Contribution of FDG-PET and MRI to improve Understanding, Detection and Differentiation of Dementia

Dissertation
zur Erlangung des akademischen Grades
Dr. rer. med.
an der Medizinischen Fakultät
der Universität Leipzig

ingereicht von: Dipl.-Psych. / Jürgen / Dukart
geboren am 08.07.1983 in Burnoe/Kasachstan

angefertigt an / in:
Max-Planck-Institut für Kognition- und Neurowissenschaften

Betreuer:
PD Dr. Matthias Schroeter
PD Dr. Karsten Müller

Beschluss über die Verleihung des Doktorgrades vom: 23.02.2011
Danksagung

Mein besonderer Dank gilt meinen beiden Betreuern Herrn PD Dr. Matthias Schroeter und Herrn PD Dr. Karsten Müller, die mir die Möglichkeit gegeben haben diese Arbeit durchzuführen und mit wertvollen Anregungen und Ratschlägen mich fachgerecht und vielseitig die gesamte Zeit unterstützt haben. Mein Dank gilt Ihnen auch dafür für das sehr aufmerksame Korrekturlesen. Insbesondere danke ich den beiden dafür, dass Sie mir die Möglichkeit gegeben haben alle Ideen, die im Rahmen dieser Arbeit entstanden sind, auch tatsächlich umzusetzen und dass Sie mich mit Ihrer Erfahrung immer tatkräftig unterstützt haben und freue mich natürlich auf die gemeinsame Arbeit in den nächsten Jahren.


Und zuletzt und vor allem danke ich meiner Familie, dass Sie mich in dieser ganzen Zeit immer und in allem nicht wissenschaftlichem unterstützt und mir damit den Weg zu dieser Arbeit erst ermöglicht haben.
# Table of Contents

Table of Contents .................................................................................................................................................. 3

Bibliographische Beschreibung ............................................................................................................................ 5

1. General introduction ........................................................................................................................................... 6

1.1. Diagnostic accuracy of Alzheimer's disease and frontotemporal lobar degeneration .............................. 7

1.2. Biomarkers in dementia ............................................................................................................................... 8

2. Intensity normalization in [F18]fluorodeoxyglucose positron emission tomography ................................. 11

2.1. Introduction .................................................................................................................................................. 11

2.2. Methods ..................................................................................................................................................... 13

2.2.1. Subjects ................................................................................................................................................ 13

2.2.2. [F18]fluorodeoxyglucose positron emission tomography imaging ....................................................... 14

2.2.3. [F18]fluorodeoxyglucose positron emission tomography data analysis .............................................. 14

2.3. Results ....................................................................................................................................................... 16

2.4. Discussion .................................................................................................................................................. 21

2.5. Conclusions ............................................................................................................................................... 24

3. Multimodal imaging in dementia using [F18]fluorodeoxyglucose positron emission tomography and magnetic resonance imaging ................................................................. 25

3.1. Introduction ............................................................................................................................................... 25

3.2. Methods ..................................................................................................................................................... 27

3.2.1. Subjects ................................................................................................................................................ 27

3.2.2. Data acquisition ................................................................................................................................ 28

3.2.3. Image processing .................................................................................................................................. 29

3.2.4. Statistical analysis ................................................................................................................................. 34

3.3. Results ....................................................................................................................................................... 36

3.3.1. Voxel-based morphometry ................................................................................................................... 37
Bibliographische Beschreibung

Dukart, Jürgen

Contribution of FDG-PET and MRI to improve Understanding, Detection and Differentiation of Dementia

Universität Leipzig, Dissertation

77 S. ¹, 100 Lit. ², 9 Abb., 8 Tab., 0 Anlagen

Referat:

The present work deals on the one hand with evaluation of methodical aspect of intensity normalization in preprocessing of [F18]fluorodeoxyglucose positron emission tomography (FDG-PET). On the other hand, in this work FDG-PET and magnetic resonance imaging (MRI) information are co-evaluated to improve understanding and diagnostic accuracy of Alzheimer’s disease (AD) and frontotemporal lobar degeneration (FTLD). Differences between both types of dementia patients and control subjects were assessed in both imaging modalities using voxel-based morphometry. Additionally, both imaging modalities were systematically compared regarding their contribution for detection and differentiation of both types of dementia using support vector machine classification. Thereby, FDG-PET and MRI whole-brain and region-of-interest information were used separately and by combining them to investigate which method provides the highest diagnostic accuracy for both, detection and differentiation of AD and FTLD.

The comparison of two reference regions for intensity normalization in FDG-PET, cerebellar and global intensity mean, revealed a remarkable differential effect of both intensity normalization procedures for detection and differentiation of dementia. While the cerebellar normalization was superior for differentiation between dementia patients and control subjects and therefore for detection of dementia, there was an opposite effect for differentiation of AD and FTLD. For this comparison global normalization was highly superior in detecting group differences.

Voxel-based morphometry resulted in a differential pattern of atrophy and hypometabolism in AD and FTLD patients, indicating different proceeding of both disorders and providing support for the assumption of different underlying mechanisms in AD and FTLD. Additionally, applying support vector machine classification on combined information from FDG-PET and MRI provided highest accuracy rates for both, detection and differentiation of AD and FTLD. This result for differentiation of AD patients and control subjects was also validated using a multicenter dataset.

The results of this study emphasize the importance of combined use and evaluation of FDG-PET and MRI to improve understanding and diagnostic accuracy of dementia.

¹ Seitenzahl insgesamt
² Zahl der im Literaturverzeichnis ausgewiesenen Literaturangaben
1. General introduction

Dementia has become a major problem in our society with an estimated 1.07 million people in Germany suffering from a moderate to severe form of dementia in 2007 and an actual rate of 244000 new cases per year (Ziegler and Doblhammer, 2009). Considering the increases in life expectancy and the accompanying trend of low birth rates in Western society in recent decades the proportion of people suffering from dementia is expected to increase further in coming decades. Worldwide, the absolute number of dementia patients is expected to increase from 25 million in 2000 to 114 million in 2050 (Wimo et al., 2003). In 2001 costs of dementia were already estimated to make up 12% of the total global cost of illness (World Health Organization, 2001). To address this considerable problem arising in our society, substantial progress is required in medical research regarding diagnostic accuracy, understanding of neural mechanisms and treatment of different dementia disorders. Although progress has been made in all of these research areas in recent decades, there is still a great deal of improvement needed to address the growing problematic of dementia in our society.

The most common type of dementia is Alzheimer’s disease (AD), with a total of about 60% of dementia patients receiving the diagnosis of this specific neurodegenerative disorder (Fratiglioni et al., 1999). Recently, revised diagnostic criteria for AD (Dubois et al., 2007) suggested the following core features for the clinical diagnosis of probable AD: gradual and progressive change in memory function reported by patients or informants over more than 6 months and objective evidence of significantly impaired episodic memory on testing, which can be isolated or associated with other cognitive changes.

A less common, although still very frequent neurodegenerative disorder, is frontotemporal lobar degeneration (FTLD). Estimates of the prevalence of FTLD vary greatly, with an overall prevalence of between 3.6 and 15 patients per 100000 in the 45- to 64-year age groups (Ratnavalli et al., 2002; Rosso et al., 2003). The results of these studies suggest that FTLD is the second most common diagnosis of dementia in individuals younger than 65. In the cohort of 85 years old, an even higher prevalence of 3% for patients with a subtype of FTLD has been reported in the total population (Gislason et al., 2003).

Due to the inhomogeneous pattern of cognitive impairment in FTLD, it has been suggested to divide FTLD into three major groups: frontotemporal dementia, semantic dementia and
progressive non-fluent aphasia (Neary et al., 1998). Frontotemporal dementia is characterized by alterations in behavior and personality, namely decline in social interpersonal conduct, impairment in regulation of personal conduct, emotional blunting, and loss of insight. Semantic dementia is described as a language disorder characterized by the following: empty fluent speech, loss of word meaning, or semantic paraphasias, a perceptual disorder characterized by impaired recognition of familiar faces or object identity, preserved perceptual matching and drawing reproduction, preserved single-word repetition, and preserved ability to read aloud and write to dictation orthographically regular words. Specific for progressive non-fluent aphasia is non-fluent spontaneous speech with at least one of the following symptoms: agrammatism, phonemic paraphasias, or anomia. Specific neural networks have been reported for each of the three clinically defined subtypes of FTLD (Schroeter et al., 2007).

The treatment of dementia syndromes is still an unsolved problem. Effective drugs have been reported in recent literature, which slow down or even briefly improve cognitive decline in AD (Wilcock et al., 2000; Doody et al., 2001; Hansen et al., 2008) or behavioral symptoms in FTLD (Ikeda et al., 2004). Although, there is still no efficient treatment providing either recovery of the effects of proceeding neurodegeneration or even a permanent stop in cognitive decline, a lot of research is currently directed to development of such therapeutic approaches. For this reason, it is highly important to provide accurate methods enabling early detection and differentiation of neurodegenerative disorders, as these are a prerequisite for an early and successful treatment.

1.1. Diagnostic accuracy of Alzheimer's disease and frontotemporal lobar degeneration

To improve the correct evaluation of a specific treatment, it is also necessary to improve diagnostic accuracy of specific dementia types. An initially incorrect group selection would substantially lower the efficacy estimation of a specific therapeutic approach. Furthermore, the correct diagnostic classification of individual cases is highly relevant for the application of the correct therapeutic treatment if this is available. For dementia syndromes, diagnostic accuracy is difficult to assess even in studies using autopsy confirmed data, given that the neuropathological standard is not the same in all studies. Another difficulty with using neuropathology as a "gold standard" for estimation of diagnostic accuracy is that healthy
subjects sometimes show similar neuropathological changes to those of AD patients (Braak and Braak, 1991). Moreover, it is well known that, in FTLD, histological changes are nonspecific. Any of the histological subtypes can be associated with the three clinical subtypes of FTLD (Hodges et al., 2004, Whitwell et al., 2005).

Studies assessing the diagnostic accuracy of AD based on clinical criteria (e.g., NINCDS-ARDA criteria [National Institute of Neurological and Communication Disorders and Stroke–Alzheimer's disease and Related Disorders Association]: McKhann et al., 1984) reported accuracy rates between 65 and 100% for detection of this specific dementia syndrome (Klatka et al., 1996; Petrovitch et al., 2001; Rascovsky et al., 2002). However, for AD, a high discrepancy was observed between sensitivity and specificity with a substantially higher sensitivity compared to specificity (Varma et al., 1999; Jagust et al., 2007). For FTLD, reported diagnostic accuracies based on clinical criteria ranged between 17 and 85% (Rascovsky et al., 2002; Knopman et al., 2005; Mendez et al., 2007). In comparison to AD syndrome, the opposite pattern between sensitivity and specificity has been reported for FTLD, with much higher specificity compared to sensitivity rates. Sensitivity ranged between 36.5 and 79% while specificity was up to 100% (Knopman et al., 2005; Mendez et al., 2007). For both types of dementia, diagnostic accuracies for early detection and differentiation with other dementia syndromes are still in need of improvement.

1.2. Biomarkers in dementia

Earlier diagnostic criteria for AD and FTLD have mainly relied on clinical and neuropsychological symptoms (McKhann et al., 1984; The Lund and Manchester Groups, 1994; Neary et al., 1998). However, recent studies reported substantial imaging abnormalities for both groups of dementia patients in different imaging modalities. Abnormal patterns in dementia patients have been observed using positron emission tomography (PET) with different ligands, magnetic resonance imaging (MRI) and single photon emission computed tomography (Rosen et al., 2002; Ishii et al., 2005; Nestor et al., 2005; Desgranges et al., 2007; Clark et al., 2008; Habeck et al., 2008; Jack et al., 2008). Accordingly, it has been suggested to integrate imaging
markers into diagnostic criteria for specific dementia syndromes (Dubois et al., 2007; Clark et al., 2008; Kipps et al., 2009).

Different methods have been proposed for the integration of such biomarkers into the diagnostic procedure ranging from basic evaluation of such images by radiologists, semiautomatic evaluation by combining automatic feature selection and evaluation by physicians, to fully automatic univariate and multivariate approaches (Hoffman et al., 2000; Santens et al., 2001, Mosconi et al., 2006; Fung and Stoeckel, 2007; Habeck et al., 2008; Klöppel et al., 2008a, b; Habert et al., 2009). The integration of such biomarkers has been shown to substantially improve diagnostic accuracy of dementia syndromes, with accuracy rates of above 90% for detection (Fung and Stoeckel, 2007; Matsunari et al., 2007; Davatzikos et al., 2008; Klöppel et al., 2008a, b) and between 84 and 89% for differentiation of AD and FTLD (Davatzikos et al., 2008; Klöppel et al., 2008a; Horn et al., 2009). Furthermore, studies comparing evaluation by physicians with multivariate approaches revealed the superiority of the multivariate methods for detection and differentiation of dementia syndromes (Klöppel et al., 2008b; Horn et al., 2009).

Evaluation of different imaging markers in neurodegenerative disorders not only provides a good instrument to improve diagnostic accuracy of those. Additionally, biomarkers might provide a new way to improve understanding of the underlying processes in different dementia syndromes and so effectively contribute to the development of new treatment alternatives. Different imaging markers provide different information regarding neural changes in neurodegenerative disorders. For example, in AD, changes in measurements provided by [F18]fluorodeoxyglucose PET (FDG-PET) have been shown to precede or at least to have a faster progression in some specific regions than those measured by structural MRI suggesting a mechanism of functional disruption in early AD ahead of actual atrophy (Chetelat et al., 2008).

A further – up to now rather theoretical – application of biomarkers for the treatment of neurodegenerative disorders might be their use for the evaluation of the efficacy of a specific therapeutic approach. Assessment of reduction of glucose hypometabolism in AD or FTLD could, for example, provide a measurement of recovery of functional activity in these disorders.

However, for the optimal use of biomarkers to improve understanding, detection and differentiation of dementia syndromes, specific methodical questions still have to be solved. Such questions range from optimization of preprocessing algorithms to enhance statistical evaluation of a specific imaging modality to an improvement of algorithms for the combination of
different imaging modalities. Addressing such questions is necessary to enable valid interpretation of imaging findings. Further methodical questions concern the optimal use of observed imaging abnormalities to increase diagnostic accuracy of specific dementia syndromes. The present work deals with such questions of optimizing preprocessing, and the combination and application of FDG-PET and MRI to improve understanding and diagnostic accuracy of AD and FTLD. Algorithms developed in this study are applied to different datasets of patients with AD, FTLD and to data of control subjects. The results obtained for the optimization of specific preprocessing steps, for the combined use of FDG-PET and MRI, and for the validation of the diagnostic procedure developed are discussed as separate studies in different chapters.
2. Intensity normalization in \([F18]\text{fluorodeoxyglucose pos}i\)\(t\)\(r\)\(on\) emit\(t\)\(o\)\(m\)\(a\)\(t\)y tomography

The following section deals with optimization of specific methodical aspects in preprocessing of FDG-PET to enable an accurate univariate and multivariate statistical evaluation. The clarification and optimization of these methodical aspects is important to enable an optimal qualitative and quantitative evaluation of FDG-PET to improve understanding, detection and differentiation of AD and FTLD. To enable this statistical evaluation of imaging data across different subjects, complex preprocessing algorithms are applied to normalize individual imaging data to a common space and to a common scale so trying to overcome differences in the interindividual anatomy and in the acquisition of the data. An important preprocessing step in the evaluation of FDG-PET data is the intensity normalization. This procedure removes differences in the absolute whole-brain intensity across different subjects so allowing a comparison of regional glucose utilization rates. However, there is still no consensus as to which reference region intensity normalization should be used.

2.1. Introduction

Research investigating the contribution of positron emission tomography with \([F18]\text{fluorodeoxyglucose (FDG-PET) in detecting dementia disorders has revealed differential patterns of a reduced regional cerebral metabolic rate for glucose in different types of dementia (Mielke et al., 1994; Salmon et al., 2000; Sakamoto et al., 2002; Jeong et al., 2005; Mosconi et al., 2006; Jagust et al., 2007; Samuraki et al., 2007). While AD patients show reduced glucose consumption in parietotemporal and posterior cingulate cortices (Ishii et al., 2001, 2005; Yakushev et al., 2008; Schroeter et al., 2009), the reduced glucose metabolism of patients with FTLD is predominately located in frontotemporal and anterior cingulate cortices (Jeong et al., 2005; Schroeter et al., 2007, 2008). Although the areas affected are sufficiently different in their anatomical distribution, it is necessary to run an intensity normalization procedure in order to compare their relative regional metabolic rate to those of healthy controls or with each other. This procedure removes
interindividually differences in the absolute whole brain intensity by relativizing the absolute metabolic rate for glucose in the whole brain to a reference area. Such differences can, for example, result due to different amounts of radiocontrast agent injected.

A review of the current literature revealed the use of the following areas for normalization of FDG-PET scans of dementia patients: the cerebral metabolic rate for glucose (CMRglc; Mielke et al., 1994; Salmon et al., 2000; Ishii et al., 2001; Herholz et al., 2002; Mosconi et al., 2004; Samuraki et al., 2007; Del Sole et al., 2008; Yakushev et al., 2008, 2009), the cerebellum (Mielke et al., 1994; Minoshima et al., 1995; Ishii et al., 2001; Santens et al., 2001; Yakushev et al., 2008, 2009), the sensorimotor cortex (Minoshima et al., 1995; Santens et al., 2001; Sakamoto et al., 2002; Yakushev et al., 2008, 2009), the center of the midpontine slice (Minoshima et al., 1995; Mosconi et al., 2006), the visual cortex (Minoshima et al., 1995; Santens et al., 2001) and cluster-based normalization to the area with the highest activity in dementia patients (Yakushev et al., 2009).

As recently pointed out by Yakushev et al. (2008), the appropriate reference area should be maximally stable in patients and in healthy controls, minimally susceptible to external physiological stimuli, unaffected by the disease of interest and reliable and easy to determine. Moreover, because recent studies have shown that the chosen reference area is also important for diagnostic accuracy (Yakushev et al., 2008, 2009), we suggest adding a further criterion to this description for practical clinical reasons: The area should allow the most accurate distinction between different clinical diagnoses. This is the case if the reference area offers a maximum contrast of the differences in glucose metabolism between patients and healthy controls or between groups of patients with different disorders, so allowing easier detection and differentiation of these.

As most studies use either intensity normalization to the CMRglc (mean metabolic rate for glucose in the whole brain) or to cerebellar glucose consumption, we compared these two methods with respect to their superiority in both detecting dementia and also in differentiating between different types of dementia. These aspects are not necessarily the same, because normalizing the data to a specific region can either increase or decrease the power of statistical tests between different groups (Ishii et al., 2001; Yakushev et al., 2008; Yakushev et al., 2009). This strongly depends on the differences between groups in the reference areas. These are not the same for the comparison of clinical patients to healthy controls or to groups of patients with different disorders. For example, an earlier study showed the cerebellar glucose uptake to be
little affected in patients with AD while the CMRgluc was significantly reduced compared to healthy controls (Kushner et al., 1987). For that reason, we hypothesize differential effects of normalization for either the diagnosis or differentiation of dementia syndromes.

2.2. Methods

2.2.1. Subjects

We analyzed FDG-PET data of 19 patients (Table 2.1) with an early stage of probable AD, 13 patients with an early stage of FTLD and 10 control subjects. Probable AD was diagnosed according to the original and revised NINCDS-ADRDA criteria (McKhann et al., 1984; Dubois et al., 2007). Diagnosis of FTLD was based on criteria suggested by Neary et al. (1998). The control group included people who visited the Day Clinic of Cognitive Neurology at the University of Leipzig with subjective memory complaints, which were not objectively confirmed by a comprehensive neuropsychological and clinical evaluation. FDG-PET for these subjects was conducted for diagnostic reasons within the clinical assessment. This control group was chosen because in clinical practice it is crucial to discriminate between these subjects and patients with an early stage of dementia. Informed consent was obtained from all subjects. The research protocol was approved by the ethics committee of the University of Leipzig, and was in accordance with the latest version of the Declaration of Helsinki.
**Table 2.1** Subject group characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>AD</th>
<th>FTLD</th>
<th>ANOVA (df, F, p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>10</td>
<td>19</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>Male / female</td>
<td>6 / 4</td>
<td>7 / 12</td>
<td>6 / 7</td>
<td>-</td>
</tr>
<tr>
<td>Age (years)</td>
<td>55.1±5.0</td>
<td>62.6±6.0</td>
<td>61.1±6.6</td>
<td>2,5.26,0.009</td>
</tr>
<tr>
<td>CDR (score)</td>
<td>0.25±0.26</td>
<td>0.87±0.47</td>
<td>0.81±0.43</td>
<td>2,7.77,0.001</td>
</tr>
<tr>
<td>CMRglc/CerMRglc ratio</td>
<td>1.20±0.06</td>
<td>1.08±0.06</td>
<td>1.11±0.08</td>
<td>2,7.52,0.002</td>
</tr>
</tbody>
</table>

Mean ± standard deviation. AD Alzheimer's disease, ANOVA Analysis of variance, CerMRglc cerebellar metabolic rate for glucose, CMRglc cerebral metabolic rate for glucose, CDR Clinical Dementia Rating Scale, FTLD Frontotemporal lobar degeneration.

2.2.2. [F18]fluorodeoxyglucose positron emission tomography imaging

All PET data were acquired on a Siemens ECAT EXACT HR+ scanner (CTI/Siemens, Knoxville, TN, USA) under a standard resting condition in 2-dimensional (2D) mode. The 2D acquisition mode was used because it allows a better quantification of the PET data due to lower scatter radiation. Sixty-three slices were simultaneously collected with an axial resolution of 5 mm full width at half maximum (FWHM) and in-plane resolution of 4.6 mm. After correction for attenuation, scatter, decay and scanner-specific dead time, images were reconstructed by filtered back-projection using a Hann-filter of 4.9 mm FWHM. The 63 transaxial slices obtained had a resolution of 128*128 voxels with an edge length of 2.45 mm.

2.2.3. [F18]fluorodeoxyglucose positron emission tomography data analysis

The resultant ECAT volume files were separated into single frames (ANALYZE-format) using the import tool from the program MRicro (http://www.sph.sc.edu/comd/rorden/micro.html) and the last three frames à 10 minutes, starting from 30 to 60 minutes post injection, of each patient were chosen for further analysis. SPM5 (Statistical Parametric Mapping software: http://www.fil.ion.ucl.ac.uk/spm/) implemented in Matlab 7.7 (MathWorks Inc., Sherborn, MA) was used for further processing and statistical analysis. Each patient's frames were spatially
realigned to minimize inter-frame motion artefacts and a mean image of these three frames was calculated for each patient (Figure 2.1).

The cerebellar metabolic rate for glucose (CerMRglc) was obtained from a volume of interest (VOI), which was defined in three adjacent slices of the superior part of the right and left cerebellar hemispheres using the program P-Mod (http://www.pmod.com/technologies/index.html) supervised by an experienced physician trained in nuclear medicine. PET image sets were spatially normalized and smoothed using a Gaussian filter of 12 mm FWHM. To make the comparison of both reference regions more reliable we repeated the preprocessing procedure with a partial volume effect correction added prior to spatial normalization of the data using the modified Müller-Gärtnert method (Müller-Gärtnert et al., 1992; Rousset et al., 1998) implemented in the PVELab software (Quarantelli et al., 2004). The intensity normalization to the cerebellum was performed using the ImCalc tool provided by SPM whereby each voxel of the preprocessed mean images was divided by the individual baseline intensity values obtained from the VOIs in the cerebellum. For normalization to CMRglc, we used proportional scaling in the SPM “global normalization” option, with global mean as the arithmetical mean of voxels above the threshold of 1/8 of the grand mean of the whole image. We ran the same voxel-based statistical analysis with both normalization methods, with and without an additional partial volume effect correction of the data. To account for group differences and for age-dependent reduction in glucose metabolism in elderly subjects (Salmon et al., 2000), age was included as a covariate in the statistical analysis.

This procedure generated t-statistics for each voxel for different contrasts and constituted a statistical parametric map for the resulting t-values. The hypometabolic areas were investigated with a threshold of p<0.001 (uncorrected) and an extent threshold of 30 voxels. To analyze the intergroup differences, we used the spatial extent of clusters (number of voxels in all significant clusters) exceeding a probability threshold on a cluster level of p<0.05 (corrected for multiple comparisons) for each contrast. This statistical threshold was selected to capture a wide range of statistical differences but also to be high enough to distinguish separate clusters of metabolic differences.

The CMRglc was compared to the CerMRglc in patient groups and controls by calculating a ratio of both metabolic rates: CMRglc/CerMRglc. Group comparisons for age, CDR (Clinical Dementia Rating Scale, Morris et al., 1993), and CMRglc/CerMRglc ratios were performed by conducting ANOVAs (analyses of variance). If an ANOVA revealed a significant between-group
effect, a Bonferroni t-test was calculated with a significance threshold of \( p < .05 \) (corrected for multiple comparisons, two-tailed). Group differences regarding sex were evaluated using a Chi-square test for independence. The statistical analysis was performed with the commercial software package SPSS 17.0 (http://www.spss.com/statistics/).

![Diagram](image)

**Figure 2.1** Schematic representation of the procedure for FDG-PET data handling and processing steps.

### 2.3. Results

The Chi-square test for independence did not reveal any statistical differences in sex between the groups \( X^2(2) = 1.43; p = .49 \). CDR scores were significantly different in the three groups
The post-hoc test revealed no difference in the mean CDR scores between both groups of dementia patients, indicating a similar level of severity of dementia syndrome \([t(30)=.37; p=1.0]\). As expected, both, early AD \([t(27)=3.86; p=.002]\) and early FTLD \([t(21)=3.571; p=.009]\) patients, had significantly higher CDR scores compared to the control group. The ANOVA also revealed a significant group difference in age. The two groups of dementia patients \([t(30)=0.67; p=1.0]\) and FTLD patients and controls \([t(21)=-2.39; p=.068]\) did not differ regarding age. There was a minor but significant difference in age between control subjects and AD patients \([t(27)=-3.36; p=.008]\). Accordingly age was included as covariate in the further analysis. Likewise, the groups differed significantly in CMRglc/CerMRglc ratios. The comparison of healthy controls with patients with the Bonferroni t-test revealed a reduced ratio in FTLD \([t(21)=2.47; p=.025]\) and AD patients \([t(27)=4.14; p=.001]\), while the difference between both groups of patients was not significant \([t(30)=-0.94; p=1.0]\).

**Table 2.2** Extension of hypometabolic brain regions for different contrasts with global and with cerebellar normalization.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Global normalization</th>
<th>Cerebellar normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cluster extent</td>
<td>Maximal p-value</td>
</tr>
<tr>
<td>AD &lt; Controls</td>
<td>7988</td>
<td>3.05E-010</td>
</tr>
<tr>
<td>FTLD &lt; Controls</td>
<td>11089</td>
<td>1.34E-013</td>
</tr>
<tr>
<td>Controls &lt; AD</td>
<td>7268</td>
<td>1.95E-009</td>
</tr>
<tr>
<td>Controls &lt; FTLD</td>
<td>7863</td>
<td>1.62E-009</td>
</tr>
<tr>
<td>FTLD &lt; AD</td>
<td>20069</td>
<td>&lt;1.0E-20</td>
</tr>
<tr>
<td>AD &lt; FTLD</td>
<td>25704</td>
<td>&lt;1.0E-20</td>
</tr>
</tbody>
</table>

The size is represented by the sum of all clusters (in voxels) which exceeded an uncorrected threshold \(p < 0.001\) on the voxel level with cluster extension \(k \geq 30\) voxels and \(p < 0.05\) (corrected) on the cluster level. AD Alzheimer’s disease, FTLD Frontotemporal lobar degeneration, n.s. not significant.

Table 2.2 and Figure 2.2 list the cluster extent (voxels) of differences detected for both types of intensity normalization. The statistical analysis after normalization to the CMRglc or to the CerMRglc revealed significant hypometabolism (Figure 2.3a) of AD patients relative to controls.
in parietal, temporal, frontal and posterior cingulate cortices. The extension of significant clusters was almost twice the volume with normalization to CerMRglc compared to normalization to CMRglc. In the comparison of FTLD patients to healthy controls, the patients showed a significantly decreased metabolic rate in frontal, temporal and anterior cingulate cortices for both reference areas, but more extended regions of hypometabolism with cerebellar normalization.

Figure 2.2  Extension of clusters with significantly reduced metabolic rate of glucose exceeding an uncorrected threshold of $p < 0.001$ on the voxel level (extent threshold $k \geq 30$ voxels) and a threshold of $p<0.05$ (corrected) on the cluster level for different contrasts with cerebral and cerebellar normalization. AD Alzheimer’s disease, FTLD Frontotemporal lobar degeneration.
Figure 2.3 a) Reduced relative glucose metabolism (p<0.001 uncorrected, extent threshold ≥30 voxels) in patients with Alzheimer’s disease (AD) and in patients with frontotemporal lobar degeneration (FTLD) relative to control subjects with normalization to the cerebellar and to the cerebral (global) metabolic rate for glucose. b) Reduced glucose metabolism in FTLD relative to AD patients and in AD relative to FTLD patients normalized to the cerebellar or to the cerebral metabolic rate for glucose. Anatomical convention.

The reverse contrasts (controls < patients) showed reduced glucose metabolism in control subjects relative to AD and FTLD patients only with global normalization (Figure 2.4). However, no significantly lower metabolism was found in the statistical analysis with normalization to the cerebellum. This increased metabolic rate of AD and FTLD patients relative to control subjects
was located in the cerebellum in both groups and additionally in the sensorimotor and auditory cortices in the AD group and in parietal, posterior temporal and occipital regions in the FTLD group.

Figure 2.4 Areas of apparent relative hypermetabolism (p<0.001 uncorrected, extent threshold ≥30 voxels) obtained with normalization to the CMRglc in patients with Alzheimer's disease (AD) and frontotemporal lobar degeneration (FTLD) relative to control subjects. No relative hypermetabolism was detected with cerebellar normalization (not shown). Anatomical convention.

Comparison between the AD and FTLD patients revealed a somewhat different pattern with regard to normalization method used (Figure 2.3b). AD patients had a significantly reduced regional CMRglc in parietal, posterior temporal and occipital cortices compared to patients with FTLD. For this contrast, the CMRglc normalization revealed much more widely extended areas of significantly reduced glucose metabolism than cerebellar normalization. A similar pattern of differentiation for both reference areas was also obtained for the reverse contrast. Regional CMRglc in FTLD relative to AD patients was substantially reduced in frontal, anterior temporal
and anterior cingulate cortices, whereas the statistical analysis with CMRglc normalization also revealed larger areas of significantly reduced metabolic rate compared with normalization to CerMRglc.

The statistical analysis with an additional partial volume effect correction of the data in the preprocessing revealed a similar pattern of hypometabolism differentiation for all comparisons depending on the reference area for normalization procedure (data not shown).

### 2.4. Discussion

The statistical analysis revealed a differential effect of both normalization procedures for comparison of dementia patients to either control subjects or to patients with a different type of dementia. These results are highly relevant for differential diagnosis of dementia. Consistent with previous studies which compared the two reference areas for AD (Ishii et al., 2001; Yakushev et al., 2008), cerebellar normalization was superior to global normalization in differentiating between controls and both groups of dementia patients. Quantitatively, the extent of areas with detected hypometabolism in patients was almost twice as large for the AD group and almost one and a half times as large for the FTLD group with cerebellar normalization relative to global normalization. This indicates a clear advantage of the first reference area for the detection of dementia. The locations of hypometabolic areas were similar to previous reports both for patients with probable AD (Ishii et al., 2001; Buchert et al., 2005) and for patients with FTLD (Jeong et al., 2005, Schroeter et al., 2007, 2008).

An opposite pattern for both reference areas was obtained in the comparison of patients with probable AD and FTLD. The contrast for hypometabolic regions in AD relative to FTLD patients showed a reduced regional CMRglc in occipital, posterior temporal and parietal cortices. Surprisingly, the extent of significant clusters was three times as large with CMRglc normalization as with normalization to CerMRglc. The reverse contrast revealed a similar tendency depending on the reference area used. Regional CMRglc in FTLD patients was reduced in frontotemporal areas and the cluster size was in fact almost five times as large with normalization to the CMRglc. To our knowledge, this is the first comparison between
qualitatively different groups of dementia patients investigating the effect of the reference area. The results of this comparison indicate that although the CMRglc does not appear to be the optimal choice as reference area for detecting dementia, it has a clear advantage in distinguishing between different dementia syndromes.

The differential effect of both reference areas can be explained regarding the CMRglc / CerMRglc ratio. This ratio did not differ significantly between both groups of patients and was lower than that of healthy controls. For this reason, normalization to the CMRglc does not remove between-group differences in the global metabolism for both patient groups, as it does in the comparison with healthy controls, but it actually increases the statistical power due to the higher CMRglc compared with CerMRglc and due to the higher signal-to-noise ratio with normalization to the CMRglc compared with the cerebellum as reference area. The higher signal-to-noise ratio with normalization to the CMRglc is a result of the greater number of voxels used for the estimation of the mean intensity in the whole brain compared with the cerebellar reference area, so reducing the effect of stochastic noise. This result has important implications for the differentiation between different types of dementia. Studies investigating the diagnostic accuracy of neuropsychological testing for the differential diagnosis of AD vs. FTLD have revealed an overall accuracy of about 90% for correctly classifying patients with AD. However, accuracy was substantially lower for correctly classifying FTLD patients, with values of only about 64% to 77% (Rascovsky et al., 2002; Walker et al., 2005). Comparing FDG-PET scans of dementia patients not only to those of healthy controls but also to the disease-specific pattern of regional CMRglc in different kinds of dementia might significantly facilitate and improve the differential diagnosis of these as has been shown by Mosconi et al. (2008). To maximally increase the efficiency of these comparison-based differentiations, it is necessary to take into account that the same reference area for intensity normalization can, on the one hand, increase the statistical power in a comparison with healthy controls. However, on the other hand it can also worsen the differentiation of groups of patients. Therefore, thorough evaluation of different reference regions with respect to the diagnostic questions is required.

The apparent relative hypermetabolism of dementia patients obtained with global normalization has been reported previously but only in AD patients (Salmon et al., 2000; Ishii et al., 2001; Samuraki et al., 2007). In our study, we also found a similar effect for FTLD patients. This hypermetabolism seems to be an artifact of global normalization, as we did not find regional hypermetabolism if the cerebellum was chosen as reference area. This finding is also supported
by the significantly different ratio of the global to the cerebellar mean between control subjects and dementia patients. Because earlier studies have shown the cerebellar metabolic rate to be preserved in dementia (Kushner et al., 1987), differences in the ratio can mainly be attributed to the reduced CMRglc in the patient groups. Therefore, intensity normalization to the CMRglc removes global uptake differences between groups, which results in method-created hypermetabolism in some relatively preserved regions in dementia patients. This result and conclusion are in accordance with observations in AD patients made in previous studies (Buchert et al., 2005; Yakushev et al., 2008). Moreover, our results are in line with recent studies using other imaging modalities that indicated a regional hyperperfusion as an artifact created by the normalization to the global mean in neurological disorders such as Parkinson’s disease (Borghammer et al., 2009). Additionally, our results may indicate that besides the cerebellum the primary cortices are unaffected in the early stages of both types of dementia.

Our study suggests applying normalization to the CerMRglc for diagnosis, and normalization to the CMRglc for differential diagnosis of dementia syndromes, at least for AD and FTLD. Moreover, both normalization approaches could be easily applied in the clinical environment. First of all, normalization to the CMRglc is already integrated into the SPM software package. Secondly, normalization to the cerebellum provides an easily determined and reliable alternative to the sensorimotor cortex which is also frequently used (Minoshima et al., 1995; Santens et al., 2001; Sakamoto et al., 2002; Yakushev et al., 2008, 2009), but which has the disadvantage that the surface appearance and structural organization varies greatly from subject to subject (White et al., 1997; Geyer et al., 1999).

Normalization to the visual cortex, which is also used, can be problematic because of it's susceptibility to visual stimulation (Newberg et al., 2005), thereby increasing the variance among the data and so reducing the sensitivity of statistical tests. A similar problem of increased variance in the data might occur with normalization to the center of the midpontine slice. Due to the small size of this VOI, extracting data from such a small region would result in a decreased signal-to-noise ratio.

Recently it has been suggested to first use CMRglc normalization to get the cluster with the highest metabolic rate in the patient group (Yakushev et al., 2009). As the next step, the absolute counts for each subject within this cluster are extracted and the FDG-PET data are normalized to these cluster-derived counts. Despite the fact that this normalization method provides a substantial increase in statistical power thus allowing good discrimination between
different groups, it has several disadvantages. For one thing, it is a circumstantial algorithm, requiring additional processing steps, which might be difficult to apply in the clinical environment. Furthermore, it is a data-driven process with limited generalizability to data of new patients, because the position of the region with the highest metabolic rate and its extent may differ substantially across patients with the same type of dementia.

2.5. Conclusions

The present study investigated the impact of normalization strategies on diagnostic accuracy in dementia. Whereas cerebellar normalization seems to be superior in early detection and diagnosis of dementia, normalization to the cerebral metabolic rate for glucose may be superior in differential diagnosis of various dementia syndromes.
3. Multimodal imaging in dementia using \([F18]\)fluorodeoxyglucose positron emission tomography and magnetic resonance imaging

Different imaging techniques have been shown to provide useful information to improve understanding of functional mechanisms and development of various neurodegenerative disorders such as AD and FTLD (Mielke et al., 1994; Salmon et al., 2000; Chetelat et al., 2008; Desgranges et al., 2007; Schroeter et al., 2007, 2009). Additionally, in recent research various biomarkers have been reported to differentiate between early stages of dementia and healthy control subjects or between different types of neurodegenerative disorders, suggesting an integration of these would improve diagnostic accuracy of dementia (Hoffman et al., 2000; Rosen et al., 2002; Diehl et al., 2004; Jeong et al., 2005; Diehl-Schmid et al., 2007; Edison et al., 2007; Fung and Stoeckel, 2007; Sabri et al., 2008; Schroeter et al., 2007, 2009). The following section investigates differential changes in FDG-PET and MRI in AD and FTLD patients and systematically compares the use of these biomarkers for detection and differentiation of both dementia syndromes.

3.1. Introduction

For the detection of dementia, accuracy rates significantly above 90% have recently been reported using univariate and multivariate statistical approaches in magnetic resonance imaging (MRI) and \([F18]\)fluorodeoxyglucose positron emission tomography (FDG-PET) (Hoffman et al., 2000; Matsunari et al., 2007; Klöppel et al., 2008a; Davatzikos et al., 2008; Fan et al., 2008; Sadeghi et al., 2008). However, the differentiation of the two most common types of dementia, namely Alzheimer's disease (AD) and frontotemporal lobar degeneration (FTLD), is still problematic. For this differentiation, accuracy rates ranging between 84 and 89% are still in need of improvement, especially due to a substantially lower sensitivity compared with specificity of actual methods (Knopman et al., 2005; Klöppel et al., 2008a; Davatzikos et al., 2008). Nevertheless, the use of biomarkers has significantly helped to improve diagnostic accuracy compared with diagnoses based solely on clinical and neuropsychological evaluation (Raskovsky et al., 2002; Dubois et al., 2007). For these reasons, recent studies have suggested
to incorporate imaging findings into criteria for diagnosis of dementia (Dubois et al., 2007; Kipps et al., 2009).

For AD patients imaging studies have shown reduced glucose consumption mainly in parietotemporal and posterior cingulate cortices (Ishii et al., 2001, 2005; Yakushev et al., 2008; Kanda et al., 2008; Schroeter et al., 2009) and structural changes in the hippocampus and entorhinal area relative to healthy controls (Van de Pol et al., 2006; Kanda et al., 2008; Schroeter et al., 2009). In FTLD patients, atrophy and reduced metabolic rate for glucose have been reported to be predominately located in the medial thalamus, amygdala and in frontotemporal and anterior cingulate cortices (Ishii et al., 2005; Jeong et al., 2005; Schroeter et al., 2007, 2008; Kanda et al., 2008).

Different univariate and multivariate approaches were proposed for the detection and differentiation of these changes to enable a more accurate detection and differential diagnosis of dementia. Univariate approaches range from the most frequently used voxel-based morphometry (VBM) at the whole-brain or region-of-interest (ROI) level (Rosen et al., 2002, 2005; Desgranges et al., 2007; Rabinovici et al., 2007; Chetelat et al., 2008) to volumetric assessment of specific regions (Perry et al., 2006) and measurement of cortical thickness (Lerch et al., 2005; Querbes et al., 2009). For multivariate differentiation of different types of dementia support vector machine classification (SVM) is used based on whole-brain voxel information (Klöppel et al., 2008a) or most frequently on ROI values (Davatzikos et al., 2008; Fung et al., 2007; Fan et al., 2008; Chaves et al., 2009; Horn et al., 2009). However, a major problem of the ROI-based approach is the limited generalizability of the trained classifier, because the ROIs are selected based on features showing a between-group differentiation in the same groups in a univariate analysis. Although ROIs selected with this method provide a good discrimination between groups used in these specific studies, they might show significantly reduced discrimination power when applied to new data sets. This could be the case if the selected regions just detect differences between groups, which are not necessarily attributed to the specific neurodegenerative disorder.

Here, we investigate the contribution of multimodal imaging using FDG-PET and high-resolution MRI for a better understanding and differentiation of AD and FTLD. We apply VBM and SVM as the most frequently used univariate and multivariate approaches to evaluate their contribution for understanding, detection and differentiation of dementia in multimodal imaging. Additionally, we apply SVM classification on data extracted from ROIs based on disorder-specific metabolic
reductions and atrophy reported in comprehensive meta-analyses investigating AD and FTLD. This method allows a better generalization of our classification algorithms to other clinical centers and ensures that only disorder-specific changes are used for SVM based discrimination. We hypothesize that common use of different imaging modalities might substantially improve early detection and differentiation of dementia.

3.2. Methods

3.2.1. Subjects

**Table 3.1** Subject group characteristics

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>AD</th>
<th>FTLD</th>
<th>ANOVA (df,F,P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>13</td>
<td>21</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Male/Female</td>
<td>7/6</td>
<td>9/12</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>53.9±6.0</td>
<td>61.1±6.7</td>
<td>60.8±6.4</td>
<td>2, 5.76, 0.006</td>
</tr>
<tr>
<td>CDR (score)</td>
<td>0.23±0.26</td>
<td>0.71±0.25</td>
<td>0.82±0.42</td>
<td>2, 13.93, 0.000</td>
</tr>
</tbody>
</table>

Mean ± standard deviation. AD Alzheimer’s disease, ANOVA analysis of variance, CDR Clinical Dementia Rating Scale, FTLD frontotemporal lobar degeneration.

We analyzed FDG-PET and T1-weighted MRI data of 21 patients (Table 3.1) with an early stage of probable AD, 14 patients with an early stage of FTLD and 13 control subjects. Probable AD was diagnosed according to the original and revised NINCDS-ADRDA criteria (McKhann et al., 1984; Dubois et al., 2007). Diagnosis of FTLD was based on criteria suggested by Neary et al. (1998). The control group included subjects who visited the Day Clinic of Cognitive Neurology at the University of Leipzig with subjective memory complaints, which were not objectively confirmed by a comprehensive neuropsychological and clinical evaluation. FDG-PET and MRI for these subjects was conducted for diagnostic reasons within the clinical assessment. This
control group was chosen because, in clinical practice, it is crucial to discriminate between these subjects and patients with an early stage of dementia. Informed consent was obtained from all subjects. The research protocol was approved by the Ethics Committee of the University of Leipzig, and was in accordance with the latest version of the Declaration of Helsinki.

3.2.2. Data acquisition

a) Magnetic resonance imaging data

For each subject, a high-resolution T1-weighted MRI scan was obtained, consisting of 128 sagittal slices adjusted to AC-PC line and a with slice thickness of 1.5mm and pixel size of 1×1mm². MRI was performed on two different 3T scanners (MedSpec 30/100, Bruker Biospin, Ettlingen Germany and Magnetom Trio, Siemens, Erlangen, Germany) using two different T1-weighted sequences (MDEFT or MP-RAGE with TR=1300ms, TI=650ms, TE=3.93ms or TE=10ms; FOV 25×25 cm²; matrix = 256×256 voxels). On the MedSpec scanner, only the MDEFT-sequence and on the Magnetom Trio scanner, either MDEFT or MP-RAGE sequences were used.

b) [F18]fluorodeoxyglucose positron emission tomography data

Each subject also underwent FDG-PET imaging either a few weeks before or after the MRI scan. All PET data were acquired on a Siemens ECAT EXACT HR+ scanner (CTI/Siemens, Knoxville, TN, USA) under a standard resting condition in 2-dimensional (2D) mode. The 2D acquisition mode was used because it allows a better quantification of the PET data due to lower scatter radiation. Sixty-three slices were simultaneously collected with an axial resolution of 5 mm full width at half maximum (FWHM) and in-plane resolution of 4.6 mm. After correction for attenuation, scatter, decay and scanner-specific dead-time, images were reconstructed by filtered back-projection using a Hann-filter of 4.9 mm FWHM. The 63 transaxial slices obtained had a matrix of 128×128 voxels with an edge length of 2.45 mm.
3.2.3. Image processing

The procedure described below has been specifically designed for this study, aiming at a most accurate co-processing of FDG-PET and MRI data to allow the comparison of quantitative and qualitative differences in the distribution of hypometabolic and atrophic regions in both groups of dementia patients (Figure 3.1). All image-processing steps and the subsequent univariate statistical analysis were carried out using the SPM5 software package (Statistical Parametric Mapping software: http://www.fil.ion.ucl.ac.uk/spm/) implemented in Matlab 7.7 (MathWorks Inc., Sherborn, MA). SVM classification was conducted with the LIBSVM software (Chang and Lin, 2001) using the Matlab interface.
**Figure 3.1** Schematic representation of the procedure for FDG-PET and MRI data handling and processing steps. FDG-PET [F18]fluorodeoxyglucose positron emission tomography, MRI magnetic resonance imaging.

a) Magnetic resonance images

The MR images were first interpolated to get an isotropic resolution of $1 \times 1 \times 1$ mm$^3$. The resultant MR images were coregistered on their respective FDG-PET images and bias corrected for inhomogeneity artifacts using the Unified Segmentation Approach described in detail by Ashburner and Friston (2005). This specific method performs a better coregistration of images from different modalities and allows a more accurate segmentation due to the bias correction. A further reason to use this approach was that the straightforward coregistration implemented in the PVElab software described later sometimes failed. We used this software for automatic partial volume correction of the FDG-PET images. The coregistered MR images were processed using the DARTEL approach (Ashburner, 2007) to enable a voxel-based morphometric comparison. This approach registers all gray matter (GM) and white matter (WM) images to an averaged-size template created from all subjects used in this study and preserves the total amount of signal from each region in the images. Subsequently, the images were smoothed using a Gaussian kernel of 12 mm FWHM. This smoothing factor, although higher than usual MR kernels, was selected based on extensive tests, because it allows the optimal coevaluation with lower resolution FDG-PET images.

b) [F18]fluorodeoxyglucose positron emission tomography data

Within the common registration with MRI data using the Unified Segmentation Approach described above, PET images were interpolated to the same voxel size as the MR images, namely $1 \times 1 \times 1$ mm$^3$. This processing does not introduce any additional noise into the PET images. However, in our experience, it substantially improves the subsequent partial volume effect (PVE) correction of voxels representing GM intensities using the modified Müller-Gärtner method (Müller-Gärtner et al., 1992; Rousset et al., 1998). Due to the interpolation, they are exactly overlaid with the MR tissue class images of the same subject obtained from the segmentation step in the PVE approach. Thus, the within-voxel correction is done only for those
voxels directly overlaying the GM structures in the MR images. Instead of smoothing the MR data to the resolution of PET data and thus loosing the exact quantitative and qualitative information of GM distribution, which is usually done in the PVE correction, the interpolation of FDG-PET preserves this information. This allows a more accurate correction of atrophy effects onto glucose utilization and so a better quantitative comparison of atrophy and hypometabolism in neurodegenerative disorders. The subsequent PVE correction including all image processing steps was done by using the automatic algorithm implemented in the PVElab software package (Quarantelli et al., 2004). Because the modified Müller-Gärtner method sets all WM voxel values to the mean WM intensity value, these regions do not contain any further valuable regional information after the PVE correction. For this reason, all voxels belonging to WM were masked using the ImCalc function in the SPM5 software package by filtering this specific intensity in the whole image. After the PVE correction, the DARTEL flow fields calculated from the MR images were applied to their respective PET images to obtain an anatomically exact overlap between GM and PET images of all subjects with modulation to preserve the total amount of signal from each region. In the same way as the MR data, the PET data were smoothed by a Gaussian kernel of 12 mm FWHM. Finally, the FDG-PET data were intensity normalized using cerebellar ROIs to account for individual differences in global PET measures. This region has been shown to be least affected in mild to moderate stages of AD (Ishii et al., 1997). Additionally, normalization to this region improves the statistical discrimination between dementia patients and control subjects in comparison to other regions reported in the literature (Yakushev et al., 2008, 2009; Dukart et al., 2010).

c) Masking

The MR and PET images obtained as described above were masked to avoid contamination by misclassified voxels. Voxels lying between WM and ventricular cerebrospinal fluid tend to be misclassified as GM voxels due to their similar intensity. The mask was obtained after extensive testing by excluding all voxels in the first and the last template created by the DARTEL approach with a probability of below 0.2 for belonging to GM and including only the voxels that exceed this threshold in both templates. This mask was applied twice: firstly prior to smoothing to avoid misclassification, and secondly, after the smoothing to avoid big edge effects. WM images were exclusively masked using the same mask to avoid overlaps between GM and WM.
voxels due to smoothing. The masked images were used for the subsequent VBM and SVM analysis of the data.

d) Region-of-interest extraction

ROI coordinates were extracted from two comprehensive, systematic and quantitative meta-analyses investigating biomarkers of AD and FTLD in MR and FDG-PET images. The meta-analyses included a total number of 1618 patients (AD/FTLD: 1351/267) and 1448 healthy control subjects (1097/351) (Schroeter et al., 2007, 2009). These meta-analyses extracted the prototypical networks of AD and FTLD by applying what is currently the most sophisticated and best-validated of coordinate-based voxel-wise meta-analyses, anatomical likelihood estimate. In the FTLD meta-analyses (Schroeter et al., 2007) only coordinates which are common to all subgroups of FTLD patients were used. Because the coordinates in both meta-analyses were reported in the Talairach space, they were transformed to MNI space according to a formula proposed by Matthew Brett (published on the Internet: http://www.mrccbu.cam.ac.uk/Imaging/Common/mnispace.shtml). DARTEL preprocessed data are registered to an averaged size template created from all subjects in this study. To transform these data to the MNI space we normalized them to an a priori MNI template in SPM by using affine-only spatial normalization. Due to the affine-only transformation, our images still differed in shape from the MNI template, so some reported coordinates were slightly outside of the anatomic regions in our imaging data. In this case, the center coordinates for the ROIs were moved slightly towards the closest point of the corresponding anatomical region reported in the meta-analyses. ROIs were selected using the 3D fill tool in the MRICron software package (http://www.sph.sc.edu/comd/rorden/mricron).
Figure 3.2 Regions of interest extracted from gray matter (left) and FDG-PET (right) data for AD and FTLD patients and used for support vector machine classification projected onto a glass brain (top) and onto an axial slice (bottom). AD Alzheimer's disease, FDG-PET [F18]fluorodeoxyglucose positron emission tomography, FTLD frontotemporal lobar degeneration, GM gray matter, ROIs regions of interest.

Separate ROI masks were created for MR and FDG-PET images based on the origin of the peak values reported in the meta-analyses. Each ROI was restricted to a sphere with a radius of 5 mm around the reported coordinate (Figure 3.2). Additionally, to increase the signal-to-noise ratio, all zero voxels and edge voxels with an intensity deviation of 13 intensity units in the MRlcron 3D fill tool were excluded from the ROI. The edge voxel restriction excludes all voxels at the edge of the smoothed GM structures within the sphere. These voxels carry much less information due to their further distance from the GM structures in the unsmoothed data and so decrease the signal-to-noise ratio in the corresponding ROI.
3.2.4. Statistical analysis

a) Voxel-based morphometry

To evaluate group differences between AD, FTLD and control subjects, voxel-wise independent sample t-tests were calculated for the whole-brain for FDG-PET, GM and WM images. Age and sex were included as covariates in all t-tests to account for age-dependent reduction in elderly subjects (Salmon et al., 2000). In order to exclude possible scanner type or sequence effects on the between group differentiation, further analyses were calculated. Firstly, the analysis was repeated twice for the subsets of MDEFT and MP-RAGE images separately. Secondly, the analysis was performed again for the subset of MR images from the Siemens scanner, both with and without the acquisition sequence being considered as an additional covariate. Hypometabolic and atrophic regions were investigated with a threshold of p<0.001 (uncorrected) on the voxel level and p<0.05 (corrected for multiple comparison) on the cluster level. To analyze the intergroup differences, we used the spatial extent of clusters (number of voxels in all significant clusters) exceeding this threshold. Due to the same resolution of FDG-PET and MR data after the interpolation the number of significant voxels can be used for a quantitative comparison of atrophy and hypometabolism in both groups.

b) Support vector machine classification

Multivariate pattern classification, as described in Klöppel et al. (2008a), was performed with a linear kernel by identifying a separating hyperplane that maximizes the distance between different clinical groups based on whole-brain or ROIs information. The cross-validation of the trained SVM was performed by using the leave-one-out method. This procedure iteratively leaves out the information of each subject and trains the model on the remaining subjects for subsequent class assignation of the person that was not included in the training procedure. This validation method enables the generalization of the trained SVM to data that have never been presented to the SVM algorithms previously. The reported accuracy is the percentage of subjects correctly assigned to the clinical diagnosis. Usually SVM classification is performed without smoothing of the data, because single voxels are assumed to contain information, for example, for prediction of future action based on functional MR images. However, in
neurodegenerative disorders single voxels are unlikely to contain generalizable information due to a limited across-subject registration of MR and FDG-PET images. Although SVM classification based on unsmoothed data has been shown to differentiate reasonably between different groups (Klöppel et al., 2008a), an additional smoothing should make this approach more reliable and generalizable to new data. To control for the effect of smoothing we ran the same whole-brain classification twice for GM, PET and for integration of GM and PET in the same vector with and without smoothing.

We performed the whole-brain SVM classification using GM, WM or FDG-PET images separately and by combining information from different modalities. For the SVM classification, all data of a subject are transformed into a vector, with information of an additional modality simply attached by extending the vector. Additionally, we repeated the whole-brain SVM classification by adding MR to FDG-PET information combining both modalities in a single image. ROI-based SVM classification was performed on data extracted from smoothed images separately for GM and FDG-PET images and also by integrating information from both modalities in a single vector. In order to reduce the number of voxels in the ROI-based classification, only nonzero voxels were included in the vector. This was done because otherwise the whole-brain SVM classification is a highly memory-consuming approach. To ensure that our classification results were not based on factors randomly discriminating between groups, we reran the whole-brain and ROI-based classification for comparison 30 times by randomly assigning all subjects to the three groups independently from the clinical diagnosis and calculating the classification accuracy by using the leave-one-out procedure described above.

c) Group comparisons

Group comparisons for age and CDR (Clinical Dementia Rating Scale, Morris, 1993) were performed by conducting ANOVAs (analyses of variance). If an ANOVA revealed a significant between-group effect, a Bonferroni t-test was calculated with a significance threshold of p<.05 (corrected for multiple comparisons, two-tailed). Group differences regarding sex were evaluated using a chi-square test for independent samples. The statistical analysis was performed with the commercial software package SPSS 17.0 (http://www.spss.com/statistics/).
3.3. Results

The chi-square test for independent samples did not reveal any statistical differences in sex between the groups [$\chi^2(2)=0.42; p=0.809$]. CDR scores differed significantly in the three groups (Table 1). The post-hoc test revealed no differences in the mean CDR scores between both groups of dementia patients indicating a similar severity of dementia syndrome [$t(33)=0.94; p=0.977$]. As expected, both early AD [$t(32)=5.36; p<0.001$] and early FTLD [$t(25)=4.35; p<0.001$] had significantly higher CDR scores compared to the control subjects. The ANOVA also revealed a significant group difference in age. The two groups of dementia patients did not differ significantly in age [$t(33)=0.16; p=1.0$]. There was a minor but significant difference between AD patients and controls [$t(32)=3.18; p=.008$] and FTLD patients and controls [$t(25)=2.86; p=.024$]. To avoid any possible effects of age and sex, both variables were included as covariates in the further analysis.
3.3.1. Voxel-based morphometry

**Figure 3.3** Extension of clusters with significantly reduced metabolic rate of glucose or gray matter intensity exceeding an uncorrected threshold of p<0.001 on the voxel level and a threshold of p<0.05 (corrected) on the cluster level for different contrasts. AD Alzheimer’s disease, FTLD frontotemporal lobar degeneration.

Figures 3.3 and 3.4, and Table 3.2 illustrate significant reduction in GM values, WM values, and FDG-PET values in AD and FTLD patients relative to control subjects. AD patients showed extensive reduction in glucose utilization in precuneal, posterior cingulate, lateral parietal, posterior temporal, and left frontal cortices (Figure 3.4). GM values were significantly reduced in left frontal, left thalamic, parahippocampal, posterior cingulate, precuneal, lateral parietal, and posterior temporal regions. WM intensity reduction was mainly restricted to the splenium of the corpus callosum and to posterior regions predominating on the left side. The extension of significant clusters for hypometabolism substantially exceeded the extension of atrophy in most affected regions except for the thalamus (Figure 3.3). In FTLD patients, GM intensity reduction was observed predominately in medial and lateral regions of the frontal lobe but also extending
to the insula, left anterior temporal lobe, hippocampus, amygdala, thalamus, putamen and bilaterally to the head of nucleus caudatus (Figure 3.4). Hypometabolic regions were mainly restricted to cortical structures in the medial and lateral frontal cortices, left insula and left temporal lobe slightly extending to the parietal cortex. Significant WM intensity reductions were detected mainly in the anterior parts of the corpus callosum, peduncular tracts and in anterior regions. All reductions in GM, WM and FDG-PET values in FTLD patients were predominantly located on the left side (Figure 3.4). The extension of atrophy exceeded the reduction of glucose utilization in most significant structures except for dorsolateral regions of the prefrontal cortex.

Figure 4 Reductions in glucose metabolism (red), gray matter intensity (blue) in top row, and white matter intensity (cyan) in bottom row in patients with Alzheimer's disease (AD, top row) and frontotemporal lobar degeneration (FTLD).

Comparing AD and FTLD patients, there was no difference in GM and WM values exceeding the significance threshold. Metabolic rate was substantially reduced in AD in comparison to
FTLD patients in posterior cingulate and precuneal regions. The reverse contrast did not reveal any significant reduction. Additional analyses for the different subgroups, separated by MRI scanner type or sequence and including the scan sequence as additional covariate did not reveal any substantial differences from the pattern described above.

3.3.2. Support vector machine classification

Table 3.3  Accuracy rates for whole-brain and ROI-based SVM classification for FDG-PET and MRI

<table>
<thead>
<tr>
<th></th>
<th>AD, FTLD and Controls</th>
<th>AD vs FTLD</th>
<th>AD vs Controls</th>
<th>FTLD vs Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM whole-brain</td>
<td>72.9%</td>
<td>80.0%</td>
<td>88.2%</td>
<td>77.8%</td>
</tr>
<tr>
<td>WM whole-brain</td>
<td>66.7%</td>
<td>74.3%</td>
<td>79.4%</td>
<td>77.8%</td>
</tr>
<tr>
<td>FDG-PET whole-brain</td>
<td>81.3%</td>
<td>82.9%</td>
<td>94.1%</td>
<td>92.6%</td>
</tr>
<tr>
<td>GM/ FDG-PET whole-brain</td>
<td>79.2%</td>
<td>82.9%</td>
<td>94.1%</td>
<td>88.9%</td>
</tr>
<tr>
<td>GM/WM/FDG-PET whole-brain</td>
<td>77.1%</td>
<td>82.9%</td>
<td>91.2%</td>
<td>85.2%</td>
</tr>
<tr>
<td>GM + FDG-PET whole-brain</td>
<td>81.3%</td>
<td>88.6%</td>
<td>91.2%</td>
<td>88.9%</td>
</tr>
<tr>
<td>GM ROIs</td>
<td>56.3%</td>
<td>60.0%</td>
<td>82.4%</td>
<td>85.2%</td>
</tr>
<tr>
<td>FDG-PET ROIs</td>
<td>75.0%</td>
<td>80.0%</td>
<td>94.1%</td>
<td>85.2%</td>
</tr>
<tr>
<td>GM/FDG-PET ROIs</td>
<td>91.7%</td>
<td>94.3%</td>
<td>100.0%</td>
<td>92.6%</td>
</tr>
</tbody>
</table>

Accuracy represents the percentage of subjects correctly assigned to the correct condition. AD Alzheimer’s disease, FDG-PET [18F]fluorodeoxyglucose positron emission tomography, FTLD frontotemporal lobar degeneration, GM gray matter, MRI magnetic resonance imaging, ROI region-of-interest, SVM support vector machine, WM white matter.
Figure 3.5 Voxels most relevant for classification of both groups of patients and control subjects in FDG-PET and MRI after SVM training. AD and FTLD vs Controls: Blue and light blue indicate decreased gray matter intensity or reduced metabolic rate that increase the likelihood of classification into a dementia group. Red and yellow indicate the opposite. AD vs FTLD: Blue and light blue indicate decreased gray matter intensity or reduced metabolic rate that increase the likelihood of classification into the AD group. Red and yellow indicate the opposite. AD Alzheimer’s disease, FTLD frontotemporal lobar degeneration. MRI magnetic resonance imaging, PET positron emission tomography.
Multivariate classification of the data using SVM at the whole-brain level revealed the best discrimination accuracy for all three groups using FDG-PET, with 81% (chance level 33%), in comparison to GM and WM information, with lowest accuracy using WM information on its own (Table 3.3). The combination of metabolism and GM values in a single image revealed a similar accuracy for differentiation of the three groups, with higher accuracy for differentiation between both types of dementia, however, with slightly lower discrimination between dementia patients and control subjects. Whole-brain SVM classification for the three groups without smoothing revealed lower accuracy rates in all classifications in comparison to differentiation based on smoothed images. The accuracy increase due to smoothing ranged between 2 (GM) and 6% (FDG-PET). Figure 3.5 displays regions that were most influential in making binary classification between the AD, FTLD and control subjects based on smoothed whole-brain information.

Accuracy based on ROIs from both meta-analyses using only GM information was substantially lower for differentiation between AD and FTLD patients in comparison with whole-brain classification. However, it was comparable to the whole-brain approach in differentiating between patients with both types of dementia and control subjects. ROIs extracted from FDG-PET data showed slightly lower discrimination accuracy compared to whole-brain information. The best accuracy rates of all SVM classifications were obtained using combined information extracted from FDG-PET and GM data.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD vs FTLD</td>
<td>94.3%</td>
<td>95.2%*</td>
<td>92.9%</td>
</tr>
<tr>
<td>AD vs Controls</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>FTLD vs Controls</td>
<td>92.6%</td>
<td>85.7%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

*Considering a correctly identified AD as a true positive. AD Alzheimer’s disease, FDG-PET [18F]fluorodeoxyglucose positron emission tomography, FTLD frontotemporal lobar degeneration, MRI magnetic resonance imaging.
This approach resulted in a classification accuracy of 92% for the differentiation of all three groups and an accuracy rate of 94% for differentiation between AD and FTLD patients. Sensitivity of this ROI-based classification ranged between 85.7% for FTLD and 100% for AD and specificity of 100% for discrimination of both types of dementia from control subjects (Table 3.4).

The classification accuracy in the 30 trials randomly assigning all subjects to the three groups resulted in a mean accuracy rate of 34±7.7% (mean ± standard deviation), ranging between 21 and 52% for the ROI-based SVM classification, and 33.7±8.2% ranging between 12 and 50% for the whole-brain classification.

3.4. Discussion

In this study we performed a multimodal comparison and discrimination of dementia patients using FDG-PET and MRI. To enable a quantitative evaluation and comparison of differences in both imaging modalities, we developed a new preprocessing algorithm. This algorithm was designed to enable an accurate anatomical registration of both modalities. All processing steps were performed as far as possible simultaneously by applying the same deformations and preprocessing parameters to both modalities of the same subject. This procedure resulted in an accurate anatomical overlap of both imaging modalities and in an accurate between-subject registration, with both images having the same voxel size and approximately the same effective smoothness to allow a direct volumetric comparison of differences to control subjects in both imaging modalities.

Previous studies have performed this comparison based mainly on qualitative differences by comparing the mean reduction of glucose and GM intensity values in specific regions with each other (Ishii et al., 2005; Chetelat et al., 2008; Kanda et al., 2008). Furthermore, in most studies, FDG-PET and MRI as biomarkers were investigated separately only for AD or FTLD patients relative to control subjects (Ishii et al., 2005; Kawachi et al., 2006; Desgranges et al., 2007). Hence, a direct comparison of the distribution of hypometabolism and atrophy in both disorders is limited due to different control groups and different processing algorithms. Only one study...
compared AD patients to frontotemporal dementia (one subtype of FTLD) in both modalities (Kanda et al., 2008). However, the main disadvantage of this study is the missing PVE correction, which strongly restricts the interpretation and comparison of the distribution and amount of hypometabolic areas relative to atrophic regions. In addition, Kanda et al. (2008) do not provide a volumetric evaluation of the amount of hypometabolism and atrophy relative to each other.

Due to the preprocessing algorithm described above our data allow the quantitative evaluation of differences between different types of dementia in different imaging modalities. By performing PVE correction we excluded that reductions in glucose metabolism can be ascribed to atrophy in these specific regions. Furthermore, the results of the VBM analysis in our study are not dependent on the MRI scanner or sequence used, as has been shown by additional analyses of different subgroups and by including the sequence as covariate. High discrimination rates in SVM, despite using different scanner types and sequences, indicate the reliability of our approach. This result is supportive of the potential of our approach to be applied to imaging data from different imaging centers and increases its potential value for clinical diagnostic applications.

3.4.1. Results of voxel-based morphometry

The statistical analysis revealed a differential pattern of atrophy and glucose hypometabolism in both types of dementia relative to control subjects. In AD, the amount of hypometabolism exceeded the amount of GM atrophy in most affected structures. This result is consistent with observations of hypometabolism exceeding atrophy in most affected areas reported in previous studies investigating AD (Matsunari et al., 2007; Samuraki et al., 2007; Chetelat et al., 2008; Kanda et al., 2008). However, the opposite pattern was observed for FTLD patients. In this group, the extent of atrophic regions was larger in comparison to hypometabolic regions relative to control subjects. Generally, the amount of GM and WM atrophy in FTLD patients substantially exceeded the amount of both in AD patients. The opposite pattern was observed for hypometabolism. Because both groups of dementia patients had a similar mild to moderate stage of dementia, these results indicate a differential development of both dementia types regarding the underlying mechanisms. In AD, at least at this early stage of dementia, the hypometabolism seems to be the predominating factor, while atrophy is less pronounced in this
group of patients. This observation indicates that changes in glucose metabolism, which is strongly connected to glutamatergic neuronal activity (Sibson et al., 1998), precede or have at least a faster mechanism in AD. The atrophy in this disease seems to be a slower process following the reduction of glucose in cortical structures. This assumption is consistent with research reporting beta amyloid, a histological feature in AD (Sisodia et al., 1990), to impair glucose transport in cortical and hippocampal neurons and in this way to contribute to neuronal degeneration in AD (Mark et al., 1997).

FTLD patients show an inverse pattern of atrophy and hypometabolism in comparison to AD patients. In this group, GM atrophy substantially exceeds the reduction of glucose metabolism in most cortical and in all subcortical structures. Furthermore, WM atrophy is much more pronounced in this group. These results suggest a different process not primarily affecting glucose consumption to explain neurodegeneration in FTLD. Extended WM atrophy also indicates a mechanism that is not restricted to GM structures, which might be supportive of the recently reported results of TAR DNA-binding protein (DNA: deoxyribonucleic acid), TDP-43, to be the major disease protein in FTLD (Davidson et al., 2007). This protein has been reported to be not only restricted to GM structures in FTLD patients but also present in frontal and temporal WM structures and brainstem (Neumann et al., 2007), which is consistent with our results of WM intensity reductions in these regions.

3.4.2. Support vector machine results

As shown above, VBM using different imaging modalities might provide a helpful instrument to understand the underlying mechanisms of different types of dementia. However, it is also very important to use this information to improve not only the understanding but also the detection and differentiation of dementia. SVM classification is a very promising tool for these purposes, as has been shown by previous studies (Klöppel et al., 2008a; Davatzikos et al., 2008; Fan et al., 2008; Chaves et al., 2009). It not only captures univariate relationships of a single voxel across all subjects but is also able to detect multivariate relationships over a large group of information, as, for example, between different structures and modalities in the brain. Furthermore, this tool provides an easy way to use this information for classifying imaging data of new subjects to a specific condition.
Here, we systematically compared different information provided by FDG-PET and MRI to enable the most accurate detection and differentiation of dementia. Although the diagnoses of our subjects were not histopathologically confirmed to be sure of assigning them to the correct condition, generally higher conformity with the clinical diagnosis, which was based on comprehensive clinical and neuropsychological testing, should also result in more accurate classification of histopathologically validated data.

The whole-brain SVM classification provided the most accurate classification using only FDG-PET information. GM and WM based classification accuracy was lower for all comparisons indicating a lower sensitivity for detection of dementia-relevant information. Nonetheless, classification based on GM, WM and FDG-PET separately or combining them revealed a discrimination accuracy which was above chance level for the correct categorization of the three groups. All classification results substantially exceeded the best classification accuracy obtained by randomly assigning all subjects to different groups. Additionally, smoothing of the data improved the classification accuracy in both imaging modalities as expected. However, in whole-brain classification noise is introduced by using a great deal of information for classification that does not differentiate between the groups. Recent comprehensive meta-analyses identified the “prototypical” networks for both disorders in both modalities using VBM (Schroeter et al., 2007, 2009). The involved regions have been shown to be affected in AD and FTLD patients most consistently in all studies investigating these disorders. By using this information, we ruled out the possibility that our classification results are dependent to our group of patients. Although this method provides lower accuracy rates for GM or FDG-PET information on their own, it shows a significantly higher discrimination rate by combining both information modalities into a single vector. This ROI-based discrimination is superior to whole-brain classification with the highest accuracy gain for the differentiation of both types of dementia, which, with 94%, is the highest differentiation rate reported up to now. Accordingly, we suggest this method as a diagnostic standard for the classification of dementia syndromes.

### 3.4.3. Conclusion and perspectives

In our study, we investigated the advantages of multimodal imaging using FDG-PET and MRI to improve understanding, detection and differentiation of dementia. AD and FTLD patients revealed a differential pattern of atrophy and hypometabolism, indicating different proceeding of
both disorders and providing support for the assumption of different underlying mechanisms in AD and FTLD. Furthermore, based on affected regions reported in previous studies, investigating AD and FTLD with univariate approaches and summarized in two meta-analyses, we applied linear SVM classification algorithm using information from both imaging modalities. Combining ROI information from FDG-PET and MRI resulted in a substantial gain in accuracy compared to whole-brain and to single modality classification for both detection and differentiation of AD and FTLD. Our results indicate that integration and combination of results from different imaging modalities might provide a new way to improve the understanding and diagnostic accuracy of these dementia disorders.
4. Validation of the support vector machine approach

In the previous section it has been shown that combining ROI information from FDG-PET and structural MRI might substantially improve diagnostic accuracy of two common types of dementia, AD and FTLD. However, as the groups used in that study were recruited from a single clinical center, the results might be less generalizable to other clinical data. The following section deals with the validation of the proposed multimodal approach using data from different clinical centers.

4.1. Introduction

In recent research, studies investigating the use of biomarkers to improve diagnostic accuracy of dementia have become more and more frequent with most of them having a focus on optimizing the use of one specific biomarker or on the comparison of different biomarkers regarding their sensitivity for specific dementia syndromes (Fung and Stoeckel, 2007; Davatzikos et al., 2008; Habeck et al., 2008; Klöppel et al., 2008a; Chaves et al., 2009; Habert et al., 2009; Horn et al., 2009). However, frequently used statistical methods such as multivariate pattern classification with SVM do not only enable automatic classification using one specific biomarker but, as has been shown before, also provide a tool to combine two or more different biomarkers within the same classification model.

The ROI approach described previously, although, less sensitive compared to whole-brain classification when using a single modality, was far superior to whole-brain classification when combined information from FDG-PET and MRI was used. Because ROIs for SVM were extracted from two comprehensive meta-analyses investigating both dementia syndromes (Schroeter et al., 2007, 2009) they were not biased to a specific dataset. Furthermore, the preprocessing algorithm suggested in this study was designed to overcome difficulties which might occur due to use of different scanner types and scanning sequences with different scaling and resolution.

To validate and to investigate the generalizability of this approach, we applied same preprocessing and classification algorithm to two different datasets. Classification accuracy
results using FDG-PET and MRI data from the Day Clinic of Cognitive Neurology at the University of Leipzig were compared to classification results obtained using the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.adni-info.org). ADNI is a free access database containing FDG-PET and MRI data beside comprehensive neuropsychological and clinical evaluation of AD patients and healthy control subjects.

4.2. Methods

4.2.1. Day Clinic database

**Table 4.1 Subject group characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Day Clinic group</th>
<th>ADNI group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>AD</td>
</tr>
<tr>
<td>Number</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>Male/Female</td>
<td>7/6</td>
<td>9/12</td>
</tr>
<tr>
<td>Age (years)</td>
<td>53.9±6.0</td>
<td>61.1±6.7*</td>
</tr>
<tr>
<td>CDR (score)</td>
<td>0.23±0.26</td>
<td>0.71±0.25*</td>
</tr>
</tbody>
</table>

Mean ± standard deviation. *significant difference to the control group from the same dataset, ~significant difference to the same diagnostic classification group from the Day Clinic dataset. AD Alzheimer's disease, ADNI Alzheimer's Disease Neuroimaging Initiative, CDR Clinical Dementia Rating Scale.

We analyzed FDG-PET and T1-weighted MRI data of 21 patients (Table 4.1) with an early stage of probable AD, 14 patients with an early stage of FTLD and 13 control subjects. Patients were recruited from the Day Clinic of Cognitive Neurology at the University of Leipzig. Probable AD was diagnosed according to the original and revised NINCDS-ADRDA criteria (McKhann et al., 1984; Dubois et al., 2007). Diagnosis of FTLD was based on criteria suggested by Neary et al.
The control group included subjects who visited the Day Clinic with subjective cognitive complaints, which were not objectively confirmed by a comprehensive neuropsychological and clinical evaluation. FDG-PET and MRI for these subjects was conducted within the clinical assessment for diagnostic reasons. This control group was chosen because, in clinical practice, it is crucial to discriminate between these subjects and patients with an early stage of dementia. Informed consent was obtained from all subjects. The research protocol was approved by the Ethics Committee of the University of Leipzig, and was in accordance with the latest version of the Declaration of Helsinki.

4.2.2. Data acquisition

a) Magnetic resonance imaging data

For each subject, a high-resolution T1-weighted MRI scan was obtained, consisting of 128 sagittal slices adjusted to AC-PC line and a with slice thickness of 1.5mm and pixel size of 1×1mm². MRI was performed on two different 3T scanners (MedSpec 30/100, Bruker Biospin, Ettlingen Germany and Magnetom Trio, Siemens, Erlangen, Germany) using two different T1-weighted sequences (MDEFT or MP-RAGE with TR=1300ms, TI=650ms, TE=3.93ms or TE=10ms; FOV 25×25 cm²; matrix = 256×256 voxels). On the MedSpec scanner, only the MDEFT-sequence and on the Magnetom Trio scanner, either MDEFT or MP-RAGE sequences were used.

b) [F18]fluorodeoxyglucose positron emission tomography data

Each subject also underwent FDG-PET imaging either a few a weeks before or after the MRI scan. All PET data were acquired on a Siemens ECAT EXACT HR+ scanner (CTI/Siemens, Knoxville, TN, USA) under a standard resting condition in 2-dimensional (2D) mode. The 2D acquisition mode was used because it allows a better quantification of the PET data due to lower scatter radiation. Sixty-three slices were simultaneously collected with an axial resolution of 5 mm full width at half maximum (FWHM) and in-plane resolution of 4.6 mm. After correction for attenuation, scatter, decay and scanner-specific dead-time, images were reconstructed by
filtered back-projection using a Hann-filter of 4.9 mm FWHM. The 63 transaxial slices obtained had a matrix of 128×128 voxels with an edge length of 2.45 mm.

4.2.3. Alzheimer’s Disease Neuroimaging Initiative database

a) Alzheimer’s Disease Neuroimaging Initiative subjects

To validate the multimodal ROI-based classification approach proposed in chapter 3 and to make it more reliable and generalizable to multicenter data we extracted MR and FDG-PET images of 28 patients with Alzheimer’s disease and 28 healthy control subjects (Table 1) from the ADNI database. The ADNI is a partnership of the National Institute of Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and non-profit organizations. Diagnosis of AD patients was based on NINCDS/ARDA criteria (McKhann et al., 1984). Exclusion criteria for the ADNI data were any significant neurological disease other than AD, history of head trauma followed by persistent neurological deficits or structural brain abnormalities, psychotic features, agitation or behavioral problems within the last three months or history of alcohol or substance abuse. For most subjects multiple follow-up FDG-PET and MR scans were available. To ensure that our approach is applicable for the early diagnosis of dementia for all subjects, data from the first FDG-PET and MR scan were used.

b) Alzheimer’s Disease Neuroimaging Initiative magnetic resonance imaging data

The MR dataset included standard T1-weighted images obtained with different scanner types using volumetric MPRAGE sequence varying in TR and TE with an in-plane resolution of 1.25 × 1.25 mm and 1.2 mm sagittal slice thickness. Only images obtained using 1.5T scanners were used in this study. All images were preprocessed as described on the ADNI website (http://www.loni.ucla.edu/ADNI/Data/ADNI_Data.shtml), including distortion correction and B1 non-uniformity correction.
c) Alzheimer’s Disease Neuroimaging Initiative [F18]fluorodeoxyglucose positron emission tomography data

All ADNI subjects also underwent FDG-PET scanning obtained with different scanner types and using one of three different protocols: 1) dynamic: a 30 minute, six frame acquisition (6 five-minute frames), with scanning from 30 to 60 min post-FDG injection; 2) static: a single-frame 30 min acquisition with scanning 30-60 min post-injection; and 3) quantitative: a 60 min dynamic protocol consisting of 33 frames, with scanning beginning at injection and continuing for 60 min. The majority of the scans in the ADNI study were acquired with the first acquisition protocol. The images further differed in resolution, orientation, voxel and image dimensions and count statistics.

4.2.4. Preprocessing of magnetic resonance and [F18]fluorodeoxyglucose positron emission tomography data

For MR and FDG-PET data the same preprocessing procedure was applied as described in detail in chapter 3. This procedure (Figure 3.1) included interpolation of both FDG-PET and MR images to an isotropic resolution of 1×1×1 mm³, bias correction for inhomogeneity artifacts for MR data, partial volume effect correction and masking of non GM voxels in FDG-PET data and spatial normalization to an averaged size template created from all subjects using the DARTEL approach (Ashburner, 2007). The same deformations calculated based on the MR template were applied to MR and to co-registered FDG-PET images. After smoothing of FDG-PET and MR images with a Gaussian kernel of 12 mm FWHM, intensity normalization of FDG-PET data to cerebellum (Dukart et al., 2010) and subsequent masking, the procedure results in an accurate anatomical overlap of both modalities. Both imaging modalities then have the same spatial orientation, resolution and effective smoothness. Individual intensity normalization to the cerebellum also accounts for initially different count statistics of FDG-PET images. However, both imaging modalities are still in the DARTEL space. To extract regional values corresponding to ROIs reported in the meta-analyses investigating AD and FTLD (Schroeter et al., 2007, 2009) all images were normalized to the MNI template using affine-only transformation.
a) Region-of-interest extraction

ROI coordinates were extracted using the MRICron 3D fill tool (http://www.sph.sc.edu/comd/rorden/mricron) as described in chapter 3 from two comprehensive, systematic and quantitative meta-analyses investigating biomarkers of AD and FTLD in MR and FDG-PET images. Although the data extracted from the ADNI data base did not include FTLD patients, ROIs reported for this group of patients were also included in the classification procedure. The inclusion of these regions not only improves the differentiation between both dementia types but also improves the differentiation of AD patients and healthy control subjects providing rather less affected reference regions for the multivariate pattern classification and so improving the classification. The meta-analyses included a total number of 1618 patients (AD/FTLD: 1351/267) and 1448 healthy control subjects (1097/351) (Schroeter et al., 2007, 2009). These meta-analyses extracted the prototypical networks of AD and FTLD by applying what is currently the most sophisticated and best-validated of coordinate-based voxel-wise meta-analyses, the anatomical likelihood estimate. In the FTLD meta-analyses (Schroeter et al., 2007), only coordinates which are common to all subgroups of FTLD patients were used. Because the coordinates in both meta-analyses were reported in the Talairach space, they were transformed to MNI space according to a formula proposed by Matthew Brett (published on the Internet: http://www.mrccbu.cam.ac.uk/Imaging/Common/mnispace.shtml).

b) Support vector machine classification

Multivariate pattern classification was performed with a linear kernel by identifying a separating linear hyperplane that maximizes the distance between different clinical groups based on ROI information. Additionally, the cost parameter for incorrect classification was optimized for each dataset. The cross-validation of the trained SVM was performed by using the leave-one-out method. The procedure iteratively leaves out the information of each subject and trains the model on the remaining subjects for subsequent class assignment of the person that was not included in the training procedure. This validation method enables the generalization of the trained SVM to data that have never been presented to the SVM algorithms previously. The
reported accuracy is the percentage of subjects correctly assigned to the clinical diagnosis. To improve the validity of this approach, separate classifiers were trained based on the dataset from the ADNI database, the dataset of AD patients and control subjects from the Day Clinic of Cognitive Neurology or on combined data from both samples. Leaving-one-out cross-validation was performed for all classifiers. Furthermore, both classifiers trained separately only on one of both datasets were applied to the data not used for the training. SVM classification was applied separately to ROIs extracted from FDG-PET and MR data and to combined information from both imaging modalities.

4.2.5. Statistical analysis

Group comparisons for age and CDR (Clinical Dementia Rating Scale, Morris, 1993) were performed by conducting Bonferroni t-tests with a significance threshold of p<.05. Group differences regarding sex were evaluated using a chi-square test for independent samples. The statistical analysis was performed with the commercial software package SPSS 17.0 (http://www.spss.com/statistics/).

4.3. Results

4.3.1. Clinical characteristics

The chi-square test for independent samples did not reveal any statistical differences in sex between AD and Controls in the ADNI [χ²(1)=0.08;p=0.771] or in the Day Clinic dataset [χ²(1)=0.39;p=0.53. There was also no significant difference in sex between corresponding groups from both datasets, namely AD patients from the ADNI and the Day Clinic database [χ²(1)=3.06;p=0.08] and control subjects from the ADNI and the Day Clinic database [χ²(1)=1.22;p=0.27]).

CDR scores differed significantly between AD patients and control subjects in the ADNI [t(54)=15.63;p<0.001] and in the Day Clinic dataset [t(32)=5.36;p<0.001] but not between both
groups of AD patients \( t(47)=1.23; p<0.22 \). The control group from the Day Clinic had slightly but significantly higher CDR scores in comparison to the control group from the ADNI database \( t(39)=3.87; p<0.001 \). Age was substantially different between both datasets, with ADNI patients \( t(47)=7.27; p<0.001 \) and control subjects \( t(39)=12.59; p<0.001 \) being older than the corresponding group from the Day Clinic. There was a minor but significant difference between AD patients and control subjects in the Day Clinic group \( t(32)=3.18; p=.008 \) but not in the ADNI dataset \( t(54)=0.22; p=.83 \).

4.3.2. Support vector machine results

**Table 4.2** Accuracy rates for ROI-based SVM classification for FDG-PET and MRI separately and for combined information

<table>
<thead>
<tr>
<th></th>
<th>FDG-PET</th>
<th>MRI</th>
<th>FDG-PET / MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADNI dataset</td>
<td>71.4%</td>
<td>66.1%</td>
<td>85.7%</td>
</tr>
<tr>
<td>Day Clinic dataset</td>
<td>94.1%</td>
<td>82.4%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Combined dataset</td>
<td>87.8%</td>
<td>75.6%</td>
<td>91.1%</td>
</tr>
<tr>
<td>therefrom ADNI</td>
<td>83.9%</td>
<td>73.2%</td>
<td>85.7%</td>
</tr>
<tr>
<td>therefrom Day Clinic</td>
<td>94.1%</td>
<td>79.4%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Accuracy represents the percentage of subjects correctly assigned to the correct condition. AD Alzheimer’s disease, ADNI Alzheimer’s Disease Neuroimaging Initiative, FDG-PET \([18F]\)fluorodeoxyglucose positron emission tomography, MRI magnetic resonance imaging, ROI region-of-interest, SVM support vector machine.

The differentiation accuracy for single modality classification using leaving-one-out cross-validation was highest with 94% using FDG-PET ROI information in the Day Clinic dataset (Table 4.2). Lowest accuracy, with 66% was obtained using MR ROI information from the ADNI database. Classification accuracy using FDG-PET information was far superior to MR-based classification in all conditions. For all comparisons, the best classification accuracy was
obtained using combined regional information from FDG-PET and MR images. For these combined information accuracy rates ranged between 86 and 100%, with lowest accuracy using only the ADNI dataset and highest accuracy using the Day Clinic dataset. The overall accuracy for the combined dataset using both modalities was 91%, with sensitivity and specificity of 86% for subjects from the ADNI database and 100% sensitivity and specificity for data from the Day Clinic (Table 4.3). The prediction for Day Clinic patients using the ADNI dataset for training was very high with 91%. The opposite comparison revealed a prediction accuracy of 84%.

Table 4.3  Differentiation rates for combined ROI information from FDG-PET and MRI

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADNI dataset</td>
<td>85.7%</td>
<td>82.1%</td>
<td>89.3%</td>
</tr>
<tr>
<td>Day Clinic dataset</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Combined dataset</td>
<td>91.1%</td>
<td>91.8%</td>
<td>90.2%</td>
</tr>
</tbody>
</table>

ADNI Alzheimer’s Disease Neuroimaging Initiative, FDG-PET [18F]fluorodeoxyglucose positron emission tomography, MRI magnetic resonance imaging, ROI region-of-interest.

4.4. Discussion

Recently, it has been shown that applying multivariate statistical methods on imaging markers might provide a substantial gain in accuracy for early detection and differentiation of dementia syndromes (Fung and Stoeckel, 2007; Davatzikos et al., 2008; Klöppel et al., 2008a; Chaves et al., 2009). However, although all studies applying multivariate statistical methods to improve diagnostic procedure have shown high detection and differentiation rates for different dementia syndromes, some of these increased accuracy rates have been obtained with a trade-off in generalizability of the proposed approaches to new datasets. This is the case because the number of features used to obtain high classification accuracy was determined by feature selection methods identifying a minimal number of regions that provide optimal separation
between different groups within the specific dataset and by using the same dataset to validate this approach (Fung and Stoeckel, 2007; Fan et al., 2008; Davatzikos et al., 2008; Gerardin et al., 2009). This method, although providing optimal classification accuracy in a specific sample, does not necessarily provide optimal classification for data from other clinical centers.

The number of selected features using this approach might be specific to the clinical group used in the study and depended on the accuracy of the clinical diagnosis in the specific clinical center. Therefore, in clinical practice, to make the diagnostic procedure more generalizable, comparable and applicable to different clinical centers, the number of features used for classification should be stable and not dependent on a specific sample.

A good alternative to this procedure is the approach applied by Klöppel et al. (2008a), namely using whole-brain information for the classification of dementia syndromes. This procedure is not biased to a specific feature selection and provides high discrimination accuracy for different dementia syndromes. A rather practical disadvantage of this method is the large number of features used for classification, which requires a large working memory and therefore specific hardware components, especially if more than one imaging modality is used for classification.

Combining information from different imaging modalities which have been reported to be sensitive biomarkers for a specific neurodegenerative disorder substantially improves detection and differentiation of dementia in comparison to single modality approaches as has been shown in chapter 3. However, this improvement combining multimodal information was only observed if ROIs information were used but not using whole-brain information from FDG-PET and MRI. The ROI approach proposed in this study is not biased to a specific dataset because ROIs were extracted from two comprehensive meta-analyses investigating AD and FTLD, which included a large number of patients from different clinical centers (Schroeter et al., 2007, 2009). The results of our study indicate that this ROI approach is applicable to new datasets and is not dependent on specific scanner types or sequences. The combination of ROI information from FDG-PET and MRI was far superior to the single-modality ROI-based classification for all comparisons. Furthermore, our study indicates that classifiers trained on data provided by ADNI, which is an open access database, result in good discrimination accuracy for new data from a single clinical center and so even increase the potential applicability of the proposed approach in the clinical environment. Best discrimination accuracy with sensitivity and specificity of about 90% was obtained for both datasets when the classifier was trained on the combined group consisting of all subjects.
4.4.1 Conclusion and perspectives

In this study, we investigated the applicability of the approach proposed in chapter 3 for detection and differentiation of dementia syndromes using multimodal information to data from different clinical centers. For this purpose we used data from the ADNI database. The results of this study are in line with the previous finding that combining FDG-PET and MRI information substantially improves detection of patients with early to moderate stages of AD. Furthermore, the method proposed in our study provides a possibility to use data from open access databases to improve clinical diagnosis of dementia in single clinical centers. Therefore, it has a high relevance for general clinical application. Additionally, our approach provides an easy method to integrate further biomarkers, if those will have been validated by future research, to improve clinical diagnosis of various dementia syndromes.
5. General discussion and outlook

As has been shown in previous chapters, the combination of biomarkers using univariate and multivariate statistical approaches might provide a useful tool for the evaluation of imaging abnormalities in different dementia syndromes and the integration of these into diagnostic procedure. However, various methodical aspects in preprocessing of the data, like intensity normalization in FDG-PET, might substantially affect the statistical outcome and the interpretation of the results (Chapter 2). For many of these questions, further methodical studies are required to investigate and to optimize their contribution to the subsequent statistical evaluation. Therefore, addressing such questions is particularly important to improve the use and combination of different imaging modalities like FDG-PET and MRI in dementia research.

5.1. Methodical aspects in preprocessing

The present work addressed the question of optimizing intensity normalization for FDG-PET data. Comparison of two different intensity normalization procedures either to the cerebellum and to the cerebral global mean revealed differential effects of both approaches for detection and differentiation of AD and FTLD. While cerebellar normalization was superior in detection of differences between dementia patients and healthy controls, normalization to the cerebral global mean was much more sensitive for differences between both groups of dementia patients. These results suggest a differential application of both normalization methods depending on the diagnostic question. Additionally, this methodical comparison demonstrated that investigation of clinically important questions like the existence of glucose hypermetabolism in specific psychiatric disorders might be biased by the choice of preprocessing procedures applied to address these questions.

Further methodical issues addressed in the present work dealt with the development of a new preprocessing algorithm to overcome difficulties resulting from the use of different scanner types and sequences within the same imaging modality and to enable the combination and common qualitative and quantitative evaluation of different imaging modalities.
Methods enabling a common evaluation of images obtained using different scanner types and sequences are a prerequisite for the use of data extracted from large datasets like the ADNI database to improve diagnostic accuracy in single clinical centers. Such datasets are usually collected using different scanner types and sequences. To use these data for diagnostic procedures by applying univariate or multivariate statistical approaches, it is necessary to overcome difficulties arising due to different resolution, measurement procedures and different intensity scaling in FDG-PET, for example. Only when these problems are solved can classifiers trained on such a database be optimally used to differentiate patients and control subjects or different groups of patients in single clinical centers.

Methodical questions regarding optimized preprocessing are directed towards enabling a qualitative and quantitative interpretation and comparison of observed signal changes in different imaging modalities. In FDG-PET the signal targeted for measurement is glucose utilization. MRI aims to measure structural brain information and so provide an estimation of atrophy in demented patients. In clinical setting, these modalities are usually evaluated independently using substantially different preprocessing algorithms, which is highly problematic for combined evaluation of both modalities as explained in more detail below.

When evaluation within a single modality is required, the preprocessing algorithm is expected to have a similar effect in all subgroups because all preprocessing steps are applied to all data in the same order. Although this is not always the case, this is one of the reasons why preprocessing algorithms are sometimes addressed to a lesser extent in clinical studies when comparing groups in a single modality. However, this differential preprocessing is an important problem for common evaluation of different imaging modalities because each single preprocessing step might have a substantial effect on the statistical outcome. In the worst case, differences observed in two different imaging modalities might be a result of the difference in the preprocessing of each modality and not related to the signal measured. For this reason, studies comparing two different imaging modalities should try to exclude effects of differential preprocessing as far as possible by making the procedure more comparable for both imaging modalities. Two good examples of such a common preprocessing of FDG-PET and MRI enabling qualitative evaluation of both imaging modalities in dementia patients are recently published studies by Chetelat et al. (2008) and Villain et al. (2008).

To overcome difficulties described above which have arisen in the present work due to combined use of different imaging modalities and due to different scanner types and sequences,
a new preprocessing algorithm was developed and applied to the data (Figure 3.1). This algorithm enabled a common statistical evaluation and interpretation of FDG-PET and MRI in AD and FTLD, revealing a differential pattern in the ratio of glucose hypometabolism and atrophy in both groups of dementia patients. Moreover, although data extracted from the ADNI database substantially differed in various parameters like quality and resolution, high diagnostic accuracy was obtained using these data after preprocessing for differentiation of AD patients and healthy controls. Nonetheless, the algorithm applied here is still in need of optimization and improvement. Some specific questions regarding the effect of an additional spatial normalization of MRI and FDG-PET data and modulation of FDG-PET images have still to be evaluated.

5.2. Voxel-based morphometry and support vector machine classification in dementia

In recent research a number of studies have reported changes in dementia patients in different imaging modalities using VBM (Mummery et al., 2000; Rosen et al., 2002; Ishii et al., 2005; Chetelat et al., 2008; Kanda et al., 2008). This method has been shown to be a sensitive tool for detecting group differences between AD and FTLD patients and control subjects and between both groups of dementia patients in specific regions. Studies applying VBM have further demonstrated that reductions in glucose metabolism and atrophy might provide a good measurement of progression in AD, thus indicating that VBM might be used for the evaluation of potential therapeutic treatment (Alexander et al., 2002; Diehl-Schmid et al., 2005; Kinkingnehun et al., 2008). However, this application of VBM in dementia, although allowing a good description and visualization of regions affected in the brain, provides only limited information to improve our understanding of the underlying mechanisms in specific dementia syndromes.

A further, more informative, application of VBM to study the underlying mechanisms in specific dementia syndromes might be the common evaluation of different imaging markers like FDG-PET and MRI. Chetelat et al. (2008) showed that qualitatively glucose hypometabolism substantially exceeded atrophy in most affected regions in AD patients, suggesting the intervention of additional hypometabolism-inducing factors in this specific disorder. Such a factor might be disconnection and amyloid deposition, resulting in genuine functional perturbations ahead of actual atrophy (Chetelat et al., 2008). Furthermore, such a multimodal
combination might provide a more differentiative insight into the progression and understanding of specific processes in AD. For example, Villain et al. (2008) found that hippocampal atrophy is specifically related to cingulum bundle disruption, which is in turn highly correlated to hypometabolism of the posterior cingulate cortex but also of the middle cingulate gyrus, thalamus, mammillary bodies and parahippocampal gyrus. They interpreted these findings as supportive evidence for the hypothesis that glucose hypometabolism of the posterior cingulate cortex not only results from local neuropathological changes, but mostly reflects the distant effect of neuronal damage in the hippocampal formation. Such disorder-related changes gained by coevaluation of FDG-PET and MRI provide a better insight into the pathological mechanisms in dementia in comparison to single-modality evaluation. They might be used for the development of new models of specific dementia syndromes and, in this way, to develop treatment alternatives.

In the present work, similar questions were addressed by common evaluation of FDG-PET and MRI using VBM (Chapter 3). The results of the study indicate the amount of glucose hypometabolism to be the predominating factor in AD, while atrophy is less pronounced in this group of patients. This observation is consistent with results described by Chetelat et al. (2007) and indicates that changes in glucose metabolism, which is strongly connected to glutamatergic neuronal activity (Sibson et al., 1998), precede or have at least a faster mechanism in AD. However, an inverse pattern of atrophy and hypometabolism was observed in FTLD in comparison to AD patients with atrophy substantially exceeding the reduction of glucose metabolism in most cortical and in all subcortical structures. Additionally, extensive WM intensity reductions were observed in patients with FTLD, mainly in frontal and temporal regions and in the brain stem. These results suggest a different process not primarily affecting glucose consumption to explain neurodegeneration in FTLD. Extended WM atrophy also indicates a mechanism that is not restricted to GM structures, which might be supportive of the recently reported results of the TAR DNA-binding protein, TDP-43, to be the major disease protein in FTLD (Davidson et al., 2007). This protein has been reported to be not only restricted to GM structures in FTLD patients but also present in frontal and temporal WM structures and the brainstem (Neumann et al., 2007), which is consistent with the reduced WM intensity in these regions.

Although it is difficult to establish causal relationships between different imaging markers and neuronal processes, this common evaluation of specific imaging modalities might be used to
generate or to test specific hypotheses to improve our understanding of mechanisms and progression in different neurodegenerative disorders.

A further application of VBM results is their use in improving diagnostic procedures in dementia. In the present work, SVM was applied to imaging features extracted from FDG-PET and MRI based on two comprehensive meta-analyses investigating AD and FTLD (Schroeter et al., 2007, 2009). Both meta-analyses reported coordinates of regions showing the largest overlap obtained by activation likelihood estimate between different studies investigating these specific disorders using VBM. In this way, it is possible to integrate the results of different VBM studies to improve single case diagnosis of patients with specific dementia syndromes. Applying SVM to combined information from FDG-PET and MRI extracted from both meta-analyses has been shown to be an effective method of improving diagnostic accuracy for detection and differentiation of AD and FTLD (Chapters 3 and 4). This approach was superior to the use of whole-brain and single modality information, especially for differentiation of AD and FTLD. Furthermore, a similarly high accuracy was obtained by applying this approach onto data extracted from the ADNI database. The classifier trained on this database provided high discrimination accuracy for subjects and patients from the Leipzig Day Clinic of Cognitive Neurology. The generalizability of such methods to single clinical centers without a lot of modifications is highly important to have a valid and easily applicable tool in clinical settings. The method proposed in this work seems to be a reliable tool for increasing the diagnostic accuracy of specific dementia syndromes. Furthermore, this SVM algorithm can be easily modified to integrate additional bio- and clinical markers if these have been shown to be sensitive and highly specific for a neurodegenerative disorder, for instance from cerebrospinal fluid analysis of beta-amyloid or tau proteins as suggested by Dubois et al. (2007). Additional integration of such markers should further improve detection and differentiation of specific dementia syndromes like AD and FTLD.

A major problem for a good estimate of the accuracy rate obtained with the SVM approach applied in this work is the lack of pathological validation of patients with both types of dementia syndromes used in this study. This might be a problem as the clinical diagnosis of dementia has been shown to be a less sensitive instrument, especially in the differentiation of AD and FTLD.

However, this approach has been validated using two different independent samples from various clinical centers. Additionally, high discrimination accuracy was obtained for both AD and
FTLD. These results indicate that this SVM method is a valid and reliable instrument for detection and differentiation of specific dementia syndromes.

The results of the present work emphasize the high potential of combined evaluation of different imaging markers to improve understanding and diagnostic accuracy of various types of dementia. The application of the multimodal classification approach described above resulted in a very high accuracy for both, detection and differentiation of AD and FTLD and is therefore of major importance for clinical application.
6. Zusammenfassung

Contribution of FDG-PET and MRI to improve Understanding, Detection and Differentiation of Dementia

Jürgen Dukart

Medizinische Fakultät, Universität Leipzig

Dissertation

Die Demenzproblematik hat in den letzten Jahrzehnten in unserer Gesellschaft immer mehr an Bedeutung gewonnen. Dies hängt größtenteils mit der weiterhin steigenden durchschnittlichen Lebenserwartung zusammen. Während in einer jüngerer Bevölkerung der Demenzdiagnose anteilmäßig nur eine geringe Bedeutung zukommt, steigt ab einem Alter von 50 bis 60 Jahren die Anzahl der Neuerkrankungen mit dem zunehmenden Alter exponentiell an. Derzeit leiden alleine in Deutschland geschätzte 1.07 Millionen Menschen an einer moderaten bis schweren Form einer Demenzerkrankung (Ziegler und Doblhammer, 2009). Prognostisch soll dabei die Anzahl von Demenzpatienten weltweit von 25 Millionen im Jahr 2000 auf rund 114 Millionen im Jahr 2050 ansteigen (Wimo et al., 2003), was die besondere Problematik dieser Erkrankungsform für unsere Gesellschaft nochmals verdeutlicht. Die häufigste Erkrankungsform mit ca. 60% stellt dabei die Alzheimerdemenz dar, bei der in den Anfangsstadien insbesondere Gedächtnisdefizite im Vordergrund stehen, die jedoch auch mit anderen kognitiven Beeinträchtigungen einhergehen können (Fratiglioni et al., 1999; Dubois et al., 2007). Die zweithäufigste neurodegenerative Demenzform im Alter ist die frontotemporale lobäre Degeneration (Ratnavalli et al., 2002; Rosso et al., 2003), die jedoch auf der Symptomebene ein weniger homogenes Muster aufweist. Bei dieser Erkrankung können in den Anfangsstadien Verhaltens- und Persönlichkeitsänderungen, Beeinträchtigung des semantischen Sprachverständnisses oder aphasische Defizite vom Broccatyp im Vordergrund stehen.
Aufgrund dieses heterogenen Musters wurde für diese Demenzform eine Aufgliederung in drei verschiedene Subtypen vorgeschlagen, für die in verschiedenen Bildgebungsmodalitäten distinkte Beeinträchtigungen neuronaler Netzwerke gezeigt werden konnten (Neary et al., 1998).

Die Früherkennung und die Unterscheidung der Alzheimererkrankung und der frontotemporalen lobären Degeneration sind jedoch nicht immer so eindeutig. Gemessen an der Histopathologie als derzeitigem „Goldstandard“ werden für die Alzheimerdemenz Genauigkeiten der klinischen Diagnose von 65 bis 100% angegeben (Klatka et al., 1996; Petrovitch et al., 2001; Rascovsky et al., 2002). Für die frontotemporale Demenz schwanken die Genauigkeitsangaben, vor allem wegen der Schwierigkeit der Abgrenzung zur Alzheimererkrankung, jedoch nur zwischen 17 und 85% (Rascovsky et al., 2002; Knopman et al., 2005; Mendez et al., 2007), was deutlich einer weiteren Verbesserung bedarf.

In der neueren Forschung wurden für beide Erkrankungsformen in unterschiedlichen Bildgebungsmodalitäten distinkte Veränderungen berichtet (Rosen et al., 2002; Ishii et al., 2005; Nestor et al., 2005; Desgranges et al., 2007; Clark et al., 2008; Habeck et al., 2008; Jack et al., 2008). Die Auswertung dieser kann zu einer Verbesserung des Verständnisses, des Fortschreitens, der Symptomatik und der Krankheitsmechanismen bei unterschiedlichen Demenzformen beitragen. Außerdem konnte gezeigt werden, dass die Hinzunahme von Bildungsdaten Informationen unter Verwendung verschiedener Methoden zu einer Verbesserung der Detektion und der Differenzierung beider Demenzformen substantiell beitragen kann (Habeck et al., 2008; Klöppel et al., 2008a, b; Habert et al., 2009).

Um eine optimale interindividuelle Bewertung und Vergleiche von Bildgebungsmarkern zu ermöglichen, müssen die Daten zuerst mit Hilfe komplexer Vorverarbeitungsmechanismen räumlich und skalenmäßig standardisiert werden. Da die einzelnen Vorverarbeitungsschritte jedoch einen massiven Einfluss auf das Ergebnis haben können, muss dieser Einfluss klar bestimmt sein, damit eine optimale Evaluation der Bildungsdaten möglich wird.

Ein solcher wichtiger Vorverarbeitungsschritt für die Auswertung von Glukoseutilizationsdaten, gemessen mit der [F18]fluorodeoxyglukose Positronenemissionstomographie (FDG-PET), ist die Intensitätsnormalisierung. Diese ermöglicht eine Skalierung von Daten relativiert an einer Referenzregion. Die Wahl dieser Referenzregion hat dabei einen starken Einfluss auf das statistische Ergebnis bei Gruppenvergleichen (Ishii et al., 2001; Yakushev et al., 2008; Yakushev et al., 2009). Allerdings gibt es bislang keinen Konsens, welche Region für die


Eine weitere Fragestellung der vorliegenden Arbeit bestand in der kombinierten Auswertung von FDG-PET und Magnetoresonanztomographie (MRT) Daten von Patienten mit Alzheimer Demenz oder der frontotemporalen lobären Degeneration. Beide Gruppen wurden in beiden Modalitäten mit einer gesunden Kontrollgruppe verglichen. Um eine gemeinsame Interpretation der modalitätsspezifischen Ergebnisse beziehungsweise eine kombinierte Verwendung beider...


In den letzten Jahren konnte außerdem für beide Bildgebungsmodalitäten separat gezeigt werden, dass die Veränderungen, die in diesen für beide Demenztypen gefunden wurden, unter Verwendung multivariater statistischer Methoden durchaus einen hohen diagnostischen Wert in Bezug auf eine Früherkennung und Differenzierung von demenziellen Syndromen haben (Davatzikos et al., 2008; Fung et al., 2007; Fan et al., 2008; Chaves et al., 2009; Horn et al., 2009). In diesen Studien wurde jedoch zumeist nur eine der Bildgebungsmodalitäten untersucht oder die Bildgebungsmodalitäten miteinander verglichen. Die multivariate Musterklassifikation, als eine Möglichkeit die Bildgebung diagnostisch zu nutzen, bietet allerdings den Vorteil, dass Informationen aus verschiedenen Modalitäten leicht miteinander kombiniert werden können. Dieses Vorgehen könnte zu einer zusätzlichen Steigerung der diagnostischen Genauigkeit der Erkennung und Unterscheidung von Demenztypen beitragen.

7. References


Kloppel, S., Stonnington, C.M., Barnes, J., Chen, F., Chu, C., Good, C.D., Mader, I., Mitchell, L.A., Patel, A.C., Roberts, C.C., Fox, N.C., Jack, C.R., Jr., Ashburner, J., Frackowiak, R.S.,


Erklärung über die eigenständige Abfassung der Arbeit


.......................................................... ..........................................................

Datum Unterschrift