Multimodal MRI-based classification of Alzheimer's disease

Vos, F. de

General introduction

1.1 Alzheimer's disease

Epidemiology

The number of people living with dementia worldwide was estimated at 50 million in 2018, and it is likely to rise to about 152 million in 2050 (World Alzheimer report 2018). Alzheimer's disease (AD) is the most common form of dementia, accounting for about two third of all dementia cases (World Alzheimer report 2018). AD is a neurodegenerative disease that is clinically characterised by memory loss and declined cognitive functioning. As the disease progresses, these behavioural characteristics get more severe, up until the point that patients are in need for 24h care. Typical AD neurodegeneration starts within the hippocampal area and later spreads out to subcortical brain regions and the medial temporal lobe (Jack et al., 1997; Pini et al., 2016). As the disease progresses further, the atrophy also extents to the frontal areas of the brain (Thompson et al., 2003).

Diagnosis

In the global action plan against dementia, the World Health Organization has specified improved diagnostics as a key area (WHO, 2017). The clinical characteristics of AD manifest relatively late in the disease, and it can only be reliably diagnosed at a stage where irreversible damage has already taken place (Jack et al., 2013). At an earlier phase, the clinical symptoms are often difficult to discriminate AD from normal ageing, or other types of dementia (Sperling et al., 2013). For this reason, these clinical symptoms cannot be used for a timely diagnosis of AD. There is need for reliable AD diagnosis at the beginning of the disease, or ideally before the onset of the disease (Frisoni et al., 2017). This would provide patients with a prognosis so they can prepare for their future trajectory. In addition, early phase AD diagnosis is important for drug research, because early phase AD patients might still be susceptible for drugs (Scheltens et al., 2016).

1.2 Biomarkers

To enable timely diagnosis of AD, there is a need for biomarkers. A number of AD biomarkers have already been proposed. Well-known biomarkers are the proteins B-amyloid and tau (Hardy & Selkoe, 2002), measured in cerebrospinal fluid (CSF) or in the brain using PET scans. The global use of these methods is still limited because of their invasive nature and the high costs associated with PET scanning. Recent

innovations have enabled measurement of tau levels via plasma, thereby reducing invasiveness and costs, but this method has not yet been validated for diagnosis in diverse clinical populations (Palmqvist et al., 2020). In addition, many magnetic resonance imaging (MRI) related biomarkers have been proposed. MRI scans visualise neurodegeneration, which is a key feature of AD. Importantly, MRI scans are non-invasive, which is advantageous for day-to-day clinical use.

Anatomical MRI

An important MRI AD biomarker is grey matter atrophy (Frisoni et al., 2010). Grey matter consists of neuronal cell bodies, and neuronal loss in AD is responsible for memory loss and cognitive decline. The location and degree of grey matter atrophy can be accurately visualised with anatomical MRI scans. Anatomical MRI scans improve AD diagnosis over cognitive testing only (Liu et al., 2011), and they are included in diagnostic criteria for the most prevalent non-Alzheimer dementias as well (McKeith et al., 2005; Rascovsky et al., 2011; Roman et al., 1993), reflecting its value for differential diagnosis. Furthermore, grey matter atrophy enables early AD diagnosis, because atrophy in the temporal lobe starts before clinical symptoms occur (Jack et al., 2013). Consequently, medial temporal lobe atrophy is used as a marker for mild cognitive impairment (MCI), the clinical phase that precedes AD, and it is used to predict which MCI patients will convert to AD (Cuingnet et al., 2011; Misra et al., 2009).

Diffusion MRI

Brain damage caused by AD is not limited to grey matter atrophy. It also comprises decreased structural integrity in the white matter (Bozzali et al., 2002; Douaud et al., 2011), which can be visualised with diffusion MRI scans. White matter consists of myelinated axons that transport neuronal signal between grey matter areas. These white matter pathways make up the brain's structural networks, and they are disrupted in AD (Mito et al., 2018). Furthermore, white matter lesions increase the risk of developing AD (Prins et al., 2004). For these reasons, diffusion MRI scans might complement anatomical MRI scans in the diagnosis of AD.

Resting state functional MRI

In addition to the structural changes, observed in both grey and white matter, AD is also characterised by changes in brain function, measured with resting state functional MRI (fMRI). These changes include decreased activation in the default mode network (Rombouts et al., 2005), altered FC between brain regions (Agosta et al., 2012; Allen et al., 2007; Binnewijzend et al., 2012), and decreased amplitude of low frequency

fluctuations (Han et al., 2011), reflecting decreased intensity of spontaneous brain activity. FC may already be altered in early stages of AD, even before the presence of brain atrophy and cognitive decline (Buckner et al., 2005; Sheline and Raichle, 2013). For these reasons, resting state fMRI scans may add complementary information to structural and diffusion MRI scans, which may further improve AD diagnosis.

1.3 Individual classification of Alzheimer's disease

Average group differences vs individual classification

The MRI AD biomarker research discussed in previous paragraphs is mostly based on average group differences, as found in case-control studies. Average group differences are not necessarily useful at the individual level. Many AD biomarkers are present in healthy ageing as well (Salat et al., 1999), and show too much overlap between AD patients and healthy elderly to accurately discriminate AD at the individual level. The focus of MRI AD biomarker research has therefore shifted from the detection of average group differences towards individual classification (Rathore et al., 2017). Individual classification studies usually combine multiple biomarkers that together yield high classification accuracy.

Statistical learning

Statistical learning is a powerful framework for individual classification studies (Klöppel et al., 2008). First, it comprises statistical methods that enable the incorporation of many features into one classification model. This is essential for MRI-based classification studies, because MRI scans yield many features. Second, it incorporates cross-validation. That is, classification models are fitted on training data, and validated on a held-out data set. Cross-validation protects for overfitted models, yielding models that better generalise to out of sample data. In this way, statistical learning enables development of MRI-based classification models that can discriminate AD at the individual level, which is necessary for clinical practice.

Multimodal MRI

Thus far, most MRI-based AD classification studies have used models with only one type of feature. For example, anatomical MRI scans were used to measure only grey matter

density features (Beheshti et al., 2016) or only hippocampal shape features (Gerardin et al., 2009). However, these different types of anatomical MRI features possess complementary information, and combining them increases AD classification accuracy (Bron et al., 2015; Wolz et al., 2011; Westman et al., 2013). Moreover, different types of MRI scans contain complementary information as well, and combining these into a multimodal MRI model improves AD classification accuracy even further (Schouten et al., 2016). In this thesis we will derive multiple types of feature from a single MRI scan, and we will combine multiple MRI modalities.

1.4 Aims and outline

The overall aim of this thesis is to develop and evaluate MRI-based models for individual AD classification. These models take MRI features as input, and yield AD probability scores as output. We will develop these models for a wide variety of MRI features, and we will compare those models on AD classification accuracy. This provides information on which MRI features are most informative for AD classification. Furthermore, we will study combinations of features to try to improve accuracy. To this end, we will use statistical learning techniques, because they enable incorporation of many features, and they focus on optimising classification accuracy. In **chapter two,** anatomical MRI scans are used to calculate multiple structural features that are thought to be informative for AD. These features will be compared on AD classification accuracy, and they will be combined into a single model. It is hypothesised that the combination outperforms the separate features. In **chapter three,** a similar approach will be used for resting state fMRI scans. We will use several approaches to calculate FC, and we will derive indirect FC measures, like FC dynamics and graph measures. Again, these features are compared on AD classification accuracy, and they will be combined into a single model. Also here, it is hypothesised that the combination outperforms the separate features.

In order to be clinically useful, AD classification models should generalise well to clinical populations. This is complicated for two reasons. First, research samples that are used to develop the models are often homogeneous, only containing clinically diagnosed AD patients and healthy elderly controls. In contrast, clinical populations are more diverse. They include patients with varying stages of AD progression, as well as MCI patients and preclinical patients that experience memory complaints. In addition, clinical populations include patients with non-AD dementia types as well. Second, MRI scans suffer from technical between-scanner variation (Ewers et al., 2006; Takao et al., 2014; Zhu et al., 2011). An AD classification model that is trained with MRI data from one scanner, is not necessarily useful for MRI data from another scanner. To be clinically useful, AD classification models should be applicable to diverse patient populations, and they should be robust to between-scanner variation. In **chapter four** we will evaluate whether MRI-based models for individual AD classification also discriminate AD in a diverse clinical population. To this end we will train AD classification models on a single-centre data set consisting of AD patients and controls, using features derived from both anatomical MRI, diffusion MRI and resting state fMRI scans. Next, we will use these models to assign AD scores to patients from a multi-centre memory clinic

data set including AD patients, MCI patients and patients with subjective memory complaints. We expect that AD patients will receive higher AD scores than MCI patients and patients with subjective memory complaints.

In **chapter five** we will evaluate MRI's predictive accuracy for future cognitive decline. To this end we use the baseline multimodal MRI scans of the memory clinic data set outlined above, to predict two-year follow-up cognitive decline.