







Master Thesis

M3 Technical Medicine Medical Imaging and Interventions

C.J.A. Tenbergen sl104543

Nuclear Medicine Department Leiden University Medical Centre

January – December 2016

Committee

Chairman Medical supervisor Technological supervisor UT Process supervisor External member Additional member Prof. dr. ir. C.H. Slump Drs. C.H.J.M. Ticheler Prof. dr. ir. C.H. Slump Drs. P.A. van Katwijk Dr. E.J.F.M. ten Berge Dr. O. Dzyubachyk

UNIVERSITY OF TWENTE.



Abstract

Background

Amyloid imaging by using 18F-florbetapir PET scans has increased the knowledge on amyloid plaques in relation to Alzheimer's disease. Besides the diagnostic value, these scans can have a potential therapeutic relevance, particularly with respect to the recently developed therapeutics aimed at modulating A β deposition. The group of MCI patients could benefit most from an early diagnosis, with regards to their possible prospective treatment. Since not all MCI progresses to AD, correct identification of these MCI patients and patients with MCI that will progress to AD is crucial. The amyloid PET scans should be assessed by a method that can provide quantitative feedback about the intervention, can measure change in amyloid burden and highlight mild effects overlooked by visual comparison.

Method

By combining information from both amyloid PET scans and individual MRI scans of 108 patients from the ADNI database, specific grey matter amyloid burden was measured. With the use of a white matter reference region, results can be compared between patients. By linking the amyloid measurements to the subject's diagnosis known in ADNI, a model was created that gives a prediction of the prospects of MCI patients.

Results

The main finding of this research is that the developed method is able to discriminate MCI patients that will convert to AD from MCI patients who will not. By a combination of multiple measures this result was achieved. The created model showed high predictive power in subjects converting from MCI to either AD or HC.

Conclusion

The developed workflow facilitates early diagnosis which yields eligible patients for drug treatment and is a start of a tool for accurately quantifying and tracking of treatment effects.

Preface

This master thesis focuses on amyloid-PET imaging and describes the development of a method to quantify the amyloid burden imaged by a specific PET tracer. The structure of this master thesis is built around chapter 2. This chapter gives a concise description of the complete research process. The preceding background in chapter 1 gives additional information on amyloid imaging and short highlights of previous research on this topic leading to the current research design. The methodology supplements in chapter 3 and 4 give additional information on methodological choices made in the paper and give an explanation of results and consequences in the design process.

Contents

Chapter 1. Background	6
Chapter 2.	9
Abstract	9
Introduction	10
Method	11
Results	В
Method 2	17
Results 2	
Discussion	20
Acknowledgements	22
Chapter 3. Methodology supplement 1	23
Image selection	
Image processing	25
White matter reference region	27
Chapter 4. Methodology supplement 2	
Statistical analysis	30
References	32

Chapter 1. Background

Modern dementia research began with the case description of Alois Alzheimer in 1906¹ mentioning the symptoms of a 51-year old patient with specific histomorphological findings. Alzheimer gives the first description of senile plaques composed by amyloid peptides (β -amyloid plaques) and neurofibrillary tangles of aggregated tau protein, processes assumed in following research to be causing the cognitive decline in Alzheimer's disease (AD). These processes lead to loss of synaptic function and eventually to neuronal cell death. These pathologic findings in the brain could for long only be encountered during autopsy, with the diagnosis being definitive only post-mortem².

With increasing knowledge about the disease pathology and progression, the AD diagnosis can be made reliably during life. Current diagnosing is based on neuro-psychological tests, including a mini-mental state examination which tests a wide range of cognitive functions and a clock-drawing test for examination of executive functioning³.

Next to neuro-psychological evaluation, there is an increasing number of tests that can be performed for diagnosing ante-mortem and in early phases of the disease. These biomarkers of AD include MRI, FDG PET scans and the detection of amyloid plaques and tau tangles in cerebrospinal fluid. In 2004, Klunk et al. introduced "Pittsburgh Compound-B" (IIC-PiB), the first PET amyloid tracer giving rise to the option of detecting amyloid plaques in the human brain by life⁴. With the half-life of IIC being 20 minutes, the tracer is not very suitable for clinical practice. Following tracers developed for this purpose are therefor all based on 18F, providing a half-life of 110 minutes, such as florbetaben, florbetapir and flutemetamol⁵.

With the ability to detect the amyloid plaques in vivo, the knowledge about these plaques and the development of the disease increases. The finding that these plaques are present in the brain long before clinical symptoms appear, has led to the formation of the amyloid hypothesis. This states that the excessive formation and deposition of insoluble fibrillary amyloid with the consequent aggregation in plaques is the primary event in AD pathogenesis⁶. The poor correlation between amyloid plaques and severity of AD and the role of the soluble non-fibrillary forms of amyloid are poorly understood^{6,7}. Though the AD neuropathology is not completely unraveled yet, the importance of amyloid plaques is universally recognized⁵.

Figure 1.1 displays a hypothetical time course of Alzheimer's disease biomarkers. Amyloid deposition, both detected in spinal fluid and with amyloid PET scan, is the first marker passing the detection threshold. This deposition induces acceleration of neurofibrillary tangles to be formed, tauopathy, and tau in central spinal fluid is the next biomarker to be detected. Finally, cognitive impairment becomes evident with a range of cognitive responses and presentations that depend on the risk profile of the patient. The region below the detection threshold shows the still unclear relation between amyloid plaques and tau tangles; which deposition occurs first and to what extent the deposition influence each other are still unanswered questions. This model show one possibility whereby tau pathology precedes the amyloid (A β) deposition in time, but only early on at a subthreshold biomarker detection threshold⁸.



Figure 1.1: from Jack et al.⁸ Model integrating Alzheimer's disease immunohistology and biomarkers. The threshold for biomarker detection of pathophysiological changes is denoted by the black horizontal line.

The visualization of amyloid plaques in the brain with the use of PET scans has led to increased knowledge about these plaques. These scans show that amyloid tracers selectively bind to amyloid aggregates found in the grey matter of human brain while there is nonspecific binding of the tracers to the white matter⁹⁻¹¹. Amyloid-PET-CT does not contain information about the border between grey and white matter. Visual qualification shown in Figure 1.2, which is based on grey-white matter contrast¹², therefore seems inadequate.



Figure 1.2: Typical negative and positive Florbetapir scans^{12,13}. Dotted lines on the negative scan (left) indicate the edge of cortical grey matter (outer line) and grey-white border (inner line). On this scan tracer uptake is clearly more intense in the white matter compared to the grey matter. On the positive scan (right), this contrast is reduced: grey matter uptake is similar to white matter uptake and the grey-white border cannot be identified.

Quantification of these scans is needed in order to surpass the positive or negative assessment and be able to specify localisation of plaques, quantify deposition progress and compare scans in time and between subjects. Quantifiable amyloid PET scans have no role in discriminating Alzheimer's disease patients from healthy subjects, since this can be done clinically. The technique should be used to indicate more subtle changes within the group of patients with mild cognitive impairment (MCI).

The clinical diagnosis of MCI is troublesome. It is known that MCI may arise from several causes, including AD and other forms of dementia, as well as depression or various physical disorders¹⁴. The diagnosis of MCI therefore does not give insight into a patient's prognosis or possible curative treatments. Early identification of the underlying cause is favourable, with the ability to predict progression to AD in the patient's interest.

Previous research into MCI amyloid burden with the tracer PiB shows a distinct time pattern of amyloid burden. Koivunen et al.¹⁵ found PiB uptake to increase during follow-up in MCI patients, while in MCI patients converted to AD the amyloid burden did not change much from the initial high level. Okello et al.¹⁶ found similar results, with MCI fast converters reaching an amyloid load plateau while slow or non-converters did not. This finding was also seen by Villemagne et al.¹⁷, stating that longitudinal assessment of amyloid burden in patients with AD showed that as AD progresses, the deposition slows towards a plateau.

This slow amyloid accumulation rate, together with the start of accumulation many years before cognitive changes, potentially yields a wide time window for intervention with anti-amyloid therapy. MCI subjects who will develop AD can be selected and, before irreversible damage to the brain is done, they can be included into therapy trials. Currently there are some symptomatic treatments that do not decelerate or prevent progression of the disease but do show modest benefit for cognition and functional ability. Disease-modifying treatments are on a rise, aimed at amyloid pathology, tauopathy or for example oxidative stress reduction^{18,19}.

While drug research is continuing, in the next chapter a method is presented to assess amyloid deposition and predict MCI prognosis. The developed workflow facilitates early diagnosis which yields eligible patients for drug treatment and is a start of a tool for accurately quantifying and tracking of treatment effects.

Chapter 2.

Quantitative amyloid-PET analysis in mild cognitive impaired patients

C.J.A. Tenbergen, for the Alzheimer's Disease Neuroimaging Initiative*

*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:

http://adni.loni.usc.edu/wp-ontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Background

Amyloid imaging by using 18F-florbetapir PET scans has increased the knowledge on amyloid plaques in relation to Alzheimer's disease. Besides the diagnostic value, these scans can have a potential therapeutic relevance, particularly with respect to the recently developed therapeutics aimed at modulating A β deposition. The group of MCI patients could benefit most from an early diagnosis, with regards to their possible prospective treatment. Since not all MCI progresses to AD, correct identification of these MCI patients and patients with MCI that will progress to AD is crucial. The amyloid PET scans should be assessed by a method that can provide quantitative feedback about the intervention, can measure change in amyloid burden and highlight mild effects overlooked by visual comparison.

Method

By combining information from both amyloid PET scans and individual MRI scans of 108 patients from the ADNI database, specific grey matter amyloid burden was measured. With the use of a white matter reference region, results can be compared between patients. By linking the amyloid measurements to the subject's diagnosis known in ADNI, a model was created that gives a prediction of the prospects of MCI patients.

Results

The main finding of this research is that the developed method is able to discriminate MCI patients that will convert to AD from MCI patients who will not. By a combination of multiple measures this result was achieved. The created model showed high predictive power in subjects converting from MCI to either AD or HC.

Conclusion

The developed workflow facilitates early diagnosis which yields eligible patients for drug treatment and is a start of a tool for accurately quantifying and tracking of treatment effects.

Introduction

Alzheimer's disease (AD) is the most common cause of dementia, showing cognitive decline in patients followed by interference with day to day functioning²⁰. With an estimation of 46.8 million patients with dementia worldwide in 2015 and the prediction of a duplication of this number every 20 years due to the ageing population^{21,22}, early diagnosis is crucial in counteracting the rise of the disease.

While the etiology of AD remains controversial, the importance of amyloid plaques is universally recognized⁵. The amyloid hypothesis states that the excessive formation and deposition of insoluble fibrillary β -amyloid (A β) with the consequent aggregation in plaques is the initiating factor in AD pathogenesis⁶. Progress in research and molecular imaging in AD has enabled detecting human brain amyloid deposition during life. By using positron emission tomography (PET), and latest developed radio ligands based on 18F, such as 18F-AV45, the amyloid imaging is possible in every clinic with a PET scanner without the need for an onsite cyclotron.

The visualization of amyloid plaques in the brain with the use of PET scans has led to increased knowledge about these plaques. These scans show that amyloid tracers selectively bind to amyloid aggregates, found in the grey matter of the human brain, while there is nonspecific binding of the tracers to the white matter⁹⁻¹¹. This may effect amyloid burden assessment when using whole-brain analyses and therefore focus should be on amyloid binding in grey matter only. However, amyloid-PET-CT does not contain the information to accurately indicate the border between grey and white matter. Visual qualification into positive and negative scans, which is based on grey-white matter contrast¹², therefore seems insufficient. Visual assessment shows adequate intra- and inter-reader agreement²³⁻²⁷, after dedicated training of the nuclear physicians, for use in routine clinical setting. Longitudinal clinical trials of amyloid progression on the other hand require a quantitative approach.

This quantification can be achieved by a PET template based approach or by adding magnetic resonance imaging (MRI) information to the process, either from individual MRI scans or from an MRI template. The advantage of individual MRI is the ability to precisely distinct grey from white matter and this information enriches the amyloid imaging. This might be particularly important in patients with cortical atrophy. Next to visually rating scans as positive or negative for amyloid pathology, additional quantification can therefore be performed of overall or local amyloid burden. The addition of MRI data has proven its value: while the PET template approach seems adequate for clinical diagnostic purposes, the MRI based analysis is more appropriate for research purposes because of its greater intergroup differences between healthy controls and subjects with Alzheimer's disease^{28,29}.

The use of the amyloid imaging biomarker has shown its diagnostic value in recent studies, elevated brain amyloid presence is associated with cognitive decline in healthy and mild cognitive impaired (MCI) subjects³⁰ and a negative amyloid PET scan in demented patients can exclude AD^{31} . Besides the diagnostic value, it can have a potential therapeutic relevance, particularly with respect to the recently developed therapeutics aimed at modulating A β deposition. The effects of these new treatments should be assessed by a method that can provide quantitative feedback about the intervention, can measure change in amyloid burden and highlight mild effects overlooked by visual comparison^{28,32}.

The group of MCI patients could benefit most from an early diagnosis, with regards to their possible prospective treatment. Since not all MCI progresses to AD, correct identification of these MCI patients and patients with MCI that will progress to AD is crucial³³.

Goal of this study is to design an automated workflow for 18F-Florbetapir (18F-AV45) amyloid PET imaging, wherein, with the use of individual MRI scans, an individual grey matter mask is used to facilitate prognosis based on amyloid burden in MCI patients.

Method

For the first part of this study the focus was on designing and optimising the workflow of segmentation of the MRI scans, coregistration of amyloid PET and MRI scans and quantification of the grey matter amyloid signal. First tests of this workflow were aimed at detecting the difference between AD patients and healthy controls. MCI patients were later on added as test of ability to detect finer differences between groups.

Alzheimer's Disease Neuroimaging Initiative

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). 59 acquisition centres are located across the United States and Canada.³⁴

18F-Florbetapir scanning is added to the imaging protocol in the second and third phase of ADNI (ADNI-GO and ADNI-2). Diagnosis in ADNI is made primarily based on clinical, cognitive and functional tests. Imaging is typically not used for diagnostics in AD³⁴. We use the diagnosis in ADNI as standard, establishing our subject groups and correlating predictive results.

Participants

First workflow developments were done with ten subjects form one centre, in order to minimize variability existing between subjects. Supplemented with subjects from two other centres with the same PET and MRI settings, a total of 21 Alzheimer's disease and 18 healthy control (HC) subjects were selected. These patients showed no diagnosis change during follow-up. Main inclusion criteria of subjects were availability of a 3T MRI scan and 18F-AV45 PET scan, with intervals less than three months.

Next to AD and HC subjects, patients indicated as having MCI are available in ADNI. Since subjects in ADNI undergo regular follow-up visits, we can track their state and corresponding signal changes. Next to subjects diagnosed as MCI during the entire follow-up period, subjects who change from MCI to HC and MCI to AD can be selected.

Scans from these three groups of subjects were selected. In case of the converted patients, scans were selected from the period before diagnosis change. For each diagnostic group, a total of 23 patients were selected from a total of 14 different centres. Within each group centres were equally represented, in order to minimise variability.

MRI and amyloid PET scans

The ADNI provides protocols for all image acquisition and for all specific scanners. The most important general protocol features are highlighted below. The complete description of the acquisition protocol can be found at the ADNI website³⁴.

MRI scans were performed with 3T MRI scanners using a 3D TI-weighted sequence with slice thickness 1.2 mm. The ADNI available magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence was chosen in order to make use of the contrast between grey and white matter.

After the intravenous administration of AV-45 (356,0 \pm 27.3 MBq), PET brain imaging was performed continuously for 20 minutes starting at 50 minutes post injection. Scans are acquired in four 5-minutes frames and corrected for attenuation and scatter by scanner-specific protocols. ADNI post-processing step 3 is used, whereby frames are coregistered, averaged into one static frame and finally standardized with regards to image and voxel size.

Visual inspection of MR images showed some scans to appear blurry due to movement. Movement within the MR image was therefore used as an exclusion factor. Wrap-around artefacts were seen in multiple patients, but assumed not troublesome if limited to skin overlapping. Visual inspection of PET images was performed as an initial exploration of the differences of amyloid pattern between AD and HC subjects.

Image processing

The FMRIB Software Library (FSL)³⁵ written by the Analysis group, FMRIB Oxford UK, was the software of choice specialized in brain image processing. PET scans were registered to MRI space, resulting transformation matrix was inverted and later on used to map MRI masks to PET space. The image processing of MRI started with brain extraction, followed by segmenting MRI using FAST and FLIRT³⁵. From the resulting grey matter probability map, a 0.5 threshold was used to create a binary grey matter mask. After mapping this MRI mask to PET, this grey matter mask was applied to the PET scan and PET signal intensities were extracted. In order to distinguish brain lobes and even smaller distinct anatomical regions, the anatomical Hammers atlas is used^{36,37}. This atlas was also registered to the PET image, by combining transformation matrices from registering PET to MRI and atlas to MRI.

Whole brain grey matter (GM) and white matter (WM) mean signal intensity and volume were measured with the latter expressed as number of non-zero pixels. Mean intensity was also calculated for the grey matter of the lobes and precuneus, and frontal and parietal lobes further divided in smaller atlas areas.

It requires normalization of regional PET activity to a reference tissue to correct for nonspecific radiotracer binding and other variations between subjects. Reference tissue was chosen to be white brain matter, as that yielded an improved accuracy in longitudinal PET data over the use of cerebellum³⁸⁻⁴⁰. Four different white matter based areas were compared in order to assess the effects of including cerebellar white matter pixels and of partial volume effects in the PET images in white matter pixels directly adjacent to grey matter. The four regions (white matter, eroded white matter, white matter without cerebellum, eroded white matter without cerebellum) were segmented on MRI and applied as binary masks to the PET. Results show that all four WM reference regions used in GM over WM ratio perform equally well; all four regions resulted in a p<0.001 for discriminating AD from HC subjects by

measuring mean intensity. The whole white matter region is chosen to use as reference in all following tests, as this largest area yields highest statistical power.

Ratio of whole brain grey matter over total white matter were noted as GM/WMmean and GM/WMvol, with additional lobe indication for the lobular grey matter over white matter ratios. Target-to-WM ratios of 13 smaller bilateral atlas areas were named after the atlas number of the right hemispherical part of the area.

Statistical analysis

As this study is a first exploration and tests the feasibility of the applied method, no power analyses were performed. Focus was put on selecting the most suitable subjects and keeping variation low by minimizing the number of used centres.

Unpaired t-tests were used for demographics analysis and initial two group comparison of AD and HC subjects. For the comparison of all five groups, ANOVA with post-hoc tests gives the information on significant difference between these groups. This test requires residuals to be normally distributed. IBM SPSS statistics version 23 was used to perform all statistical analyses⁴¹. p-Values less than 0.05 were considered as statistical significant.

Results

Demographic data

Subject demographics are given in Table 2.1. Age was recorded at the moment of the MRI scan. Since time between PET and MRI scan was one of the inclusion criteria, small differences are expected between subject groups. Healthy subjects often had their scans planned on the same day, more so than the other subjects. Follow-up time was noted to indicate stability of diagnosis, defined as time between initial base line visit and last recorded patient visit in the ADNI database. Alzheimer's disease patients had a significantly smaller follow-up period than healthy controls, possibly because of fast disease progression.

Table 2.1: Demographics of healthy controls (HC), patients suffering from Alzheimer's disease (AD),
mild cognitive impaired patients (MCI), and MCI patients converted to HC (MCI \rightarrow HC) or AD
(MCI→AD). *p<0.05, **p<0.01, ***p<0.001 compared to HC

		-			
	HC	MCI	AD	MCI→AD	MCI→HC
Number of subjects	18	23	21	23	23
Age (years ± SD)	76 (6.6)	73 (7.8)	72 (6.5)	74 (8.1)	68 (6.3)**
Gender (f/m)	11 (61%)/	6 (26%) /	15 (71%)/	6 (26%) /	15 (65%) /
	7 (39%)	17 (74%)*	6 (29%)	17 (74%)*	8 (35%)
Inter scan time	9 (17)	24 (25)*	28 (17)**	22 (27)	27 (17)**
$(days \pm SD)$					
Follow-up (months	45 (25.7)	75 (32.8)	16 (6.2)***	-	-
± SD)					

Whereas GM mean intensity and GM volume (measured as number of non-zero pixels) show subtle discrimination between HC and AD subjects, and WM mean and volume do not show discrimination at all, the ratio of these two areas significantly discriminates between HC and AD subjects, see Table 2.2.

Table 2.2: Whole brain measurements of mean signal and volume measured as number of non-zero
pixels. Discrimination is made between grey matter (GM), white matter (WM) and the ratio of these
two regions: grey matter over white matter (GM/WM).

	GM	GM	WM	WM	GM/WM	GM/WM
	mean	volume	mean	volume	mean	volume
HC subjects	3077	261423	4492	244562	0.6768	1.0723
(mean ± SD)	(1480)	(25178)	(1812)	(23261)	(0.0834)	(0.0876)
AD subjects	4050	240791	4940	244585	0.8199	0.9911
(mean ± SD)	(1107)	(29959)	(1335)	(35173)	(0.0373)	(0.0998)
t statistic	2.346	-2.305	.887	.002	6.731	-2.678
p-value	.024	.027	.381	.998	<.001	.011

This ratio of GM mean over WM mean was also successfully discriminative between AD and HC at lobe level. The overall GM mean was substituted by the mean in the GM of each lobe as indicated by the atlas. Next to the four lobes, the precuneus was as well segmented. Initial visual interpretation resulted in the finding that in frontal and precuneus area, high intensities were seen in AD patients and therefore this area is segmented separately. All lobes and precuneus showed significant differences in mean signal between AD and HC subjects, presented in Table 2.3.

Table 2.3: Lobular measurements of mean intensity of grey matter over white matter. Differences between AD and HC subjects are tested.

	Frontal	Occipital	Parietal	Temporal	Precuneus
	GM/WM	GM/WM	GM/WM	GM/WM	GM/WM
	mean	mean	mean	mean	mean
HC subjects	0.7210	0.7336	0.7284	0.6893	0.7287
(mean ± SD)	(0.0955)	(0.1268)	(0.1062)	(0.0936)	(0.1219)
AD subjects	0.9931	0.8519	0.9199	0.8234	0.9379
(mean ± SD)	(0.0489)	(0.0814)	(0.0532)	(0.0470)	(0.0648)
t statistic	8.510	3.518	6.940	5.514	6.530
p-value	<.001	.001	<.001	<.001	<.001

With these lobes to show significant differences as good as the mean signal from whole brain, this can be interpreted as supporting the thought of the mean GM truly representing the whole brain.

Since image processing and chosen measures were successful, a new selection of patients with more subtle difference was added to the test. Results of comparison between the total of five groups is displayed in Figures 2.1 and 2.2.



Figure 2.1: Whole brain measurements, expressed as ratios of mean and volume of grey matter over white matter. In panel A, the results of the HC and AD groups are shown. Panel B shows the results of the three groups of MCI patients.

Figure 2.1 displays the whole brain measurements of each patient. While the mean intensity of the white matter is always larger than the grey matter resulting in ratios smaller than 1.0, the volume ratio ranges from 0.6 to 1.35. The aforementioned discrimination between HC subjects in blue and AD subjects in yellow by whole brain measurements is visualized. The MCI to HC group (green), comparable to the group of HC subjects, shows low mean intensity ratio and high volume ratio, indicating relatively large grey matter volumes. The MCI to AD group (orange), similar to the AD group, on the other hand shows higher mean intensity ratios and lower volume ratios in comparison to the other groups. The lower volume ratios indicate smaller grey matter volumes, possibly explained by grey matter atrophy as a late phase in Alzheimer's disease progression. The MCI group (purple stars) shows large variation, both in mean signal intensity and in volume measurements, indicating the heterogeneity of this group.



Figure 2.2: Lobe measurements expressed in grey matter over white matter ratio for each group separately. Outliers are indicated. One patient of the MCI \rightarrow HC group is accountable for 3 outlier points and most outliers are found in the occipital lobe measurement.

Figure 2.2 shows the mean intensity ratios for all groups measured in the lobes and precuneus area. Again, one can see the difference between the HC and MCI to HC group on one hand and the AD and MCI to AD group on the other hand. Note that the scans from before conversion already clearly indicate the HC or AD amyloid characteristics while the clinical presentation is still unchanged at MCI. Every region of interest in the patient's brain shows the trend of increasing ratios when comparing the HC side to the AD side of the Figure. Only the temporal lobe is less explicitly discriminative and shows minor increase in the AD subject group. The MCI group shows intermediate values and a large variance in outcomes. The AD group on the other hand shows small variance in all but one (occipital lobe) measures.

Testing these measurements between all five groups for statistical significance required post-hoc tests to be performed, with the results shown in Table 2.4. Significant differences are seen in all measures acquired when comparing MCI converted to HC with AD subjects and with MCI→AD subjects and when comparing MCI→AD with HC subjects. Also discriminating AD subjects from HC and MCI subjects and MCI subjects converted to AD from MCI subjects can be done with significant differences. It is not surprising that the comparison of MCI converted to HC with HC subjects yields the lowest number of measures showing a significant difference. Results show that not every measurement results in significant differences between all groups. The distinction between all groups can therefore be made by combinations of measures.

Table 2.4: Post-hoc test results, showing differences between all groups. Green: all or all except one measure is significant, yellow: at least one measure yields significant results. Numbers refer to bilateral Hammers atlas^{36,37} areas, with additional F meaning frontal and P parietal lobe.

class	НС	MCI	AD	MCI → AD
НС				
MCI	F52,F70,F72 p<0.05			
AD	p<0.001 (ex. GM/WMvol)	p<0.05 (ex. GM/WMvol)		
MCI → AD	p<0.01	p<0.05 (ex. GM/WMvol)	GM/WMmean, fron*,P32,P60 p<0.05	
MCI → HC	GM/WMvol p<0.05	GM/WMvol, F50, P60 p<0.05	p<0.001	p<0.005

*differences between AD and MCI→AD group can be found in the frontal lobe and additionally in areas F28, F50, F54, F56, F58.

Method 2

Since in the first tests it seems that subjects can be discriminated corresponding to their diagnosis based on their amyloid imaging, the question arises if the same is possible with new subjects. Can we predict in which category the subject belongs with the help of the previously processed patients? Clinically the most important distinction where a score can help over visual interpretation is within the group of MCI patients, who will convert to AD and who will even have a normal minimal amyloid burden?

Based on the group previously formed as training set, we can make a predictive model in which measurements from a new scan can be put in and chances on each diagnosis will be the output. This model is based on multinomial logistic regression, since we have a multinomial independent variable, class, and multiple predictive variables.

Participants

New subjects are gathered in a new test group. From each group represented in training set, new patients are selected by comparable PET and MRI requisites. This selection resulted in a total of 126 patients with known diagnosis from ADNI.

Image acquisition and processing

Same scan characteristics were aspired as in the first image selection. MRI scans made with a magnetic field strength of both 1.5 and 3T were used to approach the number of included patients in the training set. Time between MRI and PET was again an important factor considered in selecting subjects. PET and MRI scans are run through the FSL processing pipeline, produced in the first part of this study.

Statistical analysis

Multinomial logistic regression has as goal to find an equation that best predicts the probability of a value of the variable Y as a function of the X variables. With this model the measurements of a new subject can be transformed into probabilities of that subject belonging to each of the categories of the dependent variable, classification outcome in this case. The outcome measures of the training set are entered into the model resulting in β values, marginal changes in odds with respect to each measure, that together can be formed into predictive equations. The created model was evaluated with the test set. This evaluates the general applicability of this model. The disadvantage of this method is that variability between the training and test set can be of influence. To neglect this factor, another round of tests was performed by using leave one out cross validation. Herein one subject is used as test set and the model is trained by the rest of the training set.

We have acquired too many measures for all of them to be entered into the model. Forward stepwise entry in SPSS therefore selects the most significant terms ending up in the model. A predictive rate of 80% correct was considered successful for this feasibility study.

Results 2

Demographics of the test set are given in Table 2.5. Again healthy subjects have the shortest interval between MRI and PET scans of the 5 groups, only now the MCI subjects show a comparable inter scan time. The difference in age noted at the time of MRI scan between healthy controls and MCI converters can be explained by our scan selection. While from the HC subjects the latest usable scans are selected from ADNI, in MCI patients the first scans in the database are used in order to select images from before clinical conversion.

p toto of compared to file					
	HC	MCI	AD	MCI → AD	MCI → HC
Number of	76	20	77	34	12
subjects	20	20	27	74	12
Age (years ± SD)	81 (4.8)	78 (8.2)	77 (10.1)	74 (7.3)***	70 (7.9)***
Sex (f/m)	12 (46%) /	11 (39%) /	10 (37%) /	7 (58%) /	18 (53%) /
	14 (54%)	17 (61%)	17 (63%)	5 (42%)	16 (47%)
Inter scan time (days ± SD)	14 (25)	16 (19)	30 (23)*	27 (21)*	31 (17)*

Table 2.5: Demographics of healthy controls (HC), patients suffering from Alzheimer's disease (AD), mild cognitive impaired patients (MCI), and MCI patients converted to HC or AD. *p<0.05, **p<0.01, ***p<0.001 compared to HC

Multiple models with the multinomial logistic regression were created. The first run used data from all the five groups in the training set as input. The measures of grey matter over white matter of the whole brain volume and mean intensity of three frontal lobe atlas areas (F28,F54,F72) were having the most significant influence on the class and therefor ended up in the model. Results of the model on the test set are shown in Table 2.6. The percentage of subjects predicted to the category corresponding to the diagnosis in ADNI is shown in the last column. There is quite some overlap between AD and MCI to AD subjects, indicating that

early scans of these MCI already show quite some similarities with AD subjects. Another remarkable result is the assigning of HC subjects, for the greater part to the MCI class.

			Model assigned class					
		HC	MCI	AD	MCI→AD	MCI→HC	%	
class	HC	2	11	3	6	4	7.7	
ADNI	MCI	4	11	1	7	5	39.3	
	AD	2	4	11	10	0	40.7	
	MCI→AD	4	4	14	8	4	23.5	
	MCI→HC	3	1	0	0	8	66.7	

Table 2.6: Number of patients assigned by the model to each class in comparison to ADNI standard.

As the most clinically interesting classification is between MCI patients converting to AD and converting to HC, a new model was created with input from the training set of these two groups. Table 2.7 shows the results of this model on the test set. These two classes seem to differ from each other and this results in a good separation of the model of these two classes. With predictive values above 80%, one can give a reliable prediction based on early scans of what the future might hold for these subjects.

Table 2.7: Subset test, results of model to distinguish MCI \rightarrow AD and MC \rightarrow HC. Measures used in the model are whole brain grey matter over white matter volume and precuneus mean over white matter mean.

		Model assi		
		MCI→AD	MCI→HC	%
class	MCI→AD	28	6	82.4
ADNI	MCI→HC	0	12	100.0

This model was also applied to MCI subjects as this group is also subject to the question of which of the MCI patients will convert to AD and which are not. Running the MCI subjects from the test set through the model resulted in a partitioning of these patients with 60% being classified as the MCI→AD group and the resulting 40% as the MCI→HC group. This indicates the MCI group being heterogeneous in amyloid burden.

Another modelling option is to divide the whole MCI group into subjects converting to AD in the follow-up period (mean follow-up until conversion is 4.0 years (0.65-8.97y)) and subjects who do convert in that time. That second group is formed by the MCI→HC and MCI group combined. This distinction was modelled by grouping the MCI patients and MCI→HC group together and put their results against those from the subjects in group MCI→AD. Results are shown in Table 2.8.

		Model assi		
		MCI→AD	MCI (→HC)	%
class	MCI	5	23	82.1
ADNI	MCI→AD	19	15	55.9
	MCI→HC	0	12	100.0

Table 2.8: Subset test, precuneus over white matter mean used.

Whereas the MCI group and MCI→HC subjects are now predicted as one would expect, the MCI→AD group is now partitioned. To rule out this result is being caused by difference in MCI heterogeneity between training and test set, leave one out cross validation is used to focus on the training set solely. Results are shown in Table 2.9 and 2.10. Again a predictive rate of over 80% is reached between MCI converters to AD and to HC. And although the splitting of MCI→AD patients is now decreased from the previous model with the use of the test set shown in Table 2.8, the overall predictive rate has not improved, from an average of 79% to 75%.

Table 2.9: Leave one out cross validation results with 2 groups. Whole brain grey matter volume and precuneus mean intensity over whole brain white matter equivalents were used to model.

		Model ass	Model assigned class			
		MCI→AD MCI→HC				
Class ADNI	MCI→AD	19	4	82.6		
	MCI→HC	4	19	82.6		

Table 2.10: leave one out cross validation results with 3 groups, input in model as 2 groups. Precuneus over white matter intensity was used in this model.

		Model assigned class		
		MCI→AD	MCI (→HC)	%
Class ADNI	MCI	6	17	73.9
	MCI→AD	14	9	60.9
	MCI→HC	2	21	91.3

Discussion

In this study we have developed a method to process amyloid PET scans with the aid of individual MRI scans in order to measure in a representative way the amyloid burden congruent with clinical diagnosis. The main finding of this research is that the developed method seems to discriminate MCI patients that will convert to AD from MCI patients who will not. By a combination of multiple measures this result was achieved. Correlation between the acquired measures is high and group sizes too small to take all the measures into account. Adding more measures than is currently done does not have additional value. Letting the forward stepwise method decide on which measures to use, decreases required manual input and effort.

Besides these side marks, the two applied methods of using a test set and perform leave one out cross-validation, do indicate high power of the model to predict subjects either converting from MCI to AD or to HC. This result would be more clinically relevant if also the group that is continuously diagnosed as MCI in ADNI could be distinguished. This test works less well than the previously mentioned two group discrimination. This smaller predictive value can be caused by heterogeneity in the MCI non-converting group. Possibly, although the MCI group had a reasonable follow-up time without diagnosis change in ADNI, some of these patients were verging on an AD diagnosis while other subjects could still have another cause of their memory complaints. Variability could also be explained by speed of conversion, where slow and fast converters show differences in amyloid burden.

Comparable research have shown that MCI patients who convert to AD have higher PIB uptake compared to those patients who remained MCI^{14–16}. Identification of patients probable to convert from MCI to AD hence seems to be possible with amyloid uptake measurements. These studies do not use an individual MRI mask to segment regions of interest, in contrast to this study. Registration of MRI and atlas to PET is done visually without strict regulations. No interpretation on how precise this registration performed with the chosen settings can be given and accuracy of registration is only assessed visually.

Wide applicability of the developed method should be aspired. Applicability on every scan protocol is partly reached, as multiple scanners and magnetic field strengths were included. With protocols outside of ADNI regulations, more steps are required. One can think of the centiloid method trying to achieve comparable results across analysis techniques and tracers⁴². That method shares the idea that one threshold will not be the desired discriminative outcome, since every analysis protocol will yield a different number and possibly will not be able to make more subtle discriminates than HC or AD. Striving for goals of the centiloid method in our image processing protocol, one can use the same training set used in this research, apply their own processing software steps and then acquire a model. Another possibility is to build an own dataset of subjects, preferably with follow-up scans, process the scans and build the model. This second option makes the method applicable to every scanning protocol.

Applicability to all amyloid tracers is another goal. Every developed tracer now requires the physicians to go through a specific training in order to get acquainted with interpreting the scans made with that specific tracer. With one semi-quantitative method applicable for all tracers, these trainings will be degraded and scans could easily be compared, regardless of used tracer. This additional research can easily be performed by applying the same method and processing steps to scans acquired with ¹¹C-PiB or other ¹⁸F tracers.

Applicability of this method to all forms of dementia is another interesting addition to this research. Further visualisation and knowledge of amyloid may also be helpful in identifying the correct type of dementia, especially in patients with an atypical presentation and early in patients' disease progression. Other forms of dementia can therefor also benefit from early detection and possible influence of amyloid imaging on treatment development. Research has shown that IIC-PiB could distinguish AD from two other forms of dementia, namely frontotemporal dementia⁴³ and vascular dementia⁴⁴, reasonably well. Though dementia with Lewy bodies shows high global cortical amyloid burden, Parkinson disease dementia does not⁴⁵. Because all these studies were performed with PiB and patients who had

profound developed disease, the value of amyloid imaging with 18F tracers in early distinguishing of dementia forms is yet to be shown.

Another possible future extension of this research is by a technical measure, namely the use of a hybrid PET/MR scanner. Interesting is to see if images and currently gained results differ from corresponding results with PET/MR scanners and if this new hybrid modality improves workflow. The idea that this research field is one of the areas the PET-MR scanner can be of additional value is mentioned in literature already. ⁴⁶⁻⁴⁸ Clinical relevancy of this additional research is low as PET/MR scanners still have to get into clinical practice. As this study already used movement within MRI scans and large inter scan time between the PET and MRI as exclusion criteria, large improvements by using a PET-MR scanner are not expected.

The main future application of amyloid imaging should be the visualisation and quantification of treatment effects. Having availability of amyloid PET scans and MRI scans and during drug therapy can render additional information as to overall effect, which regions do respond, how quickly effects are visible, and if an effect is seen in all patients or there will be a distinction between responders and non-responders. Treatment and imaging evolvements are reliant of each other and developments therefor go hand in hand.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant UOI AG024904) and DOD ADNI (Department of Defence award number W8IXWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Chapter 3. Methodology supplement 1

Image selection

Amyloid PET scans in the ADNI database are acquired as four 5-minute frames 50-70 minute post-injection. This acquisition time and moment are chosen as a save margin where florbetapir (AV45) has a stable level reached after injection. Time-activity curves in the brain of subjects with positive scans show continuous signal increase from 0 to 30 minutes post-injection, with stable values of standardized uptake values thereafter up to at least 90 minutes post-injection⁴⁹. By dividing the acquisition into separate short frames, minor movements that occur during dynamic scans can be managed by aligning or omitting frames from the analysis.

These initial scans are processed into four subsequent sets. Both the original and all preprocessed data are available from ADNI. Set 1 is the dynamic sequence of frames coregistered to frame 1 (rigid body translation + rotation). After this coregistration, all frames are averaged into a single static frame, resulting in Set 2. Set 3 provides two additional steps of processing: transformation into a standardized orientation and grid (160x160x96 grid with 1.5mm³ voxels) and intensity normalization (scaling) using an atlas-defined cerebellar grey matter reference region (set 3). Pre-processed Set 4 is exactly the same as Set 3, but additionally a scanner specific smoothing is applied to achieve an isotropic resolution of 8 mm FWHM (set 4).⁵⁰

We chose to use the processed Set 3 in our study. With the idea of eventually performing our own amyloid scanning instead of using the database in mind, we thought of what image would then be put into the workflow. Set 3 seems most likely as a standard orientation of all images from one scanner is expected. Since in Set 4 the resolution matching approach can only decrease image resolution and will smooth away potentially high-resolution signals, this final processing step does not seem of additional value. The four image sets are visualized in Figure 3.1.



Figure 3.1: Screenshot of DICOM viewer showing the four separate pre-processed sets. Panel A shows the first set, as the sequence of 436 images (4 subsequent sets of 109 frames). Pre-processing Step 2 results in panel B, with the four subsequent sets averaged into one set of 109 frames (green box). The image in panel C is resized into a standard grid with changed slice thickness to 1.5 mm (before was 2.0 mm, blue box). Panel D shows the result of Step 4, the resolution is decreased to the 24

In table 3.1 and 3.2 the characteristics of the 108 used PET and MRI scans are noted. As subjects were selected from a total of 14 centres to form the training set, scan characteristics are varying.

Manufacturer	Number of	Matrix size Voxel size		Slice thickness	Number of
	subjects		(mm)	(mm)	slices
Philips	12	128 x 128	2.0 X 2.0	2.00	90
GE	9	128 x 128	2.0 X 2.0	4.25	35
	9	128 x 128	2.0 X 2.0	3.27	47
	6	128 x 128	1.9 X 1.9	3.27	47
Siemens	42	336 x 336	1.0 X 1.0	2.03	109
	27	128 x 128	2.6 x 2.6	2.43	63
	3	256 x 256	1.2 X 1.2	1.22	207

Table 3.1: PET characteristics of the train set.

Table 3.2: MRI characteristics of the train set.

Manufacturer	Number	Field	Matrix size	Voxel size	Slice	Number of
	of	strength		(mm)	thickness	slices
	subjects	(T)			(mm)	
Philips	36	3	256 x 256	1.00 X 1.00	1.2	170
GE	26	3	256 x 256	1.02 X 1.02	1.2	196
	7	1.5	256 x 256	0.94 X 0.94	1.2	180
	4	1.5	256 x 256	0.94 X 0.94	1.2	166
Siemens	32	3	240 x 256	1.00 X 1.00	1.2	176
	2	1.5	192 X 192	1.25 X 1.25	1.2	160
	1	3	240 x 256	1.05 X 1.05	1.2	176

Image processing

In previous research reports we came across two software packages commonly used for comparable automatic brain segmentation and registration. We decided to test both FSL³⁵ (FMRIB Software Library) and SPM⁵¹ (statistical parameter mapping): load our images into both, perform the initial steps and pick one to use from that point on. The decision on a particular software package was made on necessary manual input and ability to visualise every step. Reorientation of the MRI scans to a standard view was easily automated and registration to PET was thereby facilitated. SPM required manual input as the angles between PET and MRI were too large for simple automatic registration. Since FSL furthermore offered the possibility to view the outcome of every step of the process, FSL was the package of choice for offering easy implementation of an automatic processing workflow.

After the initial downloading of the scans in DICOM format and checking of the DICOM headers, scans were converted to NIFTI format in order to read into FSL. The image processing of MRI starts with brain extraction, which has to be performed before tissue segmentation. The brain extraction tool (BET) in FSL segments the image into brain and non-brain by using a deformable mesh model that evolves to fit the brain's surface by the application of a set of locally adaptive model forces, described in detail by Smith⁵².

Several parameters are available for optimizing BET performance. The most successful parameter combination was the bias field correction and neck clean-up together with a fractional intensity threshold of 0.1, achieving the most robust brain extraction in the first set of subjects. This combination of parameters was encountered in the paper by Popescu et al.⁵³ optimizing the brain extraction tool on TI MR images in multiple sclerosis. Although we did not perform an equivalent quantitative analysis of BET parameters, we can expect our TI images and required outcome to be similar to the described.

After the brain is extracted from background and non-relevant structures, we can continue with segmentation of the brain itself into grey and white matter and cerebrospinal fluid. This segmentation is encoded in FAST (FMRIB's automated segmentation tool) which, next to segmentation, also performs a correction for spatial intensity variations. The FAST algorithm is based on a hidden Markov random field model and associated expectation-maximization algorithm (details can be found in the paper by Zhang et al.⁵⁴). Both probabilistic and deterministic segmentations are possible to obtain. We used the probabilistic grey and white matter maps and applied a threshold of 0.5 to create a binary grey and white matter masks.

Before the masks can be applied to the PET images, MRI and PET image sets need to be coregistered. PET and MRI images resembled each other most after the first processing step of MRI, reoriented to FSL standard space, which is also the space of the PET scans. FMRIB's linear image registration tool (FLIRT) was used to perform the affine registration of PET to reoriented MRI. The resulting transformation matrix was inverted and later on used to map MRI masks to PET space. After mapping this MRI mask to PET, this grey matter mask was applied to the PET scan and PET signal intensities were extracted. The process of the grey matter mask is represented schematically in Figure 3.2. Since the information on amyloid burden is in the PET and the MRI is used as a tool, the created MRI masks should be mapped to the PET data. Visually the affine registration performed well and no non-rigid registration was needed, confirming the assumption that no change in position of structures happened between the two scans and that variance in head positioning can be overcome by rigid registration.



Figure 3.2: Flowchart of image processing steps in FSL. MRI processing and registration to PET is shown. The second process of coregistering the atlas to PET is shown in the background.

In order to distinguish brain lobes and even smaller distinct regions, the anatomical Hammers atlas is used^{36,37}. This atlas is chosen over for example the widely used automatic anatomical labelling (AAL) atlas from Montreal Neurological Institute⁵⁵, since the latter is based on a single subject and Hammers based their atlas on 30 subjects. Although those subjects were healthy and in average 31 years of age, the number of subjects results in some anatomical variation included in the atlas. The Hammers atlas was also registered to PET space, by combining transformation matrices from registering PET to MRI and atlas to MRI. For this registration the MRI set after brain extraction is used as this view resembled the atlas which is also brain tissue only. After registration, atlas masks of several anatomical regions of interest were superimposed over the grey matter PET region.

Rescaling of the PET scans was added as processing step, since some centres saved their scans with an added rescale slope which resulted in very small (<0.001) intensity values. In order to be able to perform calculations on PET scan intensities, all scans were processed and their rescale slopes were set to 1 before processing in FSL was performed.



Figure 3.3: Example of image processing outcome in sagittal, coronal and transversal views. The original PET data is overlaid with the masks for the four lobes. These masks are obtained by applying atlas based masks on the grey-matter-only PET. Frontal lobe is depicted in red, parietal lobe in green, occipital lobe in blue and temporal lobe in yellow.

White matter reference region

Regional PET activity has to be normalized to a reference tissue to account for nonspecific radiotracer binding and other variations between subjects. First amyloid PET quantification studies used cerebellum as a whole or the cerebellar grey matter as reference region, since fibrillary amyloid deposits targeted by the 18F-florbetapir are very rarely observed^{56,57}. However variability observed in longitudinal progression measures using the cerebellar reference is not congruent with pathological or biological expected values⁵⁸. Some seek the cause of cerebellar reference variability in the position close to the edge of field of view and scatter correction errors³². Others emphasize on the rigorous correction performed on PET to ensure uniformity within the FOV and find biological factors more likely to cause the cerebellar variability⁵⁸.

In recent studies^{38-40,58} researchers have examined the feasibility of alternative reference regions for amyloid PET. In all studies, the use of subcortical white matter normalization was

found to improve the accuracy of longitudinal PET data more strongly than tested grey matter normalization. Landau et al.³³ adds that the use of eroded subcortical white matter results in a cortical change that was more physiologically plausible while Shokouhi et al.⁵³ find a higher inverse correlation to CSF amyloid measures compared to grey matter normalization and more likely to increase over time. The white matter normalization increases the power to detect amyloid burden change opposed to cerebellar normalization found by Chen et al.³⁹.

One possible reason for the success of white matter reference region is that it is located in the same slice as cortical target regions and therefore is less susceptible to differences in scatter correction between superior and inferior planes³². This proximity to grey matter, on the other hand, induces the risk of spill-over of signal from white matter into the small grey matter region. As we want to use subcortical white matter as a reference region, four different white matter based regions are segmented and compared: white matter, eroded white matter, white matter without cerebellum and eroded white matter without cerebellum.

In Figure 3.4 these different regions are visualized. When tested in 39 subjects, 21 AD and 18 HC, the four different reference region did slightly influence the ratios of mean intensity measured as whole brain grey matter over white matter regions. When testing for the differences between the AD and HC subjects, mean intensity ratios showed p<0.001, irrespective of the reference region used. Choice of the reference region did not decrease the discriminative power of the mean intensity ratio of whole brain and also lobular grey matter measurements.



Figure 3.4: Visualization of four white matter based reference regions. All colours summed show the original whole brain white matter region, eroded white matter is represented by red and blue areas together. The white matter region without cerebellum is composed of yellow and blue area and blue alone depicts the smallest option which is eroded and cerebellum stripped of.

This result led to the decision to use the whole white matter region as a reference. That region, as the largest option, yields the highest amount of counts and therefore highest statistical power. Figure 3.5 shows the effect of using the reference region in whole brain volume and intensity measures. Although white matter shows no distinction between HC and AD subjects (mean p>0.3, volume p>0.9), the ratios of white matter over grey matter show far less variability and overlap (mean p<0.001, volume p<0.01) than the grey matter measurements alone (mean and volume p<0.03). All subsequent measurements were expressed as the ratio of mean intensity of target cortical region over whole brain white matter

reference region. Such normalization has the positive effect that the unit of mean intensity is omitted and no attention has to be paid to subjects weight or total administered dose as otherwise needs to be done in measurements of standardized uptake values.





Figure 3.5: Measurements of 21 AD and 19 HC subjects. Grey matter volume and intensity measurements in panel A show moderate distinction between subject classes. Panel B shows complete overlap of white matter measurements. Grey and white matter combined, resulting in ratios in panel C, show good discrimination between subject groups by mean intensity ratio.

Chapter 4. Methodology supplement 2

Statistical analysis

To answer the question if new subjects can be classified into diagnostic groups based on their amyloid imaging, we made a predictive model in which measurements from a new scan can be entered and chances on each diagnosis will be the output. This model is based on multinomial logistic regression, since we have a multinomial independent variable, class, and multiple predictive variables, the amyloid signal ratios from different brain areas. The goal of multinomial logistic regression is to find an equation that optimally predicts the probability of a value of the variable Y as a function of the variables X. You can then measure the independent variables on a new subject and estimate the probability of the subject having a particular value of the dependent variable. The main null hypothesis states that there is no relationship between the variables X and the variable Y; the Y values predicted from the multiple logistic regression equation are no closer to the actual Y values than one would expect by chance. The null hypothesis for each X variable is that adding that variable to the model does not improve the fit of the equation more than expected by chance. The regression model should be used for suggesting patterns in your data rather than rigorous hypothesis testing.

The multinomial logistic regression assumes that each observations are independent. Since we have single time-point measurements from all separate subjects, this assumption is met. Another assumption is that there should be no multicollinearity. Multicollinearity occurs when you have two or more independent variables that are highly correlated with each other. This leads to problems with understanding which variable contributes to the explanation of the dependent variable and technical issues in calculating a multinomial logistic regression. This requirement of the data cannot be completely fulfilled, as whole brain grey matter measurements are quite correlated to lobular grey matter signals and variability within patients between the lobes is also not very large. Despite these remarks, we still expect the regression model to work with some independent variables such as whole brain volume and small atlas areas.^{59,60}

Finally, the model should be fitted correctly. Neither overfitting nor underfitting should occur. To prevent overfitting, one needs to have several times as many observations as independent variables, with a factor of ten being commonly accepted. A common approach to ensure all meaningful variables are included is to use a stepwise method to estimate the logistic regression. One of the stepwise methods is forward entry, whereby the most significant term is iteratively added to the model until none of the stepwise terms left out of the model would have a statistically significant contribution if added to the model. Forward stepwise, another method, uses the model that is selected by the forward entry method as a starting point. Then the algorithm alternates between elimination of the least significant stepwise term in the model and forward entry on the terms left out of the model. This continues until no terms meet the entrance or removal criteria.⁴¹

The implementation of the model is schematically explained with corresponding equations. The useful outcome of multinomial logistic regression is given as an intercept value

and β values. These values are always rendered relative to a reference group, one of the categories of Y. The β corresponding to one of the X variables represents the marginal change in log odds with respect to that X. The exponent of β n is the amount by which the relative risk for one of the categories to the reference category is multiplied when variable Xn in increased by one.

$$log(odds) = b_0 + b_1 X_1 + b_2 X_2$$

$$p = \frac{odds}{1 + odds} = \frac{e^{a + b_1 X_1 + b_2 X_2}}{1 + e^{a + b_1 X_1 + b_2 X_2}} = \frac{1}{1 + e^{-(a + b_1 X_1 + b_2 X_2)}}$$

With the aforementioned equations one can calculate the probability that a subject with values X can be categorized into each of the independent variable Y. As an example the calculation is shown in the case of 3 categories for the independent variable Y with category 1 as a reference. The β values are denoted with two numbers, the first being the category and the second the corresponding X.

$$p(2ref1) = \frac{1}{1 + e^{-(b_{20} + b_{21}X_{21} + b_{22}X_{22})}} \qquad p(3ref1) = \frac{1}{1 + e^{-(b_{30} + b_{31}X_{31} + b_{32}X_{32})}}$$

In order to convert these relative chances to absolute chances, the first step is to normalize the relative chances. This is done by multiplying the relative chances with 1/(1-relative chance). This results in relative chance of the reference group being equal to 1.

$$\|p(2ref1)\| = p(2ref1)\frac{1}{1-p(2ref1)} \qquad \|p(3ref1)\| = p(3ref1)\frac{1}{1-p(3ref1)}$$

The chances are now expressed relatively to a chance on reference category of 1. So the relative chances relate to each other by the following expression:

$$1 : \|p(2ref1)\| : \|p(3ref1)\|$$

This means that now we can express the likelihood of each category without the reference group in the equation:

$$p1 = \frac{1}{1 + norm_{p(2ref_1)} + norm_{p(3ref_1)}} \quad p2 = \frac{norm_{p(2ref_1)}}{1 + norm_{p(2ref_1)} + norm_{p(3ref_1)}} \quad p3 = \frac{norm_{p(3ref_1)}}{1 + norm_{p(2ref_1)} + norm_{p(3ref_1)}}$$

These expressions give the chance of a subject with corresponding measurements being assigned to each of the categories. This matches the goal of the regression model. We can either use these probabilities for each category or look only at the category with the highest probability. In this work the last option was used and for each subject in the test group the category with the highest probability is noted and compared to the standard category from ADNI.

References

- 1. Strassnig, M. & Ganguli, M. About a peculiar disease of the cerebral cortex. By Alois Alzheimer, 1907. *Psychiatry* 30–33 (2005).
- 2. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* **82**, 239–259 (1991).
- 3. Nederlands Huisartsen Genootschap. NHG-Standaard Dementie. Available at: https://www.nhg.org/standaarden/volledig/nhg-standaard-dementie.
- 4. Klunk, W. E. *et al.* Imaging Brain Amyloid in Alzheimer 's Disease with Pittsburgh Compound-B. *Ann. Neurol.* **55**, 306–319 (2004).
- 5. Barthel, H., Seibyl, J. & Sabri, O. The role of positron emission tomography imaging in understanding Alzheimer's disease. *Expert Rev. Neurother.* **15**, 395–406 (2015).
- 6. Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
- 7. Alves, L., Correia, A. S. A., Miguel, R., Alegria, P. & Bugalho, P. Alzheimer's Disease: A Clinical Practice-Oriented Review. *Front. Neurol.* **3**, 1–20 (2012).
- 8. Jack, C. R. *et al.* Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* **12**, 207–216 (2013).
- 9. Fodero-Tavoletti, M. T. *et al.* Characterization of PiB binding to white matter in Alzheimer disease and other dementias. *J. Nucl. Med.* **50**, 198–204 (2009).
- 10. Vandenberghe, R. *et al.* 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: A phase 2 trial. *Ann. Neurol.* **68**, 319–329 (2010).
- 11. Lister-James, J. *et al.* Florbetapir F-18: A Histopathologically Validated Beta-Amyloid Positron Emission Tomography Imaging Agent. *Semin. Nucl. Med.* **41**, 300–304 (2011).
- 12. Eli Lilly and Company. *Amyvid prescribing information*. **1**, (2012).
- 13. Yang, L., Rieves, D. & Ganley, C. Brain Amyloid Imaging FDA Approval of Florbetapir F18 Injection. *N. Engl. J. Med.* **367**, 885–887 (2012).
- 14. Forsberg, A. *et al.* PET imaging of amyloid deposition in patients with mild cognitive impairment. *Neurobiol. Aging* **29**, 1456–1465 (2008).
- 15. Koivunen, J. *et al.* Amyloid PET imaging in patients with mild cognitive impairment: a 2-year follow-up study. *Neurology* **76**, 1085–90 (2011).
- 16. Okello, A. *et al.* Conversion of amyloid positive and negative MCI to AD over 3 years: an 11C-PIB PET study. *Neurology* **73**, 754–760 (2009).
- Villemagne, V. L. *et al.* Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *Lancet Neurol.* 12, 357–367 (2013).
- 18. Yiannopoulou, K. G. & Papageorgiou, S. G. Current and future treatments for Alzheimer's disease. *Ther. Adv. Neurol. Disord.* **6**, 19–33 (2013).

- 19. Anand, R., Gill, K. D. & Mahdi, A. A. Therapeutics of Alzheimer's disease: Past, present and future. *Neuropharmacology* **76**, 27–50 (2014).
- 20. McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7**, 263–269 (2011).
- 21. Alzheimer Nederland. Neemt het aantal mensen met dementie toe of af? (2014). Available at: http://www.alzheimernederland.nl/nieuws/onderzoek/2014/februari/aantal-mensen-met-dementie.aspx.
- 22. Prince, M. *et al.* World Alzheimer Report 2015: The Global Impact of Dementia An analysis of prevalence, incidence, cost and trends. *Alzheimer's Dis. Int.* 84 (2015). doi:10.1111/j.0963-7214.2004.00293.x
- 23. Clark, C. M. *et al.* Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: A prospective cohort study. *Lancet Neurol.* **11**, 669–678 (2012).
- 24. Zwan, M. D. *et al.* Comparison of Simplified Parametric Methods for Visual Interpretation of 11C-Pittsburgh Compound-B PET Images. *J. Nucl. Med.* **55**, 1305–1307 (2014).
- 25. Ng, S. *et al.* Visual assessment versus quantitative assessment of 11C-PIB PET and 18F-FDG PET for detection of Alzheimer's disease. *J. Nucl. Med.* **48**, 547–552 (2007).
- 26. Tolboom, N. *et al.* Molecular imaging in the diagnosis of Alzheimer's disease: visual assessment of [11C]PIB and [18F]FDDNP PET images. *J. Neurol. Neurosurg. Psychiatry* **81**, 882–884 (2010).
- 27. Suotunen, T. *et al.* Visual assessment of [11C]PIB PET in patients with cognitive impairment. *Eur. J. Nucl. Med. Mol. Imaging* **37**, 1141–1147 (2010).
- 28. Edison, P. *et al.* Comparison of MRI based and PET template based approaches in the quantitative analysis of amyloid imaging with PIB-PET. *Neuroimage* **70**, 423–433 (2013).
- 29. Saint-Aubert, L. *et al.* Comparison between PET template-based method and MRIbased method for cortical quantification of florbetapir (AV-45) uptake in vivo. *Eur. J. Nucl. Med. Mol. Imaging* **41**, 836–43 (2014).
- 30. Lim, Y. Y. *et al.* Disruption of cholinergic neurotransmission exacerbates Aβ-related cognitive impairment in preclinical Alzheimer's disease. *Neurobiol. Aging* **36**, 2709–2715 (2015).
- 31. Nordberg, A. The use of amyloid imaging in clinical praxis: a critical review. *Clin. Transl. Imaging* **3**, 7–11 (2015).
- 32. Schmidt, M. E. *et al.* The influence of biological and technical factors on quantitative analysis of amyloid PET: Points to consider and recommendations for controlling variability in longitudinal data. *Alzheimer's Dement.* **11**, 1050–1068 (2015).
- 33. Anandh, K. R., Sujatha, C. M. & Ramakrishnan, S. A Method to Differentiate Mild Cognitive Impairment and Alzheimer in MR Images using Eigen Value Descriptors. *J. Med. Syst.* **40**, 25 (2016).

- 34. University of Southern California. Alzheimer's Disease Neuroimaging Initiative. Available at: adni.loni.usc.edu.
- 35. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* 62, 782–790 (2012).
- 36. Hammers, A. *et al.* Statistical neuroanatomy of the human inferior frontal gyrus and probabilistic atlas in a standard stereotaxic space. *Hum. Brain Mapp.* **28**, 34–48 (2007).
- 37. Hammers, A. *et al.* Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum. Brain Mapp.* **19**, 224–247 (2003).
- 38. Landau, S. M. *et al.* Measurement of Longitudinal β-Amyloid Change with 18F-Florbetapir PET and Standardized Uptake Value Ratios. *J. Nucl. Med.* **56**, 567–574 (2015).
- 39. Chen, K. *et al.* Improved Power for Characterizing Longitudinal Amyloid-β PET Changes and Evaluating Amyloid-Modifying Treatments with a Cerebral White Matter Reference Region. *J. Nucl. Med.* **56**, 560–566 (2015).
- 40. Brendel, M. *et al.* Improved longitudinal [18F]-AV45 amyloid PET by white matter reference and VOI-based partial volume effect correction. *Neuroimage* **108**, 450–459 (2015).
- 41. Armonk NY: IBM Corp. IBM Corp. IBM SPSS Statistics for Windows.
- 42. Klunk, W. E. *et al.* The Centiloid Project: Standardizing quantitative amyloid plaque estimation by PET. *Alzheimer's Dement.* **11**, 1–15 (2015).
- 43. Rabinovici, G. D. *et al.* Amyloid vs FDG-PET in the differential diagnosis of AD and FTLD. *Neurology* **77**, 2034–2042 (2011).
- 44. Lee, J. H., Kim, S. H. & Kim, G. H. C-Pittsburgh compound B Identification of pure subcortical vascular dementia using 11 C-Pittsburgh compound B. (2012). doi:10.1212/WNL.ob013e318221acee
- 45. Gomperts, S. N. *et al.* Imaging amyloid deposition in lewy body diseases. *Neurology* **71**, 903–910 (2008).
- 46. Drzezga, A., Barthel, H., Minoshima, S. & Sabri, O. Potential Clinical Applications of PET/MR Imaging in Neurodegenerative Diseases. *J. Nucl. Med.* **55**, 47S–55S (2014).
- 47. Barthel, H., Schroeter, M. L., Hoffmann, K.-T. & Sabri, O. PET/MR in Dementia and Other Neurodegenerative Diseases. *Semin. Nucl. Med.* **45**, 224–233 (2015).
- 48. Jena, A., Renjen, P., Taneja, S., Gambhir, A. & Negi, P. Integrated ¹⁸ Ffluorodeoxyglucose positron emission tomography magnetic resonance imaging (¹⁸ F-FDG PET/MRI), a multimodality approach for comprehensive evaluation of dementia patients: A pictorial essay. *Indian J. Radiol. Imaging* **25**, 342 (2015).
- 49. Joshi, a. D. *et al.* Performance Characteristics of Amyloid PET with Florbetapir F 18 in Patients with Alzheimer's Disease and Cognitively Normal Subjects. *J. Nucl. Med.* **53**, 378–384 (2012).
- 50. Jagust, W. J. *et al.* The Alzheimer's Disease Neuroimaging Initiative 2 PET Core: 2015. *Alzheimer's Dement.* **11**, 757–771 (2015).

- 51. Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
- 52. Smith, S. M. Fast robust automated brain extraction. *Hum. Brain Mapp.* **17**, 143–155 (2002).
- 53. Popescu, V. *et al.* Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis. *Neuroimage* **61**, 1484–1494 (2012).
- 54. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *Tmi* **20**, 45–57 (2001).
- 55. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).
- 56. Klunk, W. E. *et al.* The binding of 2-(4'-methylaminophenyl)benzothiazole to postmortem brain homogenates is dominated by the amyloid component. *J. Neurosci.* 23, 2086–2092 (2003).
- 57. Joachim, C. L., Morris, J. H. & Selkoe, D. J. Diffuse senile plaques occur commonly in the cerebellum in Alzheimer's disease. *Am. J. Pathol.* **135**, 309–19 (1989).
- 58. Shokouhi, S. *et al.* Reference tissue normalization in longitudinal (18)F-florbetapir positron emission tomography of late mild cognitive impairment. *Alzheimers. Res. Ther.*8, 2 (2016).
- 59. Statistics Solutions. Assumptions of Logistic Regression. (2016). Available at: http://www.statisticssolutions.com/assumptions-of-logistic-regression/.
- 60. Lund Research Ltd. Muktinomial Logistic Regression using SPSS Statistics. (2013). Available at: https://statistics.laerd.com/spss-tutorials/multinomial-logistic-regressionusing-spss-statistics.php.