld, J. C., Vajda, A., Pender, N., Hutchinson, S., Donaghy, C., McLaughlin, R. L., Hardiman, O. & Bede, P. (2020). "Switchboard" malfunction in motor neuron diseases: Selective pathology of thalamic nuclei in amyotrophic lateral sclerosis and primary lateral sclerosis. *NeuroImage. Clinical*, *27*, 102300. https://doi.org/10.1016/j.nicl.2020.102300

Chipika, Rangariroyashe H, Siah, W. F., McKenna, M. C., Li Hi Shing, S., Hardiman, O. & Bede, P. (2020). The presymptomatic phase of amyotrophic lateral sclerosis: are we merely scratching the surface? *Journal of Neurology*. https://doi.org/10.1007/s00415-020-10289-5

Chipika, Rangariroyashe Hannah, Finegan, E., Li Hi Shing, S., Hardiman, O. & Bede, P. (2019). Tracking a Fast-Moving Disease: Longitudinal Markers, Monitoring, and Clinical Trial Endpoints in ALS. *Frontiers in Neurology*, *10*, 229. https://doi.org/10.3389/fneur.2019.00229

Chouinard-Decorte, F., McKay, D. R., Reid, A., Khundrakpam, B., Zhao, L., Karama, S., Rioux, P., Sprooten, E., Knowles, E., Kent, J. W., Curran, J. E., Göring, H. H. H., Dyer, T. D., Olvera, R. L., Kochunov, P., Duggirala, R., Fox, P. T., Almasy, L., Blangero, J., … Glahn, D. C. (2014). Heritable changes in regional cortical thickness with age. *Brain Imaging and Behavior*, *8*(2), 208–216. https://doi.org/10.1007/s11682-014-9296-x

Christidi, F., Karavasilis, E., Rentzos, M., Kelekis, N., Evdokimidis, I. & Bede, P. (2018). Clinical and Radiological Markers of Extra-Motor Deficits in Amyotrophic Lateral Sclerosis. *Frontiers in Neurology*, *9*, 1005. https://doi.org/10.3389/fneur.2018.01005

Christidi, F., Karavasilis, E., Rentzos, M., Velonakis, G., Zouvelou, V., Xirou, S., Argyropoulos, G., Papatriantafyllou, I., Pantolewn, V., Ferentinos, P., Kelekis, N., Seimenis, I., Evdokimidis, I. & Bede, P. (2019). Hippocampal pathology in amyotrophic lateral sclerosis: selective vulnerability of subfields and their associated projections. *Neurobiology of Aging*, *84*, 178–188. https://doi.org/10.1016/j.neurobiolaging.2019.07.019

Christidi, F., Karavasilis, E., Rentzos, M., Velonakis, G., Zouvelou, V., Xirou, S., Argyropoulos, G., Papatriantafyllou, I., Pantolewn, V., Ferentinos, P., Kelekis, N., Seimenis, I., Evdokimidis, I. & Bede, P. (2020). Neuroimaging data indicate divergent mesial temporal lobe profiles in amyotrophic lateral sclerosis, Alzheimer's disease and healthy aging. *Data in Brief*, *28*, 104991. https://doi.org/10.1016/j.dib.2019.104991

Christidi, F., Karavasilis, E., Velonakis, G., Ferentinos, P., Rentzos, M., Kelekis, N., Evdokimidis, I. & Bede, P. (2018). The Clinical and Radiological Spectrum of Hippocampal Pathology in Amyotrophic Lateral Sclerosis. *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.00523

Christidi, F., Zalonis, I., Kyriazi, S., Rentzos, M., Karavasilis, E., Wilde, E. A. & Evdokimidis, I. (2014). Uncinate fasciculus microstructure and verbal episodic memory in amyotrophic lateral sclerosis: a diffusion tensor imaging and neuropsychological study. *Brain Imaging and Behavior*, *8*(4), 497–505. https://doi.org/10.1007/s11682-013-9271-y

Clark, M., Huang, C., Bageac, D., Danielian, L., Smallwoed, R. & Floeter, M. K. (2017). Neuroimaging changes in the first 5 years of symptoms in patients with primary lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *18*, 216–217. https://doi.org/10.1080/21678421.2017.1374605/008

Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., Elavsky, S., Marquez, D. X., Hu, L. & Kramer, A. F. (2006). Aerobic Exercise Training Increases Brain Volume in Aging Humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *61*(11), 1166–1170. https://doi.org/10.1093/gerona/61.11.1166

Cosottini, M., Giannelli, M., Siciliano, G., Lazzarotti, G., Michelassi, M. C., Del Corona, A., Bartolozzi, C. & Murri, L. (2005). Diffusion-Tensor MR Imaging of Corticospinal Tract in Amyotrophic Lateral Sclerosis and Progressive Muscular Atrophy. *Radiology*, *237*(1), 258–264. https://doi.org/10.1148/radiol.2371041506

Costello, E., Rooney, J., Pinto-Grau, M., Burke, T., Elamin, M., Bede, P., McMackin, R., Dukic, S., Vajda, A., Heverin, M., Hardiman, O. & Pender, N. (2021). Cognitive reserve in amyotrophic lateral sclerosis (ALS): a population-based longitudinal study. *Journal of Neurology, Neurosurgery & Psychiatry*, *92*(5), 460–465. https://doi.org/10.1136/jnnp-2020-324992

Craig-Schapiro, R., Fagan, A. M. & Holtzman, D. M. (2009). Biomarkers of Alzheimer's disease. *Neurobiology of Disease*, *35*(2), 128–140. https://doi.org/10.1016/j.nbd.2008.10.003

Dale, A. M., Fischl, B. & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

de Borba, F. C., Querin, G., França, M. C. & Pradat, P.-F. (2020). Cerebellar degeneration in adult spinal muscular atrophy patients. *Journal of Neurology*, *267*(9), 2625–2631. https://doi.org/10.1007/s00415-020-09875-4

De Maindreville, A., Bedos, L. & Bakchine, S. (2015). Systemic Sarcoidosis Mimicking a Behavioural Variant of Frontotemporal Dementia. *Case Reports in Neurological Medicine*, *2015*, 1–4. https://doi.org/10.1155/2015/409126

de Vries, B. S., Spreij, L. A., Rustemeijer, L. M. M., Bakker, L. A., Veldink, J. H., van den Berg, L. H., Nijboer, T. C. W. & van Es, M. A. (2019). A neuropsychological and behavioral study of PLS. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *20*(5–6), 376–384. https://doi.org/10.1080/21678421.2019.1620284

den Heijer, T., van der Lijn, F., Koudstaal, P. J., Hofman, A., van der Lugt, A., Krestin, G. P., Niessen, W. J. & Breteler, M. M. B. (2010). A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. *Brain*, *133*(4), 1163–1172.

https://doi.org/10.1093/brain/awq048

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S. & 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. https://doi.org/10.1016/j.neuroimage.2006.01.021

Devos, D., Moreau, C., Kyheng, M., Garçon, G., Rolland, A. S., Blasco, H., Gelé, P., Timothée Lenglet, T., Veyrat-Durebex, C., Corcia, P., Dutheil, M., Bede, P., Jeromin, A., Oeckl, P., Otto, M., Meininger, V., Danel-Brunaud, V., Devedjian, J.-C., Duce, J. A. & Pradat, P. F. (2019). A ferroptosis-based panel of prognostic biomarkers for Amyotrophic Lateral Sclerosis. *Scientific Reports*, *9*(1), 2918. https://doi.org/10.1038/s41598-019-39739-5

Di Fede, G., Catania, M., Atzori, C., Moda, F., Pasquali, C., Indaco, A., Grisoli, M., Zuffi, M., Guaita, M. C., Testi, R., Taraglio, S., Sessa, M., Gusmaroli, G., Spinelli, M., Salzano, G., Legname, G., Tarletti, R., Godi, L., Pocchiari, M., … Giaccone, G. (2019). Clinical and neuropathological phenotype associated with the novel V189I mutation in the prion protein gene. *Acta Neuropathologica Communications*, *7*(1), 1. https://doi.org/10.1186/s40478-018-0656-4

Dickie, E. W., Anticevic, A., Smith, D. E., Coalson, T. S., Manogaran, M., Calarco, N., Viviano, J. D., Glasser, M. F., Van Essen, D. C. & Voineskos, A. N. (2019). Ciftify: A framework for surface-based analysis of legacy MR acquisitions. *NeuroImage*, *197*, 818–826. https://doi.org/10.1016/j.neuroimage.2019.04.078

Draganski, B., Geisler, P., Hajak, G., Schuierer, G., Bogdahn, U., Winkler, J. & May, A. (2002). Hypothalamic gray matter changes in narcoleptic patients. *Nature Medicine*, *8*(11), 1186–1188. https://doi.org/10.1038/nm1102-1186

Du, A.-T., Schuff, N., Kramer, J. H., Rosen, H. J., Gorno-Tempini, M. L., Rankin, K., Miller, B. L. & Weiner, M. W. (2006). Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. *Brain*, *130*(4), 1159–1166. https://doi.org/10.1093/brain/awm016

Dukic, S., McMackin, R., Buxo, T., Fasano, A., Chipika, R., Pinto-Grau, M., Costello, E., Schuster, C., Hammond, M., Heverin, M., Coffey, A., Broderick, M., Iyer, P. M., Mohr, K., Gavin, B., Pender, N., Bede, P., Muthuraman, M., Lalor, E. C., … Nasseroleslami, B. (2019). Patterned functional network disruption in amyotrophic lateral sclerosis. *Human Brain Mapping*, *40*(16), 4827–4842. https://doi.org/10.1002/hbm.24740

Echaniz-Laguna, A., Rousso, E., Anheim, M., Fleury, M., Cossée, M. & Tranchant, C. (2005). L'amyotrophie bulbaire et spinale liée au chromosome X : une étude clinique, neurophysiologique et moléculaire de 12 patients issus de 4 familles. *Revue Neurologique*, *161*(4), 437–444. https://doi.org/10.1016/S0035-3787(05)85073-8

Eisen, A., Kiernan, M., Mitsumoto, H. & Swash, M. (2014). Amyotrophic lateral sclerosis: a long preclinical period? *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(11), 1232–1238. https://doi.org/10.1136/jnnp-2013-307135

El Mendili, M. M., Querin, G., Bede, P. & Pradat, P.-F. (2019). Spinal Cord Imaging in Amyotrophic Lateral Sclerosis: Historical Concepts—Novel Techniques. *Frontiers in Neurology*, *10*. https://doi.org/10.3389/fneur.2019.00350

Elamin, M., Bede, P., Byrne, S., Jordan, N., Gallagher, L., Wynne, B., O'Brien, C., Phukan, J., Lynch, C., Pender, N. & Hardiman, O. (2013). Cognitive changes predict functional decline in ALS: A population-based longitudinal study. *Neurology*, *80*(17), 1590–1597. https://doi.org/10.1212/WNL.0b013e31828f18ac

Elamin, Marwa, Bede, P., Montuschi, A., Pender, N., Chio, A. & Hardiman, O. (2015). Predicting prognosis in amyotrophic lateral sclerosis: a simple algorithm. *Journal of Neurology*, *262*(6), 1447–1454. https://doi.org/10.1007/s00415-015-7731-6

Elamin, Marwa, Pinto-Grau, M., Burke, T., Bede, P., Rooney, J., O'Sullivan, M., Lonergan, K., Kirby, E., Quinlan, E., Breen, N., Vajda, A., Heverin, M., Pender, N. & Hardiman, O. (2017). Identifying behavioural changes in ALS: Validation of the Beaumont Behavioural Inventory (BBI). *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *18*(1–2), 68–73. https://doi.org/10.1080/21678421.2016.1248976

Feigin, V. L., Vos, T., Nichols, E., Owolabi, M. O., Carroll, W. M., Dichgans, M., Deuschl, G., Parmar, P.,

Brainin, M. & Murray, C. (2020). The global burden of neurological disorders: translating evidence into policy. *The Lancet Neurology*, *19*(3), 255–265. https://doi.org/10.1016/S1474-4422(19)30411-9

Feron, M., Couillandre, A., Mseddi, E., Termoz, N., Abidi, M., Bardinet, E., Delgadillo, D., Lenglet, T., Querin, G., Welter, M.-L., Le Forestier, N., Salachas, F., Bruneteau, G., del Mar Amador, M., Debs, R., Lacomblez, L., Meininger, V., Pélégrini-Issac, M., Bede, P., … de Marco, G. (2018). Extrapyramidal deficits in ALS: a combined biomechanical and neuroimaging study. *Journal of Neurology*, *265*(9), 2125–2136. https://doi.org/10.1007/s00415-018-8964-y

Ferraro, P. M., Agosta, F., Riva, N., Copetti, M., Spinelli, E. G., Falzone, Y., Sorarù, G., Comi, G., Chiò, A. & Filippi, M. (2017). Multimodal structural MRI in the diagnosis of motor neuron diseases. *NeuroImage: Clinical*, *16*, 240–247. https://doi.org/10.1016/j.nicl.2017.08.002

Filippi, M., Horsfield, M. A., Adèr, H. J., Barkhof, F., Bruzzi, P., Evans, A., Frank, J. A., Grossman, R. I., McFarland, H. F., Molyneux, P., Paty, D. W., Simon, J., Tofts, P. S., Wolinsky, J. S. & Miller, D. H. (1998). Guidelines for using quantitative measures of brain magnetic resonance imaging abnormalities in monitoring the treatment of multiple sclerosis. *Annals of Neurology*, *43*(4), 499–506. https://doi.org/10.1002/ana.410430414

Finegan, E., Chipika, R. H., Li Hi Shing, S., Doherty, M. A., Hengeveld, J. C., Vajda, A., Donaghy, C., McLaughlin, R. L., Pender, N., Hardiman, O. & Bede, P. (2019). The clinical and radiological profile of primary lateral sclerosis: a population-based study. *Journal of Neurology*, *266*(11), 2718–2733. https://doi.org/10.1007/s00415-019-09473-z

Finegan, E., Chipika, R. H., Shing, S. L. H., Hardiman, O. & Bede, P. (2019). Primary lateral sclerosis: a distinct entity or part of the ALS spectrum? *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *20*(3–4), 133–145. https://doi.org/10.1080/21678421.2018.1550518

Finegan, E., Hi Shing, S. L., Chipika, R. H., McKenna, M. C., Doherty, M. A., Hengeveld, J. C., Vajda, A., Donaghy, C., McLaughlin, R. L., Hutchinson, S., Hardiman, O. & Bede, P. (2020). Thalamic, hippocampal and basal ganglia pathology in primary lateral sclerosis and amyotrophic lateral sclerosis: Evidence from quantitative imaging data. *Data in Brief*, *29*, 105115. https://doi.org/10.1016/j.dib.2020.105115

Finegan, E., Li Hi Shing, S., Chipika, R. H., Doherty, M. A., Hengeveld, J. C., Vajda, A., Donaghy, C., Pender, N., McLaughlin, R. L., Hardiman, O. & Bede, P. (2019). Widespread subcortical grey matter degeneration in primary lateral sclerosis: a multimodal imaging study with genetic profiling. *NeuroImage: Clinical*, *24*, 102089. https://doi.org/10.1016/j.nicl.2019.102089

Finegan, E., Li Hi Shing, S., Siah, W. F., Chipika, R. H., Chang, K. M., McKenna, M. C., Doherty, M. A., Hengeveld, J. C., Vajda, A., Donaghy, C., Hutchinson, S., McLaughlin, R. L., Hardiman, O. & Bede, P. (2020). Evolving diagnostic criteria in primary lateral sclerosis: The clinical and radiological basis of "probable PLS." *Journal of the Neurological Sciences*, *417*, 117052. https://doi.org/10.1016/j.jns.2020.117052

Finegan, E., Shing, S. L. H., Chipika, R. H., Chang, K. M., McKenna, M. C., Doherty, M. A., Hengeveld, J. C., Vajda, A., Pender, N., Donaghy, C., Hutchinson, S., McLaughlin, R. L., Hardiman, O. & Bede, P. (2021). Extra-motor cerebral changes and manifestations in primary lateral sclerosis. *Brain Imaging and Behavior*. https://doi.org/10.1007/s11682-020-00421-4

Finegan, E., Siah, W. F., Shing, S. L. H., Chipika, R. H., Chang, K. M., McKenna, M. C., Doherty, M. A., Hengeveld, J. C., Vajda, A., Donaghy, C., Hutchinson, S., McLaughlin, R. L., Hardiman, O. & Bede, P. (2020). Imaging and clinical data indicate considerable disease burden in 'probable' PLS: Patients with UMN symptoms for 2–4 years. *Data in Brief*, *32*, 106247. https://doi.org/10.1016/j.dib.2020.106247

Fischl, B, Sereno, M. I. & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207. https://doi.org/10.1006/nimg.1998.0396

Fischl, Bruce. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., Storsve, A. B., Tamnes, C. K., Sala-Llonch, R., Due-Tønnessen, P., Bjørnerud, A., Sølsnes, A. E., Håberg, A. K.,

Skranes, J., Bartsch, H., Chen, C. H., Thompson, W. K., Panizzon, M. S., Kremen, W. S., … Walhovd, K. B. (2015). Development and aging of cortical thickness correspond to genetic organization patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(50), 15462–15467. https://doi.org/10.1073/pnas.1508831112

Fjell, A. M. & Walhovd, K. B. (2010). Structural brain changes in aging: courses, causes and cognitive consequences. *Reviews in the Neurosciences*, *21*(3), 187–221. https://doi.org/10.1515/revneuro.2010.21.3.187

Floeter, M. K. & Gendron, T. F. (2018). Biomarkers for Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Associated With Hexanucleotide Expansion Mutations in C9orf72. *Frontiers in Neurology*, *9*, 1063. https://doi.org/10.3389/fneur.2018.01063

Foerster, B. R., Carlos, R. C., Dwamena, B. A., Callaghan, B. C., Petrou, M., Edden, R. A. E., Mohamed, M. A., Welsh, R. C., Barker, P. B., Feldman, E. L. & Pomper, M. G. (2014). Multimodal MRI as a diagnostic biomarker for amyotrophic lateral sclerosis. *Annals of Clinical and Translational Neurology*, *1*(2), 107–114. https://doi.org/10.1002/acn3.30

Foerster, B. R., Dwamena, B. A., Petrou, M., Carlos, R. C., Callaghan, B. C., Churchill, C. L., Mohamed, M. A., Bartels, C., Benatar, M., Bonzano, L., Ciccarelli, O., Cosottini, M., Ellis, C. M., Ehrenreich, H., Filippini, N., Ito, M., Kalra, S., Melhem, E. R., Pyra, T., … Pomper, M. G. (2013). Diagnostic accuracy of diffusion tensor imaging in amyotrophic lateral sclerosis: a systematic review and individual patient data meta-analysis. *Academic Radiology*, *20*(9), 1099–1106. https://doi.org/10.1016/j.acra.2013.03.017

Fried-Oken, M., Mooney, A. & Peters, B. (2015). Supporting communication for patients with neurodegenerative disease. *NeuroRehabilitation*, *37*(1), 69–87. https://doi.org/10.3233/NRE-151241

Fukutomi, H., Glasser, M. F., Zhang, H., Autio, J. A., Coalson, T. S., Okada, T., Togashi, K., Van Essen, D. C. & Hayashi, T. (2018). Neurite imaging reveals microstructural variations in human cerebral cortical gray matter. *NeuroImage*, *182*(May 2017), 488–499. https://doi.org/10.1016/j.neuroimage.2018.02.017

Gajamange, S., Raffelt, D., Dhollander, T., Lui, E., van der Walt, A., Kilpatrick, T., Fielding, J., Connelly, A. & Kolbe, S. (2018). Fibre-specific white matter changes in multiple sclerosis patients with optic neuritis. *NeuroImage: Clinical*, *17*(September 2017), 60–68. https://doi.org/10.1016/j.nicl.2017.09.027

Gaser, C. & Schlaug, G. (2003). Brain Structures Differ between Musicians and Non-Musicians. *The Journal of Neuroscience*, *23*(27), 9240–9245. https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003

Genc, S., Tax, C. M. W., Raven, E. P., Chamberland, M., Parker, G. D. & Jones, D. K. (2020). Impact of b -value on estimates of apparent fibre density. *Human Brain Mapping*, *41*(10), 2583–2595. https://doi.org/10.1002/hbm.24964

Geser, F., Martinez-Lage, M., Robinson, J., Uryu, K., Neumann, M., Brandmeir, N. J., Xie, S. X., Kwong, L. K., Elman, L., McCluskey, L., Clark, C. M., Malunda, J., Miller, B. L., Zimmerman, E. A., Qian, J., Van Deerlin, V., Grossman, M., Lee, V. M.-Y. & Trojanowski, J. Q. (2009). Clinical and Pathological Continuum of Multisystem TDP-43 Proteinopathies. *Archives of Neurology*, *66*(2). https://doi.org/10.1001/archneurol.2008.558

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Essen, D. C. Van, Jenkinson, M. & Hcp, W. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, *80*, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Goethe, J. W. (2015). Prelude on Stage. In A. S. Kline (Translator) & E. Delacroix (Illustrator) (Eds.), *Faust: Parts I & II*. CreateSpace Independent Publishing Platform.

Gómez-Sancho, M., Tohka, J., Gómez-Verdejo, V. & Alzheimer's Disease Neuroimaging Initiative. (2018). Comparison of feature representations in MRI-based MCI-to-AD conversion prediction. *Magnetic Resonance Imaging*, *50*, 84–95. https://doi.org/10.1016/j.mri.2018.03.003

Gordon, P. H., Cheng, B., Katz, I. B., Pinto, M., Hays, A. P., Mitsumoto, H. & Rowland, L. P. (2006). The natural history of primary lateral sclerosis. *Neurology*, *66*(5), 647–653.

https://doi.org/10.1212/01.wnl.0000200962.94777.71

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M. & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Grollemund, V., Chat, G. Le, Secchi-Buhour, M.-S., Delbot, F., Pradat-Peyre, J.-F., Bede, P. & Pradat, P.-F. (2020). Development and validation of a 1-year survival prognosis estimation model for Amyotrophic Lateral Sclerosis using manifold learning algorithm UMAP. *Scientific Reports*, *10*(1), 13378. https://doi.org/10.1038/s41598-020-70125-8

Grollemund, V., Le Chat, G., Secchi-Buhour, M.-S., Delbot, F., Pradat-Peyre, J.-F., Bede, P. & Pradat, P.-F. (2021). Manifold learning for amyotrophic lateral sclerosis functional loss assessment. *Journal of Neurology*, *268*(3), 825–850. https://doi.org/10.1007/s00415-020-10181-2

Grollemund, V., Pradat, P.-F., Querin, G., Delbot, F., Le Chat, G., Pradat-Peyre, J.-F. & Bede, P. (2019a). Machine Learning in Amyotrophic Lateral Sclerosis: Achievements, Pitfalls, and Future Directions. *Frontiers in Neuroscience*, *13*, 135. https://doi.org/10.3389/fnins.2019.00135

Grollemund, V., Pradat, P.-F., Querin, G., Delbot, F., Le Chat, G., Pradat-Peyre, J.-F. & Bede, P. (2019b). Machine Learning in Amyotrophic Lateral Sclerosis: Achievements, Pitfalls, and Future Directions. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00135

Hardiman, O, Doherty, C. P., Elamin, M. & Bede, P. (2016). *Neurodegenerative Disorders: A Clinical Guide* (S. C. H. N. Y. D. L. © S. I. P. S. 2016. Springer International Publishing (Ed.)).

Hardiman, Orla, Al-Chalabi, A., Chio, A., Corr, E. M., Logroscino, G., Robberecht, W., Shaw, P. J., Simmons, Z. & van den Berg, L. H. (2017). Amyotrophic lateral sclerosis. *Nature Reviews. Disease Primers*, *3*, 17071. https://doi.org/10.1038/nrdp.2017.71

Harper, L., Barkhof, F., Scheltens, P., Schott, J. M. & Fox, N. C. (2014). An algorithmic approach to structural imaging in dementia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *85*(6), 692–698. https://doi.org/10.1136/jnnp-2013-306285

Harris, J. M., Gall, C., Thompson, J. C., Richardson, A. M. T., Neary, D., du Plessis, D., Pal, P., Mann, D. M. A., Snowden, J. S. & Jones, M. (2013). Sensitivity and specificity of FTDC criteria for behavioral variant frontotemporal dementia. *Neurology*, *80*(20), 1881–1887. https://doi.org/10.1212/WNL.0b013e318292a342

Harrison Dening, K., Sampson, E. L. & De Vries, K. (2019). Advance care planning in dementia: recommendations for healthcare professionals. *Palliative Care: Research and Treatment*, *12*, 117822421982657. https://doi.org/10.1177/1178224219826579

He, Y., Chen, Z. J. & Evans, A. C. (2007). Small-World Anatomical Networks in the Human Brain Revealed by Cortical Thickness from MRI. *Cerebral Cortex*, *17*(10), 2407–2419. https://doi.org/10.1093/cercor/bhl149

Herculano-Houzel, S., Watson, C. & Paxinos, G. (2013). Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. *Frontiers in Neuroanatomy*, *7*(OCT), 1–14. https://doi.org/10.3389/fnana.2013.00035

Hodges, J. R., Mitchell, J., Dawson, K., Spillantini, M. G., Xuereb, J. H., McMonagle, P., Nestor, P. J. & Patterson, K. (2010). Semantic dementia: demography, familial factors and survival in a consecutive series of 100 cases. *Brain*, *133*(1), 300–306. https://doi.org/10.1093/brain/awp248

Hojjati, S. H., Ebrahimzadeh, A., Khazaee, A., Babajani-Feremi, A. & Alzheimer's Disease Neuroimaging Initiative. (2017). Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM. *Journal of Neuroscience Methods*, *282*, 69–80. https://doi.org/10.1016/j.jneumeth.2017.03.006

Hollander, M., Wolfe, D. A. & Chicken, E. (2014). *Nonparametric Statistical Methods*. John Wiley & Sons.

Househam, E. & Swash, M. (2000). Diagnostic delay in amyotrophic lateral sclerosis: what scope for improvement? *Journal of the Neurological Sciences*, *180*(1–2), 76–81. https://doi.org/10.1016/S0022-510X(00)00418-4

Hudziak, J. J., Albaugh, M. D., Ducharme, S., Karama, S., Spottswood, M., Crehan, E., Evans, A. C. & Botteron, K. N. (2014). Cortical Thickness Maturation and Duration of Music Training: Health-

Promoting Activities Shape Brain Development. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(11), 1153-1161.e2. https://doi.org/10.1016/j.jaac.2014.06.015

Hüfner, K., Binetti, C., Hamilton, D. A., Stephan, T., Flanagin, V. L., Linn, J., Labudda, K., Markowitsch, H., Glasauer, S., Jahn, K., Strupp, M. & Brandt, T. (2010). Structural and functional plasticity of the hippocampal formation in professional dancers and slackliners. *Hippocampus*, n/a-n/a. https://doi.org/10.1002/hipo.20801

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Research*, *163*(2), 195–205. https://doi.org/10.1016/0006-8993(79)90349-4

Huttenlocher, P. R. & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology*, *387*(2), 167–178. https://doi.org/10.1002/(sici)1096-9861(19971020)387:2<167::aid-cne1>3.0.co;2-z

Huttenlocher, P. R., De Courten, C., Garey, L. J. & Van der Loos, H. (1982). Synaptic development in human cerebral cortex. *International Journal of Neurology*, *16–17*, 144–154. http://www.ncbi.nlm.nih.gov/pubmed/6765658

Ince, P., Clark, B., Holton, J., Revesz, T. & Wharton, S. (2008). Chapter 13: Diseases of movement and system degenerations. In J. Greenfield, S. Love, D. Louis & D. Ellison (Eds.), *Greenfield's neuropathology* (8th editio, p. 947).

Ince, P. G., Evans, J., Knopp, M., Forster, G., Hamdalla, H. H. M., Wharton, S. B. & Shaw, P. J. (2003). Corticospinal tract degeneration in the progressive muscular atrophy variant of ALS. *Neurology*, *60*(8), 1252–1258. https://doi.org/10.1212/01.WNL.0000058901.75728.4E

Jenkinson, M., Bannister, P., Brady, M. & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. http://www.ncbi.nlm.nih.gov/pubmed/12377157

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jeurissen, B., Leemans, A., Tournier, J. D., Jones, D. K. & Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping*, *34*(11), 2747–2766. https://doi.org/10.1002/hbm.22099

Johannesen, S., Huie, J. R., Budeus, B., Peters, S., Wirth, A. M., Iberl, S., Kammermaier, T., Kobor, I., Wirkert, E., Küspert, S., Tahedl, M., Grassinger, J., Pukrop, T., Schneider, A., Aigner, L., Schulte-Mattler, W., Schuierer, G., Koch, W., Bruun, T. H., … Bogdahn, U. (2021). Modeling and Bioinformatics Identify Responders to G-CSF in Patients With Amyotrophic Lateral Sclerosis. *Frontiers in Neurology*, *12*(March), 1–15. https://doi.org/10.3389/fneur.2021.616289

Jurkowski, M. P., Bettio, L., K. Woo, E., Patten, A., Yau, S.-Y. & Gil-Mohapel, J. (2020). Beyond the Hippocampus and the SVZ: Adult Neurogenesis Throughout the Brain. *Frontiers in Cellular Neuroscience*, *14*. https://doi.org/10.3389/fncel.2020.576444

Kassubek, J., Juengling, F. D. & Sperfeld, A.-D. (2007). Widespread white matter changes in Kennedy disease: a voxel based morphometry study. *Journal of Neurology, Neurosurgery & Psychiatry*, *78*(11), 1209–1212. https://doi.org/10.1136/jnnp.2006.112532

Kennedy, D. (1998). Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cerebral Cortex*, *8*(4), 372–384. https://doi.org/10.1093/cercor/8.4.372

Kiernan, J. A. & Hudson, J. (1994). Frontal lobe atrophy in motor neuron diseases. *Brain*, *117*(4), 747–757. https://doi.org/10.1093/brain/117.4.747

Kim, J. P., Kim, J., Park, Y. H., Park, S. B., Lee, J. S., Yoo, S., Kim, E.-J., Kim, H. J., Na, D. L., Brown, J. A., Lockhart, S. N., Seo, S. W. & Seong, J.-K. (2019). Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. *NeuroImage: Clinical*, *23*, 101811. https://doi.org/10.1016/j.nicl.2019.101811

Kito, Y., Kazui, H., Kubo, Y., Yoshida, T., Takaya, M., Wada, T., Nomura, K., Hashimoto, M., Ohkawa, S., Miyake, H., Ishikawa, M. & Takeda, M. (2009). Neuropsychiatric symptoms in patients with idiopathic normal pressure hydrocephalus. *Behavioural Neurology*, *21*(3), 165–174. https://doi.org/10.3233/BEN-2009-0233

Koedam, E. L. G. E., Lehmann, M., Van Der Flier, W. M., Scheltens, P., Pijnenburg, Y. A. L., Fox, N., Barkhof,

F. & Wattjes, M. P. (2011). Visual assessment of posterior atrophy development of a MRI rating scale. *European Radiology*, *21*(12), 2618–2625. https://doi.org/10.1007/s00330-011-2205-4

Kong, L., Herold, C. J., Zöllner, F., Salat, D. H., Lässer, M. M., Schmid, L. A., Fellhauer, I., Thomann, P. A., Essig, M., Schad, L. R., Erickson, K. I. & Schröder, J. (2015). Comparison of grey matter volume and thickness for analysing cortical changes in chronic schizophrenia: a matter of surface area, grey/white matter intensity contrast, and curvature. *Psychiatry Research*, *231*(2), 176–183. https://doi.org/10.1016/j.pscychresns.2014.12.004

Kotagal, V., Lorincz, M. T. & Bohnen, N. I. (2012). A Frontotemporal Dementia-like Syndrome Mimicking Postpartum Depression Detected by 18F Fluorodeoxyglucose Positron Emission Tomography. *Clinical Nuclear Medicine*, *37*(9), e223–e224. https://doi.org/10.1097/RLU.0b013e31824440a1

Kozareva, D. A., Cryan, J. F. & Nolan, Y. M. (2019). Born this way: Hippocampal neurogenesis across the lifespan. *Aging Cell*, *18*(5). https://doi.org/10.1111/acel.13007

Kraemer, M., Buerger, M. & Berlit, P. (2010). Diagnostic problems and delay of diagnosis in amyotrophic lateral sclerosis. *Clinical Neurology and Neurosurgery*, *112*(2), 103–105. https://doi.org/10.1016/j.clineuro.2009.10.014

Lai, T.-H., Liu, R.-S., Yang, B.-H., Wang, P.-S., Lin, K.-P., Lee, Y.-C. & Soong, B.-W. (2013). Cerebral involvement in spinal and bulbar muscular atrophy (Kennedy's disease): A pilot study of PET. *Journal of the Neurological Sciences*, *335*(1–2), 139–144. https://doi.org/10.1016/j.jns.2013.09.016

Laird, N. M. & Ware, J. H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics*, *38*(4), 963–974.

Last, N., Tufts, E. & Auger, L. E. (2017). The Effects of Meditation on Grey Matter Atrophy and Neurodegeneration: A Systematic Review. *Journal of Alzheimer's Disease : JAD*, *56*(1), 275–286. https://doi.org/10.3233/JAD-160899

Lauterbur, P. C. (1973). Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature*, *242*(5394), 190–191. https://doi.org/10.1038/242190a0

Lazar, S. W., Kerr, C. E., Wasserman, R. H., Gray, J. R., Greve, D. N., Treadway, M. T., McGarvey, M., Quinn, B. T., Dusek, J. A., Benson, H., Rauch, S. L., Moore, C. I. & Fischl, B. (2005). Meditation experience is associated with increased cortical thickness. *Neuroreport*, *16*(17), 1893–1897. https://doi.org/10.1097/01.wnr.0000186598.66243.19

Lebouteux, M.-V., Franques, J., Guillevin, R., Delmont, E., Lenglet, T., Bede, P., Desnuelle, C., Pouget, J., Pascal-Mousselard, H. & Pradat, P.-F. (2014). Revisiting the spectrum of lower motor neuron diseases with snake eyes appearance on magnetic resonance imaging. *European Journal of Neurology*, *21*(9), 1233–1241. https://doi.org/10.1111/ene.12465

Lee, J. S., Park, Y. H., Park, S., Yoon, U., Choe, Y., Cheon, B. K., Hahn, A., Cho, S. H., Kim, S. J., Kim, J. P., Jung, Y. H., Park, K.-C., Kim, H. J., Jang, H., Na, D. L. & Seo, S. W. (2019). Distinct Brain Regions in Physiological and Pathological Brain Aging. *Frontiers in Aging Neuroscience*, *11*. https://doi.org/10.3389/fnagi.2019.00147

Li Hi Shing, S., Lope, J., Chipika, R. H., Hardiman, O. & Bede, P. (2021). Extra-motor manifestations in post-polio syndrome (PPS): fatigue, cognitive symptoms and radiological features. *Neurological Sciences*. https://doi.org/10.1007/s10072-021-05130-4

Li Hi Shing, S., Lope, J., McKenna, M. C., Chipika, R. H., Hardiman, O. & Bede, P. (2021). Increased cerebral integrity metrics in poliomyelitis survivors: putative adaptation to longstanding lower motor neuron degeneration. *Journal of the Neurological Sciences*, *424*, 117361. https://doi.org/10.1016/j.jns.2021.117361

Li Hi Shing, S., McKenna, M. C., Siah, W. F., Chipika, R. H., Hardiman, O. & Bede, P. (2021). The imaging signature of C9orf72 hexanucleotide repeat expansions: implications for clinical trials and therapy development. *Brain Imaging and Behavior*. https://doi.org/10.1007/s11682-020-00429-w

Li, M., Meng, Y., Wang, M., Yang, S., Wu, H., Zhao, B. & Wang, G. (2017). Cerebral gray matter volume reduction in subcortical vascular mild cognitive impairment patients and subcortical vascular dementia patients, and its relation with cognitive deficits. *Brain and Behavior*, *7*(8), e00745. https://doi.org/10.1002/brb3.745

Li, Q., Zhao, Y., Chen, Z., Long, J., Dai, J., Huang, X., Lui, S., Radua, J., Vieta, E., Kemp, G. J., Sweeney, J.

A., Li, F. & Gong, Q. (2019). Meta-analysis of cortical thickness abnormalities in medication-free patients with major depressive disorder. *Neuropsychopharmacology*. https://doi.org/10.1038/s41386-019-0563-9

Lian, C., Liu, M., Zhang, J. & Shen, D. (2020). Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis Using Structural MRI. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(4), 880–893. https://doi.org/10.1109/TPAMI.2018.2889096

Liu, G., Tan, S., Peng, K., Dang, C., Xing, S., Xie, C. & Zeng, J. (2019). Network change in the ipsilesional cerebellum is correlated with motor recovery following unilateral pontine infarction. *European Journal of Neurology*. https://doi.org/10.1111/ene.13974

Losito, L., Gennaro, L., Lucarelli, E. & Trabacca, A. (2020). Brain MRI abnormalities in a child with spinal muscular atrophy type II. *Acta Neurologica Belgica*. https://doi.org/10.1007/s13760-020-01524-x

Luders, E., Toga, A. W. & Thompson, P. M. (2014). Why size matters: differences in brain volume account for apparent sex differences in callosal anatomy: the sexual dimorphism of the corpus callosum. *NeuroImage*, *84*, 820–824. https://doi.org/10.1016/j.neuroimage.2013.09.040

Lule, D., Diekmann, V., Anders, S., Kassubek, J., Kübler, A., Ludolph, A. C. & Birbaumer, N. (2007). Brain responses to emotional stimuli in patients with amyotrophic lateral sclerosis (ALS). *Journal of Neurology*, *254*(4), 519–527. https://doi.org/10.1007/s00415-006-0409-3

Lule, D., Diekmann, V., Muller, H.-P., Kassubek, J., Ludolph, A. C. & Birbaumer, N. (2010). Neuroimaging of multimodal sensory stimulation in amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *81*(8), 899–906. https://doi.org/10.1136/jnnp.2009.192260

Lulé, D. E., Müller, H.-P., Finsel, J., Weydt, P., Knehr, A., Winroth, I., Andersen, P., Weishaupt, J., Uttner, I., Kassubek, J. & Ludolph, A. C. (2020). Deficits in verbal fluency in presymptomatic C9orf72 mutation gene carriers-a developmental disorder. *Journal of Neurology, Neurosurgery, and Psychiatry*, *91*(11), 1195–1200. https://doi.org/10.1136/jnnp-2020-323671

Ma, X., Li, Z., Jing, B., Liu, H., Li, D. & Li, H. (2016). Identify the Atrophy of Alzheimer's Disease, Mild Cognitive Impairment and Normal Aging Using Morphometric MRI Analysis. *Frontiers in Aging Neuroscience*, *08*. https://doi.org/10.3389/fnagi.2016.00243

Mackenzie, I. R. A. & Briemberg, H. (2020). TDP-43 pathology in primary lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *21*(sup1), 52–58. https://doi.org/10.1080/21678421.2020.1790607

Maguire, E A, Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(8), 4398–4403. https://doi.org/10.1073/pnas.070039597

Maguire, Eleanor A., Woollett, K. & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, *16*(12), 1091–1101. https://doi.org/10.1002/hipo.20233

Marcus, D. S., Harms, M. P., Snyder, A. Z., Jenkinson, M., Wilson, J. A., Glasser, M. F., Barch, D. M., Archie, K. A., Burgess, G. C., Ramaratnam, M., Hodge, M., Horton, W., Herrick, R., Olsen, T., McKay, M., House, M., Hileman, M., Reid, E., Harwell, J., … Van Essen, D. C. (2013). Human Connectome Project informatics: Quality control, database services, and data visualization. *NeuroImage*, *80*, 202–219. https://doi.org/10.1016/j.neuroimage.2013.05.077

Marcus, D. S., Harwell, J., Olsen, T., Hodge, M., Glasser, M. F., Prior, F., Jenkinson, M., Laumann, T., Curtiss, S. W. & Van Essen, D. C. (2011). Informatics and Data Mining Tools and Strategies for the Human Connectome Project. *Frontiers in Neuroinformatics*, *5*(June), 1–12. https://doi.org/10.3389/fninf.2011.00004

Martin, S. B., Covell, D. J., Joseph, J. E., Chebrolu, H., Smith, C. D., Kelly, T. H., Jiang, Y. & Gold, B. T. (2007). Human experience seeking correlates with hippocampus volume: Convergent evidence from manual tracing and voxel-based morphometry. *Neuropsychologia*, *45*(12), 2874–2881. https://doi.org/10.1016/j.neuropsychologia.2007.05.009

McFarquhar, M., Mckie, S., Emsley, R., Suckling, J., Elliott, R. & Williams, S. (2016). NeuroImage Multivariate and repeated measures ( MRM ): A new toolbox for dependent and multimodal group-level neuroimaging data. *NeuroImage*, *132*, 373–389.

https://doi.org/10.1016/j.neuroimage.2016.02.053

McKeon, A., Marnane, M., O'Connell, M., Stack, J. P., Kelly, P. J. & Lynch, T. (2007). Potassium Channel Antibody–Associated Encephalopathy Presenting With a Frontotemporal Dementia–like Syndrome. *Archives of Neurology*, *64*(10), 1528. https://doi.org/10.1001/archneur.64.10.1528

McMahon, P. M., Araki, S. S., Sandberg, E. A., Neumann, P. J. & Gazelle, G. S. (2003). Cost-Effectiveness of PET in the Diagnosis of Alzheimer Disease. *Radiology*, *228*(2), 515–522. https://doi.org/10.1148/radiol.2282020915

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S. & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757–757. https://doi.org/10.1038/431757a

Meeter, L. H. H., Steketee, R. M. E., Salkovic, D., Vos, M. E., Grossman, M., McMillan, C. T., Irwin, D. J., Boxer, A. L., Rojas, J. C., Olney, N. T., Karydas, A., Miller, B. L., Pijnenburg, Y. A. L., Barkhof, F., Sánchez-Valle, R., Lladó, A., Borrego-Ecija, S., Diehl-Schmid, J., Grimmer, T., … Van Swieten, J. C. (2019). Clinical value of cerebrospinal fluid neurofilament light chain in semantic dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(9), 997–1004. https://doi.org/10.1136/jnnp-2018-319784

Meier, J. M., van der Burgh, H. K., Nitert, A. D., Bede, P., de Lange, S. C., Hardiman, O., van den Berg, L. H. & van den Heuvel, M. P. (2020). Connectome-Based Propagation Model in Amyotrophic Lateral Sclerosis. *Annals of Neurology*, *87*(5), 725–738. https://doi.org/10.1002/ana.25706

Mendonça, R. H., Rocha, A. J., Lozano-Arango, A., Diaz, A. B., Castiglioni, C., Silva, A. M. S., Reed, U. C., Kulikowski, L., Paramonov, I., Cuscó, I., Tizzano, E. F. & Zanoteli, E. (2019). Severe brain involvement in 5q spinal muscular atrophy type 0. *Annals of Neurology*, *86*(3), 458–462. https://doi.org/10.1002/ana.25549

Menke, R. A. L., Proudfoot, M., Talbot, K. & Turner, M. R. (2018). The two-year progression of structural and functional cerebral MRI in amyotrophic lateral sclerosis. *NeuroImage. Clinical*, *17*, 953–961. https://doi.org/10.1016/j.nicl.2017.12.025

Menzler, K., Belke, M., Wehrmann, E., Krakow, K., Lengler, U., Jansen, A., Hamer, H. M., Oertel, W. H., Rosenow, F. & Knake, S. (2011). Men and women are different: diffusion tensor imaging reveals sexual dimorphism in the microstructure of the thalamus, corpus callosum and cingulum. *NeuroImage*, *54*(4), 2557–2562. https://doi.org/10.1016/j.neuroimage.2010.11.029

Miller, T., Cudkowicz, M., Shaw, P. J., Andersen, P. M., Atassi, N., Bucelli, R. C., Genge, A., Glass, J., Ladha, S., Ludolph, A. L., Maragakis, N. J., McDermott, C. J., Pestronk, A., Ravits, J., Salachas, F., Trudell, R., Van Damme, P., Zinman, L., Bennett, C. F., … Ferguson, T. A. (2020). Phase 1–2 Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *New England Journal of Medicine*, *383*(2), 109–119. https://doi.org/10.1056/NEJMoa2003715

Mitsumoto, H., Ulug, A. M., Pullman, S. L., Gooch, C. L., Chan, S., Tang, M.-X., Mao, X., Hays, A. P., Floyd, A. G., Battista, V., Montes, J., Hayes, S., Dashnaw, S., Kaufmann, P., Gordon, P. H., Hirsch, J., Levin, B., Rowland, L. P. & Shungu, D. C. (2007). Quantitative objective markers for upper and lower motor neuron dysfunction in ALS. *Neurology*, *68*(17), 1402–1410. https://doi.org/10.1212/01.wnl.0000260065.57832.87

Mitsumoto, Hiroshi, Brooks, B. R. & Silani, V. (2014). Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *The Lancet. Neurology*, *13*(11), 1127–1138. https://doi.org/10.1016/S1474-4422(14)70129-2

Mueller, C., Hussl, A., Krismer, F., Heim, B., Mahlknecht, P., Nocker, M., Scherfler, C., Mair, K., Esterhammer, R., Schocke, M., Wenning, G. K., Poewe, W. & Seppi, K. (2018). The diagnostic accuracy of the hummingbird and morning glory sign in patients with neurodegenerative parkinsonism. *Parkinsonism & Related Disorders*, *54*, 90–94. https://doi.org/10.1016/j.parkreldis.2018.04.005

Müller, H.-P., Del Tredici, K., Lulé, D., Müller, K., Weishaupt, J. H., Ludolph, A. C. & Kassubek, J. (2020). In vivo histopathological staging in C9orf72-associated ALS: A tract of interest DTI study. *NeuroImage: Clinical*, *27*, 102298. https://doi.org/10.1016/j.nicl.2020.102298

Müller, H.-P., Gorges, M., Kassubek, R., Dorst, J., Ludolph, A. C. & Kassubek, J. (2018). Identical patterns of cortico-efferent tract involvement in primary lateral sclerosis and amyotrophic lateral sclerosis:

A tract of interest-based MRI study. *NeuroImage: Clinical*, *18*, 762–769. https://doi.org/10.1016/j.nicl.2018.03.018

Müller, H.-P., Turner, M. R., Grosskreutz, J., Abrahams, S., Bede, P., Govind, V., Prudlo, J., Ludolph, A. C., Filippi, M. & Kassubek, J. (2016). A large-scale multicentre cerebral diffusion tensor imaging study in amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(6), 570–579. https://doi.org/10.1136/jnnp-2015-311952

Müller, H.-P., Unrath, A., Huppertz, H.-J., Ludolph, A. C. & Kassubek, J. (2012). Neuroanatomical patterns of cerebral white matter involvement in different motor neuron diseases as studied by diffusion tensor imaging analysis. *Amyotrophic Lateral Sclerosis*, *13*(3), 254–264. https://doi.org/10.3109/17482968.2011.653571

MUQIT, M. M. K. (2001). "Hot cross bun" sign in a patient with parkinsonism secondary to presumed vasculitis. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*(4), 565–566. https://doi.org/10.1136/jnnp.71.4.565

Nasseroleslami, B., Dukic, S., Broderick, M., Mohr, K., Schuster, C., Gavin, B., McLaughlin, R., Heverin, M., Vajda, A., Iyer, P. M., Pender, N., Bede, P., Lalor, E. C. & Hardiman, O. (2019). Characteristic Increases in EEG Connectivity Correlate With Changes of Structural MRI in Amyotrophic Lateral Sclerosis. *Cerebral Cortex*, *29*(1), 27–41. https://doi.org/10.1093/cercor/bhx301

Nelles, M., Block, W., Träber, F., Wüllner, U., Schild, H. H. & Urbach, H. (2008). Combined 3T Diffusion Tensor Tractography and 1 H-MR Spectroscopy in Motor Neuron Disease. *American Journal of Neuroradiology*, *29*(9), 1708–1714. https://doi.org/10.3174/ajnr.A1201

Nichols, T. E. & Holmes, A. P. (2001). Nonparametric Permutation Tests For Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping*, *25*(July), 1–25. https://doi.org/10.1002/hbm.1058

Nishimura, Y. & Isa, T. (2009). Compensatory Changes at the Cerebral Cortical Level after Spinal Cord Injury. *The Neuroscientist*, *15*(5), 436–444. https://doi.org/10.1177/1073858408331375

Nishio, Y., Nakano, Y., Matsumoto, K., Hashimoto, M., Kazui, H., Hirono, N., Ishii, K. & Mori, E. (2003). Striatal infarcts mimicking frontotemporal dementia: a case report. *European Journal of Neurology*, *10*(4), 457–460. https://doi.org/10.1046/j.1468-1331.2003.00628.x

Omer, T., Finegan, E., Hutchinson, S., Doherty, M., Vajda, A., McLaughlin, R. L., Pender, N., Hardiman, O. & Bede, P. (2017). Neuroimaging patterns along the ALS-FTD spectrum: a multiparametric imaging study. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *18*(7–8), 611–623. https://doi.org/10.1080/21678421.2017.1332077

Pannek, K., Fripp, J., George, J. M., Fiori, S., Colditz, P. B., Boyd, R. N. & Rose, S. E. (2018). Fixel-based analysis reveals alterations is brain microstructure and macrostructure of preterm-born infants at term equivalent age. *NeuroImage: Clinical*, *18*, 51–59. https://doi.org/10.1016/j.nicl.2018.01.003

Paterson, R. W., Slattery, C. F., Poole, T., Nicholas, J. M., Magdalinou, N. K., Toombs, J., Chapman, M. D., Lunn, M. P., Heslegrave, A. J., Foiani, M. S., Weston, P. S. J., Keshavan, A., Rohrer, J. D., Rossor, M. N., Warren, J. D., Mummery, C. J., Blennow, K., Fox, N. C., Zetterberg, H. & Schott, J. M. (2018). Cerebrospinal fluid in the differential diagnosis of Alzheimer's disease: clinical utility of an extended panel of biomarkers in a specialist cognitive clinic. *Alzheimer's Research & Therapy*, *10*(1), 32. https://doi.org/10.1186/s13195-018-0361-3

Pecheva, D., Tournier, J.-D., Pietsch, M., Christiaens, D., Batalle, D., Alexander, D. C., Hajnal, J. V, Edwards, A. D., Zhang, H. & Counsell, S. J. (2019). Fixel-based analysis of the preterm brain: Disentangling bundle-specific white matter microstructural and macrostructural changes in relation to clinical risk factors. *NeuroImage. Clinical*, *23*, 101820. https://doi.org/10.1016/j.nicl.2019.101820

Pereira, A. C., Huddleston, D. E., Brickman, A. M., Sosunov, A. A., Hen, R., McKhann, G. M., Sloan, R., Gage, F. H., Brown, T. R. & Small, S. A. (2007). An in vivo correlate of exercise-induced neurogenesis in the adult dentate gyrus. *Proceedings of the National Academy of Sciences*, *104*(13), 5638–5643. https://doi.org/10.1073/pnas.0611721104

Perneczky, R. (2019). Dementia prevention and reserve against neurodegenerative disease. *Dialogues in Clinical Neuroscience*, *21*(1), 53–60. http://www.ncbi.nlm.nih.gov/pubmed/31607780

Perry, D. C., Brown, J. A., Possin, K. L., Datta, S., Trujillo, A., Radke, A., Karydas, A., Kornak, J., Sias, A. C., Rabinovici, G. D., Gorno-Tempini, M. L., Boxer, A. L., De May, M., Rankin, K. P., Sturm, V. E., Lee, S.

E., Matthews, B. R., Kao, A. W., Vossel, K. A., … Seeley, W. W. (2017). Clinicopathological correlations in behavioural variant frontotemporal dementia. *Brain*, *140*(12), 3329–3345. https://doi.org/10.1093/brain/awx254

Pieper, C. C., Konrad, C., Sommer, J., Teismann, I. & Schiffbauer, H. (2013). Structural changes of central white matter tracts in Kennedy's disease - a diffusion tensor imaging and voxel-based morphometry study. *Acta Neurologica Scandinavica*, *127*(5), 323–328. https://doi.org/10.1111/ane.12018

Piguet, O., Hornberger, M., Shelley, B. P., Kipps, C. M. & Hodges, J. R. (2009). Sensitivity of current criteria for the diagnosis of behavioral variant frontotemporal dementia. *Neurology*, *72*(8), 732–737. https://doi.org/10.1212/01.wnl.0000343004.98599.45

Pinheiro, J., Bates, D., DebRoy, S. & D, S. (2020). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-150. *R Core Team*. https://doi.org/https://cran.r-project.org/package=nlme

Pioro, E. P., Turner, M. R. & Bede, P. (2020). Neuroimaging in primary lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *21*(sup1), 18–27. https://doi.org/10.1080/21678421.2020.1837176

Ponti, G., Obernier, K. & Alvarez-Buylla, A. (2013). Lineage progression from stem cells to new neurons in the adult brain ventricular-subventricular zone. *Cell Cycle*, *12*(11), 1649–1650. https://doi.org/10.4161/cc.24984

Pradat, P.-F., Bruneteau, G., Munerati, E., Salachas, F., Le Forestier, N., Lacomblez, L., Lenglet, T. & Meininger, V. (2009). Extrapyramidal stiffness in patients with amyotrophic lateral sclerosis. *Movement Disorders : Official Journal of the Movement Disorder Society*, *24*(14), 2143–2148. https://doi.org/10.1002/mds.22762

Pradat, P.-F. & El Mendili, M.-M. (2014). Neuroimaging to investigate multisystem involvement and provide biomarkers in amyotrophic lateral sclerosis. *BioMed Research International*, *2014*, 467560. https://doi.org/10.1155/2014/467560

Proudfoot, M., Bede, P. & Turner, M. R. (2019). Imaging Cerebral Activity in Amyotrophic Lateral Sclerosis. *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.01148

Prudlo, J., Bißbort, C., Glass, A., Grossmann, A., Hauenstein, K., Benecke, R. & Teipel, S. J. (2012). White matter pathology in ALS and lower motor neuron ALS variants: a diffusion tensor imaging study using tract-based spatial statistics. *Journal of Neurology*, *259*(9), 1848–1859. https://doi.org/10.1007/s00415-012-6420-y

Qin, Q., Tang, Y., Dou, X., Qu, Y., Xing, Y., Yang, J., Chu, T., Liu, Y. & Jia, J. (2021). Default mode network integrity changes contribute to cognitive deficits in subcortical vascular cognitive impairment, no dementia. *Brain Imaging and Behavior*, *15*(1), 255–265. https://doi.org/10.1007/s11682-019-00252-y

Querin, G., Bede, P., El Mendili, M. M., Li, M., Pélégrini-Issac, M., Rinaldi, D., Catala, M., Saracino, D., Salachas, F., Camuzat, A., Marchand-Pauvert, V., Cohen-Adad, J., Colliot, O., Le Ber, I. & Pradat, P. (2019). Presymptomatic spinal cord pathology in c9orf72 mutation carriers: A longitudinal neuroimaging study. *Annals of Neurology*, *86*(2), 158–167. https://doi.org/10.1002/ana.25520

Querin, G., Bede, P., Marchand-Pauvert, V. & Pradat, P.-F. (2018). Biomarkers of Spinal and Bulbar Muscle Atrophy (SBMA): A Comprehensive Review. *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.00844

Querin, G., El Mendili, M.-M., Bede, P., Delphine, S., Lenglet, T., Marchand-Pauvert, V. & Pradat, P.-F. (2018a). Multimodal spinal cord MRI offers accurate diagnostic classification in ALS. *Journal of Neurology, Neurosurgery, and Psychiatry*, *89*(11), 1220–1221. https://doi.org/10.1136/jnnp-2017-317214

Querin, G., El Mendili, M.-M., Bede, P., Delphine, S., Lenglet, T., Marchand-Pauvert, V. & Pradat, P.-F. (2018b). Multimodal spinal cord MRI offers accurate diagnostic classification in ALS. *Journal of Neurology, Neurosurgery & Psychiatry*, *89*(11), 1220–1221. https://doi.org/10.1136/jnnp-2017-317214

Querin, G., El Mendili, M.-M., Lenglet, T., Behin, A., Stojkovic, T., Salachas, F., Devos, D., Le Forestier, N., Del Mar Amador, M., Debs, R., Lacomblez, L., Meininger, V., Bruneteau, G., Cohen-Adad, J., Lehéricy, S., Laforêt, P., Blancho, S., Benali, H., Catala, M., … Pradat, P.-F. (2019). The spinal and

cerebral profile of adult spinal-muscular atrophy: A multimodal imaging study. *NeuroImage. Clinical*, *21*, 101618. https://doi.org/10.1016/j.nicl.2018.101618

Querin, G., Lenglet, T., Debs, R., Stojkovic, T., Behin, A., Salachas, F., Le Forestier, N., Amador, M. del M., Lacomblez, L., Meininger, V., Bruneteau, G., Laforêt, P., Blancho, S., Marchand-Pauvert, V., Bede, P., Hogrel, J.-Y. & Pradat, P.-F. (2018). The motor unit number index (MUNIX) profile of patients with adult spinal muscular atrophy. *Clinical Neurophysiology*, *129*(11), 2333–2340. https://doi.org/10.1016/j.clinph.2018.08.025

Raaphorst, J., van Tol, M.-J., Groot, P. F. C., Altena, E., van der Werf, Y. D., Majoie, C. B., van der Kooi, A. J., van den Berg, L. H., Schmand, B., de Visser, M. & Veltman, D. J. (2014). Prefrontal involvement related to cognitive impairment in progressive muscular atrophy. *Neurology*, *83*(9), 818–825. https://doi.org/10.1212/WNL.0000000000000745

Radakovic, R., Stephenson, L., Colville, S., Swingler, R., Chandran, S. & Abrahams, S. (2016). Multidimensional apathy in ALS: validation of the Dimensional Apathy Scale. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(6), 663–669. https://doi.org/10.1136/jnnp-2015-310772

Raffelt, D. A., Tournier, J. D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R. & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-based analysis. *NeuroImage*, *144*, 58–73. https://doi.org/10.1016/j.neuroimage.2016.09.029

Raffelt, D., Tournier, J. D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., Salvado, O. & Connelly, A. (2012). Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*, *59*(4), 3976–3994. https://doi.org/10.1016/j.neuroimage.2011.10.045

Rajagopalan, V. & Pioro, E. P. (2015). Comparing brain structural MRI and metabolic FDG-PET changes in patients with ALS-FTD: 'the chicken or the egg?' question. *Journal of Neurology, Neurosurgery & Psychiatry*, *86*(9), 952–958. https://doi.org/10.1136/jnnp-2014-308239

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M.-L., Rosen, H., … Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*(9), 2456–2477. https://doi.org/10.1093/brain/awr179

Rasmussen, H., Hellzen, O., Stordal, E. & Enmarker, I. (2019). Family caregivers experiences of the pre-diagnostic stage in frontotemporal dementia. *Geriatric Nursing*, *40*(3), 246–251. https://doi.org/10.1016/j.gerinurse.2018.10.006

Reuter, M., Rosas, H. D. & Fischl, B. (2010). Highly accurate inverse consistent registration: a robust approach. *NeuroImage*, *53*(4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020

Robinson, P. N., Mungall, C. J. & Haendel, M. (2015). Capturing phenotypes for precision medicine. *Cold Spring Harbor Molecular Case Studies*, *1*(1), a000372. https://doi.org/10.1101/mcs.a000372

Rohrer, J. D., Lashley, T., Schott, J. M., Warren, J. E., Mead, S., Isaacs, A. M., Beck, J., Hardy, J., de Silva, R., Warrington, E., Troakes, C., Al-Sarraj, S., King, A., Borroni, B., Clarkson, M. J., Ourselin, S., Holton, J. L., Fox, N. C., Revesz, T., … Warren, J. D. (2011). Clinical and neuroanatomical signatures of tissue pathology in frontotemporal lobar degeneration. *Brain*, *134*(9), 2565–2581. https://doi.org/10.1093/brain/awr198

Sale, M. V., Reid, L. B., Cocchi, L., Pagnozzi, A. M., Rose, S. E. & Mattingley, J. B. (2017). Brain changes following four weeks of unimanual motor training: Evidence from behavior, neural stimulation, cortical thickness, and functional MRI. *Human Brain Mapping*, *38*(9), 4773–4787. https://doi.org/10.1002/hbm.23710

Salgado, A. J. (2020). *Handbook of Innovations in Central Nervous System Regenerative Medicine*. Academic Press.

Salimi-Khorshidi, G., Smith, S. M. & Nichols, T. E. (2011). Adjusting the effect of nonstationarity in cluster-based and TFCE inference. *NeuroImage*, *54*(3), 2006–2019. https://doi.org/10.1016/j.neuroimage.2010.09.088

Salvatore, C. & Castiglioni, I. (2018). A wrapped multi-label classifier for the automatic diagnosis and prognosis of Alzheimer's disease. *Journal of Neuroscience Methods*, *302*, 58–65. https://doi.org/10.1016/j.jneumeth.2017.12.016

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B. & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*, *28*(9), 3095–3114. https://doi.org/10.1093/cercor/bhx179

Schneider, A., Krüger, C., Steigleder, T., Weber, D., Pitzer, C., Laage, R., Aronowski, J., Maurer, M. H., Gassler, N., Mier, W., Hasselblatt, M., Kollmar, R., Schwab, S., Sommer, C., Bach, A., Kuhn, H.-G. & Schäbitz, W.-R. (2005). The hematopoietic factor G-CSF is a neuronal ligand that counteracts programmed cell death and drives neurogenesis. *The Journal of Clinical Investigation*, *115*(8), 2083–2098. https://doi.org/10.1172/JCI23559

Schuster, C., Elamin, M., Hardiman, O. & Bede, P. (2016). The segmental diffusivity profile of amyotrophic lateral sclerosis associated white matter degeneration. *European Journal of Neurology*, *23*(8), 1361–1371. https://doi.org/10.1111/ene.13038

Schuster, Christina, Elamin, M., Hardiman, O. & Bede, P. (2015). Presymptomatic and longitudinal neuroimaging in neurodegeneration--from snapshots to motion picture: a systematic review. *Journal of Neurology, Neurosurgery, and Psychiatry*, *86*(10), 1089–1096. https://doi.org/10.1136/jnnp-2014-309888

Schuster, Christina, Hardiman, O. & Bede, P. (2016). Development of an Automated MRI-Based Diagnostic Protocol for Amyotrophic Lateral Sclerosis Using Disease-Specific Pathognomonic Features: A Quantitative Disease-State Classification Study. *PloS One*, *11*(12), e0167331. https://doi.org/10.1371/journal.pone.0167331

Schuster, Christina, Hardiman, O. & Bede, P. (2017). Survival prediction in Amyotrophic lateral sclerosis based on MRI measures and clinical characteristics. *BMC Neurology*, *17*(1), 73. https://doi.org/10.1186/s12883-017-0854-x

Schuster, Christina, Kasper, E., Machts, J., Bittner, D., Kaufmann, J., Benecke, R., Teipel, S., Vielhaber, S. & Prudlo, J. (2013). Focal thinning of the motor cortex mirrors clinical features of amyotrophic lateral sclerosis and their phenotypes: a neuroimaging study. *Journal of Neurology*, *260*(11), 2856–2864. https://doi.org/10.1007/s00415-013-7083-z

Schwarz, C. G., Gunter, J. L., Wiste, H. J., Przybelski, S. A., Weigand, S. D., Ward, C. P., Senjem, M. L., Vemuri, P., Murray, M. E., Dickson, D. W., Parisi, J. E., Kantarci, K., Weiner, M. W., Petersen, R. C. & Jack, C. R. (2016). A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity. *NeuroImage: Clinical*, *11*, 802–812. https://doi.org/10.1016/j.nicl.2016.05.017

Seo, S. W., Ahn, J., Yoon, U., Im, K., Lee, J.-M., Tae Kim, S., Ahn, H.-J., Chin, J., Jeong, Y. & Na, D. L. (2010). Cortical Thinning in Vascular Mild Cognitive Impairment and Vascular Dementia of Subcortical Type. *Journal of Neuroimaging*, *20*(1), 37–45. https://doi.org/10.1111/j.1552-6569.2008.00293.x

Shafto, M., Tyler, L. K., Dixon, M., Taylor, J. R., Rowe, J. B., Cusack, R., Calder, A. J., Marslen-Wilson, W. D., Duncan, J., Dalgleish, T., Henson, R. N., Brayne, C. & Matthews, F. E. (2014). The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurology*, *14*, 204. https://doi.org/10.1186/s12883-014-0204-1

Shapiro, L. A., Ng, K., Zhou, Q.-Y. & Ribak, C. E. (2009). Subventricular zone-derived, newly generated neurons populate several olfactory and limbic forebrain regions. *Epilepsy & Behavior*, *14*(1), 74–80. https://doi.org/10.1016/j.yebeh.2008.09.011

Shaw, M. E., Abhayaratna, W. P., Sachdev, P. S., Anstey, K. J. & Cherbuin, N. (2016). Cortical Thinning at Midlife: The PATH Through Life Study. *Brain Topography*, *29*(6), 875–884. https://doi.org/10.1007/s10548-016-0509-z

Sheskin, D. J. (2003). *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press.

Shimoda, K., Kimura, M., Yokota, M. & Okubo, Y. (2015). Comparison of regional gray matter volume abnormalities in Alzheimer's disease and late life depression with hippocampal atrophy using VSRAD analysis: A voxel-based morphometry study. *Psychiatry Research: Neuroimaging*, *232*(1), 71–75. https://doi.org/10.1016/j.pscychresns.2015.01.018

Shing, S. L. H., Chipika, R. H., Finegan, E., Murray, D., Hardiman, O. & Bede, P. (2019). Post-polio Syndrome: More than just a lower motor neuron disease. *Frontiers in Neurology*, *10*(JUL). https://doi.org/10.3389/fneur.2019.00773

Shivamurthy, V. K. N., Tahari, A. K., Marcus, C. & Subramaniam, R. M. (2015). Brain FDG PET and the Diagnosis of Dementia. *American Journal of Roentgenology*, *204*(1), W76–W85. https://doi.org/10.2214/AJR.13.12363

Singh, G., Samavedham, L., Lim, E. C.-H., Alzheimer's Disease Neuroimaging Initiative & Parkinson Progression Marker Initiative. (2018). Determination of Imaging Biomarkers to Decipher Disease Trajectories and Differential Diagnosis of Neurodegenerative Diseases (DIsease TreND). *Journal of Neuroscience Methods*, *305*, 105–116. https://doi.org/10.1016/j.jneumeth.2018.05.009

Snowden, J., Neary, D. & Mann, D. (2007). Frontotemporal lobar degeneration: clinical and pathological relationships. *Acta Neuropathologica*, *114*(1), 31–38. https://doi.org/10.1007/s00401-007-0236-3

Sowell, E. R., Peterson, B. S., Kan, E., Woods, R. P., Yoshii, J., Bansal, R., Xu, D., Zhu, H., Thompson, P. M. & Toga, A. W. (2007). Sex Differences in Cortical Thickness Mapped in 176 Healthy Individuals between 7 and 87 Years of Age. *Cerebral Cortex*, *17*(7), 1550–1560. https://doi.org/10.1093/cercor/bhl066

Spasov, S., Passamonti, L., Duggento, A., Liò, P. & Toschi, N. (2019). A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *NeuroImage*, *189*, 276–287. https://doi.org/10.1016/j.neuroimage.2019.01.031

Sperfeld, A.-D., Bretschneider, V., Flaith, L., Unrath, A., Hanemann, C. O., Ludolph, A. C. & Kassubek, J. (2005). MR-Pathologic Comparison of the Upper Spinal Cord in Different Motor Neuron Diseases. *European Neurology*, *53*(2), 74–77. https://doi.org/10.1159/000084650

Spinelli, E. G., Agosta, F., Ferraro, P. M., Querin, G., Riva, N., Bertolin, C., Martinelli, I., Lunetta, C., Fontana, A., Sorarù, G. & Filippi, M. (2019). Brain MRI shows white matter sparing in Kennedy's disease and slow-progressing lower motor neuron disease. *Human Brain Mapping*, hbm.24583. https://doi.org/10.1002/hbm.24583

Spreadbury, J. H. & Kipps, C. (2019). Measuring younger onset dementia: What the qualitative literature reveals about the 'lived experience' for patients and caregivers. *Dementia*, *18*(2), 579–598. https://doi.org/10.1177/1471301216684401

Stadelmann, C. (2011). Multiple sclerosis as a neurodegenerative disease: pathology, mechanisms and therapeutic implications. *Current Opinion in Neurology*, *24*(3), 224–229. https://doi.org/10.1097/WCO.0b013e328346056f

Steenwijk, M. D., Geurts, J. J. G., Daams, M., Tijms, B. M., Wink, A. M., Balk, L. J., Tewarie, P. K., Uitdehaag, B. M. J., Barkhof, F., Vrenken, H. & Pouwels, P. J. W. (2016). Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain : A Journal of Neurology*, *139*(Pt 1), 115–126. https://doi.org/10.1093/brain/awv337

Steinacker, P., Semler, E., Anderl-Straub, S., Diehl-Schmid, J., Schroeter, M. L., Uttner, I., Foerstl, H., Landwehrmeyer, B., von Arnim, C. A. F., Kassubek, J., Oeckl, P., Huppertz, H.-J., Fassbender, K., Fliessbach, K., Prudlo, J., Roßmeier, C., Kornhuber, J., Schneider, A., Volk, A. E., … Otto, M. (2017). Neurofilament as a blood marker for diagnosis and monitoring of primary progressive aphasias. *Neurology*, *88*(10), 961–969. https://doi.org/10.1212/WNL.0000000000003688

Sterr, A., Dean, P. J. A., Vieira, G., Conforto, A. B., Shen, S. & Sato, J. R. (2013). Cortical thickness changes in the non-lesioned hemisphere associated with non-paretic arm immobilization in modified CI therapy. *NeuroImage: Clinical*, *2*, 797–803. https://doi.org/10.1016/j.nicl.2013.05.005

Strong, M. J., Abrahams, S., Goldstein, L. H., Woolley, S., Mclaughlin, P., Snowden, J., Mioshi, E., Roberts-South, A., Benatar, M., HortobáGyi, T., Rosenfeld, J., Silani, V., Ince, P. G. & Turner, M. R. (2017). Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *18*(3–4), 153–174. https://doi.org/10.1080/21678421.2016.1267768

Swift, I. J., Sogorb-Esteve, A., Heller, C., Synofzik, M., Otto, M., Graff, C., Galimberti, D., Todd, E., Heslegrave, A. J., van der Ende, E. L., Van Swieten, J. C., Zetterberg, H. & Rohrer, J. D. (2021). Fluid biomarkers in frontotemporal dementia: past, present and future. *Journal of Neurology, Neurosurgery & Psychiatry*, *92*(2), 204–215. https://doi.org/10.1136/jnnp-2020-323520

Tahedl, M. (2020). Towards individualized cortical thickness assessment for clinical routine. *Journal of Translational Medicine*, 1–12. https://doi.org/10.1186/s12967-020-02317-9

Tahedl, M., Chipika, R. H., Lope, J., Li Hi Shing, S., Hardiman, O. & Bede, P. (2021). Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use. *Journal of Neurology*. https://doi.org/10.1007/s00415-020-10368-7

Thomas, A. G., Dennis, A., Rawlings, N. B., Stagg, C. J., Matthews, L., Morris, M., Kolind, S. H., Foxley, S., Jenkinson, M., Nichols, T. E., Dawes, H., Bandettini, P. A. & Johansen-Berg, H. (2016). Multi-modal characterization of rapid anterior hippocampal volume increase associated with aerobic exercise. *NeuroImage*, *131*, 162–170. https://doi.org/10.1016/j.neuroimage.2015.10.090

Toda, T., Parylak, S. L., Linker, S. B. & Gage, F. H. (2019). The role of adult hippocampal neurogenesis in brain health and disease. *Molecular Psychiatry*, *24*(1), 67–87. https://doi.org/10.1038/s41380-018-0036-2

Tournier, J., Yeh, C., Calamante, F. & Cho, K. (2008). Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *NeuroImage*, *42*, 617–625. https://doi.org/10.1016/j.neuroimage.2008.05.002

Traxinger, K., Kelly, C., Johnson, B. A., Lyles, R. H. & Glass, J. D. (2013). Prognosis and epidemiology of amyotrophic lateral sclerosis: Analysis of a clinic population, 1997-2011. *Neurology: Clinical Practice*, *3*(4), 313–320. https://doi.org/10.1212/CPJ.0b013e3182a1b8ab

Trojsi, F., Di Nardo, F., Caiazzo, G., Siciliano, M., D'Alvano, G., Ferrantino, T., Passaniti, C., Ricciardi, D., Esposito, S., Lavorgna, L., Russo, A., Bonavita, S., Cirillo, M., Santangelo, G., Esposito, F. & Tedeschi, G. (2020). Hippocampal connectivity in Amyotrophic Lateral Sclerosis (ALS): more than Papez circuit impairment. *Brain Imaging and Behavior*. https://doi.org/10.1007/s11682-020-00408-1

Trojsi, F., Sorrentino, P., Sorrentino, G. & Tedeschi, G. (2018). Neurodegeneration of brain networks in the amyotrophic lateral sclerosis-frontotemporal lobar degeneration (ALS-FTLD) continuum: evidence from MRI and MEG studies. *CNS Spectrums*, *23*(6), 378–387. https://doi.org/10.1017/S109285291700075X

Tucholka, A., Fritsch, V., Poline, J.-B. & Thirion, B. (2012). An empirical comparison of surface-based and volume-based group studies in neuroimaging. *NeuroImage*, *63*(3), 1443–1453. https://doi.org/10.1016/j.neuroimage.2012.06.019

Turner, M. R., Barohn, R. J., Corcia, P., Fink, J. K., Harms, M. B., Kiernan, M. C., Ravits, J., Silani, V., Simmons, Z., Statland, J., van den Berg, L. H. & Mitsumoto, H. (2020). Primary lateral sclerosis: consensus diagnostic criteria. *Journal of Neurology, Neurosurgery & Psychiatry*, *91*(4), 373–377. https://doi.org/10.1136/jnnp-2019-322541

Unrath, A., Müller, H.-P., Riecker, A., Ludolph, A. C., Sperfeld, A.-D. & Kassubek, J. (2010). Whole brain-based analysis of regional white matter tract alterations in rare motor neuron diseases by diffusion tensor imaging. *Human Brain Mapping*, NA-NA. https://doi.org/10.1002/hbm.20971

van der Burgh, H. K., Westeneng, H.-J., Meier, J. M., van Es, M. A., Veldink, J. H., Hendrikse, J., van den Heuvel, M. P. & van den Berg, L. H. (2019). Cross-sectional and longitudinal assessment of the upper cervical spinal cord in motor neuron disease. *NeuroImage: Clinical*, *24*, 101984. https://doi.org/10.1016/j.nicl.2019.101984

van der Burgh, H. K., Westeneng, H.-J., Walhout, R., van Veenhuijzen, K., Tan, H. H. G., Meier, J. M., Bakker, L. A., Hendrikse, J., van Es, M. A., Veldink, J. H., van den Heuvel, M. P. & van den Berg, L. H. (2020). Multimodal longitudinal study of structural brain involvement in amyotrophic lateral sclerosis. *Neurology*, *94*(24), e2592–e2604. https://doi.org/10.1212/WNL.0000000000009498

van der Graaff, M. M., Lavini, C., Akkerman, E. M., Majoie, C. B., Nederveen, A. J., Zwinderman, A. H., Brugman, F., van den Berg, L. H., de Jong, J. M. B. V. & de Visser, M. (2010). MR Spectroscopy Findings in Early Stages of Motor Neuron Disease. *American Journal of Neuroradiology*, *31*(10), 1799–1806. https://doi.org/10.3174/ajnr.A2217

van der Graaff, M. M., Sage, C. A., Caan, M. W. A., Akkerman, E. M., Lavini, C., Majoie, C. B., Nederveen, A. J., Zwinderman, A. H., Vos, F., Brugman, F., van den Berg, L. H., de Rijk, M. C., van Doorn, P. A., Van Hecke, W., Peeters, R. R., Robberecht, W., Sunaert, S. & de Visser, M. (2011). Upper and extra-motoneuron involvement in early motoneuron disease: a diffusion tensor imaging study. *Brain*, *134*(4), 1211–1228. https://doi.org/10.1093/brain/awr016

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K. & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: an overview. *NeuroImage*, *80*(5),

62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

Verstraete, E., Turner, M. R., Grosskreutz, J., Filippi, M. & Benatar, M. (2015). Mind the gap: The mismatch between clinical and imaging metrics in ALS. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *16*(7–8), 524–529. https://doi.org/10.3109/21678421.2015.1051989

Vogel, T. (1972). Significance of pneumencephalography in the diagnosis of traumatic brain lesions. *Der Nervenarzt*, *43*(8), 412–417. http://www.ncbi.nlm.nih.gov/pubmed/5069865

Wallner, S., Peters, S., Pitzer, C., Resch, H., Bogdahn, U. & Schneider, A. (2015). The Granulocyte-colony stimulating factor has a dual role in neuronal and vascular plasticity. *Frontiers in Cell and Developmental Biology*, *3*, 48. https://doi.org/10.3389/fcell.2015.00048

Wang, S., Zhang, Y., Liu, G., Phillips, P. & Yuan, T.-F. (2016). Detection of Alzheimer's Disease by Three-Dimensional Displacement Field Estimation in Structural Magnetic Resonance Imaging. *Journal of Alzheimer's Disease : JAD*, *50*(1), 233–248. https://doi.org/10.3233/JAD-150848

Wang, X., Gerken, M., Dennis, M., Mooney, R., Kane, J., Khuder, S., Xie, H., Bauer, W., Apkarian, A. V. & Wall, J. (2010). Profiles of Precentral and Postcentral Cortical Mean Thicknesses in Individual Subjects over Acute and Subacute Time-Scales. *Cerebral Cortex*, *20*(7), 1513–1522. https://doi.org/10.1093/cercor/bhp226

Way, C., Pettersson, D. & Hiller, A. (2019). The 'Hot Cross Bun' Sign Is Not Always Multiple System Atrophy: Etiologies of 11 Cases. *Journal of Movement Disorders*, *12*(1), 27–30. https://doi.org/10.14802/jmd.18031

Welsh, R. C., Jelsone-Swain, L. M. & Foerster, B. R. (2013). The utility of independent component analysis and machine learning in the identification of the amyotrophic lateral sclerosis diseased brain. *Frontiers in Human Neuroscience*, *7*, 251. https://doi.org/10.3389/fnhum.2013.00251

Westeneng, H.-J., Debray, T. P. A., Visser, A. E., van Eijk, R. P. A., Rooney, J. P. K., Calvo, A., Martin, S., McDermott, C. J., Thompson, A. G., Pinto, S., Kobeleva, X., Rosenbohm, A., Stubendorff, B., Sommer, H., Middelkoop, B. M., Dekker, A. M., van Vugt, J. J. F. A., van Rheenen, W., Vajda, A., … van den Berg, L. H. (2018). Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *The Lancet Neurology*, *17*(5), 423–433. https://doi.org/10.1016/S1474-4422(18)30089-9

Westeneng, H.-J., Verstraete, E., Walhout, R., Schmidt, R., Hendrikse, J., Veldink, J. H., van den Heuvel, M. P. & van den Berg, L. H. (2015). Subcortical structures in amyotrophic lateral sclerosis. *Neurobiology of Aging*, *36*(2), 1075–1082. https://doi.org/10.1016/j.neurobiolaging.2014.09.002

Westeneng, H.-J., Walhout, R., Straathof, M., Schmidt, R., Hendrikse, J., Veldink, J. H., van den Heuvel, M. P. & van den Berg, L. H. (2016). Widespread structural brain involvement in ALS is not limited to the C9orf72 repeat expansion. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(12), 1354–1360. https://doi.org/10.1136/jnnp-2016-313959

Westfall, P. H., Young, S. S. & Wright, S. P. (1993). On Adjusting P-Values for Multiplicity. *Biometrics*, *49*(3), 941. https://doi.org/10.2307/2532216

Whitwell, J. L., Weigand, S. D., Boeve, B. F., Senjem, M. L., Gunter, J. L., DeJesus-Hernandez, M., Rutherford, N. J., Baker, M., Knopman, D. S., Wszolek, Z. K., Parisi, J. E., Dickson, D. W., Petersen, R. C., Rademakers, R., Jack, C. R. & Josephs, K. A. (2012). Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. *Brain*, *135*(3), 794–806. https://doi.org/10.1093/brain/aws001

Whitwell, Jennifer L., Jack, C. R., Parisi, J. E., Knopman, D. S., Boeve, B. F., Petersen, R. C., Dickson, D. W. & Josephs, K. A. (2011). Imaging Signatures of Molecular Pathology in Behavioral Variant Frontotemporal Dementia. *Journal of Molecular Neuroscience*, *45*(3), 372. https://doi.org/10.1007/s12031-011-9533-3

Whitwell, Jennifer L., Josephs, K. A., Rossor, M. N., Stevens, J. M., Revesz, T., Holton, J. L., Al-Sarraj, S., Godbolt, A. K., Fox, N. C. & Warren, J. D. (2005). Magnetic Resonance Imaging Signatures of Tissue Pathology in Frontotemporal Dementia. *Archives of Neurology*, *62*(9), 1402. https://doi.org/10.1001/archneur.62.9.1402

WHO. (2001). *International classification of functioning, disability and health*. World Health Organization.

WHO. (2021). *International statistical classification of diseases and related health problems, tenth revision* (W. H. Organization (Ed.); 2nd editio). https://www.cdc.gov/nchs/icd/icd10cm.htm

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., Duggirala, R. & Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *NeuroImage*, *53*(3), 1135–1146. https://doi.org/10.1016/j.neuroimage.2009.12.028

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060

Winklewski, P. J., Sabisz, A., Naumczyk, P., Jodzio, K., Szurowska, E. & Szarmach, A. (2018). Understanding the Physiopathology Behind Axial and Radial Diffusivity Changes—What Do We Know? *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.00092

Wirth, A. M., Khomenko, A., Baldaranov, D., Kobor, I., Hsam, O., Grimm, T., Johannesen, S., Bruun, T.-H., Schulte-Mattler, W., Greenlee, M. W. & Bogdahn, U. (2018). Combinatory Biomarker Use of Cortical Thickness, MUNIX, and ALSFRS-R at Baseline and in Longitudinal Courses of Individual Patients With Amyotrophic Lateral Sclerosis. *Frontiers in Neurology*, *9*, 614. https://doi.org/10.3389/fneur.2018.00614

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R. & others. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysio*, *106*, 1125–1165. https://doi.org/jn.00338.2011 [pii]\r10.1152/jn.00338.2011

Younes, K., Lepow, L. A., Estrada, C. & Schulz, P. E. (2018). Auto-antibodies against P/Q- and N-type voltage-dependent calcium channels mimicking frontotemporal dementia. *SAGE Open Medical Case Reports*, *6*, 2050313X1775092. https://doi.org/10.1177/2050313X17750928

Young, P. N. E., Estarellas, M., Coomans, E., Srikrishna, M., Beaumont, H., Maass, A., Venkataraman, A. V., Lissaman, R., Jiménez, D., Betts, M. J., McGlinchey, E., Berron, D., O'Connor, A., Fox, N. C., Pereira, J. B., Jagust, W., Carter, S. F., Paterson, R. W. & Schöll, M. (2020). Imaging biomarkers in neurodegeneration: Current and future practices. *Alzheimer's Research and Therapy*, *12*(1), 1–17. https://doi.org/10.1186/s13195-020-00612-7

Yunusova, Y., Plowman, E. K., Green, J. R., Barnett, C. & Bede, P. (2019). Clinical Measures of Bulbar Dysfunction in ALS. *Frontiers in Neurology*, *10*. https://doi.org/10.3389/fneur.2019.00106

Zarei, M., Ibarretxe-Bilbao, N., Compta, Y., Hough, M., Junque, C., Bargallo, N., Tolosa, E. & Martí, M. J. (2013). Cortical thinning is associated with disease stages and dementia in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *84*(8), 875–881. https://doi.org/10.1136/jnnp-2012-304126

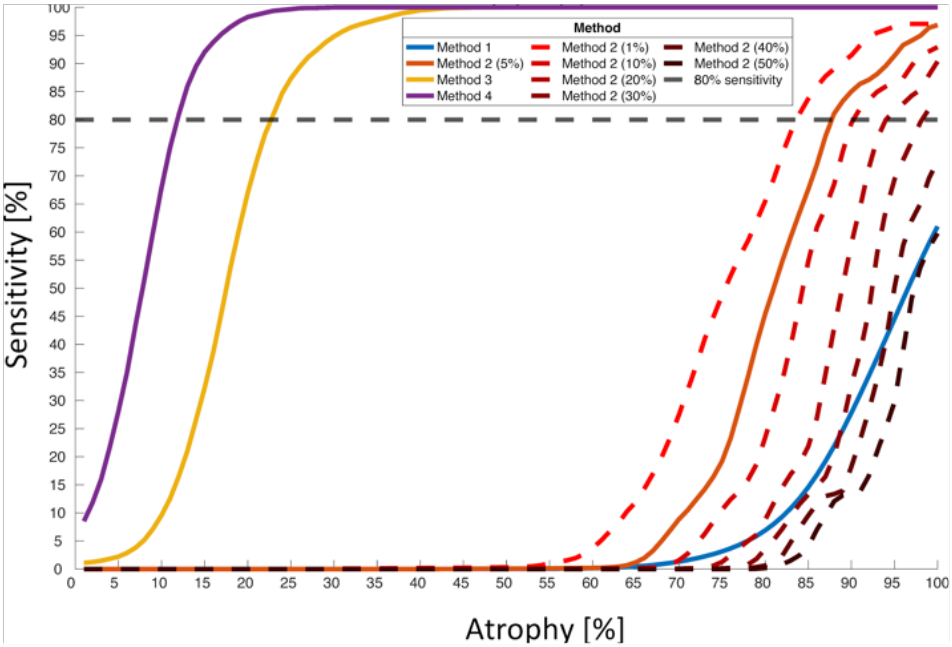
Zhou, J., Greicius, M. D., Gennatas, E. D., Growdon, M. E., Jang, J. Y., Rabinovici, G. D., Kramer, J. H., Weiner, M., Miller, B. L. & Seeley, W. W. (2010). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain : A Journal of Neurology*, *133*(Pt 5), 1352–1367. https://doi.org/10.1093/brain/awq075

Zilles, K., Armstrong, E., Schleicher, A. & Kretschmann, H. J. (1988). The human pattern of gyrification in the cerebral cortex. *Anatomy and Embryology*, *179*(2), 173–179. https://doi.org/10.1007/BF00304699

Zoccolella, S., Beghi, E., Palagano, G., Fraddosio, A., Samarelli, V., Lamberti, P., Lepore, V., Serlenga, L., Logroscino, G. & SLAP registry. (2006). Predictors of delay in the diagnosis and clinical trial entry of amyotrophic lateral sclerosis patients: a population-based study. *Journal of the Neurological Sciences*, *250*(1–2), 45–49. https://doi.org/10.1016/j.jns.2006.06.027

# 10 Appendix

## 10.1 Supplementary Figure 1




Supplementary Figure 1. Cumulative sensitivity relative to the degree of simulated atrophy (across vertices/brain regions), comparison between the four tested methods and different thresholds for method 2. In method 2, a label was defined "atrophic" if a certain percentage of its vertices yielded $p_{FWER} < 0.05$. Here, the results for thresholds 1%, 5% (which is shown in the main text), 10%, 20%, 30%, 40% and 50% are displayed.

## 10.2 Supplementary Table 1

Supplementary Table 1. Cumulative sensitivity calculations for different thresholds for method 2 (in method 2, a label was defined "atrophic" if a certain percentage of each label's vertices yielded $p_{FWER} < 0.05$).

| | 1% | 5% | 10% | 20% | 30% | 40% | 50% |
|---|---|---|---|---|---|---|---|
| **Degree of atrophy required for detection of atrophy in 80% of cases (sensitivity)*** | 84% | 88 % | 90% | 94% | 98% | n.a. | n.a. |

** Note that lower values of atrophy suggest more sensitive methods, since they detect less pronounced atrophy. Abbreviations: n.a. = not available*

## 10.3 Supplementary Table 2

Supplementary Table 2. For methods 2,3 and 4, cumulative sensitivity was defined based on the degree of simulated atrophy a method required to sensitively detect 80% (method 2: 88%, method 3: 23%, method 4: 12%). However, regional sensitivity varied for that degree of atrophy. This table indicates which labels showed < 80% sensitivity for each method's "crucial" degree of atrophy, along with the regional sensitivity detected for that degree of atrophy.

| Method 2 (88 % simulated atrophy) | | Method 3 (23% simulated atrophy) | | Method 4 (12% simulated atrophy) | |
|---|---|---|---|---|---|
| Labelname | Regional sensitivity | Labelname | Regional sensitivity | Labelname | Regional sensitivity |
| L_cuneus | 75.22% | L_caudalanteriorcingulate | 0.7965 | L_caudalanteriorcingulate | 0.6903 |
| L_entorhinal | 40.71% | L_entorhinal | 0.4867 | L_cuneus | 0.7434 |
| L_inferiortemporal | 0.00% | L_lingual | 0.6549 | L_entorhinal | 0.5664 |
| L_lingual | 0.00% | L_parahippocampal | 0.4779 | L_isthmuscingulate | 0.7965 |
| L_middletemporal | 9.73% | L_parstriangularis | 0.6195 | L_lingual | 0.6726 |
| L_parahippocampal | 2.65% | L_rostralanteriorcingulate | 0.4956 | L_parahippocampal | 0.3540 |
| L_pericalcarine | 0.00% | L_frontalpole | 0.0973 | L_paracental | 0.7788 |
| L_rostralanteriorcingulate | 20.35% | L_temporalpole | 0.2389 | L_parsorbitalis | 0.7788 |
| L_temporalpole | 0.00% | L_transversetemporal | 0.0177 | L_pericalcarine | 0.6549 |
| R_inferiortemporal | 0.00% | R_bankssts | 0.6460 | L_rostralanteriorcingulate | 0.7788 |
| R_lingual | 0.00% | R_caudalmiddlefrontal | 0.5310 | L_frontalpole | 0.5487 |
| R_parahippocampal | 0.88% | R_entorhinal | 0.6195 | L_temporalpole | 0.5664 |
| R_pericalcarine | 0.00% | R_parsorbitalis | 0.2743 | L_transversetemporal | 0.7080 |
| R_rostralanteriorcigulate | 76.99% | R_parstriangularis | 0.7788 | R_caudalanteriorcingulate | 0.7522 |
| R_frontalpole | 76.11% | R_rostralanteriorcigulate | 0.2478 | R_cuneus | 0.6726 |
| R_temporalpole | 44.25% | R_frontalpole | 0.2920 | R_entorhinal' | 0.5310 |
| | | R_temporalpole | 0.0619 | R_isthmuscingulate | 0.7788 |
| | | R_transversetemporal | 0.2389 | R_lingual | 0.7522 |
| | | | | R_parahippocampal' | 0.4867 |
| | | | | R_parsorbitalis | 0.7788 |
| | | | | R_pericalcarine | 0.5752 |
| | | | | R_rostralanteriorcigulate | 0.6283 |
| | | | | R_frontalpole | 0.4956 |
| | | | | R_temporalpole | 0.3540 |
| | | | | R_transversetemporal | 0.7080 |

*Abbreviations: R = right, L = left*