CHAPTER 5

Summary and general discussion
SUMMARY

This thesis focused on ways to assist the clinician in the challenges of differential diagnosis, prognosis and survival estimates. The aim of this thesis was to improve care for individuals with (incipient) neurodegenerative diseases, by finding ways to optimize the use of tests for differential diagnosis and prognosis. In the coming paragraphs, I will summarize the most important findings of this thesis.

Operationalization of structural MRI in daily clinical practice

The first part of this thesis focused on operationalization of structural MRI scans. Given the challenges in applying MRI scans for diagnosis and prognosis in individual patients, this thesis set out to extract more information from MRI scans. Chapter 2.1 assessed the value of well-known visual ratings scales. For this I examined the associations of age and diagnosis with visual ratings of MTA, PA, GCA and WMH and investigated their clinical value, in patients with dementia due to AD, MCI and SCD. Linear regression analyses showed main effects of age and diagnosis and an interaction age*diagnosis for MTA, PA and GCA. For MTA, GCA and PA, diagnostic value was best in patients <65 years (optimal cut-off: ≥1), while WMH did not discriminate. This led to the conclusion that visual ratings of atrophy and WMH are differently affected by age and diagnosis, requiring an age-specific approach in clinical practice. Their diagnostic value seems strongest in younger patients. Chapter 2.2 studied whether these visual MRI ratings can also be estimated computationally. A set of imaging biomarkers was automatically extracted from T1-weighted and FLAIR images, and a regression model was developed for estimating visual rating scale values from a combination of imaging biomarkers. The computed ratings produced results that corresponded well with visual ratings, further facilitating the use of these markers in clinical practice. Chapter 2.3 investigated which features in structural MRI scans could best distinguish Alzheimer’s disease (AD), frontotemporal dementia (FTD), vascular dementia (VaD), and dementia with Lewy bodies (DLB), and controls. An extensive set of features quantifying volumetric and morphometric characteristics was extracted from T1 images, and vascular characteristics from FLAIR images. From this study, I concluded that automatic MRI quantification methods are able to distinguish between the main types of dementia where different quantification methods provide complementary information.

Clinical decision support tools in the memory clinic

The second part of this thesis examined if we can combine the standard tests used in memory clinics and the abovementioned MRI features, to aid in diagnostics. For this a clinical decision support systems (CDSS) was developed. Chapter 3.1 describes this clinical decision support tool, designed for the differential diagnosis of different types of dementia. It combines information obtained from multiple diagnostic tests such as neuropsychological tests, MRI features (as described in Chapter 2.2 and 2.3) and CSF biomarkers. Here, I evaluated how the
tool performs in differentiating between controls with subjective cognitive decline (SCD), AD, VaD, FTD and DLB. The tool tested was highly accurate in separating the different dementias and controls from each other. Accuracy was highest for VaD and lowest for DLB. **Chapter 3.2** assessed how a precursor of this tool can aid clinicians to integrate biomarker evidence (CSF and MRI) to support AD diagnosis in patients with MCI. From this study, I concluded that the ability of the tool to identify Alzheimer pathophysiology is comparable to individual biomarkers. The tool has the advantage that it assigns likelihood to all patients, regardless of missing or conflicting data, allowing clinicians to integrate biomarker data in daily practice. **Chapter 3.3** studied whether the clinical decision support tool (as developed in **Chapter 2.3** and **3.1**) can also aid in identifying SCD subjects at risk for clinical progression, using data on demographics, cognitive test, CSF and automated MRI markers I found that combining diagnostic tests in a clinical decision support tool helps to identify cognitively normal individuals at risk of progression. The prediction model is particularly accurate in identifying subjects who are likely to remain stable. Classifiers such as this tool could aid clinicians in interpreting diagnostic test results and discuss results of these tests with cognitively normal individuals who present with SCD at a memory clinic.

**Survival in neurodegenerative diseases**

The final part of this thesis investigated whether we can translate clinical data obtained in the context of diagnosis to estimate survival in neurodegenerative diseases. **Chapter 4.1** set out to answer the question whether markers with proven diagnostic value also have prognostic value. This chapter aimed to identify determinants of mortality in patients with AD. I found that, in this relatively young sample of AD patients, disease-related determinants are associated with an increased risk of mortality, while neither comorbidity nor APOE genotype had any prognostic value. Finally, **Chapter 4.2** studied survival in a memory clinic based cohort with a relatively young age and representation of different dementia types. I concluded median survival time in dementia in patients with dementia is short, six years. When comparing younger (≤65 years) and older (>65 years) patients, median survival time hardly differed. Despite all efforts for better care and cure, and a general increased life time expectancy, survival in young onset patients has not improved since 2000.
GENERAL DISCUSSION

Combining and weighing biomarkers and diagnostic tests in clinical practice is difficult. The aim of this thesis was to improve care for patients with neurodegenerative diseases, by finding ways to optimize the use of tests for differential diagnosis and prognosis. In the previous chapters I have presented several steps that can aid the clinician in translating diagnostic test results to individual patients. In these final paragraphs, I will place the key findings of this thesis, as summarized below, in the context of current literature, focusing on methodological considerations, clinical implications and suggestions for future research.

Methodological considerations

This thesis focused on ways to guide the clinician in the challenges of daily practice, by translating diagnostic tests results to individual patients. The used approach, as described in the previous chapters, has several strengths and limitations.

To be able to use the results presented in this thesis in other clinical settings, it is essential that findings are robust and generalizable. The following strengths add to this. First, we used a

KEY FINDINGS OF THIS THESIS:

1) Visual ratings of atrophy and white matter hyperintensities are differently affected by age and diagnosis, requiring an age-specific approach in clinical practice. Their diagnostic value seems strongest in younger patients.
2) Visual ratings scales can be computed automatically. These automatically derived ratings perform as accurate as their visual counterparts in differential diagnosis.
3) A combination of automated MRI features has adequate accuracy to discriminate between the main types of dementia.
4) A clinical decision support tool is able to combine and visualize MRI features with cognitive tests and CSF biomarkers, and aids in the differential diagnosis of dementia.
5) An earlier version of this tool is useful to combine and weigh biomarkers to provide evidence for underlying Alzheimer pathology in MCI.
6) A clinical decisions support tool can be used to identify which individuals with SCD will remain stable and can thus guide the clinician in who to follow and who to reassure.
7) In dementia due to AD, disease-related determinants are associated with an increased risk of mortality, while neither comorbidity nor APOE genotype had any prognostic value.
8) Overall median survival in patients with dementia is short, only six years, even in young onset patients.
harmonized diagnostic protocol to analyze all patients in a similar manner. The patients were selected from the same memory clinic\textsuperscript{1, 2}, while multi-centers cohorts were used for external validation (in Chapter 2.2 and 3.3). The large majority of CSF analyses were performed in the Neurochemistry laboratory of the VUmc, minimizing inter-laboratory variability.\textsuperscript{3} Moreover, the large sample size of patients per study and the broad age spectrum ranging from 45 to 95 increased the robustness of our results. Also, follow-up duration in the studies examining prognosis was considerable, making the outcome measure more reliable, although it cannot be ruled out that stable subjects progressed after the studied timeframe.\textsuperscript{4-9} Since this thesis aimed to study ways to optimize diagnostic tests in daily practice, it is also essential that the presented methods can deal with missing data. To enable this, we used data that was typical of memory clinics, which is commonly varied and incomplete, including a large variety of image quality, enabling us to study methods to improve diagnostics in data representative of data used by clinicians. Moreover, the developed clinical decision support system (CDSS) can deal with incomplete data since it takes into account all available information and does not exclude subjects based on incomplete data.\textsuperscript{10-13} In Chapter 3.1, 3.2 and 3.3 we repeated the analyses for the subgroup of patients with complete data, finding similar results as for the total group. This encourages us that also in the patients studied the tool can deal with missing data. Finally, it is important that results are not only robust but also that they are generalizable.\textsuperscript{14, 15} In Chapter 2.2 and and 3.3 we were able to externally validate our developed models in independent cohorts. We found a similar accuracy in Chapter 2.2, however performance was lower in Chapter 3.3. This latter finding does not necessarily reflect lack of generalizability, but rather that the validation cohorts may lack sufficient follow-up to warrant reliable prediction. Another possible explanation is the difference in availability of biomarkers. Biomarkers are strongly related to clinical progression, with VUmc having a higher proportion available than the other cohorts.\textsuperscript{1, 16-18}

Some limitations also warrant discussion. First, for all studies, the clinical assessment (based on clinical information, cognitive tests, MRI scans and CSF biomarkers) was regarded as the gold standard. This thesis only considered the core diagnosis and disregarded information on underlying mixed pathologies. However, the situation is not as straightforward in clinical practice where mixed diagnoses are common, and the diagnosis of some of the underlying neurodegenerative disease can be only confirmed in autopsy.\textsuperscript{19-23} Especially in elderly patients, with comorbid small vessel disease, atrophy might also be the result of white matter hyperintensities (WMH) or hippocampal sclerosis and not of amyloid pathology.\textsuperscript{24, 25} Due to this I might have selected patients that have been misclassified with AD or another type of dementia. However, Chapter 2.1 found a similar degree of WMH in all elderly subjects, regardless of diagnosis, diminishing the importance of specifying the etiology as mixed or not. Also, in the clinical work-up clinicians were not blinded for biomarker and MRI results, which might have introduced circular reasoning. However, all diagnoses were made in our
multidisciplinary consensus meeting, in which the clinical characteristics of the patient and the cognitive profile on neuropsychological testing is leading.

Second, although the total sample sizes were considerable, some subgroups were small. In Chapter 2.2 the number of patients with the most severe grades of atrophy and WMH was small limiting the construction and validation of the model for these severe grades. In Chapter 2.2, 2.3, 3.1 and 4.1 I included subjects with FTD but did not differentiate between the subtypes of FTD, because this would have resulted in too small subgroups. The language variants are likely to be easier to classify for imaging analyses (in Chapter 2.2, 2.3 and 3.1) due to highly specific pattern of atrophy, which is completely different as compared to the pattern seen in the behavioral variant of FTD. Also, it cannot be ruled out that the language variants of FTD have a different disease progression and survival (in Chapter 4.1) as compared to the behavioral variant of FTD.

A third limitation could be the use of SCD as controls in the reference population in all Chapters, except 3.3. As shown in Chapter 3.3, SCD is a population at risk, where a minority of the patients has underlying Alzheimer pathology. However, if the SCD sample is indeed an enriched population, then the ability of the described methods to detect pathology would be underestimated. This implies the results would have been better if ‘healthy controls’ had been used. Also, underlying neurodegeneration can also not be excluded in “healthy controls”, as it is known that roughly one third of normal elderly harbors AD pathology and vascular pathology also increases with increasing age (Chapter 2.1). Furthermore, the comparison of AD or any other type of dementia with SCD patients is a clinically relevant comparison, as this is the differential diagnosis that a clinician has to make every day.

Finally, a potential limitation could be that the population studied was from a tertiary memory clinic, which hampers generalizability. Especially for Chapter 4.1 and 4.2 the results may not accurately reflect survival time for dementia in general. However, this data provides an important extension of existing literature, that is most often restricted to older patients with unspecified types of dementia. In Chapter 2.2 and 3.3 we were able to externally validate our developed models in independent cohorts, to increase variability.

Clinical implications

Based on this thesis several recommendations for clinical practice can be made. Let’s go back to the patients that I introduced at the beginning of this thesis, and who are summarized in figure 1. The three patients all illustrate challenges that clinicians face in interpreting test results of an individual patient, in the context of both diagnosis and prognosis.

Chapter 2.1, 2.2 and 2.3 studied ways to operationalize methods to extract more information from MRI scans for clinical use, in fairly simple ways, by applying age-related cut-offs and
using software that automatically analyses the scans. These methods can be applied in daily clinical practice, independent of the level of experience and could increase the standard of (diagnostic) care. However, a clinical diagnosis is not based on MRI findings alone, but is based on a combination of different types of data.\textsuperscript{2, 45} This is confirmed by the case of patient A, where only MRI does not provide enough information to make a diagnosis.

**Patient A**
Patient A is a 73 year old male, with a history of a transient ischemic attacks. He visits our memory clinic because he is experiencing memory and word finding problems since 3 years. On neurological examination we find only subtle rigidity. On the MMSE, a global screening test for memory impairment, he scores 28/30. Cognitive testing revealed impairment in memory, naming and executive functions, and discrete slowness. MRI showed only mild atrophy (MTA and GCA 1) and mild vascular damage (Fazekas 1). No CSF is obtained.

**Patient B**
Patient B is a 62 year old female who comes to our clinic with memory complaints. She has a MMSE of 27/30 and performs within normal limits on the cognitive tests, yet scores are slightly lower than expected based on age and educational level. Furthermore, on the MRI she has evident hippocampal atrophy (MTA 1 and 2). CSF testing showed normal biomarkers.

**Patient C**
Patient C is a 70 year old female, who is diagnosed in our memory clinic with dementia due to Alzheimer’s disease. She has no other diseases and took no medications. Her MMSE is low, 19/30, with low scores on additional cognitive tests. Her MRI scan however, shows very little atrophy (MTA 0.5/ 0), with little white matter hyperintensities (Fazekas 1). No CSF is obtained.

**Figure 1.** Three cases of patients from the VUmc Alzheimer Center to illustrate challenges clinicians face in daily practice.

We therefore need to combine the visual ratings and the automatic MRI quantification methods with information on demographics, cognitive tests and CSF biomarkers. Translating these different diagnostic tests to an individual patient is however not straightforward. \textbf{Chapter 3.1} presented a clinical decision support system (CDSS) as a method to combine and weigh data.\textsuperscript{12, 13}

In figure 2 we see the results after we have entered the data, including the raw MRI images, of patient A to the developed clinical decision tool. For this patient A, the tool points towards DLB, although there is considerable overlap with AD, as can be seen in the graph on the bottom left. On clinical follow-up this patient had a positive amyloid-PET scan as well as a positive DAT-scan, meaning the patient has DLB with comorbid Alzheimer pathology.\textsuperscript{46}
Figure 2. Results of PredictND tool for patient A, resulting in DSI and fingerprint.

In case of MCI or SCD, the challenge is not the differential diagnosis, but to discriminate who will remain stable and who will show progressive cognitive decline. The MCI-criteria of the National Institute on Aging-Alzheimer’s Association (NIA-AA) allow the clinician to use biomarkers to assign a likelihood of underlying AD pathophysiology and thus predict who will remain stable and who will progress to dementia. Again, translating these criteria to clinical practice is a challenge due to a lack of operationalization. Chapter 3.2 showed an earlier version of the tool, to overcome these challenges for using biomarker results in daily clinical practice, as it allows for missing, conflicting and borderline (ab)normal biomarkers (which occurred in a substantial part of subjects in the sample). The combination of a fitness and relevance value of each biomarker measurement allows the clinician to evaluate available biomarker evidence and use the NIA-AA research criteria in a balanced way. In this way, both professionals and MCI patients obtain more certainty concerning the cause of the syndrome and a risk estimate of future decline.

Most individuals presenting with SCD at a memory clinic are “worried-well”, only a small proportion of these patients suffers from preclinical AD. Former studies have shown that, CSF and MRI markers, and to a lesser extent cognitive tests, are associated with cognitive decline in SCD. How to translate these findings to clinical practice remains unclear. This is illustrated by patient B, who has SCD, but scores lower on cognitive tests results than expected. Her MRI scan shows moderate hippocampal atrophy, yet CSF biomarkers are
normal. Should I reassure her or follow her? In Chapter 3.3 a clinical decision support tool combined demographics, cognitive tests, automatic MRI methods and CSF biomarkers to estimate disease progression. In figure 3 we see the results when entering her data to the tool.

**Figure 3.** Results of PredictND tool for patient B, resulting in DSI and fingerprint.

For patient B, the tool provides a disease state index of 0.33, which means the patient looks more like a patient who remains stable. So, despite slightly abnormal cognitive tests and mild hippocampal atrophy, the otherwise normal biomarkers result in a low DSI, meaning I can reassure this patient and follow-up is not needed.

**Chapter 3.1, 3.2 and 3.3** showed that it is feasible to extract and combine information from routine diagnostic tests into a measures that can be used within a clinical decision support tool, supporting clinicians to identify the correct underlying etiology or prognosis. The tool is a data-driven tool that can use all available information from a specific population to fit the classification model. It is preferably trained on center-specific data, but it can also be successfully trained using other available datasets assuming they are sufficiently similar. This means the tool can also be implemented in daily practice in smaller clinics, possibly using a less extensive evaluation, and is not limited to be used in specialized centers. Clinical decision support tools might be especially useful for unexperienced clinicians and will equalize the treatment of patients regardless of in which hospital they are diagnosed. However, the tool is not magic, and accuracy is not perfect. Yet the purpose of these CDSSs is not to replace the clinician, but to assist the clinician mainly by visualizing the test results and
to provide an overview. These tools can provide a first step in taking personalized medicine to a next level, by visualizing how the patient’s data relates to certain underlying etiologies or prognosis. Recent research has shown that patients would like to be actively involved in decisions about diagnostic and prognostic testing, but they feel they often lack important information on the implication of the tests.\textsuperscript{62, 64}

Finally, we move to Chapter 4.1 and 4.2 for clinical implications for survival. Patient C illustrates the question patients often ask their clinician: ‘what can I expect in terms of disease progression’ and ‘how much time do I have until I die?’ Clinicians are however hardly able to answer these questions and predict the course of the disease for the individual patient. More knowledge is relevant to provide patient C with an accurate prognosis, which is important for advanced care planning. Chapter 4.1 showed that we cannot predict disease progression per se, but that on group level those with higher age and more severe AD-related determinants show shortest survival. Patient C is fairly young, has no co-morbidities, scores low on cognitive testing and has little atrophy on MRI. In Chapter 4.2 on group level patients with AD aged 65-75 year have a median survival is 6.7 (6.2-7.2) years, which is considerably shorter than the general Dutch population.\textsuperscript{65} This means for patient C, as visualized in figure 4, that the AD process itself, as reflected by neuropsychology, MRI and CSF biomarkers, has prognostic value in terms of mortality as well. This knowledge enables timely dialogue on prognosis, even in patient C who seems otherwise healthy.

\textbf{Figure 4} Prognostic factors and median survival that, on group based level, could be relevant for patient C.

Clinicians should be aware that neurodegenerative diseases causing dementia are lethal, especially in young onset patients. In these patients, it is important to enable timely discussions on advanced care planning, even if they are not frail or the ‘would you be surprised if this patient dies within the next year’ question is negative.\textsuperscript{66, 67} Finally, it is conceivable that
SUMMARY AND GENERAL DISCUSSION

over the years changes in patient care, e.g., the introduction of novel criteria or increased awareness resulting in earlier diagnosis resulted in longer survival. If diagnosis would be made at an earlier stage of the disease without influencing the actual disease course, this would introduce ‘lead time’ bias. Chapter 4.2 showed that despite increased awareness, median survival of these younger patients, has not changed between 2000 and 2014. The fear that increased awareness would lead to (too) early diagnosis with ensuing too many years lived with a diagnosis of a disease that can as yet not be cured, seems ungrounded.

**Future perspectives**

Based on the findings described in this thesis, several future perspectives emerge. Chapter 2.2, 2.3 and 3.1 showed that clinical decision support tools are helpful, but not yet optimal. Although patients with AD and VaD can be identified with high accuracy, accuracy for DLB and FTD is much lower. For DLB this is probably due to frequent occurring comorbid Alzheimer pathology, a non-specific pattern of atrophy on MRI and ‘Alzheimer light’ in CSF biomarkers. It is not without reason that the revised criteria for DLB emphasize the importance of the presence of other clinical features, like hallucinations, parkinsonism and REM-sleep disorders. The ability of a clinical decision support tools, to identify patients with DLB more accurate might increase by adding these features to the tool. FTD recognition might be hampered by subtle changes on regular cognitive testing and overlap in temporal atrophy on MRI with AD. Other clinical features, mostly behavioral and language changes, are much more important in the clinical profile of patients with FTD. Adding these features to the tool might improve recognition of FTD. Furthermore, elaborating automatic imaging methods could aid in pattern recognition for FTD. Asymmetry features (both left-right and anterior-posterior) might provide information for the discrimination of FTD from other groups. In addition, functional imaging modalities, such as PET, SPECT, and fMRI, have proven to produce complementary information for the differential diagnostics of dementias, whereas diffusion weighted MRI can be used to quantify white matter damage.

For the use of clinical decision support tools in the prognosis for MCI and SCD, it would first be of importance to find out how the tool is able to aid in identifying underlying AD pathology when applying the novel A (amyloid deposition)-T (pathologic tau)- N (neurodegeneration) classification as introduced in the novel research framework from the NIA-AA. One can argue the added value will probably be the same since the issues of missing data (for example with no CSF or PET scan the A and T criteria cannot be scored) or conflicting data (for example if there is atrophy on the MRI but a normal total tau value in CSF) remain present. Future research should also focus on what clinicians need and want from CDSSs. This thesis showed that a CDSS can aid in combining and weighing data, and visualize the results in a comprehensive way. The question remains how this helps the clinician. To evaluate the contribution of the tool to daily clinical practice, the developed tool is currently tested in a prospective study in several European memory clinics. In this prospective study, a data set containing a complete
set of data (neuropsychological tests, CSF sample, genetic biomarkers and MRI) is collected for all patients. Also, the change in diagnosis and confidence in the diagnosis before and after using the tool is recorded. However, this study was performed in highly specialized memory clinics, while clinicians with less experience and less resources might profit more from tools like CDSS and automatic MRI methods. A feasibility pilot and validation study in peripheral memory clinics would be a next step to investigate this further.

Finally, Chapter 4.1 and 4.2 showed we can estimate median survival and have identified factors that relate to mortality in AD. However, I was not able to predict the disease duration, per patient. This also guides us to future research. For both the patient and the clinician is it important to be provide accurate information on the disease course and duration. Therefor we should find better determinants to be able to predict prognosis in dementia. Novel studies could focus on for example neurofilaments in CSF, inflammatory markers in serum, MRI markers that predict brain age, or wearables that can observe the activities pf the patient at home.81–83

In conclusion, the main findings of this thesis were:

1) Structural MRI scans can be operationalized by applying visual ratings scales, when using an age-specific approach. Furthermore, the visual ratings can be computed automatically, and a combination of automated MRI features is able to discriminate between the main types of dementia.

2) A clinical decision support tool is able to combine and visualize MRI features with cognitive tests and CSF biomarkers, and aids in the differential diagnosis of dementia. This tool is also useful in providing evidence for underlying Alzheimer pathology in MCI and to identify individuals with SCD who will remain stable and can thus guide the clinician in who to follow and who to reassure.

3) Finally, in dementia due to AD, disease-related determinants are associated with an increased risk of mortality, while neither comorbidity nor APOE genotype had any prognostic value. Overall median survival in patients with dementia is short, only six years, even in young onset patients.

To further improve care for individuals with (incipient) neurodegenerative diseases, the studies described in this thesis, could be extended by expanding the used features in clinical decision support tools and by finding better determinants for prognostication. The findings of this thesis should be validated and replicated in prospective multicenter cohort studies.