MRI and Other Biomarkers in Early MS

PhD Thesis

Einar August Høgestøl

Faculty of Medicine

UNIVERSITY OF OSLO

Oslo, 2020

MRI and Other Biomarkers in Early MS

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-695-9

Cover: Hanne Baadsgaard Utigard.
Photo cover: Amalie Huth Hovland
Print production: Reprosentralen, University of Oslo.

Acknowledgements

Initially, there was just an idea in my head of what a doctoral thesis would encompass, yet ultimately it has become so much different from what I had expected. A wise person once told me that 'There is no specific PhD path, it is made by the PhD students walking it'. I found my research period both an intellectually liberation and a school of life in itself. I have not counted all the hours spent on the aspects besides my data, subjects, manuscripts, or thesis, but they are also part of my doctoral research path to become an improved version of myself. The days and nights (yes, doctoral candidates also work at night) have not been wasted, and I will use all my experiences in my future challenges, both at work and in life in general.

First, I would like to thank my main supervisor Hanne Flinstad Harbo for her guidance in recent years. She hired me for this project straight from a neurosurgical residency and hopefully saw some potential. Her energy and knowledge have been and continue to be a source of inspiration. I appreciate all the time she spent on pulling apart my manuscripts, followed by her helping me to shorten sentences, and to be clear and precise when putting the manuscripts together again. Her interpersonal skills and leadership experiences have hopefully in some sense been transferred to me. I also would like to thank my co-supervisor Gro Owren Nygaard for sacrificing her lunch break to convince me to start my doctoral research on multiple sclerosis. Her guidance and knowledge of the patients and the data have been invaluable. She has sacrificed many hours of her leisure time to answer my complex, yet naïve questions. I also have to thank my other co-supervisor Lars Tjelta Westlye for accepting and including me in his cutting-edge research environment at the Norwegian Centre for Mental Disorders Research (NORMENT). His technical insights and intellectual span are impressive and I have tried to comprehend the most of our conversations. The way he improved some sentences in our co-authored manuscripts still baffles me.

I am grateful to all of the patients who made time to visit our research barrack to redo the same tests yet one more time. I spent the most part of my first year of research getting to know all of their stories and rediscovering their neurological symptoms. Their enthusiasm for research inspired me to be responsible for taking care of their valuable research data.

I have been fortunate to be included in two excellent research groups. First and foremost is the Multiple Sclerosis Research Group at the University of Oslo and Oslo University Hospital (where it is part of the Neuroscience Research Unit), which is led by Elisabeth Gulowsen Celius. This research group constitutes a perfect blend of different people with diverse scientific backgrounds and frequently hosts students of all kinds from all parts of

the world. The list of all group members is long and I appreciate all moments spent on group meetings, lunches, seminars and other social activities with every one of them. In particular, I wish to mention Anna, Ingvild, Ina, Synne, Sigrid, Dan, Heidi, Cecilia, Stine Marit, Piotr, Mona, Christian, Pål, Chiara, Marte, and Line. Special thanks are owed to the three persons with whom I shared an office with at Domus Medica 4, namely Sandra Pillar Henriksen, Steffan Bos-Haugen, and Tone Berge. I still share an office with Tone, hopefully for many years to come.

With regard to the second research group, at NORMENT, I have appreciated all the morning meetings at 10 am, which enabled me to work in silence in the common office for almost two hours before anyone else arrived. This is a unique and cutting-edge research group with many dedicated and bright researchers. I have appreciated all hours spent in the common office and the computer cluster room, even though the heating was lacking during the cold winters. I have to thank the persons who have assisted me the most and inspired me to learn ever more detailed statistics, namely Christine, Ida, Luigi, Knut, Geneviève, Martina, Torgeir, Torbjørn, Nathalia, and Kristine. Special thanks are owed to the technical computer geniuses Dag Alnæs and Tobias Kaufmann for careful guidance and great teaching capabilities (and patience).

The remaining list of persons whom I wish to thank is to long for the space available and therefore I have to offer a general ‘thank you’. All help has been greatly appreciated and they have all played a part in enabling me to take this research path and hopefully successfully defend my thesis when the day comes.

Last and most importantly, I thank my family, who mean the world to me. My parents, my brothers and my in-laws have been helpful in so many ways. Although they have admitted that their understanding of my scientific work is sparse, they have been encouraging and helpful in all arenas during the years I have spent on the research. I owe special thanks to my father-in-law, Dag Bøgeberg, who worked wonders with illustrations in times of need.

I am fortunate and grateful to be the father of two perfect children, Eva and Olav, both of whom are unique. They have been a necessary source of distraction during my most intense working periods and a continuous to remind of what matters most. The single person who has been there every minute for the past 15 years is my fantastic wife Kaja. Words cannot encompass her importance. I am truly grateful for her encouragements when times have been tough, and even though she does not acknowledge it, she has have contributed in so many ways to follow me along my doctoral research path from the beginning until the end.

Table of Contents

List of Figures

List of Tables

Abbreviations

ARMSSS	Age Related Multiple Sclerosis Severity Score
ASL	Arterial Spin-Labelling
BAG	brain age gap
BBB	blood–brain barrier
BDI-II	Beck Depression Inventory – II
BICAMS	Brief International Cognitive Assessment in Multiple Sclerosis
BMI	body mass index
BOLD	blood oxygen level dependent
brain-PAD	brain-predicted age difference
BRB	Brief Repeatable Battery
BVMT-R	Brief Visuospatial Memory Test – Revised
CIS	clinically isolated syndrome
CNS	central nervous system
CombiWISE	Combinatorial Weight-adjusted Disability Score
COWAT	Controlled Oral Word Association Test
CSF	cerebrospinal fluid
CT	computerized tomography
CVLT-II	California Verbal Learning Test – II
D-KEFS CWIT	Delis-Kaplan Executive Function System Colour-Word Interference Test
DCE	dynamic contrast-enhanced
DIS	dissemination in space
DIT	dissemination in time
DMN	default mode network
DMT	disease-modifying therapy
EAE	experimental autoimmune encephalomyelitis
EDSS	Expanded Disability Status Scale
FLAIR	fluid-attenuated inversion recovery
fMRI	functional MRI
FOV	field of view
FSL	FMRIB Software Library
FSMC	Fatigue Scale for Motor and Cognitive Functions
FSPGR	fast spoiled gradient-echo
FSS	Fatigue Severity Scale
GM	grey matter
GRAPPA	GeneRalized Autocalibrating Partial Acquisition
HC	healthy controls
HLA	human leukocyte antigen
HSCT	hematopoietic stem cell transplantation
ICA	independent component analysis
IFN	interferon
IMSGC	International MS Genetics Consortium
JC virus	John Cunningham virus
LME	linear mixed-effects
MACFIMS	Minimal Assessment of Cognitive Function in MS
MAGNIMS	Magnetic Resonance Imaging in MS
MFIS	Modified Fatigue Impact Scale
MMS	Mini Mental Status

MRI	magnetic resonance imaging
MS	multiple sclerosis
MS-DSS	Multiple Sclerosis Disease Severity Scale
MSFC	Multiple Sclerosis Functional Composite
MSSS	Multiple Sclerosis Severity Score
NAWM	normal-appearing white matter
NEDA	no evidence of disease activity
NfL	neurofilament
NORMENT	Norwegian Centre for Mental Disorders Research
OCB	oligoclonal band
OCT	optical coherence Tomography
PASAT	Paced Auditory Serial Addition Test
PCA	principal component analysis
PPMS	primary progressive MS
RCFT	Rey Complex Figure Test and Recognition Trial
RIS	radiologically isolated syndrome
ROI	regions of interest
RRMS	relapsing-remitting form
rs-fMRI	resting state functional MRI
SDMT	Symbol Digit Modalities Test
SF-12	Short Form Health Survey (12-items)
SPMS	secondary progressive MS
TE	echo time
TR	repetition time
VEP	visual evoked potentials
WASI	Wechsler's Abbreviated Scale of Intelligence
WM	white matter
XGBoost	extreme gradient boosting

Norsk sammendrag

MS er en kronisk nevrologisk sykdom som rammer sentralnervesystemet og medfører varierende grad av uførhet. MS er den vanligste årsaken til uførhet blant unge voksne foruten ulykker, og blir typisk diagnostisert i tidlig voksen alder. De siste 25 årene har vi gjennomgått et paradigmeskifte innen MS behandling, der utviklingen av effektive bremsemedisiner har medført en betydelig bedring av prognose for de med MS. I Norge i dag lever over 12 000 personer med MS, omtrent halvparten av disse mottar behandling rettet mot å bremse sykdomsprogresjonen. Det har vært et økende fokus på tidlig diagnostikk og tidlig effektiv behandling, mens det også i de siste par årene har vært rettet spesielt fokus mot persontilpasset behandling.

MR har over de siste tiårene fått en stadig større betydning for både diagnostikk og oppfølging av personer med MS. Tidlig på 80-tallet ble det vist at MR maskiner hadde en overlegen evne sammenliknet med konvensjonelle CT maskiner, til å oppdage de typiske lesjonene som kan oppstå i hjernen og ryggmargen hos personer med MS. Store forskningsgrupper har siden 80-tallet hatt stort fokus på MR forskning innen MS. Det er den visuelle undersøkelsen av MR bildene gjort av en erfaren radiolog eller nevroradiolog, som gir vurderingen av MR bildene som utføres som ledd i utredning og oppfølging av personer med MS. I tillegg så er det de siste årene, gjennom MAGNIMS konsortiet, blitt definert gitte kriterier for spesielt diagnostiseringen av MS. Det er imidlertid et kjent paradoks at de typiske MR forandringene man ser hos MS pasienter ikke nødvendigvis korrelerer med graden av uførhet hos pasienten, også kjent som det «klinisk-radiologiske paradokset».

I mitt doktorgradsarbeid har jeg undersøkt ulike MR markører hos personer med MS som er i tidlig fase av sykdommen. Vi har fulgt 76 nydiagnostiserte personer med MS, fra kort tid etter diagnose og fem år frem i tid. Pasientene har blitt undersøkt i forskningssammenheng ved totalt tre anledninger. Både klinisk nevrologisk undersøkelse, kognitive tester, MR undersøkelse, selvrapporteringsskjemaer, forskningsundersøkelser, øyeundersøkelser, blodprøver og genetiske analyser har blitt utført.

I min første artikkel benyttet vi data fra første undersøkelsestidspunkt med en funksjonell MR sekvens og undersøkte sammenhenger med depressive symptomer og grad av fatigue. Vi undersøkte spesifikt hvordan aktiviteten i hjernens hvilemodusnettverk ble påvirket av grad av fatigue og depressive symptomer. I et forsøk på å dele opp effektene av de ulike symptomene, benyttet vi en prinsipalkomponent-analyse på våre data. Vi fant at det var en sammenheng mellom økte symptomer av både fatigue og depressive symptomer med en

overaktivering av hjernens hvilemodusnettverk. I tillegg fant vi, ved hjelp av prinsipalkomponent-analysen, en sammenheng for en undergruppe av studiedeltakerne der liten byrde av depressive symptomer og høy grad av fatigue var assosiert med økt aktivitet i hjernens hvilemodusnettverk.

I min andre artikkel etablerte vi en maskinlæringsmodell for estimering av hjernealder ved hjelp av 1118 MR mål fra strukturelle sekvenser, basert på 3208 friske kontroller, til å kunne appliseres på hele vårt longitudinelle MS materiale. Ved hjelp av ytterligere 235 matchede friske kontroller beregnet vi, for hver MR undersøkelse fra MS deltakerne, om hjernen tilsvarte en yngre, lik eller eldre hjerne enn tilsvarende den kronologiske alderen til deltakerne med MS. Vi laget en egen estimeringsmodell for hele hjernen, i tillegg til syv estimeringsmodeller for spesielle hjerneområder for å kunne undersøke også de regionale forskjellene i hjernealder hos personene med MS. Våre resultater viste at personer med MS som i gjennomsnitt var 40 år, ved hjelp av våre modeller for å estimere hjernealder i snitt var 4.4 år eldre enn deres kronologiske alder. I tillegg viste vi at ved hjelp av de longitudinelle dataene kunne estimere at hjernealdringen i gjennomsnitt var økt med 41% sammenliknet med forventet kronologisk aldring. Vi fant også at det var store regionale forskjeller i estimert hjernealder for personene med MS, der lillehjernen og subkortikale områder syntes å eldes tidligere og mer markant.

I min tredje artikkel benyttet vi de allerede etablerte hjernealderestimatene til deltakerne med MS, og undersøkte om vi fant noen sammenhenger med de kognitive testene som ble utført i forløpet av hele studien. Vi fant signifikante forbedringer av de kognitive testresultatene over tid, som også i stor grad kan tilskrives læringseffekter. Vi undersøkte de overordnete kognitive domene for eksekutive funksjoner, hukommelse og prosesseringshastighet. Vi fant en signifikant sammenheng der tregere prosesseringshastighet var assosiert med økt hjernealder. Til sammenlikning, fant vi også at reduserte hastigheter for en annen test for prosesseringshastighet var signifikant assosiert med mindre volumer av thalamus. Talamus er kjent for å være et område i hjernen som ofte viser sammenheng mellom størrelsen av talamus og spesielt kognitive funksjonstester, noe vi også fant i våre resultater.

I alle mine artikler har vi funnet flere mulige MR markører som potensielt kan benyttes i en klinisk setting for å bedre hjernehelsen til personer med MS. Videre studier er nødvendig for å kunne bekrefte og utvide våre funn.

1. Preface

An investment in knowledge pays the best interest
(Benjamin Franklin, "The Way to Wealth", 1758)

The greater our knowledge increases, the more our ignorance unfolds
(John F. Kennedy, "We choose to go to the Moon…", 1962)

The technological advances in recent decades have enabled researchers to exploit increasingly more of the data acquired from magnetic resonance imaging (MRI) scans of the human brain. However, to translate this vast body of information into clinical use in order to improve the brain health of people with multiple sclerosis (MS) has proven a major challenge.

I faced the challenge of translating MRI data to something meaningful during the period when I conducted my research as a doctoral candidate from December 2015 to 2020 at the Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway. This thesis is based on my investigation of different MRI markers in the brain, specifically brain connectivity based on functional MRI and estimated brain age using state-of-the-art machine learning methods.

I hope you enjoy reading this thesis.

Einar August Høgestøl

Oslo, 19 March 2020

2. Scientific environment

Institute of Clinical Medicine
Faculty of Medicine
University of Oslo
Oslo, Norway

Department of Neurology
Oslo University Hospital
Oslo, Norway

Multiple Sclerosis Research Group
University of Oslo & Oslo University Hospital
Oslo, Norway

Norwegian Centre for Mental Disorders Research
University of Oslo & Oslo University Hospital
Oslo, Norway

Neuroscience Research Unit
Oslo University Hospital & University of Oslo
Oslo, Norway

3. List of publications included

Paper I (published)

Einar A. Høgestøl, Gro O. Nygaard, Dag Alnæs, Mona K. Beyer, Lars T. Westlye,* Hanne F. Harbo*. Symptoms of fatigue and depression is reflected in altered default mode network connectivity in multiple sclerosis. PLOS ONE, 2019 April, doi:10.1371/journal.pone.0210375

Paper II (published)

Einar A. Høgestøl, Tobias Kaufmann, Gro O. Nygaard, Mona K. Beyer, Piotr Sowa, Jan E. Nordvik, Knut Kolskår, Geneviève Richard, Ole A. Andreassen, Hanne Flinstad Harbo, Lars T. Westlye. Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis. Frontiers in Neurology, 2019, April, doi:10.3389/fneur.2019.00450

Paper III (manuscript)

Einar A. Høgestøl, Gro O. Nygaard, Petter E. Emhjellen, Tobias Kaufmann, Mona K. Beyer, Piotr Sowa, Ola A. Andreassen, Elisabeth G. Celius, Nils I. Landrø, Lars T. Westlye, Hanne F. Harbo. Preprint: Brain age estimation is a sensitive marker of processing speed in the early course of multiple sclerosis. bioRxiv, 2019, May, doi:10.1101/651521

*The authors contributed equally

4. Introduction

4.1 The history of MS

One of the first recorded multiple sclerosis (MS) relapses was that of Lidwina van Shiedam, who fell while skating on 2 February 1396, when she was 15 years old (1). The fall itself and her broken rib are not the interesting part of the record, but rather the subsequent neurological symptoms of walking difficulties, pain in her teeth, and headache in the wake of the aforementioned injury, which are well documented. Before her death on 14 April 1433, Lidwina van Shiedam had experienced several periods of increased neurological symptoms, but also periods with partial improvements were recorded. Those periods with improved walking and vision were thought to have been due to contact with 'God' and 'the Angels', and she was therefore later made a Saint by the Church (1).

Jean Cruveilhier and Robert Carswell were the first persons to describe and illustrate MS, in 1835 and 1838 respectively (2). Later, the renowned Dr Jean-Martin Charcot named the disease '*sclerose en plaques*' after finding plaque in the post-mortem brain of a young female who had previously been treated for tremor and difficulty speaking, and thereafter he linked the pathological findings to the neurological symptoms (3). The name of the disease describes the early descriptions of pathological changes with multiple hard lesions (sclerosis) in the central nervous system (CNS). The myelin sheath covering the axons was discovered in 1878 (4). In the 19th century, a new study concerning the aetiology of MS shifted the understanding of MS as a viral infection to MS developing because of an immune response when Thomas Rivers successfully induced experimental autoimmune encephalomyelitis (EAE) in rodents and monkeys (4, 5).

In 1959, it was discovered that the B and T-cell lymphocytes were closely associated with EAE, and later the human leukocyte antigen (HLA) locus was linked to MS in 1972 (6-11). When it was found that relatives of patients with MS were more often affected by the HLA risk factor than were the general population (12), the race for the genetic underpinnings intensified and finally led to the first genome-wide association study in MS, in 2007 (13). Today, the updated genetic MS map includes 238 genetic variants that are associated with MS (Fig. 1) (14). As history has revealed, many important research milestones have enlightened us on the 'why' and 'how' questions of MS.

However, those who treat people with MS see how debilitating the disease can be despite extensive MS research efforts during the past centuries. The next research breakthroughs will hopefully improve overall MS care by shedding light on myelin repair

drugs, personalized treatment approaches, improved biomarkers, and the application of artificial intelligence utilizing big data (15).

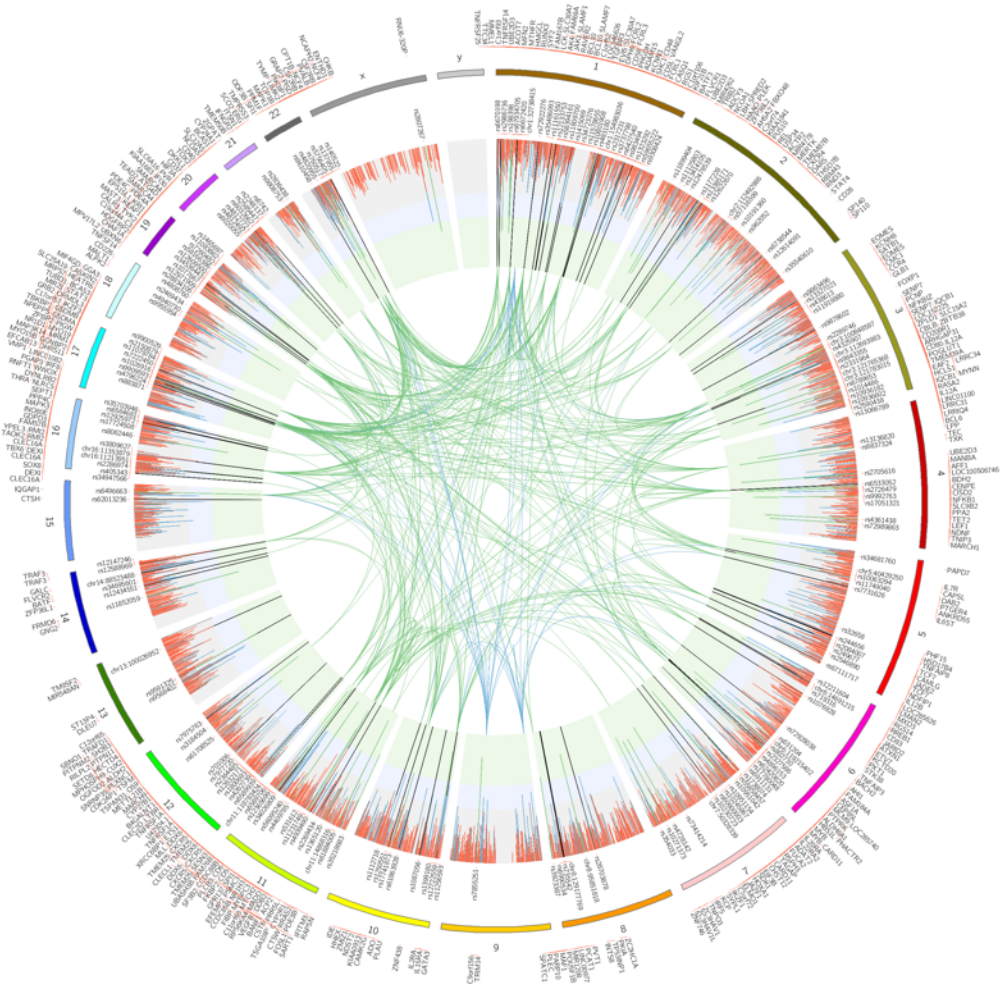

Fig. 1. The current genetic map of MS. The map, from 2019, is based on 47,351 MS subjects and 68,284 healthy controls. Reproduced with permission from the International MS Genetics Consortium (14), in accordance with International Public License CC-BY-ND 4.0.

4.2 Epidemiology

In 2013, MS International Federation estimated that about 2.3 million humans in the world had MS, resulting in a global prevalence of 33 per 100,000 (16). The MS International Federation expects to have a comprehensive update of its atlas of MS finished in 2020.

Norway is among the countries with the highest reported prevalence of MS, with 203 per 100,000 persons, although this is still lower than the prevalence in the Orkney Islands, Scotland, which is 403 per 100,000 (17, 18). Fig. 2 shows how the reported incidences of MS vary worldwide. Thus, both the prevalence and incidences of MS vary geographically around the world.

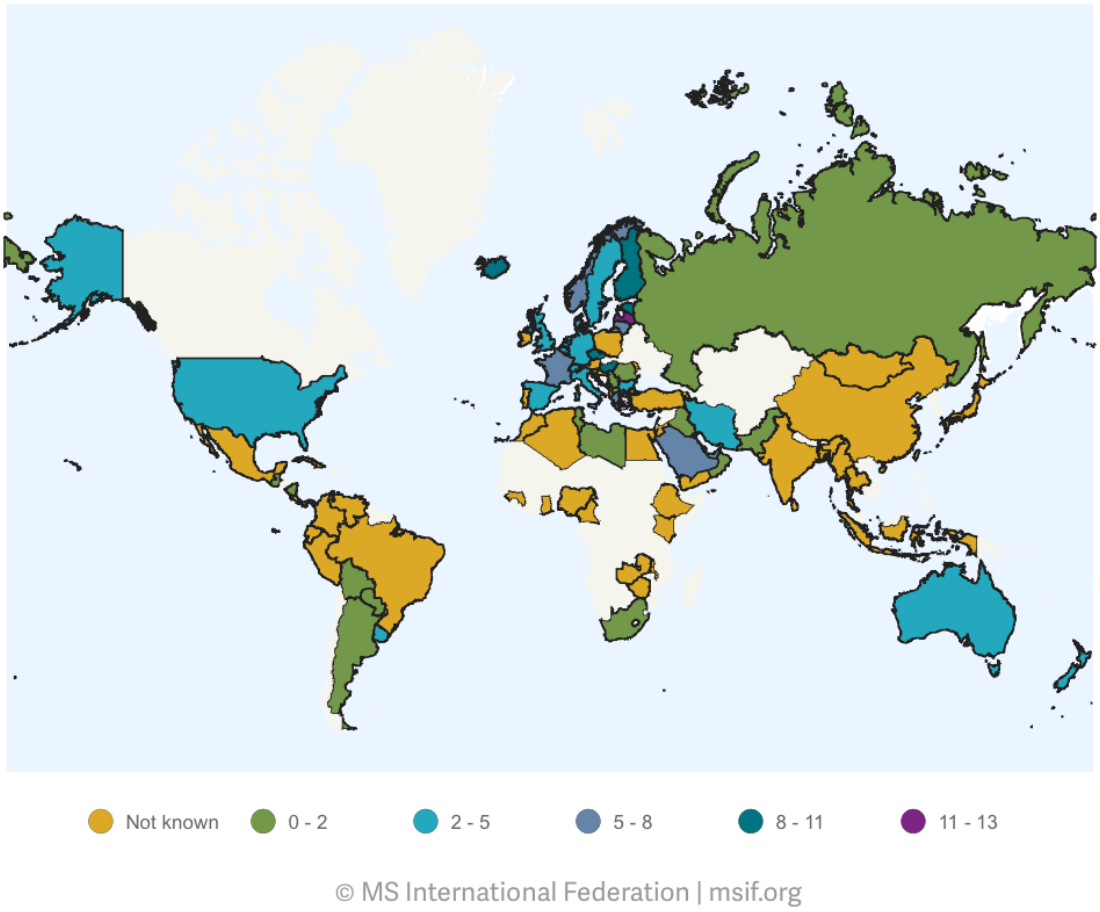

Fig. 2. Worldwide incidences of people with MS in 2013. The estimates were based on reported numbers for the years 2008 to 2012 inclusive. The colour codes indicate incidences per 100,000 inhabitants. Reproduced with permission from the Multiple Sclerosis International Federation (16).

Earlier demographic data showed increased MS susceptibility in white adults (19), but this finding has been challenged by recent reports of increased prevalence in black adults living in the USA (20, 21).

Although the latitude has not been reported as having an effect on MS prevalence in Norway (18), previous epidemiologic reports have provided evidence of a gradient of increasing MS risk from the Southern Hemisphere to the Northern Hemisphere (22).

Furthermore, two publications, dating from 2010 and 2011, report findings of latitude effects in some parts of the world, but not in others (23, 24).

An increasing female-to-male ratio has been reported for people with MS, from 1.4:1 in 1955 to 2.3:1 in 2000 (22). The main reason for the increased gender ratio is believed to be increased smoking by females in the 20th century, while reproductive factors are hypothesized to contribute to the gender ratio (22).

4.3 Risk factors in MS

To date, a specific trigger for MS is unknown, although several risk factors have been identified as contributing to MS susceptibility without being the cause of the development of MS (Fig. 3) (25). The relationship between MS and exposure to sunlight has been extensively investigated (26, 27). Several studies have demonstrated a relationship between the prevalence of MS and the disease activity of MS, serum vitamin D levels, exposure to ultraviolet (UV) sunlight, and sun exposure (27-31).

The optimal blood levels of 25-OH vitamin D for MS patients is still subject to debate, although a higher level than considered optimal for the general population is widely accepted in the latest studies (32-34). The resulting vitamin D serum levels for people with MS is a complex interaction of several factors, such as the expected variation over time, exposure to sunlight, immunoregulatory effects, and intake through diet (26, 32). New large randomized controlled trials on vitamin D supplements in MS will hopefully provide more knowledge of how to improve MS care with vitamin D supplements.

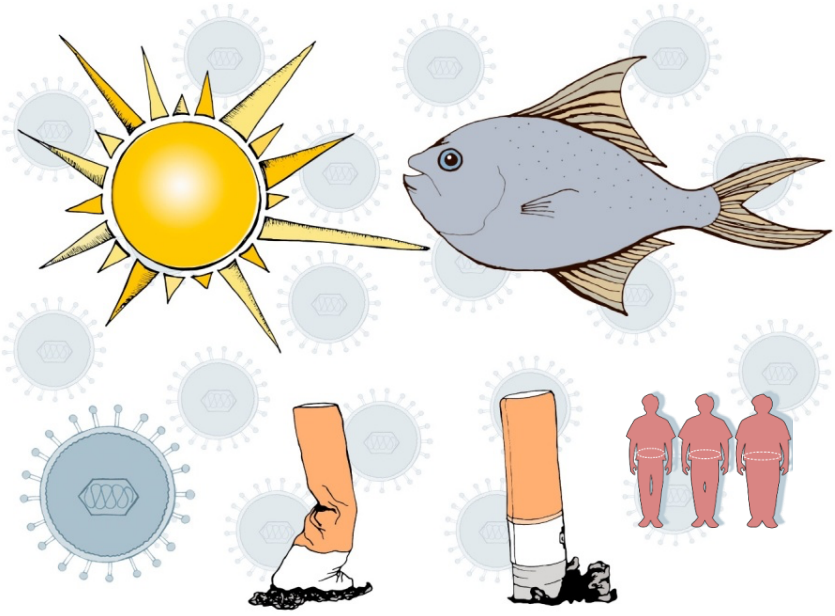

Fig. 3. Some of the environmental factors that contribute to the risk of MS. Known factors that affect the risk of developing MS include exposure to sunlight, viruses such as the Epstein-Barr virus, smoking, and obesity. Source: Multiple Sclerosis Research Group, Oslo University Hospital.

Smoking cigarettes increases the risk of developing MS (increased odds ratio of 1.54) (26, 35-38). There is also moderate evidence indicating an accelerated disease course for MS for people who smoke (26, 35).

Obesity (body mass index (BMI) > 30) in early life increases a person's risk of developing MS by a factor of two (37). Further research is needed to validate to what extent the increased risk observed in young obese persons is still evident after accounting for lower vitamin D levels (37).

The Epstein-Barr virus is considered a strong risk factor for the development of MS, but the mechanisms are not known (37, 39-42). In an article published in 2016, Ascherio et al. estimate that almost two-thirds of future MS cases could potentially be avoided by reducing smoking, Epstein Barr Virus (EBV). infection, and obesity in the young, and by taking vitamin D supplements (37). There is now a collective body of evidence that should encourage societies to focus on preventing these modifiable risk factors for MS (43).

The most recent International MS Genetics Consortium (IMSGC) genetic study found that 238 genetic variants were significantly associated with MS, each of which very modestly contributes to MS susceptibility (14) (Fig. 1). A large contribution to the genetic risk is conferred by genes encoding molecules in the immune system (*HLA-DRB1*15:01*, *IL2* and

IL7R). To understand the complex interplay of these genetic variants and other omics data, advanced mathematical models have been and are continuing to be applied to large MS datasets, for instance the ongoing collaborations MultipleMS (44) and Sys4MS (45). These projects use a systems biology approach to disentangle the underlying biology in MS.

Pathway analyses now enable identification of the most functionally associated gene per region and cell type, which can open new methods to integrate past and future genetic data in order to elaborate the functional aspects of disease aetiology and possibly discover new MS phenotypes (46). Also, systems biology approaches in conjunction with all available genetic data are proving useful to identify and interpret the genetic underpinnings of MS (46).

4.4 Pathophysiology in MS

MS is assumed to be an immune mediated chronic inflammatory disease that affects the CNS (47). MS is believed to be triggered by the environmental risk factors mentioned in the preceding section (Section 4.3) in a complex interplay in genetically susceptible people (47, 48). The pathophysiological process of MS includes a temporary destruction of the blood–brain barrier (BBB), multifocal inflammation, demyelination, oligodendrocyte loss, reactive gliosis, and axonal degeneration (49, 50). Immune cells migrate across the BBB. Macrophage and microglial activation induce proinflammatory cytokine, oxygen and nitric oxide radicals, leading to demyelination and possible axonal loss (50).

4.4.1 Immune pathophysiology

Immune cells enter the CNS through the blood vessels of the BBB, the subarachnoid space (SAS) and the choroid plexus (Fig. 4) (38). In MS relapses, cells from both the innate and adaptive immune systems, such as CD4+ and CD8+ T cells, B cells and myeloid cells, infiltrate CNS parenchyma around post-capillary venules to cross through the BBB (38). These immune cells, together with locally activated microglia and astrocytes, contribute to oligodendrocyte injury, demyelination and neuroaxonal injury through cell contact-dependent mechanisms and the secretion of soluble factors (38). In late MS, the episodic infiltration of immune cells into the CNS is reduced. Mechanisms contributing to tissue injury in late MS include neurodegeneration (neuroaxonal, astrocyte and oligodendrocyte damage), due to acute or chronic oxidative stress promoted by innate and adaptive immune cell activation, mitochondrial dysfunction, extracellular free iron accumulation, loss of myelin trophic support, hypoxia, altered glutamate homeostasis, and a pro-inflammatory environment, with

involvement of cytotoxic factors and complement activation (38). Chronic inflammation is potentially mediated both by ongoing CNS compartmentalized inflammation involving meningeal immune cell infiltrates (e.g. B cells) that can form lymphoid-like structures and by local CNS innate cells (e.g. microglia) (38).

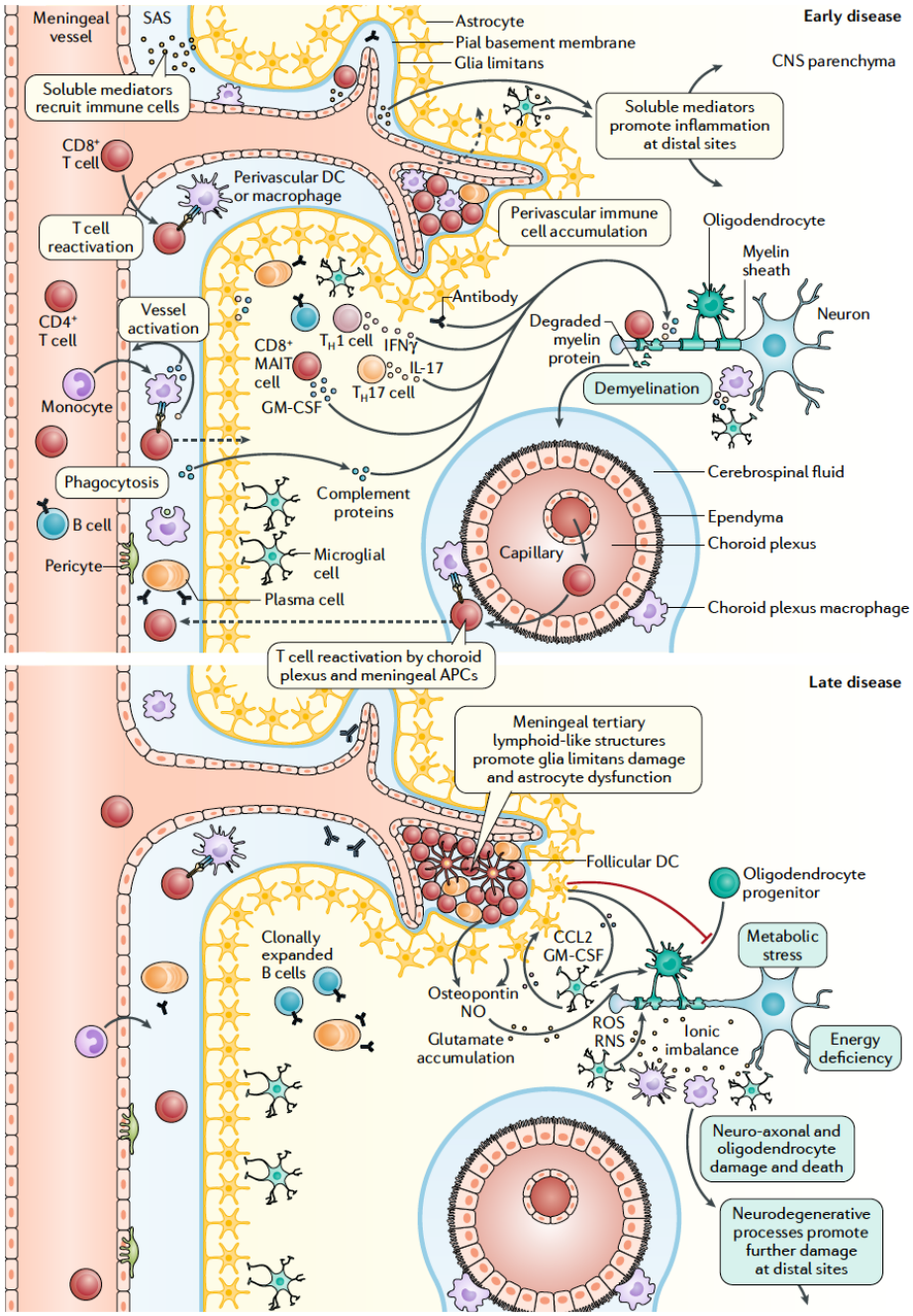

Fig. 4. Overview of the dysregulation of the immune system in both early and late MS within the CNS. Reproduced from Filippi et al. (2018) (38), with permission from Springer Nature Ltd., obtained 8 August 2019.

4.4.2 White matter lesions

The pathological hallmark of MS is the perivenular inflammatory lesions, which in turns lead to demyelinating MS lesions (51). In the early phases of MS, it is common to find demyelinating lesions, with lymphocyte infiltration by major histocompatibility complex (MHC) Class I restricted $CD8^+$ T cells, $CD20^+$ B cells, but also plasma cells, activated microglia, macrophages, and large, reactive astrocytes. This leads to oligodendrocyte destruction and demyelination (48). In the early phases of MS, the axons themselves are usually preserved, although if the disease progresses, the patient with MS will develop irreversible axonal damage.

The progressive phenotypes of MS tend to be characterized more by inactive CNS lesions. These lesions are commonly sharply delineating, hypocellular and have defined demyelination, decreased axonal thickness, reactive astrocyte gliosis, fluctuating microglial activation isolated in periplaque white matter, and a lower quantity of lymphocytes than active lesions (38). The discrimination between active and inactive lesions in the progressive phenotypes is not clear-cut. Also, inflammatory mechanisms may play an important role in the immune pathophysiology in both primary progressive MS (PPMS) and secondary progressive MS (SPMS). There is no distinct histological difference between the MS phenotypes (48). Active and mixed (both active and inactive) lesions account for up to 57% of all lesions in the progressive phenotypes of MS (38). Active lesions in PPMS and SPMS seem to be associated with a more progressive disease course (38).

Furthermore, with the use of advanced imaging techniques and microscopic analyses, the macroscopic non-affected white matter (WM), referred to as normal-appearing WM (NAWM), exhibits features of both widespread inflammation and neuroaxonal damage (38). NAWM damage is more pronounced in the progressive forms of MS and includes macrophage infiltration, widespread microglial activation, gliosis, small round cell infiltration, axonal degeneration, and demyelination (38).

4.4.3 Grey matter lesions

Cortical demyelination in the forebrain and cerebellum that is already evident from the pre-clinical stage of the disease process can also be seen in autopsies of people with MS (38, 52, 53). Cortical demyelination is more prominent in the progressive phenotypes of MS. Grey matter (GM) lesions can occur in the deep GM nuclei and the GM in the spinal cord. Cortical lesions are usually located in the cortical sulci and in the deep invaginations on the surface of

the brain, often associated with inflammatory infiltrates of the meninges (38). Studies indicate that lesions in the cortex are driven by pro-inflammatory mediators from the meninges or from the cerebrospinal fluid (CSF). The presence of reduced synaptic density in MS patients is also, evident in normal-appearing GM in autopsies, which is suggestive of an independent finding in addition to focal demyelination in the cortex. Lesions in the cortex show less contrast enhancement in MRI scans, less oedema, a reduced degree of inflammation, and more efficient myelin repair, which indicates the presence of unique processes in the grey matter (38, 54).

Cortical lesions within the GM can be divided into four different subtypes: Type I lesions are located at the cortico-subcortical border affecting both GM and WM; Type II lesions are small perivenular intracortical lesions that do not affect either WM or the pial surface of the brain; Type III lesions (most frequent in patients with MS, and related to meningeal inflammatory infiltrates) extend inward from the subpial layers of the cortex; and Type IV lesions stretch through the whole thickness of the cortex but without bridging over to the WM (38, 55-57).

It has been demonstrated that remyelination is far more common in GM lesions (up to 90%) than in WM lesions (40–50%) (38). The extent to which a lesion is remyelinated has been found associated with several factors, including age, lesion location, disease duration, axonal integrity, and the existence of oligodendrocyte progenitor cells (38). Furthermore, remyelination is more common in the RRMS than in the progressive phenotypes (38). Remyelination, which occurs both in MS and in other neuroinflammatory and neurodegenerative diseases, has recently been in focus as a candidate process for targeted treatment (58). Recent studies indicate that optimal remyelination requires death of pro-inflammatory microglia followed by the replacement of microglia in a pro-regenerative state (58). Specifically, one study highlighted the CNS region-specific consequences of interferon (IFN) signalling in the WM, where it was found that positively regulated Type 1 IFN signalling promoted microglia repopulation (58).

4.5 MS diagnosis

George A. Schumacher and other leading MS researchers reached consensus on the diagnostic criteria and definition of MS in 1965 (59). The criteria for diagnosing MS were revised in 1983 by Charles M. Poser and colleagues (60), and updated in 2001, 2005, 2010, and 2017 (Fig. 5) (61-64).

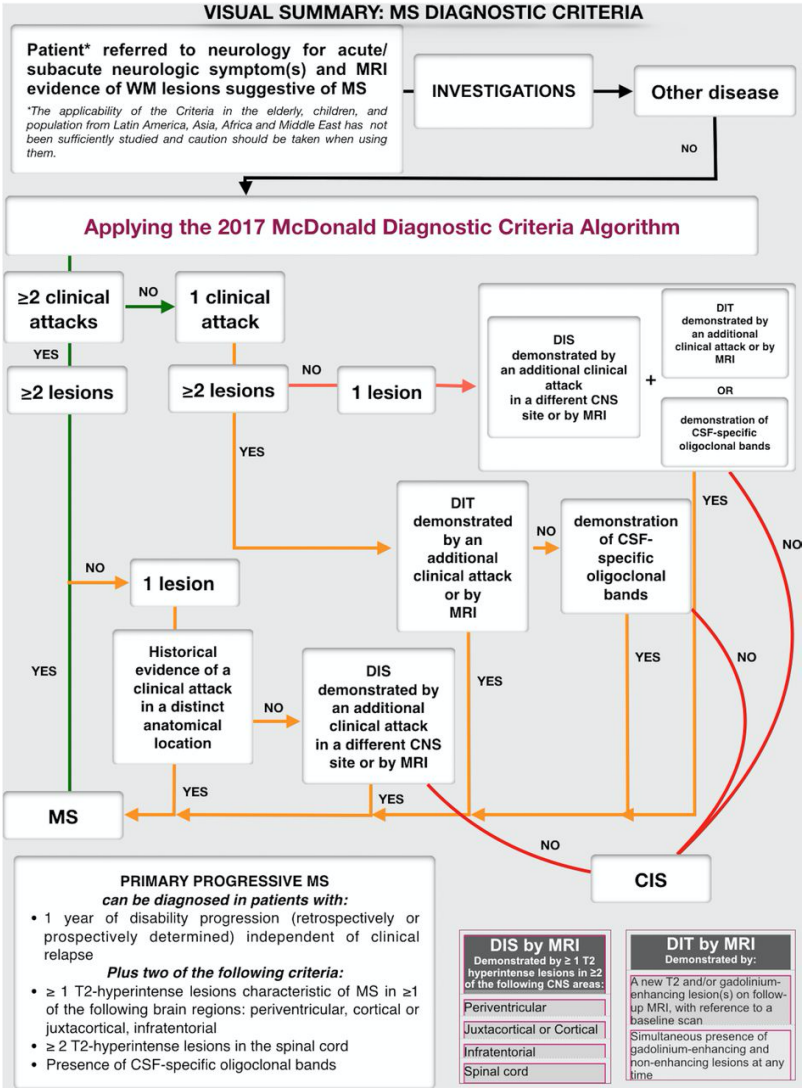

Fig. 5. Overview of the practical implications of the 2017 revisions to the diagnostic criteria for MS. Lesion(s) refers to an area of hyperintensity on a T2-weigthed or proton-density weighted MRI scan that is at least 3 mm in longitudinal axis. Reproduced from De Angelis et al. (2019) (65), with permission from BMJ Publishing Group Ltd., obtained 12 August 2019.

As summarized by De Angelis et al. in 2019 (65), the main requirements for the diagnosis of MS should be based on the following:

- Objective clinical evidence of CNS involvement
- Evidence of lesions disseminated in time (DIT) and space (DIS)
- Exclusion of other conditions that could better explain the findings.

The 2017 revision of the MS criteria established unmatched CSF oligoclonal bands (OCBs) as a validation of dissemination in time (Table 1). When a patient presents at a clinic with a

clinically isolated syndrome (CIS), the presence of OCBs in CSF will act as a marker for dissemination in time. This in turn will allow both for an earlier diagnosis of MS if all other requirements for the MS diagnosis are met, and for the start of early and effective treatment (66). In previous revisions of the MS diagnostic criteria, a distinction was made between symptomatic and asymptomatic MRI lesions, but this was removed in the 2017 revision. In addition, MRI lesions in the cortex were added as a possible area that would be considered as dissemination in space (67).

Table 1. Overview of the 2017 revisions to the diagnostic criteria for MS. Reproduced from Filippi et al. 2018 (38), with permission from Springer Nature Ltd., obtained 8 August 2019.

The 2017 revised criteria for the diagnosis of MS

Relapsing–remitting MS

- At least two clinical relapses and objective clinical evidence on neurological examination of at least two lesions with distinct anatomical location, or at least two clinical relapses and objective clinical evidence of one lesion and clear-cut historical evidence of a prior relapse involving a lesion in a distinct anatomic location
- At least two clinical relapses and objective clinical evidence of one lesion; in addition, DIS should be demonstrated by either a second clinical relapse implicating a different CNS site or using MRI[a]
- One clinical relapse and objective clinical evidence of two or more lesions; in addition, DIT should be demonstrated by a second clinical relapse, or using MRI[b] or demonstration of cerebrospinal fluid-specific OCBs
- One clinical relapse and objective clinical evidence of one lesion; in addition, DIS should be demonstrated by a second clinical relapse implicating a different CNS site or using MRI, whereas DIT should be demonstrated by a second clinical relapse, or using MRI or demonstration of cerebrospinal fluid-specific OCBs

Primary progressive MS

A disease course characterized by progression from onset, 1 year of disability progression (retrospectively or prospectively determined) independent of clinical relapse and two of the following criteria:

- One or more T2-hyperintense lesions in at least one area in the brain characteristic of MS (periventricular, cortical and/or juxtacortical or infratentorial)
- Two or more T2-hyperintense lesions in the spinal cord with no distinction between symptomatic or asymptomatic lesions
- Demonstration of cerebrospinal fluid-specific OCBs

CNS, central nervous system; DIS, dissemination in space; DIT, dissemination in time; MS, multiple sclerosis; OCB, oligoclonal band. [a]One or more T2-hyperintense lesions in at least two of four areas of the CNS (periventricular, juxtacortical (by combining cortical or juxtacortical lesions), infratentorial and spinal cord lesions), with the removal of the distinction between symptomatic and asymptomatic lesions. [b]Simultaneous presence of gadolinium-enhancing and non-enhancing lesions at any time, with the removal of the distinction between symptomatic and asymptomatic lesions, or a new T2-hyperintense and/or gadolinium-enhancing lesion on follow-up MRI with reference to a baseline scan irrespective of the timing of the baseline MRI.

MRI images and their assessments were discussed as part of the diagnostic criteria for MS in the early 1980s, as a potential paraclinical tool for clinicians (60). The pivotal role of MRI in the diagnosis and follow-up of people with MS has since been confirmed. Since 1990, the Magnetic Resonance Imaging in MS (MAGNIMS) network and others have made major contributions to defining the central role of MRI in both the diagnosis of MS and monitoring

its treatment (68).

MAGNIMS' consensus guidelines are partly incorporated into the 2017 revised diagnostic criteria for MS (also known as the McDonald guidelines) (69). Notable discrepancies between these two hallmark publications in terms of MRI are inclusion of the optical nerve as a location for DIS and the requirement of > 3 periventricular lesions (Fig. 6). The revised McDonald guidelines state that incidental periventricular lesions in persons > 50 years of age or with established vascular risk factors should first be evaluated for other diagnoses than MS. The authors encourage increased research efforts to validate visual evoked potentials (VEPs), optical coherence tomography (OCT) and MRI in fulfilling either the DIS or DIT criteria in support of an MS diagnosis.

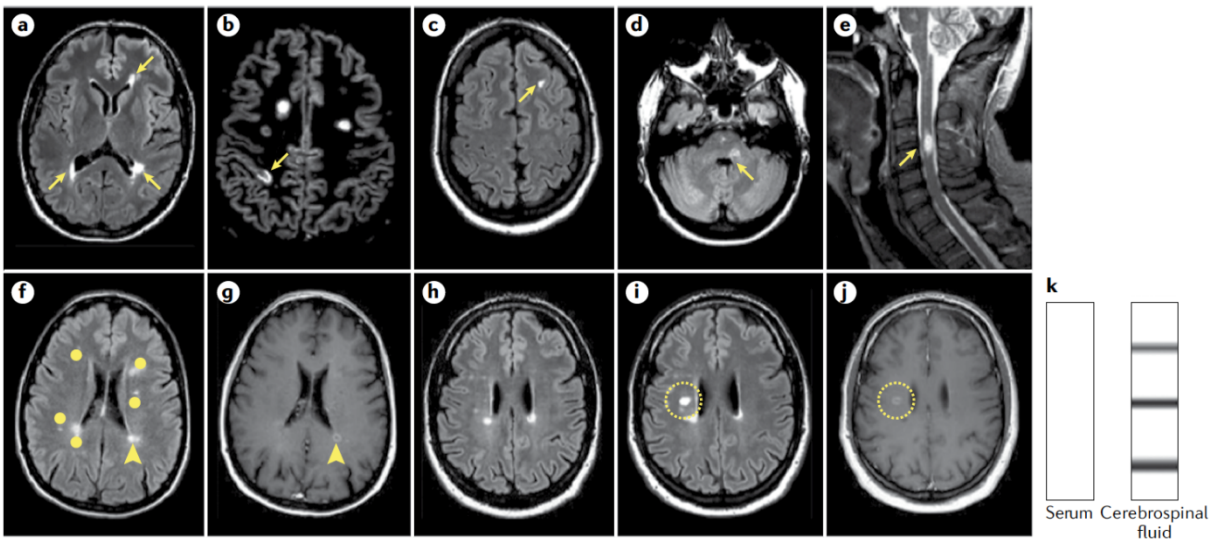

Fig. 6. Different subtypes of MS shown in the 2017 diagnostic criteria for MS.

a–e – Dissemination in space (DIS) can be illustrated by > 1 T2 hyperintense lesion in > 2 typical areas of the CNS (arrows); a – Periventricular lesions, b–c – Cortical or juxtacortical lesions to define juxtacortical involvement, d – Infratentorial lesions, e – A spinal cord lesion, f–j – Dissemination in time (DIT) illustrated by the simultaneous presence of gadolinium-enhancing and non-enhancing lesions – a new finding can be demonstrated by a simultaneous presence of gadolinium-enhancing and non-enhancing lesions (f–g) at any time, a new T2-hyperintense and/or gadolinium-enhancing lesion on follow-up MRI compared with a baseline MRI (h–j), irrespective of the timing of the baseline MRI, or the presence of cerebrospinal fluid-specific oligoclonal bands (OCBs), which are not visible in the serum (k). Several white matter lesions are visible on the fluid-attenuated inversion recovery (FLAIR) sequence (f): one shows enhancement (arrowhead) on the post-contrast sequence (g), whereas the majority are non- enhancing (dots). Compared with the baseline FLAIR sequence (h), a new T2 hyperintense and gadolinium-enhancing lesion (circled) are visible on follow-up FLAIR (i) and on post-contrast sequences (j). Reproduced from Filippi et al. 2018 (38), with permission from Springer Nature Ltd., obtained 8 August 2019

MRI has undergone extensive technical and software advances since 2010. The implementation of ultra-high-field 7.0T imaging increases the amount of visible details, provides higher resolution and enables visualization of the association between lesions and the vasculature (e.g. central vein sign) (69).

As previously mentioned (in Section 4.5) one of the requirements in a diagnostic setting of MS is the exclusion of other conditions (70). MS mimics have become easier to differentiate as the specificity of MRI assessment has improved. One such MS mimic is neuromyelitis optica spectrum disorders (NMOSDs), which share many of the common symptoms and findings as MS (71). Recent advances in defining MRI characteristics and in serum antibodies have improved the diagnostic guidelines for these disorders (71). In addition, elderly people with non-specific white matter lesions are often diagnosed with vascular lesions after thorough examination (70).

4.6 Natural history of MS

At diagnosis, people with MS are commonly in the age range 20–40 years (40, 48). The 2017 revision of the McDonald criteria for diagnosis of MS enables earlier MS diagnosis because the presence of OCBs can fulfil the DIT criteria (48, 64). A study reported that the prevalence of 'asymptomatic' and unknown MS was 0.1% (40).

In clinical settings, an MS relapse becomes acute or subacute over the course of hours to days and persists for > 24 hours (38, 48). The clinical features of a relapse usually persist over several weeks and the patient gradually recovers with or without persisting symptoms. A true relapse should occur in the absence of infection, fever or encephalopathy (48). A pseudo-relapse is seen when new or known neurological symptoms occur in the presence of infection, fever or reduced general health. Patients with a suspected pseudo-relapse with new neurological symptoms often undergo MRI and clinical examination to rule out true relapses.

Lublin and colleagues describe four different MS phenotypes (72). The most common MS phenotype is the relapsing-remitting form (RRMS), which is seen in 80–90% of the people diagnosed with MS (73). The natural history of RRMS patients has gradually changed in recent decades. The median time from MS diagnosis until requiring a cane to walk (expanded disability status scale (EDSS) of > 6), has recently increased to above 20 years (40, 73). Usually, the disease course for people with RRMS gradually develops into a SPMS phenotype characterized by slowly progression of their disability (Fig. 7) (72). The life expectancy of people with MS has increased by 20 years over the past 70 years (73). In Norway, a recent population study found an expected shorter life expectancy of 7 years for

people with MS compared with the general population (74). The improvements in the disease course of people with RRMS are thought to be due to a complex interplay of implemented guidelines for care and follow-up of MS patients, the increase in available disease modifying therapies (DMTs), the possibility of establishing an earlier diagnosis of MS, earlier treatment with highly efficient DMTs, and improved diagnostics (38, 48, 66, 75).

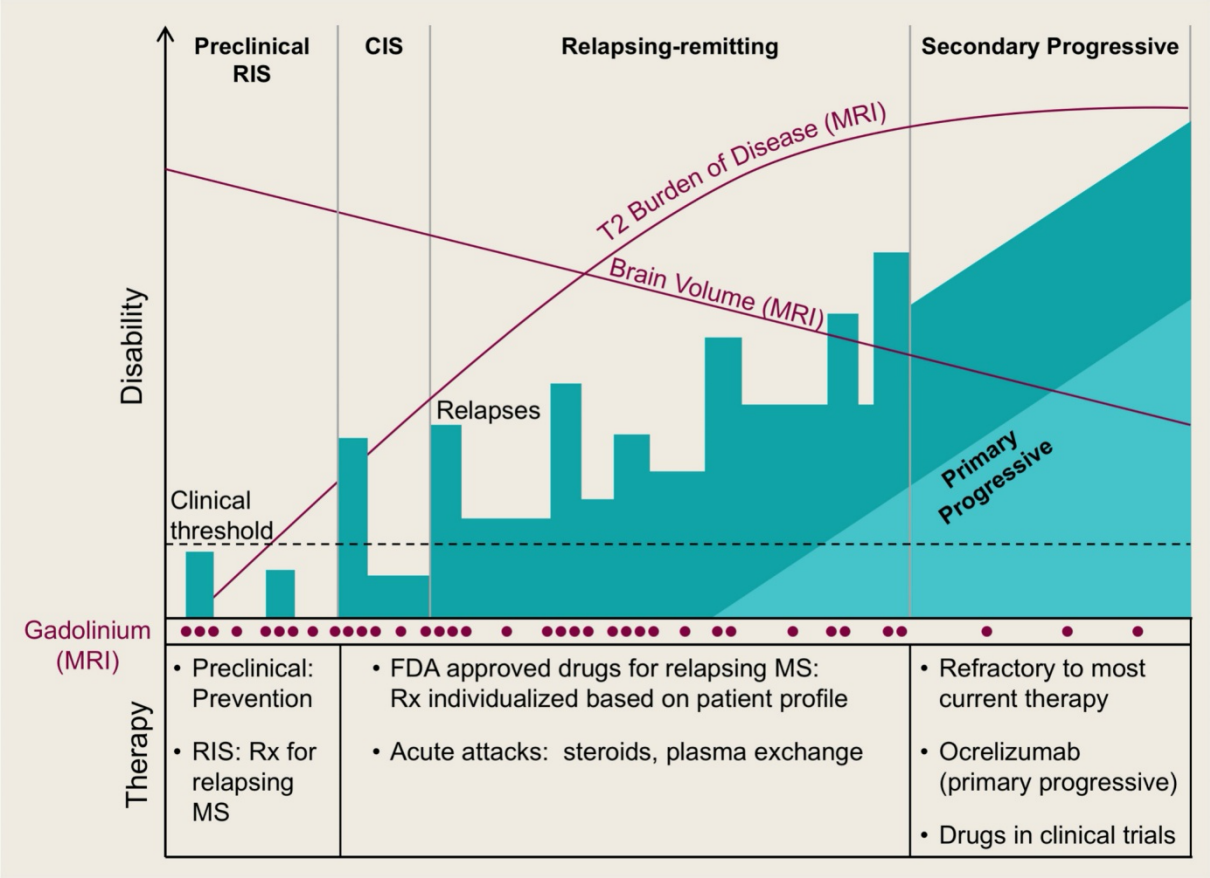

Fig. 7. The different stages of multiple sclerosis. MS often starts with a preclinical disease course. The first clinical symptom can be termed a clinically isolated syndrome, but can in some instances fulfil the revised McDonald criteria for diagnosis of MS. The relapsing-remitting disease course is illustrated by episodes of increased disability, followed by an improvement in function. The MRI markers of MS are illustrated with visible gadolinium-enhancing lesions, the accumulation of T2 lesions and brain atrophy. Reproduced from Baecher-Allen et al. (2018) (47), with permission from Elsevier Inc., obtained 5 August 2019.

The PPMS phenotype affects a minority of MS patients (5–20%), where the onset and follow-up are typically constituted by accumulation of progressive disability over time (48, 76). People with PPMS usually have one dominant neuronal system affected, described as unique

patterns such as progressive spastic paraparesis, progressive visual loss, cognitive deficit, and sensory or cerebellar ataxia (38, 77).

The terms ‘benign MS’ and ‘malignant MS’ have previously been used to describe the severity of the retrospective progression of the disease course (72). However, these illustrative terms have since been discarded from the revised McDonald guidelines for the diagnosis of MS.

Since 2013, the disease course has been considered as either active or inactive, as defined by clinical or imaging features (76). In addition, the 2010 revision of the McDonald guidelines for the diagnosis of MS implemented the importance of the progression of disability over time as an important factor (76). The differences between the 1996 and 2013 MS subtypes and phenotypes are illustrated in Fig. 8.

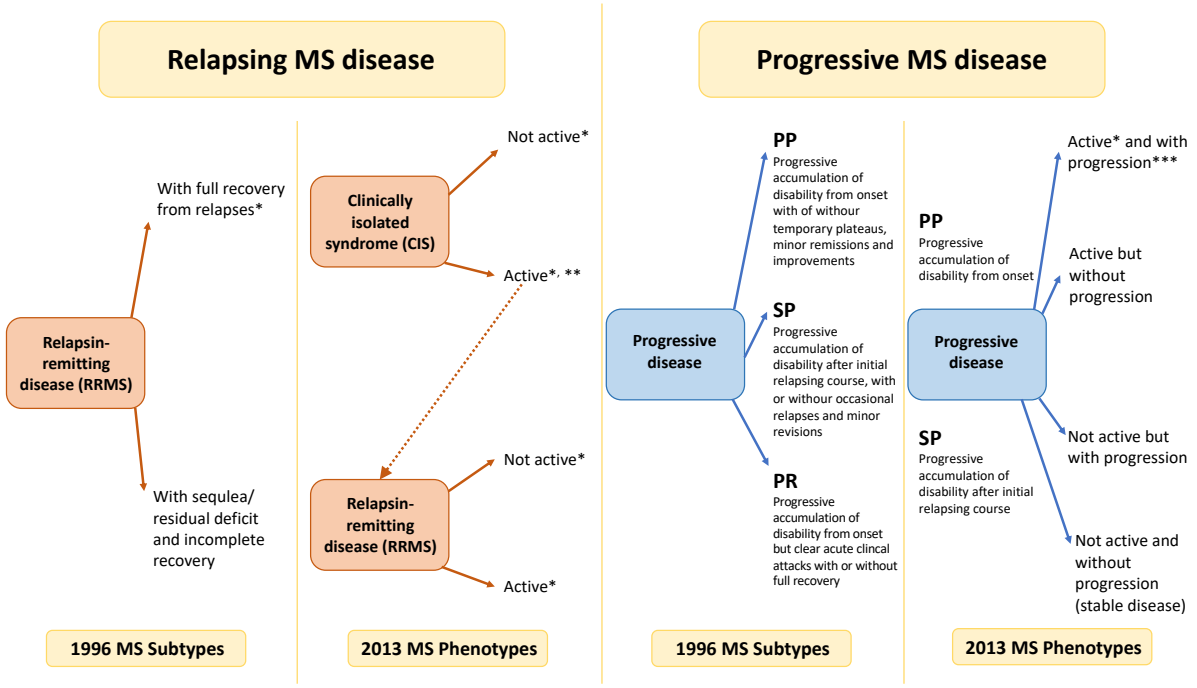

Fig. 8. Overview of the definitions of the clinical course in MS and the change from 1996 to 2013.
* Activity is determined by clinical relapses assessed at least annually and/or MRI activity (contrast-enhancing lesions; new and unequivocally enlarging T2 lesions). ** CIS, if subsequently clinically active and fulfilling current MS diagnostic criteria, becomes RRMS. *** Progression is evaluated clinically, at least annually. CIS – clinically isolated syndrome; MS – multiple sclerosis, PP – primary progressive; PR – progressive relapsing; RRMS – relapsing-remitting MS; SP – secondary progressive. Source: Multiple Sclerosis Research Group, Oslo University Hospital and based on Lublin et al. (2014) (76).

With the increasing availability of MRI scanners in the Western world there has been an increase in the incidental discovery of asymptomatic subjects with apparent MRI findings

suggestive of MS, and these findings are referred to as radiologically isolated syndrome (RIS) (38, 48). Typically, individuals with RIS are referred for MRI due to headaches or head injury. When investigated, a substantial proportion of people with RIS have shown signs of cognitive impairment and brain volume loss, which has led to the hypothesis that the neurodegenerative and inflammation processes are present even in the pre-clinical stages of MS. It is estimated that up to 34% of people with RIS will develop a clinical attack within five years (38, 48).

4.7 Symptoms in MS

In theory, the clinical presentation of MS would be unique to each patient based on the location(s) of the demyelinating lesion(s) (Fig. 9). The disease typically starts with a clinical attack. If the requirements for DIS and DIT are present, the diagnosis of MS can be set after a thorough examination to rule out other possible conditions, based on the 2017 revised criteria for the diagnosis of MS (64). MS often presents with an episode of optic neuritis, when typically there is found partial or complete visual loss in one eye with a central scotoma, loss of colour vision and most often associated with pain or pressure behind the eye that worsens with movement of the affected eye (38, 78-81). Approximately 70% of all MS patients experience an episode with optic neuritis during their disease course, and 25% of MS patients will present with an episode of optic neuritis (38). It is estimated that for 34–75% of people with optic neuritis, MS will develop as a certain diagnosis within the next 10–15 years after onset of the optic neuritis (38).

The initial clinical attack can, as previously mentioned in Section 4.6, be located in all areas of the CNS, and present as either typical or atypical findings. It has been estimated that motor deficits (pyramidal signs, paresis and spasticity) are the first MS-symptom in 30–40% of MS cases, while sensory symptoms (paraesthesia, Lhermitte sign, loss of vibration and sensation of joint position, and reduced perception of light touch and pain) are the initial symptoms in as many as 43% of people with MS (38). Examples of less typical presenting signs of MS are epileptic seizures, acute urinary retention, trigeminal neuralgia, photophobia, headache, bulbar weakness, and severe pain.

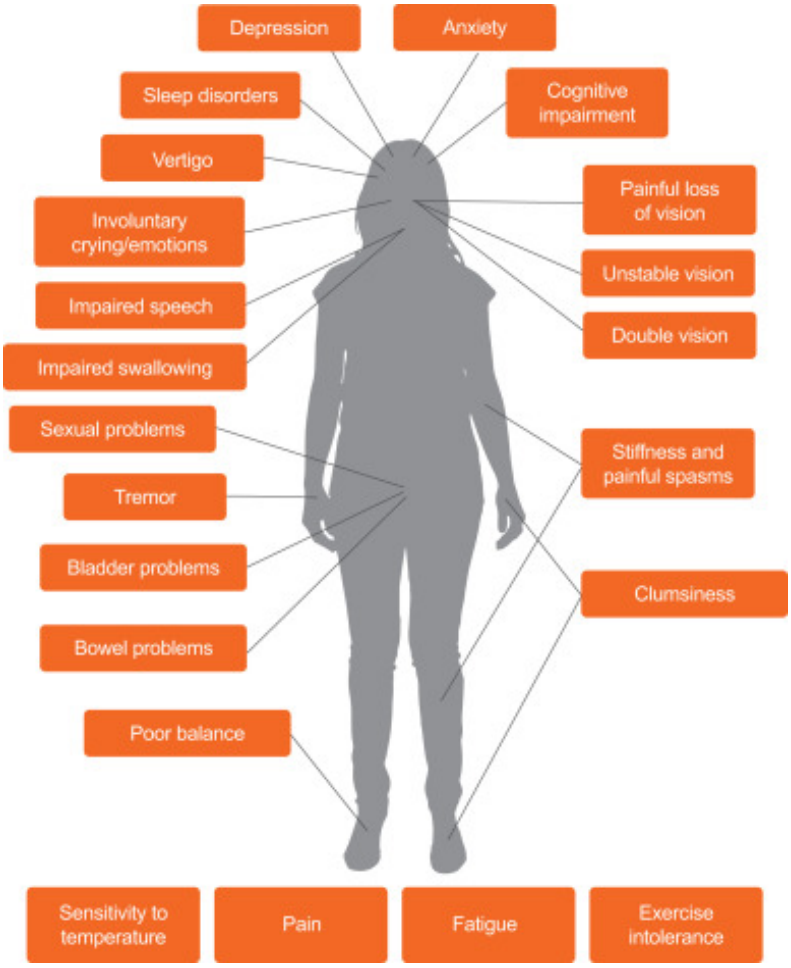

Fig. 9. Common symptoms in MS. Reproduced from Giovannoni et al. 2016 (66), with permission from Oxford Pharma Genesis Ltd. © 2016, obtained 8 August 2019.

During the course of MS, usually additional neurological symptoms accumulate. Up to 70% of people with MS experience symptoms from the brainstem and cerebellum (impairment of eye movements, diplopia, ataxia and gait imbalance, dysmetria, impaired speech, and dysphagia). In total, 34–99% of MS patients experience neurological symptoms of sphincter or sexual dysfunction (urinary urgency, hesitancy, urge incontinence, constipation, faecal incontinence, erectile dysfunction, and impotence) (38).

Other common neurological symptoms related to MS are sleep disorders (affecting up to 54% of people with MS), cognitive impairment (evident in 40–70% of people with MS), affective abnormalities (affecting up to 67% of people with MS), and pain (reported by up to 43% of MS patients) (Fig. 9) (38).

For a large proportion of people with MS, their former neurological symptoms resurface when their body temperature is increased, either during warm summer days or

during episodes of fever. This experience of MS disease worsening is referred to as the Uhthoff's phenomenon.

A focused awareness of all the invisible symptoms of MS and their impact on quality of life for the majority of the MS population is important. Examples of invisible symptoms are fatigue, cognitive deficits, sphincter and sexual dysfunctions, pain, spasms, dizziness, visual loss, depression, heat sensitivity, brain fog, and swallowing issues (82).

4.7.1 Measuring disability in MS

In order to evaluate people with MS and systematically report their disability status, John F. Kurtzke produced the disability rating scale in 1955 (83), which later was renamed the Disability Status Scale (DSS). There have been several modifications to the DSS, but since 1983 the Expanded Disability Status Scale (EDSS) system has remained unchanged and accepted as the most widely used grading system for MS disability, based on a standard clinical neurological examination (84, 85) (Table 2). The EDSS grading tool scores the disability of MS patients on a scale from 0 to 10, where 0 is normal and 10 is death due to MS (Table 2). Eight subscales, called functional system (FS) scores, are used to cover different parts of the CNS commonly affected by MS: pyramidal, cerebellar, brain stem, sensory, bowel & bladder, visual, cerebral or mental, and other functions. All FS groups, except 'other functions', are given a grade from 0 (normal function) to 5 or 6 (complete impairment).

EDSS is the most common tool for assessing clinical disability in people with MS. Especially for the lower parts of the scale (< 4.0), the clinical neurological examination is important to investigate and uncover specific symptoms from all parts of the CNS (84). EDSS scores between 4.5 and 7 reflect the gait ability, while scores > 7 are mostly dependent on the subject's ability to execute regular activities of daily living unaided by others.

Several methods for improving the EDSS have been discussed and aspects of the scale have been criticized (66). First, the scoring may vary depending on the subjective nature of the clinical neurological examination and the dynamics of MS, especially on the lower parts of the scale, namely scores < 4.0 (85). Second, the steps on the scale do not reflect the true nature of the gradual disease progression observed. Third, the importance of upper body function, mood, cognitive function, quality of life, and fatigue are not addressed in a proportional manner for EDSS scores > 4 (66).

Table 2. Overview of the expanded disability status scale and the scoring. Adapted from Kurtzke (1983) (84) and the Multiple Sclerosis Trust (86).

Expanded Disability Status Scale (EDSS)	
Score	***Description***
0	No disability
1.0	No disability, minimal signs in one FS
1.5	No disability, minimal signs in more than one FS
2.0	Minimal disability in one FS
2.5	Mild disability in one FS or minimal disability in two FS
3.0	Moderate disability in one FS, or mild disability in three or four FS. No impairment to walking
3.5	Moderate disability in one FS and more than minimal disability in several others. No impairment to walking
4.0	Significant disability but self-sufficient and up and about some 12 hours a day. Able to walk without aid or rest for 500m
4.5	Significant disability but up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance. Able to walk without aid for 300m
5.0	Disability severe enough to impair full daily activities and ability to work a full day without special provisions. Able to walk without aid or rest for 200m
5.5	Disability severe enough to preclude full daily activities. Able to walk without aid or rest for 100m
6.0	Requires a walking aid – cane, crutch, etc. – to walk about 100m with or without resting
6.5	Requires two walking aids – pair of canes, crutches, etc. – to walk about 20m without resting
7.0	Unable to walk beyond approximately 5m even with aid. Essentially restricted to wheelchair; though wheels self in standard wheelchair and transfers alone. Up and about in wheelchair some 12 hours a day
7.5	Unable to take more than a few steps. Restricted to wheelchair and may need aid in transfering. Can wheel self but cannot carry on in standard wheelchair for a full day and may require a motorised wheelchair
8.0	Essentially restricted to bed or chair or pushed in wheelchair. May be out of bed itself much of the day. Retains many self-care functions. Generally has effective use of arms
8.5	Essentially restricted to bed much of day. Has some effective use of arms retains some self-care functions
9.0	Confined to bed. Can still communicate and eat
9.5	Confined to bed and totally dependent. Unable to communicate effectively or eat/swallow
10.0	Death due to MS
Scoring is based on an examination by a neurologist. A functional system (FS) represents a network of neurons in the brain with responsibility for particular tasks. Each FS is scored on a scale of 0 (no disability) to 5 or 6 (more severe disability). The EDSS is based on measures of impairment in eight FS: pyramidal – muscle weakness or difficulty moving limbs cerebellar – ataxia, loss of balance, coordination or tremor brainstem – problems with speech, swallowing and nystagmus sensory – numbness or loss of sensations bowel and bladder function visual function - problems with sight cerebral functions - problems with thinking and memory other EDSS steps 1.0 to 4.5 refer to people with MS who are able to walk without any aid EDSS steps 5.0 to 9.5 are defined by the impairment to walking.	

4.7.2 Cognitive deficits in MS

Epidemiological research have shown that up to 70% of all subjects with MS have cognitive deficits, which often are associated with reduced quality of life, psychiatric symptoms and the ability to participate in work-related and social activities (87). Most frequently, the cognitive domains of information processing speed and memory are affected in MS patients. In addition, reduction in verbal fluency, visuospatial processing and executive functions are common domains associated with MS (88-90). Symptoms of cognitive deficits may be evident even early in the disease course of MS, and only show weak associations with physical impairment (91, 92).

4.7.3 Depressive symptoms in MS

People with MS are generally two to three times more likely to acquire a depressive disorder than the general population, with the lifetime prevalence as high as 50% (93-95). Depression is a robust determinant of quality of life, along with other associated symptoms such as fatigue and anxiety (95). Studies of MS have revealed a complex interaction in which both cognitive deficits and fatigue are overlapping features in MS patients with depression (95). The basis for depression in MS is not fully understood, yet studies suggest a multifactorial etiology, including common psychological, biological (structural damage to frontoparietal brain regions and consequently functional disconnection within the hippocampal-centred network) and social factors (95-101). Based on current evidence, people with both MS and a depressive disorder benefit from standard treatment of depression such as antidepressant medication and cognitive behavioural therapy (95).

4.7.4 Fatigue in MS

The most common symptom in a MS population is fatigue, which affects up to 80–95% of all people with MS (102-104). Fatigue is most often described as a feeling of excessive tiredness or exhaustion (103). There is a distinction between the central and peripheral origin of fatigue (105). Peripheral fatigue is often explained as performance fatigability or muscle fatigability. Whereas peripheral fatigue is often relieved by rest, this is not usually the case for central fatigue (105). The exact cause of fatigue is unknown, and fatigue is very common in other disorders such as cancer, and autoimmune and psychiatric disorders (105). Some studies show that MS fatigue could be sustained due to a dysfunction of cortico-subcortical connections in the CNS (38, 106).

Quantitative neuroimaging studies have found associations between fatigue and damage to selected grey matter and white matter structures in subjects with MS (106, 107). Further, the importance of damage to the cortico-striato-thalamo-cortical loop is supported by associations between fatigue and general damage to the thalamus, selected cerebral cortical areas and the striatum (106). Also, disconnections between deep GM structures and the cortex, such as the cortico-thalamic or cortico-striatal tracts, may contribute to the development of MS (106). Damage to the corpus callosum have been associated with fatigue, which suggests that a disconnection between the left and right cortico-striato-thalami-cortical areas is potentially crucial in the development of fatigue (106, 108). Studies of resting state functional MRI (rs-fMRI) in RRMS found that MS subjects showed evidence of altered

functional connectivity in the posterior cingulate cortex and anterior cingulate cortex (108, 109).

4.8 Conventional neuroimaging in MS

The clinical adaptation of MRI on a clinical platform was released in 1980. MRI of the brain and spinal cord was mentioned as a possible paraclinical tool for MS in 1983 (60). In the early 1980s, MRI of the brain was shown to be far superior in terms of sensitivity to detect MS lesions than the established computerized tomography (CT) brain scans (110). Finally, MRI was included in the 2001 revised criteria for MS diagnosis (60, 61). Today, MRI assessment is deeply integrated into the diagnostic criteria for MS (64, 69, 111, 112). Neuroradiological evaluation of MS is currently based on a thorough visual inspection of MRI scans and compared with previous MRI scans if possible. Using validated scales, neuroradiologists can measure the degree of atrophy with the global cortical atrophy (GCA) scale, a validated 4-point scale used to estimate widening of sulci and gyral volume loss by visual evaluation (113). Standard MRI acquisition protocols have been proposed to address carefully harmonizing MRI acquisitions and assessments from people with MS in regards to the visible number of lesions, new lesions, contrast-enhancing lesions, the locations of the lesions, and brain atrophy (114). Also, to distinguish MS more clearly from its clinical and imaging imitators, an MRI scan provides a sensitive tool for use by trained and experienced neuroradiologists (70).

4.8.1 Understanding MRI physics

A standard MRI system comprises a control centre with the screens for visual interface, a separate room for the electronics, power and cooling system, and a third room for the machine itself, with the board to lie on surrounded by a main magnet coil (made of a super-conducting metal alloy), three gradient coils (representing the three orthogonal directions, x, y, z, for a functioning coordinate system), shim coils, and an integral radio frequency transmitter coil (115). In clinical practice, the main magnet usually maintains a stable magnetic field of either 1.5T or 3T, although some research scanners achieve a magnet field of 7T or even >10T (115).

The main signal used for MRI originates from the hydrogen nuclei, where data derived from the relaxation after manipulating the spin, direction and the time to normalization of the protons in the hydrogen nuclei are used to generate different images (115). A T1-weighted

MRI sequence is generated by the differences in tissue T1 relaxation time, and show fat as bright and intense, whereas CSF is typically dark (115). In addition, there is the possibility to add a gadolinium-enhancing contrast agent and to do another T1-weighted scan to enhance the visualization of newly acquired lesions. T2-weighted images are generated by T2 relaxation time and typically show CSF as bright (115). To distinguish more clearly between normal and pathological CNS signals, a fluid-attenuated inversion recovery (FLAIR) sequence is often applied to suppress the CSF signal (115). The FLAIR sequence enables T1- and T2-weighted contrast to delineate lesions and subtle tumours more clearly (115).

Since 1990, it has been possible to use an imaging surrogate marker using the signal generated by changes in oxygen concentration in the brain in order to capture regional changes in brain metabolism (116). This blood oxygen level dependent (BOLD) signal derives from changes in the magnetic field covering the haemoglobin, depending on the levels of oxygen within the cells. The BOLD signal is dependent on the regional blood flow and the local blood oxygen level, and constitutes the basis for functional MRI (fMRI). FMRI sequences are usually task based, whereby a person performs either mental tasks or specific motor tasks within the scanner or while in a resting state (rs-fMRI), when the subject lies still with their eyes open.

New MRI sequences with improved sensitivity and specificity with regards to certain tissues or disorders have been subject to vast research efforts to show specific utility and superior performance (54, 117).

4.8.2 Gadolinium-enhancing contrast agents in MS

Gadolinium-enhanced MRI is a sensitive method used to visualize active MS lesions (118). Gadolinium serves as a marker for BBB disruption and is associated with the histologically inflammatory state of lesion evolution (118). New MS lesions typically exhibit contrast enhancement for two to six weeks, and very rarely for more than six months. The different patterns (ring-like, nodular) of enhancement are considered to be of clinical importance (118).

An article published in 2014 reported increased signal intensity in the globus pallidus and dentate nucleus on T1-weighted MRI images without contrast in subjects with MS who had undergone MRI scans with gadolinium-based contrast agents (119). As a consequence, the scientific community, health institutions, industries, and regulatory agencies worldwide have changed their recommendations to exclude linear gadolinium contrast agents and instead recommend the use of macrocyclic gadolinium agents in MRI (120). A recent longitudinal study that included both healthy subjects and participants with MS found that gadolinium

deposition (globus pallidus, dentate nucleus and thalamus) in early MS was associated with past gadodiamide exposure (a linear gadolinium-based contrast agent, Omniscan ©), but without radiological or clinical correlates of more aggressive disease (120).

Other severe adverse events, such as nephrogenic systemic fibrosis and acute allergic reactions, warrant specific guidelines on restricting the use of gadolinium-enhancing contrast agents in MS and other disorders (120-123). Suggestions for revising the MRI diagnostic guidelines constitute a more targeted approach to the use of contrast agents (in the diagnostic setting of MS, when a clinical relapse is suspected or in specific cases with known complex MRI findings) (124).

4.9 Experimental neuroimaging in MS

Despite important MRI correlations with MS, there is still heavy reliance on the visual assessment by a neuro-radiologist of the MRI scans in a clinical setting (69). Some MRI features, such as the central vein sign, cortical lesions, brain volume, and atrophy, are already incorporated in a research setting when evaluating MS patients on a group level (125). The application of reliable quantitative MRI markers for atrophy and lesions in a clinical setting are not feasible in individual patients, due to the lack of standardized methods, validation studies and technical improvements (117, 125).

The main target for experimental MRI studies in MS is to be able to predict future response to treatment and future disability progression (54, 117, 126). One study used high-dimensional statistical models incorporating a wide multiplicity of imaging features, and the researchers found that their models outperformed the standard conventional models in detecting an imaging response to treatment in MS (126). In 2018, MAGNIMS produced a positional paper for unravelling treatment response in a real-world MS setting (117). In a separate expert review advanced MRI techniques are discussed as a tool to monitor both therapeutic and rehabilitative treatment in MS, where recent advances in MRI research are described as fundamental progress towards personalized medicine and individual treatment decisions (54).

There are already available third-party services offering automated quantitative brain MRI analyses, such as MSmetrix and NeuroQuant (with LesionQuant for MS) (127). Furthermore, some studies that used data procured by these third-party services have been published showing comparable results with standard MRI assessment, and raise the question as to when these methods for automatic quantification of MRI brain scans can be implemented in a clinical setting (128, 129). Studies have highlighted that improved MRI

technology will lead to improved visual interpretation of suspected MS lesions by allowing consideration of the central vein sign, lesional rims and subpial demyelination (114).

Many large international collaborations and institutions are involved in experimental neuroimaging research, such as UK Biobank, the ENIGMA (Enhancing Neuro Imaging Genetics through Meta Analysis) Consortium, the Human Connectome Project, the Norwegian HUNT project, and the ABCD cohort.

4.10 Treating MS

The search for a cure for MS has not yet been successful, although many DMTs have been successfully developed and implemented into standard MS care (47). There is strong evidence that time to diagnosis and time to treatment start, are essential factors in the long-term prognosis of MS (66). Moreover, early treatment with potent agents within the time of MS onset has been shown to reduce disease progression (130). For this reason, an international guideline for MS care from the Delphi Panel (21 neurologists from 19 countries) incorporates these goals to reduce disease activity in MS and improve MS care (75). When treating MS patients, it is necessary to consider all features of the disease in terms of vitamin D supplements, smoking cessation, physical activity, and other lifestyle factors for healthy brains (75). People with MS can have a variety of neurological symptoms derived from accumulated CNS lesions, such as urinary and faecal incontinence, depressive symptoms, fatigue, muscle spasms, pain, dizziness, and cognitive deficits (75). Also, when deficits occur, it is necessary to consider the possibility of rehabilitation, equipment to aid specific deficits (e.g. a wheelchair or a car with the possibility to rely only on the hands when driving), personal assistance, residence in a nursing home, physiotherapy, and referral to other medical specialties if appropriate (75).

4.10.1 Disease-modifying therapies in MS

From the first successful studies of the use of Interferon beta-1b to slow the rate of MS progression in the early 1990s (131, 132) to the most recently implemented MS drug Ocrelizumab (133, 134), an increasing number of people with MS have received therapy designed to decrease the long-term destruction of their CNS. Ocrelizumab, as the first DMT ever, is also a treatment option approved for people with PPMS with disease activity (134).

Treatment options in MS have specific mechanisms for targeting the immune system, including modulation of immune activation, alteration of immune cell trafficking and depletion of immune cell populations (47, 135). DMTs have different modes of

administration, as subcutaneous injections (interferons), tablets (Fingolimod, Dimethyl fumarate, Glatirameracetate) or infusions (Natalizumab, Alemtuzumab, Rituximab, Ocrelizumab), and are to be taken at different time intervals, ranging from twice per day to once every sixth months (136). Also, there are different approaches to modulate the immune system, either by continuous suppression or by induction therapy (Cladribine, Alemtuzumab, hematopoietic stem cell transplantation (HSCT)).

In Norway, a section of the national health system, called Hospital Procurement (Sykehusinnkjøp), is appointed to negotiate with the pharmaceutical companies on the prices of some selected groups of drugs. Each year, Hospital Procurement receives discounted prices from the pharmaceutical companies providing MS drugs and update the national guidelines for MS treatment. The appointed MS specialist group in Norway provides counsel and discusses future treatments for MS (Table 3).

Table 3. Current MS therapy guidelines in Norway. Adapted from the Sykehusinnkjøp guidelines for 2019 (137).

Overview of current Norwegian MS guidelines for DMT as of 19 March 2020		
Group 1: High efficacy	**Group 2: Oral**	**Group 3: Injectables**
1. Cladribin tablets	1. Teriflunomide	1. Glatirameracetate
	2. Dimethyl fumarate	2. Interferon beta-1b
		3. Peginterferon beta-1a
		4. Interferon beta-1a
Others:		
Natalizumab – as of 1 December 2019, not approved, for economic reasons		
Fingolimod – as of 1 December 2019, not approved, for economic reasons		
Rituximab – off-label, has to be registered in the national MS registry		
Alemtuzumab – currently not advised as the first treatment of choice, or with comorbidities		
HSCT – as a study participant or in special circumstances		

DMT – disease-modifying therapy, HSCT – hematopoietic stem cell transplantation.

Due to a recent safety review of Alemtuzumab by the European Medicines Agency (EMA), in Norway there is careful selection of the people who start this treatment (138, 139). Also, Norway is one of the few Western countries where Ocrelizumab has not been taken into use

and, as of 1 December 2019, Natalizumab and Fingolimod are similarly not available for new MS treatment, based on a national assessment of all aspects of MS treatment (140).

Rituximab is a widely used off-label chimeric anti-CD20 monoclonal antibody drug in many neurological disorders, but there are no current standardized treatment guidelines (141). The topic of B-cell depletion in MS has been extensively discussed since the unblinding of the Phase II trial of Rituximab study in September 2006 and its implications for MS treatment (47, 142, 143). Due to the growing use of Rituximab in Sweden (144), and more recently in some parts of Norway, the increased use of Rituximab has enabled large register-based studies in a real-world population to challenge current standards for defining the model of DMT use in MS (141, 145).

4.10.2 Hematopoietic stem cell transplantation in MS

In MS, a 'reset' of the immune system by autologous HSCT has proven superior to conventional DMT for people with an aggressive disease course in MS (135). The intense HSCT regime is associated with high levels of morbidity and previous reports of mortality (135). Initial prospective cohort studies and registry-based data were positive towards HSCT as an option for younger patients with MS, although premature menopause for women and other side-effects were significant (135).

A recent study with an international multicentre, blinded, randomized trial of 110 MS patients who had undergone HSCT revealed less disease progression, fewer relapses, reduced MRI lesion load, a higher share of patients maintaining no evidence of disease activity (NEDA, characterized by the coexisting features of no relapse, no clinical progression and no new MRI changes), and increased quality of life (146). However, several severe aspects of undergoing HSCT are highlighted as problematic (146). It is still not clear which subgroup of MS patients benefits most from HSCT, although younger MS patients with severe progression and failure to maintain NEDA status when receiving DMT seems a possible target group (147).

Large international studies with the aim of comparing HSCT with DMT are currently ongoing. Among them is RAM-MS, an international study aiming to include 100 MS patients, which is organized from Bergen, Norway, in collaboration with other European countries (148). RAM-MS aims to report the differences in MS outcome measures between HSCT and currently available high-efficacy DMTs by randomizing eligible patients into one of the groups.

4.10.3 Myelin repair in MS

Despite all the advances made in DMT in recent decades, currently there is no effective therapy for the inevitable axonal degeneration and neuronal loss occurring after a clinical relapse in MS (149). Many researchers therefore rightfully point to the implementation of regenerative treatment approaches in MS as the next goal in MS pharmacotherapy (150). Already, many agents have shown potential in promoting oligodendrocyte precursor cells (clemastine, opicinumab, biotin, simvastatin, quetiapin, and anti-GNbAC) (150). Also, possible new targets for promoting remyelination have recently been proposed, such as promoting microglia necroptosis and repopulation to boost remyelination (58). All findings have led to recommendations for future translational and clinical research to establish new targets for myelin reconstruction in MS.

4.11 Markers and prognostic factors in MS

Markers associated with conversion from CIS to MS and disability progression are diverse and are found within the environmental, genetic, clinical, laboratory, and imaging domains (38). Female gender is associated with higher risk of conversion from CIS to MS, while male gender is associated with worse prognosis in people with MS (38). MRI features associated with poor prognosis are the presence of spinal cord lesions, new T2 lesions within the first five years of MS diagnosis, infratentorial lesions, and higher T2-hyperintense lesion volume and number (38). In addition, increased levels of the John Cunningham virus (JC virus) are known to raise the risk of developing progressive multifocal leukocencephalopathy (PML) for patients taking Natalizumab and Dimethyl fumarate (151).

Recent advances in sensitive blood analyses have made it possible to measure neurofilament light (NfL) chains reliably also in peripheral blood, most commonly using an ultrasensitive singe-molecule array called Simoa® (by Quanterix®) (152, 153). NfL is a segment of the axonal cytoskeleton and is released into the CSF following axonal damage, and proportionally transferred over to the peripheral blood (47, 152-154). Serum NfL (sNfL) is increased in people with MS compared with healthy controls (HCs), positively associated with T2 and gadolinium enhancing lesions in the brain and spinal cord, increased in MS patients with recent relapses or worsening of disability, and decreased with longer treatment duration with DMT (153). Already in 2017, sNfL was suggested as a biomarker to evaluate tissue damage and the effects of DMT in patients with MS (153).

4.11.1 No evidence of disease activity

In 2014, a new composite endpoint in MS was proposed, based on experiences within rheumatology (155). The reasons for suggesting a new endpoint in MS care were the low relapse rates in many clinical trials and a call for a more sensitive measure (155). Therefore, the concept of NEDA was put forward, originally termed ‘disease activity free status’ (DAFS). Later, more levels of NEDA were added to encompass more domains. Currently, the most used NEDA measure contains the following requirements: (1) no relapses, (2) no confirmed disability progression measured by EDSS scores, (3) no new or enlarging T2-weighted CNS lesions, and (4) no new gadolinium-enhancing CNS lesions. The aforementioned NEDA classification is usually termed NEDA-3, whereas NEDA-2 lacks the MRI aspect and NEDA-4/5 additionally encompass normalizing brain atrophy loss and normalizing of CSF NfL levels (47).

5. Thesis aims

My overall objective for this thesis was to investigate how state-of-the-art MRI post-processing methods could be used in a sample of early MS subjects to give new insights into disease pathophysiology and possibly serve as an imaging marker for disease activity. To achieve the objectives for my thesis, I aimed at exploring the complete five-year Oslo longitudinal MS sample (Fig. 10). If possible, we hoped to use our results to gain valuable new pathophysiological insights to MS and to encourage future studies in order to improve MS brain health.

The MS sample is based upon a prospectively collected sample of subjects with MS, originally funded as part of doctoral research work by my co-supervisor Gro O. Nygaard. As part of my own research for my doctoral thesis I have examined the same subjects at time point 3 (Fig. 10). From the MS sample at this third time point we have collected blood samples, MRI scans, self-report questionnaires and investigated the subjects by cognitive testing, eye examinations, neurological testing, and other experimental tests.

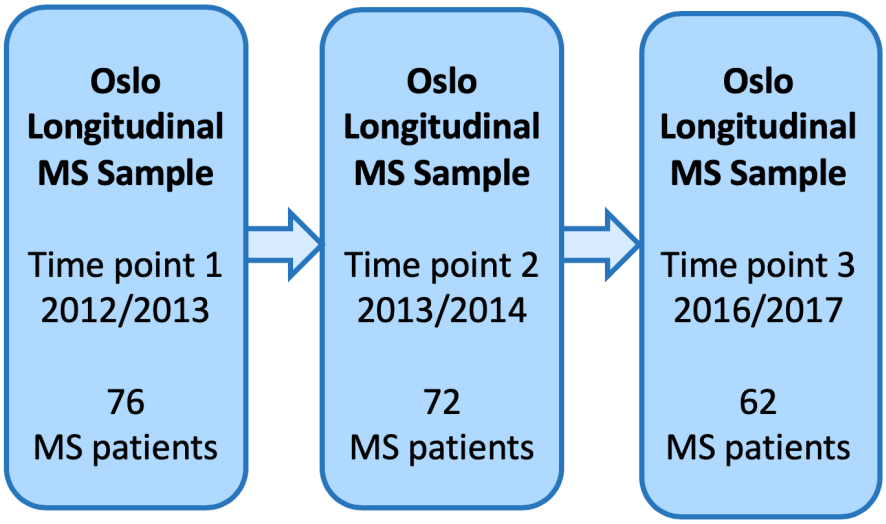

Fig. 10. An overview of the prospectively collected Oslo longitudinal MS sample.

My specific research objectives were:

1. To investigate how the default mode network connectivity is affected by depressive symptoms and symptoms of fatigue in subjects with early MS, using cross-sectional rs-fMRI data from time point 1.

2. To establish a method for brain age estimation in a longitudinal sample of MS subjects and to compare the brain age estimates from the MS subjects with a matched sample of healthy controls. An additional aim was to explore possible associations between clinical variables and both global and regional brain age estimates in a sample of early MS subjects.

3. To evaluate the longitudinal cognitive performance by the MS subjects, to assess whether there were any associations between the cognitive test performance and brain age estimates in the longitudinal MS sample, and to compare established structural MRI features with brain age in terms of sensitivity to cognitive performance.

6. Summary of the thesis

6.1 Paper I (Study 1)

'Symptoms of fatigue and depression is reflected in altered default mode network connectivity in multiple sclerosis'

In Study 1, reported in Paper I, we only used cross-sectional data from time point 1 to investigate how the brain connectivity, as measured with resting-state functional MRI (rs-fMRI), correlated with depressive symptoms and fatigue. We specifically investigated a brain network termed the default mode network (DMN), which is known to be a critical hub for both flow and integration of information in the brain. Using principal component analysis (PCA), we tried to disentangle some of the complicated interactions between fatigue and depressive symptoms. Our findings suggest that increasing symptoms of both fatigue and depression are associated with over-activation of the DMN. In addition, the PCA revealed a subgroup of subjects with a low burden of depressive symptoms and high scores on fatigue, which were associated with increased connectivity in the DMN. This suggests that there might have been different subgroups within our MS sample, an observation that was often reflected upon in the follow-up of MS patients. Our findings do not point to any specific aetiology underlying the DMN connectivity disturbances, and we only considered variations within the MS sample itself without comparing them to healthy controls.

6.2 Paper II (Study 2)

'Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis'

Study 2, which is reported in Paper II, included data from all time points. Paper II is the first paper to establish the method of brain age estimation in MS a. Using advanced machine learning methods, based on a training set of brain scans from 3208 healthy subjects, we trained a model to provide accurate estimations of brain age, based on 1118 features from an ordinary structural brain MRI scan. We included a global brain age estimate and seven regional brain age estimates to capture the morphological pattern of brain aging in our subjects. We found that MS patients aged 40 years had, on average, brains that were 4.4 years older than the brains of matched healthy controls. In addition, we found a 41% accelerated brain aging process over the five-year observational period compared with the expected chronological brain aging process.

6.3 Paper III (Study 3)

'Brain age estimation is a sensitive marker of processing speed in the early course of multiple sclerosis'

Study 3, which is reported in Paper III, investigated the relationship between brain age and cognitive performance in our complete dataset. It is known that specific cognitive deficits are evident also in the early course of MS, and mainly affect memory and information processing speed. We found a significant association between slower information processing speed and increased brain age estimates in our sample. As a comparison, we also investigated the thalamus in association with the cognitive tests, and found that reduced performance using a different test for information processing was significantly associated with smaller thalamus volumes.

In all our studies combined, we found several possible MRI markers that could potentially be applied in a clinical setting to improve the brain health of patients with MS. Future studies are needed to confirm and elaborate on our findings.

7. Considerations regarding materials and methods

We conducted clinical quantitative research, applying appropriate statistical methods to assess whether or not our preformed null-hypotheses could be abandoned (156). For this kind of research to be valid, we had to consider a number of several aspects, including the following: the statistical methods should report appropriate effects sizes including measures of uncertainty, the measures used should be relevant and correct, and the sample tested should encompass the suited composition to enable us to test our research hypotheses (157).

7.1 Study design

In the prospective cohort of MS patients, we selected our potential study subjects in advance, based on our hospital registries, so that our study sample would exhibit a high level of internal validity (if the sample reflected the population investigated). A good study design will mean that the results will have the potential of being transferable, in our case to other MS populations, which is referred to as external validity (157).

The cumulative data available from the MS subjects were prospectively collected over a course of five years, including baseline, first and second follow-up (Fig. 10 and Fig. 11). We aimed at including subjects early in their MS disease course, thus limiting the time since diagnosis to less than four years.

The three studies conducted as part of the research for this thesis were all observational in terms of not exposing the subjects to experimental treatments and not interfering with their MS treatment. However, in Study 2 (Paper II) and Study 3 (Paper III) we included results from specific study-related examinations that were not part of a standard neurological examination for MS patients. At all three time points, we included self-report questionnaires to capture known confounders and other possible features that might have been of interest; we used the Short Form Health Survey (12 items) (SF-12), the Beck Depression Inventory, version II (BDI-II), the Fatigue Severity Scale (FSS), and a collection of demographic and environmental questions.

Study 1 had a cross-sectional design, mainly because it was a project that I could begin at the same time as I was examining the subjects for the second follow-up in the study. For Study 2 and Study 3, the full longitudinal dataset was utilized. For Study 1 we did not have any healthy control (HC) datasets available, mainly due to ethical restrictions relating to the HC dataset used in studies done previously based on the MS dataset (158, 159). For Study 2 we had a cross-sectional HC dataset for the second follow-up, carefully matched for the 3T

MRI scans. For Study 3 we used some of the acquired MRI features from Study 2 to investigate longitudinal cognitive data. Hence, although the MRI data were correlated with HC, we did not have a control group for the cognitive data. The cognitive data have been compared with HC at baseline, as previously mentioned in this paragraph (159), although the same ethical restrictions made it impossible to use the longitudinal HC data from the same dataset.

With regard to internal validity, there are many possible confounders (157). In the case of Study 1, it might be questioned how the self-report questionnaires for depressive symptoms and fatigue truly reflected the subjects' clinical symptoms. There are known disadvantages, especially in the case of the FSS (160), for which there are issues in capturing severe fatigue-related disability (161), and many overlapping questions concerning other disorders and symptoms with the method for reporting depressive symptoms using BDI-II (162).

7.2 Patients and controls

7.2.1 Patient recruitment and follow-up

As previously mentioned, the prospective cohort was founded by my co-supervisor Gro O. Nygaard as part of her doctoral thesis work. Inclusion criteria were established to capture the main part of the RRMS population that was being treated with DMT (RRMS diagnosis in the period 2009–2012, age range 18–50 years, affiliated with the Department of Neurology, Oslo University Hospital, Ullevål). As shown in Fig. 11, only nine patients declined to participate (11%). In addition, some patients were not contacted due to finite MRI slots (21%), and a large share of the initial MS subjects discovered in the hospital registries (28%) fulfilled at least one exclusion criteria (uncertain diagnosis, lack of fluency in Norwegian, neurological or psychiatric disease, drug abuse, previous head trauma, pregnancy, or previous adverse reaction to MRI contrast agents).

We had to consider the fact that 29% of the subjects (31 of 108 eligible subjects) either denied to participate or were not included due to low availability of the MRI scanner and that might have contributed to selection bias in our data. For the longitudinal data with two follow-up investigations, the missing subjects also had to be considered as a selection bias. From the first and second follow-up, the subjects mainly stated five reasons for not being able to contribute: (1) pregnancy or young children at home, (2) they had moved to another city or country, (3) they were hospitalized, (4) lack of time, (5) or they simply did not want to participate in the study. We were very flexible in the second follow-up in terms of

adapting our calendars so that the subjects could continue to be part of the study, and even examined them at the weekends if necessary.

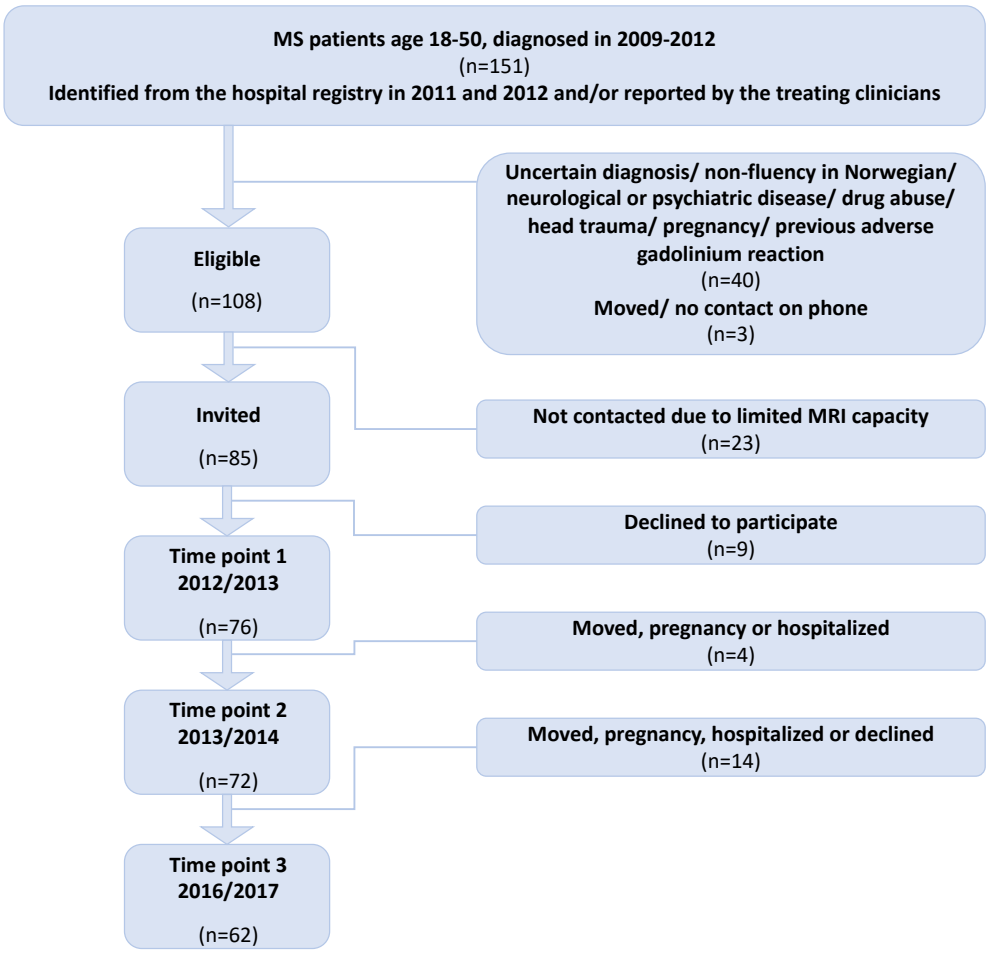

Fig. 11. Patient selection and follow-up at all time points. Adapted from Nygaard et al. (2015) (158).

7.2.2 Healthy controls and training set

For Study 2, we included both a HC group and a training set for the machine learning analyses and brain age adjustments. In the case of the HCs, NORMENT (Norwegian Centre for Mental Disorders Research) gave us access to make a careful selection of an age and gender matched selection from their in-house collected HC samples (TOP and StrokeMRI). We selected 235 HCs (mean age 40.8±7.3 years, 72% female) to match our 58 MS subjects (mean age 40.7±7.6 years, 72% female) with 3T MRI scan at the second follow-up. The two groups were recruited to two parallel studies after stratified random selection from Norway's National Population Register or from newspaper advertisements (163, 164). Exclusion criteria

included a history of neurological or psychiatric disease, estimated IQ (intelligence quotient) < 70, and current medication significantly affecting the nervous system (165).

The training set for the brain age estimation model was carefully set up to fulfil certain requirements: (1) the included datasets should be easily accessible in terms of ethical consent and intellectual rights, (2) the inclusion of as many subjects as possible in the age range 20-50 years, and (3) a wide age range to ensure a reliable prediction model across the life span. In total, we included MRI scans from 3208 HCs aged >12 years (mean age 47.5±19.8 years, 54% female) in the range 12–95 years. The samples were obtained from several publicly available datasets (Fig. 12) and processed in the same MRI pipeline as the HC and MS scans.

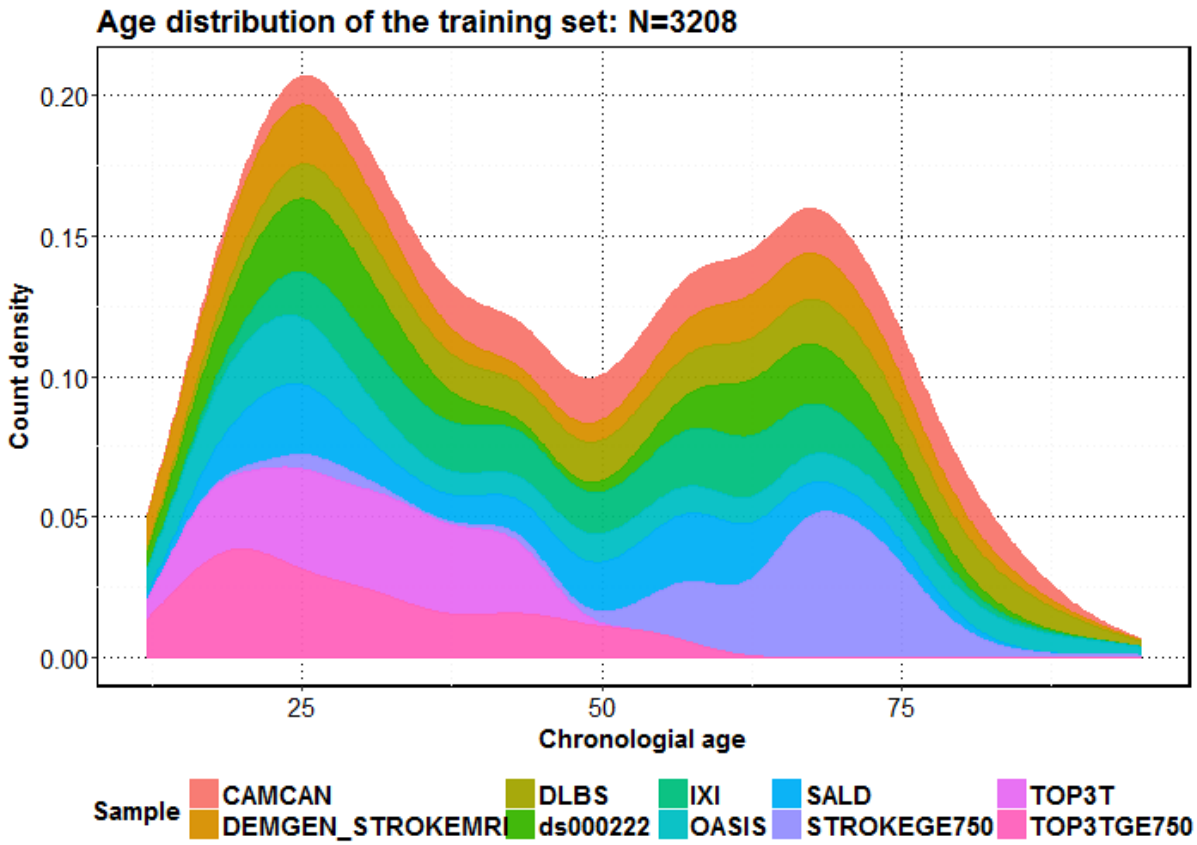

Fig. 12. Overview of the age distribution for the training set and the different samples included. Information concerning the sample cohorts (CAMCAN, DLBS, IXI, OASIS, SALD, STROKEGE750, STROKEMRI, TOP3T, TOP3TGE750, and ds000222) is available in an article by Kaufman et al. (2019) (166).

7.3 MS diagnosis

All MS patients included at baseline were originally diagnosed at the time by experienced clinicians at the Department of Neurology, Oslo University Hospital (OUH), Ullevål between 2009 and 2012, who used the 2005 or 2010 revised McDonald diagnostic criteria for MS (62,

63). During the second follow-up, two of the subjects were classified as RRMS instead of CIS. These changes in MS phenotype were due to the revision of the MS diagnostic criteria in 2017 (64).

The revisions to the diagnostic criteria for MS have been made when the research community has achieved sufficiently more novel insights to be able to shorten the delay from disease onset to diagnosis, yet at the same time maintain the high standards of sensitivity and specificity previously achieved. In sum, the improvements in the diagnostic criteria during our longitudinal study would have made it possible back at baseline to include more subjects with RRMS phenotype. Some subjects who lacked DIT were mainly classified as CIS, especially before the 2010 revisions the diagnostic criteria for MS. Therefore, it might be hypothesized that with the same parameters, a similar inclusion made today with the 2017 revisions would include more MS subjects with shorter disease duration and lower disability.

7.4 Clinical assessments

During all three time points, only three investigators were involved in the clinical neurological assessment of the MS subjects. All investigators were Neurostatus certified to adhere to the then current research standards (167). This was a strength of the longitudinal study, as it lessened the risk of investigator bias, and to some extent evaded the question of inter-rater variability. Across all time points we achieved very high intra-rater variability, as the median EDSS was 2.0 at all time points (range 0-6). As described in detail in the Introduction (Section 4.7.1), EDSS is the most common tool for assessing clinical disability in people with MS. Especially for the lower parts of the scale (< 4.0), the clinical neurological examination is important in order to investigate and uncover specific symptoms from all parts of the CNS (84).

7.4.1 Neurological assessment

For all study visits, we estimated the EDSS score after thorough and focused neurological examinations. The EDSS score has many limitations; an example of a normal distribution of patients with MS is shown in Fig. 13, which shows the bimodal distribution of the EDSS scores in a general MS population (85). By implication, Fig. 13 shows that the rate of increase in EDSS is highly dependent on the EDSS score at baseline. Some researchers have discussed possible limitations regarding whether the scale is responsive to changes in disability over

time in patients, especially in their early course of MS. Also, several aspects such as cognitive function, energy level, quality of life, and mood, are not included in the FS of the EDSS.

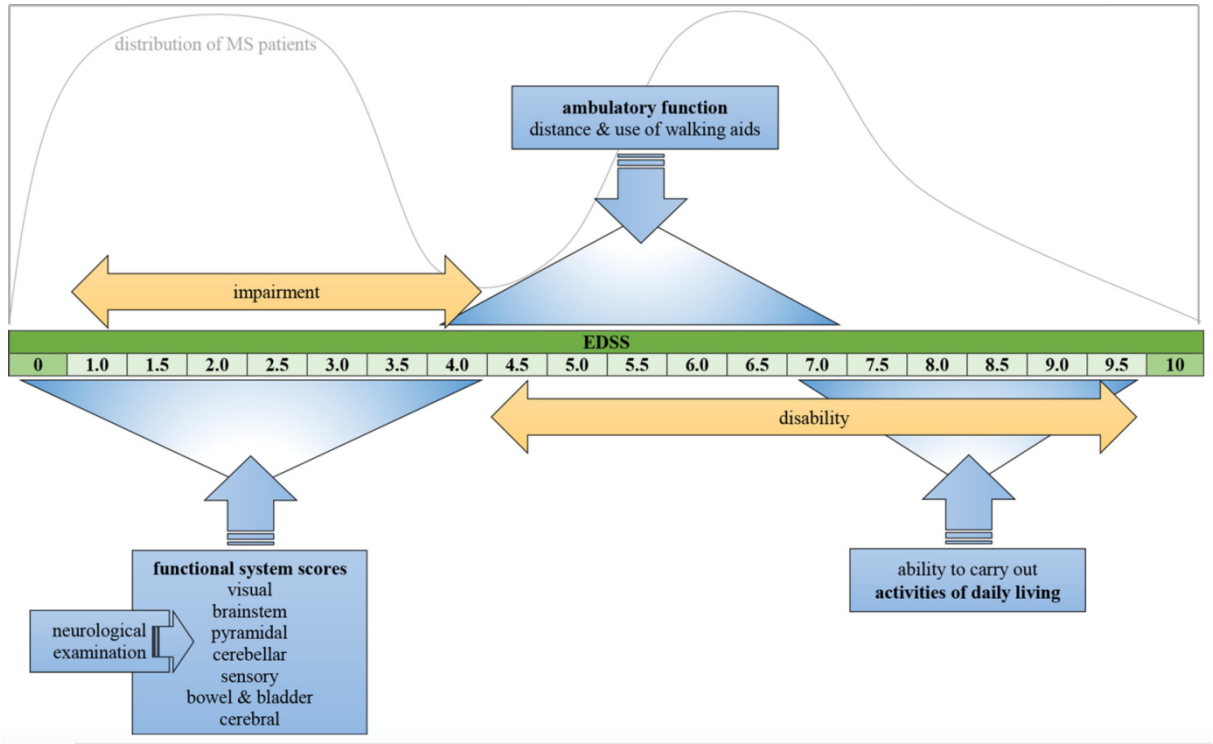

Fig. 13. Schematic overview of the Expanded Disability Status Scale (EDSS). The different neurological symptoms contributing to the different parts of the EDSS eventually determine the final EDSS score. The distribution of patients with MS is shown according to their EDSS score. Reproduced from van Munster et al. (2017) (85), with permission from Springer Nature ltd, obtained 5 September 2019.

Both the Multiple Sclerosis Severity Score (MSSS) (168) and the Age Related Multiple Sclerosis Severity Score (ARMSSS) (169) have been suggested as replacements for the EDSS. While MSSS encompasses the speed of the increase in EDSS, ARMSSS switches age for disease duration. In addition, many other scales and scores have been investigated and validated in larger datasets as superior to EDSS, such as the Multiple Sclerosis Functional Composite (MSFC) (including walking speed, hand function, processing speed, and working memory) (170, 171) and the MS Disease Severity Scale (MS-DSS), the Combinatorial Weight-adjusted Disability Score (CombiWISE) which switches age with disease duration (172). Yet, EDSS is still the single most used MS-disability scale in both the clinical and research setting.

7.4.2 Cognitive assessment

For many years cognitive impairment was not a prioritized research area or clinical focus in MS care, although this situation shifted after a correlation between MRI findings and neuropsychological testing in MS was published in 1989 (173). Not long thereafter, Rao et al. introduced the Brief Repeatable Battery (BRB) for both clinical and research purposes in MS (87, 174). Later, the Minimal Assessment of Cognitive Function in MS (MACFIMS) was introduced and used for many years (175). More recently, in 2011, a large international consensus group lead by Dawn Langdon proposed a simple and robust assembly of cognitive tests called the Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS), performed in just 15 minutes per subject (176, 177). The BICAMS test battery has been validated though international validation studies in a great number of countries (178), including Italy (179), Canada (180), Belgium (181), Japan (182), and recently also in Norway (183).

When we prepared the cognitive test battery before the first assessment in our MS sample, we relied on support from clinical neuropsychologists to ensure that the test battery was comprehensive. However, for the two follow-up investigations we reduced the number of cognitive tests. The reduction in cognitive tests was made as a consequence of feedback from our test subjects. In close collaboration with clinical neuropsychologists, we aimed to reduce the overall strain of research tests imposed on the subjects participation in the study. An overview of our complete cognitive tests is shown in Fig. 14.

Time point 1 (2012/2013)	***Time point 2 (2013/14)***	***Time point 3 (2016/17)***
• **SDMT**	• **SDMT**	• **SDMT**
• **CVLT-II**	• **CVLT-II**	• **CVLT-II**
• **BVMT-R**	• **BVMT-R**	• **BVMT-R**
• **PASAT 3"**	• **D-KEFS Colour-word Interference Test ("Stroop")**	• **PASAT 3"**
• **D-KEFS Colour-word Interference Test ("Stroop")**	• **COWAT**	• **D-KEFS Colour-word Interference Test ("Stroop")**
• **COWAT**	• Attention Network Task (ANT)	• **COWAT**
• MMS	• Antisaccade Test	
• WASI Vocabulary and Matrix Reasoning	• Rey CFT	
• Attention Network Task (ANT)		
• Antisaccade Test		
• Visual N-back Task		

Fig. 14. Overview of the complete cognitive test battery for all time points in the MS sample. Tests marked in bold font were performed throughout the study period, except for PASAT, which was not performed at time point 2. BVMT-R – Brief Visuospatial Memory Test – Revised, COWAT – Controlled

Oral Word Association Test, CVLT-II – California Verbal Learning Test – II, D-KEFS CWIT – Delis-Kaplan Executive Function System Colour-Word Interference Test, MMS – Mini Mental Status, PASAT – Paced Auditory Serial Addition Test, RCFT – Rey Complex Figure Test and Recognition Trial, SDMT – Symbol Digit Modalities Test, WASI – Wechsler's Abbreviated Scale of Intelligence.

In previous publications on the Oslo Longitudinal MS sample authors report comparisons of the cognitive test results from the first two time points, when they also used healthy control samples to compare specific cognitive domains. Especially for the California Verbal Learning Test – II (CVLT-II), they discuss how both the MS and HC samples performed better than the norms for the same data (158). Both samples might have been confounded by selection bias, since they were both recruited from young, urban, well-educated, mainly female persons.

In Study 3 we focused on the complete longitudinal MS dataset and therefore only including the cognitive tests with longitudinal data (shown in bold font in Fig. 14). We included all longitudinal cognitive data to ensure that we would not have any effect of publishing bias. The Paced Auditory Serial Addition Test (PASAT) 3 seconds version, for which the total number of correctly performed calculations is reported, was the only test not performed at all assessments, since we did not include it at time point 2 (156). We excluded the Attention Network Task (ANT) due to complexity of the raw data and the time it would require to process the data in a proper way. The results of PASAT, the Symbol Digit Modalities Test (SDMT) (184), the Colour-Word Interference Test (CWIT) subtest from the Delis-Kaplan Executive Function System (D-KEFS) (185), with time to completion of the Colour Naming and Word Reading conditions, were assessed to evaluate information processing speed and working memory.

To evaluate verbal memory we used the Norwegian version of the CVLT-II (186, 187), with the total number of correctly recalled words in each condition. Visuospatial memory was evaluated using the Brief Visuospatial Memory Test – Revised (BVMT-R) (188), with the total number of points earned in the three immediate recall conditions. Verbal fluency was assessed using the Controlled Oral Word Association Test (COWAT) (189). Lastly, executive functions were measured using the D-KEFS CWIT, with time to completion of the Inhibition and Inhibition/Switching conditions. We did not have a comparable healthy control dataset for this longitudinal dataset, for to the same reasoning as stated earlier (Section 7.1), namely missing ethical approval and consent to share data with us.

The SDMT is known as a hallmark test for processing speed and working memory in MS, where each of the numbers 1–9 is correlated with an abstract symbol in a symbol-number

key (184). The test subject is able to see the key at all times, and is supposed to tell the corresponding number for as many symbols as possible in the course of 90 seconds, where only correct numbers are counted (184). As suggested by the updated BICAMS recommendations, we changed the SDMT instructions before the last assessment so that the subjects performed the SDMT test orally. The purpose was to eliminate any confounding effect of neurological disability in terms of writing performance. The changed instruction might have improved our SDMT results for time point 3, although no significant difference was found in the test performance of the SDMT between the last two assessments (time point 2 and 3). However, neurological deficits affecting speech have been found to affect also the oral version of the SDMT (190).

The SDMT has shown satisfactory to excellent reliability in persons with and without MS when tested over a wide span of time intervals in several different countries (191). Based on the test results, studies have shown that the SDMT is superior compared to other tests of processing speed in terms of differentiating HCs from people with MS (191), where sensitivity and specificity of 82% and 60% respectively have been found for SDMT (192). Additionally, the SDMT has been found to be a sensitive test for longitudinal changes in processing speed (193-195). Due to its psychometric abilities, the SDMT has been recommended as the single preferred surrogate measure of processing speed (191), as a study endpoint (196), and for screening and monitoring of longitudinal cognitive functioning in subjects with MS (197). Standard MRI measures have been found to have a moderate to strong correlation between the SDMT in a meta-analysis of a mixed MS sample, especially T2 lesion volume ($r = -0.45$, $p < 0.001$) and atrophy ($r = -0.54$, $p < 0.001$) (198). The researchers in the same study also favoured the SDMT over PASAT using binomial tests.

The list of words used in the Norwegian version of the CVLT-II (199) has been shown to encompass the most important psychometric strengths as the original version (200). In a report published in 2014 the Norwegian Directory of Health (Helsedirektoratet) concludes that the CVLT-II constitutes adequate abilities as a measure of verbal learning and memory (201). For all longitudinal cognitive tests, we used alternative forms and tests whenever possible to avoid practice-related effects in a retest setting. For example, in the BVMT-R the six geometric figures were different at all time points. However, tests on healthy subjects have shown similar adversity for the same figures in comparison with different figures (188). Although the lists and figures in the cognitive tests were different, the set-up and instructions were the same, thus making practice effects evident, as seen in the case of Study 3.

Differences between countries have been found with respect to the use of the BICAMS test battery, and therefore any systematic effects on the resulting performance cannot be excluded (183, 202). When assessing the cognitive data, we used the raw test scores from all cognitive tests, mainly because this had been evaluated as the most appropriate approach in a longitudinal setting. We corrected for confounding effects of age, gender, disease duration, and educational level at time point 1 without any significant effects in the MS sample. In any sense, the use of raw scores can be beneficial when striving to avoid overcorrection of resulting cognitive tests.

7.4.3 Assessment of fatigue and depression through self-report questionnaires

For all time points, we used the self-report questionnaire Fatigue Severity Scale (FSS) (160), with nine subscores to capture the different dimensions of fatigue. FSS was already well established and had been shown to be associated with structural and brain connectivity changes (107, 108). In addition, we collected the completed Norwegian version of BDI-II questionnaires from the subjects, with a total of 21 subscores to encompass various features of depression (162). In Study 1, FSS mean score ≥ 4 was categorized as clinically significant fatigue, while BDI-II sum score ≥ 14 was categorized as clinically significant depression (160, 162, 203). Both tools are well-established in the context of capturing the symptoms of MS comorbidity with fatigue and depressive symptoms. The reasons for including the questionnaires and especially not a more comprehensive tool for depressive symptoms or the Fatigue Scale for Motor and Cognitive Functions (FSMC) were pragmatic and were decided prior to the first assessment. The FSMC has been shown to distinguish better between symptoms of motor and cognitive fatigue in MS compared with FSS (108, 204). To evaluate the feasibility of the FSS and FSMC in our MS cohort, we collected the FSMC scores at the second time point. However, the results are not yet published. A recent study using longitudinal assessment of fatigue with the Modified Fatigue Impact Scale (MFIS) found evidence of biologically meaningful phenotypes of fatigue in MS (205-207). With regard to BDI-II, the participants in the HC dataset had already used this questionnaire, and therefore to be able to compare the BDI-II data from subjects in our MS sample we adhered to their plan and have performed the BDI-II at all time points. There were some incomplete data in terms of the questionnaires. For data from time point 3, I contacted the subjects with missing data and was able to update the information by filling in all missing data in the questionnaires. The fact that the study subjects filled in the self-report questionnaires at home might have

influenced their answers to some degree, depending on both their home environment and the time they needed to complete the task.

7.5 MRI acquisition

All subjects in the five-year longitudinal MS sample were scanned using the same MRI scanner at all three time points throughout the period. However, there were several possible confounders with respect to the acquisition of images from the scanner. First, degradation of the magnet might have influenced the images longitudinally and caused deterioration in the quality of the images. Second, software upgrades might have influenced the acquisition and resulting data quality. Third, the positioning of the subjects inside the scanner and the positioning of the acquisition window might have influenced the resulting output data, which were especially relevant when analysing the rs-fMRI data in cases where the field of view did not cover the complete area of interest. Fourth, when collecting longitudinal patient data, the following have to be considered as possible confounders: changes in the state and health of the subjects during the course of the follow-up, hydration status, how rested the subjects were, what time of day their scan was done, comorbidities, gender, genetics, and their state of mind in terms of rs-fMRI.

7.5.1 Siemens 1.5T scanner

All of our MS subjects were scanned at up to three time points between January 2012 and August 2017 in a combined clinical and study setting to avoid unnecessary MRI scans for the participants. The same 1.5T scanner (Avanto, Siemens Medical Solutions, Erlangen, Germany) equipped with a 12-channel head coil was used throughout the study period. Minor software updates were installed whenever necessary during that period. Structural MRI data were collected using a 3D T1-weighted MPRAGE (Magnetization Prepared Rapid Gradient Echo) sequence, with the following parameters: TR (repetition time)/TE (echo time)/flip angle/voxel size/FOV (field of view)/slices/scan time/matrix/time to inversion = 2400 ms/3.61 ms/8°/1.20 × 1.25 × 1.25 mm/240/160 sagittal slices/7:42 minutes/192 × 192/1000 ms. The MPRAGE sequence was kept identical during the scanning period. FLAIR (fluid-attenuated inversion recovery), T2 and pre- and post-gadolinium 3D T1 sequences were attained and used for neuroradiological evaluation (158). For the rs-fMRI sequence in Study 1 we used a T_2* weighted echo-planar imaging (EPI) sequence with the following parameters: TR/TE flip angle/voxel size/FOV =3000 ms/70 ms/90°/3.44 × 3.44 × 4 mm/220, descending

acquisition, GeneRalized Autocalibrating Partial Acquisition (GRAPPA) acceleration factor = 2), 28 transversally oriented slices, no gap, with a scan time of 7 minutes and 30 seconds, yielding 150 volumes. Three dummy volumes were collected to avoid T1 saturation effects. At time point 3 we did not perform gadolinium contrast enhancing sequences on the 1.5T scanner as this had been done just two days prior to the use of the 3T scanner, to avoid redundant exposure to the gadolinium contrast agents.

7.5.2 GE 3T scanner

A total of 58 MS subjects were scanned using a 3T GE 750 Discovery MRI scanner with a 32-channel head coil at time point 3 between August 2016 and June 2017, two days prior the 1.5T scan. HCs were scanned solely on the 3T scanner at one time point to provide cross-sectional data. Structural MRI data were collected using a 3D high-resolution IR (inversion recovery)-prepared FSPGR (fast spoiled gradient echo) T1-weighted sequence (3D BRAVO) with the following parameters: TR/TE/flip angle/voxel size/FOV/slices/scan time =8.16 ms/3.18 ms/12°/1 × 1 × 1 mm/256 × 256 mm/188 sagittal slices/4:42 minutes. The reason for performing the two scans just two days apart at the last assessment was to enable direct comparison of scanner effect when estimating brain age. We knew from previous work that the scanner effect on the resulting brain age estimates would be substantial (166), and we wanted to exploit the unique opportunity to investigate the topic in the specific study setting.

7.5.3 Training set using multiple MRI scanners

The training set comprised several local or publicly available MRI datasets (CAMCAN, DLBS, IXI, OASIS, SALD, STROKEGE750, STROKEMRI, TOP3T, TOP3TGE750, and ds000222), both to provide detailed imaging data from 3208 healthy participants across a wide age range and to provide a robust machine learning model for estimating age (see Fig. 12, in Section 7.2.2). The training set consisted of MRI scans from several different MRI scanners and magnet strengths in order to provide a generalizable estimation model to perform well across most other datasets.

7.6 Image analysis

All image analyses were performed at NORMENT, using established international pipelines for the desired MRI sequences and processing methods.

7.6.1 Post-processing pipeline

FreeSurfer is a free open source software package widely used in imaging research across the whole neuroscience research field. The software enables the transformation of raw structural brain MRI data for both volume-based and surface-based analysis (208, 209). The FMRIB Software Library (FSL) is used to analyse functional and diffusion MRI imaging data relating to the brain (210-212), while also taking into account the structural background, for which in the MS field a specific integrated package called SienaX is used (211, 213). We used mainly FreeSurfer and FSL for our imaging analysis, based on previous experience and the current state-of-the-art methods.

7.6.2 FreeSurfer analysis and lesion filling

The structural features used in all three studies were acquired using the T1-weighted scans to provide cortical reconstruction and volumetric segmentation using FreeSurfer 5.3 (208). To optimize our longitudinal dataset in Study 2 and Study 3, these images were processed with the longitudinal stream in FreeSurfer (214). For each subject we created an unbiased within-subject template space and image using robust inverse consistent registration (215). Then we performed a set of processing steps for skull stripping, Talairach transforms, atlas registration, parcellations, and the creation of spherical surface maps based on the individual scans accounting for the within-subject template to increase power and reliability (214). Manual quality control of the MRI scans was performed by me, after thorough training, to edit or exclude scans according to a given set of criteria. Lesion filling using the lesion filling tool in FSL (212) was performed, utilizing already acquired lesion masks based on Cascade (216). The Cascade method achieved high sensitivity (90%) and specificity (99.5%) compared with manual delineation of white matter lesions (216). The lesion masks were assessed and edited by a trained neuroradiologist. Replacement of white matter lesions in MS with computed brain tissue has been shown to improve tissue class measurements (217).

7.6.3 Processing of the rs-fMRI data

Our fMRI analysis for Study 1 is described in detail in Paper I. Briefly, we used an established method with the FMRI Expert Analysis Tool (FEAT) Version 6.00, from FSL (211, 212). We corrected for head motion, removed non-brain tissue, and examined for image artefacts or bad coverage of the desired brain areas (218-220). The subjects' fMRI volumes were registered to the Montreal Neurological Institute (MNI) 152 standard template using the

T1-weighted scan as an intermediate, which had the non-brain tissue removed prior to the use of procedures for automated volumetric segmentation in FreeSurfer (208).
Later, single-session independent component analysis (ICA) was performed for all recorded data using Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) (221). The single-session ICA were submitted to FIX, which attempts to auto-classify ICA components into "good" vs "bad" components (222), to check, improve and distinguish between noise compartments and signal. Data cleaning and consideration were performed by me and the third author of Paper I.

The processed MNI-conformed rs-fMRI datasets were submitted to temporal concatenation group independent component analysis (gICA) using MELODIC (221) with a model order of 30. The group level spatial components were then used as spatial repressors against the original rs-fMRI datasets to estimate subject-specific components and associated time series (dual regression (223)). After careful consideration, we had to exclude eight different group ICA components. The second group ICA component, encompassing the regions of the renowned DMN including the PCC (posterior cingulate cortex), angular gyrus and mPFC (medial prefrontal cortex), was thresholded at $z > 4$ and used as a mask for extracting the mean DMN connectivity value from the subject-specific Dual Regression maps. Due to our limited data, we considered it most efficient to combine all of the involved subregions of the DMN into one subregion for the analyses in Study 1.

7.6.5 Brain age estimation

Brain age estimation is done using machine learning on a large training set of MRI data from HCs to develop a model that can accurately estimate the individual age from brain scans (166, 224-227). Utilizing sensitive measures of MRI-based brain morphometry, brain age estimation provides a robust and novel imaging-based biomarker with potential to yield insights into similarities and differences of disease pathophysiology across brain disorders (166, 227-229). In Study 2 and Study 3 we applied the extreme gradient boosting (XGBoost) package in R (230) to train a prediction model for each sex for the whole brain, using the training sample described in Section 7.5.3 (learning rate eta = 0.01, optimal number of rounds determined in a nested cross-validation loop within the training set, other parameters as default). To encompass the regional specificity of the brain, we used the lobe parcellation labels from FreeSurfer (208) to identify features that overlapped with a given lobe and made unique prediction models for the insula, cingulate, parietal, temporal, frontal, occipital, and subcortical/cerebellar brain regions. XGBoost was chosen on the basis of its resource

efficiency and demonstrated superior performance in previous machine learning competitions, and due to our data being highly monotonic (231). A 10-fold cross-validation of the resulting brain age estimates within the training set validated the good performance of the model and generalizability, with correlation between age and estimated age on r=0.91, before applying the model on our MS data.

To capture the structural integrity of the whole brain, we used an established multimodal cortical parcellation method to extract cortical thickness, volume and area for 180 regions of interest (ROI) per hemisphere (232). We also extracted a set of standard outputs from cerebellar, subcortical and cortical summary statistics (233). This yielded a total set of 1118 structural brain imaging features, 360 each for cortical thickness, volume, and area, and 38 from cerebellar, subcortical and cortical summary statistics.

Imaging-based brain age has been shown to be reliable both within and between MRI scanners (166). Our MS dataset comprised only two different MRI scanners, which allowed us to estimate the scanner coefficient by pooling the data into linear mixed-effects (LME) models, for the whole brain and all regional estimation models in Study 2. The specific scanner coefficients between the 1.5T and 3T scanners in our sample for all brain age estimation models are listed in Table 4. Systematic difference between scanners is a known confounder in imaging research, and therefore it is important to account for such differences in imaging research that incorporate several MRI scanners (234).

Table 4. The calculated scanner coefficients for the brain age estimation models between the 1.5T and 3T MRI scanner.

1.5T versus 3T scanner coefficients for MS patients		
Brain region	**Brain age difference 1.5T vs. 3T**	**Correlation (r)**
Full brain	9.69	0.78
Occipital	3.34	0.68
Temporal	10.20	0.67
Frontal	10.60	0.74
Parietal	9.21	0.72
Cingulate	-6.08	0.74
Insula	-0.43	0.71
Subcortical/Cerebellar	-3.54	0.86

Using linear regressions, we removed the common variance with age, age^2 and sex to account for confounding factors before submitting the residualized version of brain age to further analyses (235). We estimated brain age and calculated the difference between chronological

age and brain age for all subjects, defined as the brain age gap. Several approaches to brain age estimation utilize information from a variety of brain regions (e.g. the hippocampus, subcortical structures, and grey matter and white matter (163)) or from MRI sequences (e.g. T1, T2, diffusion tensor imaging (DTI), and functional MRI) to inform the estimation model (226, 228). It has already been shown that an older-appearing brain is associated with many diseases across both the neurodegenerative and the neurodevelopmental domains, and with worse outcomes in terms of a number of aspects of disability, functioning and quality of life (166). To summarize, the quantitative information derived from brain MRI can be compressed into a single individual measure by using machine learning methods to estimate brain age, comprising advanced imaging information from the whole brain to convey complex data into a single output (227, 236).

7.7 Statistical considerations and illustrations (Figures)

We used R (R Core Team, Vienna, 2018), MATLAB Version 9.2 (The MathWorks Inc., Natick, MA, 2017) and IBM SPSS Statistics 25 (SPSS, Chicago, IL) for statistical analyses. SPSS was mainly used to organize the data for all three studies. MATLAB was used in Study 1 in the early phases of my doctoral research, with assistance from co-author Dag Alnæs, for the statistics and generating the two Figures in Paper I. For the revision of Paper I, I did the remaining statistical analyses with R. For Study 2 and Study 3, R was used for the statistical analyses and to generate all Figures. The main Figures included in Papers II and III were made using the 'ggplot2' (237) and 'cowplot' packages in R, and adapted to visualize our results in a meaningful way.

7.7.1 Descriptive statistics and normality tests

We inspected histograms, Q-Q plots and calculated Pearson's correlations to assess the quality of the data, to check for normal distribution, and to address outliers. For the descriptive statistics, we used the 'summary', 'sd', 'view', and 'lm'. To assess the reliability of data across time we computed the intraclass correlation coefficient (ICC) using the R package 'irr'.

For Study 2, we first explored the commonly accepted 'MatchIt' package in R to select the matched HC, but found it was not as selective and precise as we had aimed for (238). A newly designed script made for R by an international research colleague in the Sys4MS collaboration was explored. This script focuses on matching the gender distribution

first and then later the age distribution, thus maximizing the included HC compared with MatchIt.

Table 5. Overview of the MS-MRI datasets at time point 3 and the matched HC in Study 2.

	Time point 3, 1.5T scanner	Time point 3, 3T scanner	Healthy controls, 3T scanner
Number	**n=58**	**n=58**	**n=235**
Age (SD)	40.8 (7.3)	40.8 (7.3)	40.8 (7.6)
Female (%)	72%	72%	72%

7.7.2 Principal component analysis

A principal component analysis (PCA) utilizes an orthogonal transformation to transform data of possibly correlated features into several principal components that are linearly uncorrelated. The resulting principal components are then reported back in a hierarchical manner, where the first principal component explains the largest portion of the variance in the original data. Every succeeding principal component then has to be orthogonal to the already reported principal components. The number of principal components with enough statistical power to be used in further analyses varies greatly, depending on the size and structure of the data.

We used PCAs in Study 1 and Study 3. For Study 1 we used an in-house MATLAB script from NORMENT. For Study 3 we used the R package 'factoextra' and the embedded R package 'stats'. For Study 1, all the BDI-II and FSS subscores for all subjects were included in the PCA. We also included seven subjects who were missing rs-fMRI data, to increase the statistical power of the resulting principal components. Based on the size of the MS cohort and the interpretation of the component loading coefficients, we included the two highest ranked PCA components for further analyses. The PCA also reported individual component subject scores for each of the two highest ranked principal components. There were no missing data among the data included for Study 1.

In Study 3 we performed a PCA based on data from all 13 cognitive tests for all subjects at all time points, imputing all missing data with the median value for the specific test at the specific time point to increase the statistical power of the PCA analysis.

7.7.3 Effect sizes

The clinical usefulness of p-values are of limited value, and therefore researchers are encouraged to report effect sizes too whenever possible (239). To improve statistical interpretation of the significance of our results, we included Cohen's f^2 estimates (Study 1)

and Cohen's d estimates (Study 2) to address the resulting effect sizes (240). For Study 1, we used MATLAB to estimate Cohen's f^2, while for Study 2 we used the R package 'cohen.d' to estimate Cohen's d. The formulae for both estimates are:

$$\textbf{Cohen's } f^2 = \frac{R^2}{1 - R^2}$$, where R^2 is the squared multiple correlation.

$$\textbf{Cohen's d} = \frac{t(n_1 + n_2)}{\sqrt{n_1 n_2}\sqrt{df}}$$, based on estimated contrast t-statistics.

7.7.4 Linear mixed-effects models and longitudinal analyses

To analyse the longitudinal data in Study 2 and Study 3, we applied LME models to be able to account for the intrasubject correlations in a longitudinal setting (241-243). The LME models accounted for both fixed and random effects, while also enabling us to incorporate and make use of all the available information in the data. The statistical method with LME models allows for a robust approach to assess the relative importance of general versus individual-specific contributions to observed variability in data. All LME models accounted for age, age^2, sex, and scanner (244), and the models were run using the R package 'lme4'.

In Study 2, the annualized rate of change in brain age gap (BAG) was calculated by dividing the total change in BAG by the time interval between the time points. Using this set-up, 0 rate of brain aging would indicate that the rate of chronological and estimated brain aging is similar. Compared with chronological aging, positive and negative values of brain aging correspond to accelerated and decelerated brain aging, respectively. We included the longest possible time interval between time points. Eight MS subjects were excluded due to lack of longitudinal data. Also, to evaluate different rates of brain aging between two groups, we extracted the coefficients from regular linear models in which BAG was used as the dependant variable, while also accounting for age, age^2, sex, and DMT as independent variables. Annual brain atrophy was extracted by comparing the estimated total brain volume from the FreeSurfer output (BrainSegVolNotVent) between time points.

7.7.5 Type 1 and Type 2 errors

Discussions regarding the significance level of 0.05 have been frequent and important (239). Also, the clinical relevance of a statistically significant finding needs to be accounted for and elaborated with confidence intervals, effect sizes and discussion of possible explanations.

Non-significance is often mistaken for no effect, and there is a need to emphasize estimates and their uncertainties in open science (239).

Typically, a Type 1 error in the context of performing a hypothesis test is understood as discarding the null-hypothesis even though the hypothesis is true. A Type 2 error would correspond to accepting a null-hypothesis even though it is wrong (156). All results reported in this thesis were based upon a significance level of 0.05. This p-value corresponds to a 5% chance that reported significant results were due to chance, even though the results were not true. To counter sample size issues in our data, we strived for robust methods with good precision.

In Study 2, we adjusted all resulting p-values using false discovery rate (FDR) procedures implemented in the R package 'p.adjust' to control for multiple testing (245). To adjust for multiple testing in Study 3, we calculated the degree of independence between the resulting cognitive data by making a 13 × 13 correlation matrix based on the Pearson's correlation between all pairwise combinations of the cognitive data. Utilizing the ratio of observed eigenvalue variance to its theoretical maximum, the estimated equivalent number of independent traits in our analyses was 9.0 (246). To control for multiple testing, our significance threshold was therefore adjusted accordingly from 0.05 to 5.6×10^{-3} (246).

8. Ethical Considerations

8.1 Ethical approvals

The research project with the Oslo Longitudinal MS sample on which this thesis is based was carried out in accordance with the recommendations of the Regional Committee for Medical and Health Research Ethics – South East Norway (REC South East Norway). The subjects received updated written and oral information regarding the project, and later all subjects gave written informed consent in accordance with the Declaration of Helsinki prior to the assessments of the second follow-up visit (247). We regularly update our research protocols, participant information and consent forms to meet the demands set forth by REC South East Norway.

Data collection, storage and usage were done in a collaborative environment, in which we adhered to all imposed precautions to secure proper handling. Overall, we gave priority to ensuring proper ethical treatment of every subject included in the project. All concerns before, during and after the project were discussed and addressed accordingly.

8.2 Clinical and cognitive assessments

In the prospective cohort, all of our study subjects were invited for a second follow-up. Based on previous experience, we sought to reduce unnecessary visits and time-consuming examinations to a minimum. We made a structured pipeline for the follow-up, meaning a maximum of two visits, with the main part of the examinations happening on the first visit during daytime. We examined up to three subjects during the first study visit, which involved the following: taking blood samples, filling out self-report questionnaires, clinical examination, MRI scans, eye-examinations, and specific research tasks such as a walking test and a cognitive test battery. To ensure that subjects would continue to participate in the project, we kept in close contact and sent mobile messages (SMSs) to remind them about a planned research visit a few days ahead of the scheduled visit.

All subjects received the same invitation, and those who did not respond were contacted by telephone in the hope that we would secure a response and could plan for their follow-up visit in the project. In total, 14 subjects were unable to attend the second follow-up examination, with valid reasons in all cases. Since fatigue is such an important feature of the MS disease, we tried to make individual adjustments whenever necessary, to ensure that all eligible subjects would attend the follow-up study. We also arranged transport by taxi, both between locations at the hospital and back and forth to their home address if subjects reported

difficulties with walking from one test location to another. Lastly, we made sure to give positive responses throughout the two planned study visits, to ensure a positive experience for all our test subjects.

When we discovered clinically important findings, for example with MRI reports of new lesions, new neurological symptoms, experience of reduced cognitive function, or high levels of depressive symptoms, we reported back to the treating clinician so that the findings would be considered shortly after being discovered.

8.3 MRI safety

Today, both standard and research-related MRI assessments are regarded as safe. However, there are several minor concerns and some possible major risk factors that need to be taken into account. The main concern is the strong magnetic field produced by the scanner's magnets, which can lead to magnetic objects causing damage if the safety regulations are not followed. Also, MS subjects can experience heat-related Uhthoff's phenomenon due to the induced heating of the subjects caused by the radio frequency pulse during some MRI scan sequences. A prolonged research MRI scan causes increased discomfort in comparison with a shorter standard MRI assessment, as well as prolonged exposure to the noise generated by the MRI scanner. For some patients, the prolonged exposure might lead to increased claustrophobia, with resulting fear and anxiety. In our case, we tried to accommodate the subjects with established claustrophobia by giving them shorter pauses during their MRI scan. Some subjects with severe claustrophobia also took prescribed drugs beforehand to reduce the level of stress and discomfort they experienced during the MRI scan. Only one subject declined to participate in the project due to severe claustrophobia.

There has been an increasing focus on sustained gadolinium enhancement in the brain following serial MRI assessments (Section 4.8.2). At Oslo University Hospital we use macrocyclic gadolinium injections only, thus significantly lowering the occurrence of gadolinium accumulation. There is a trend towards excluding gadolinium in most follow-up MRI assessments, although this is still not implemented as a guideline in our clinical practice. In addition, the 3T MRI scans were done as part of the routine clinical assessments and were to be included in a European MS project for which gadolinium sequence was a requirement. Two subjects did not receive gadolinium injection due to previously reported allergic reactions to gadolinium. For the second follow-up of our MS sample, the 1.5T MRI assessment was performed only two days after the first 3T MRI assessment. However, this

MRI protocol was kept to a minimum, in addition to not including gadolinium injection. All MRI images acquired as part of the project were reviewed by an experienced neuroradiologist and included in the patient journals so that they would be available to the treating neurologist and other health care workers.

8.4 Drawing blood

Before starting the project, we made sure to have a standard operating procedure for drawing blood, to ensure the safety and efficacy for all subjects. We were allowed to take nine tubes of full blood, in total 40 ml of blood, at the third time point. The blood samples were taken by me, successfully for all but one subject. The amount of blood is considered to be a low volume, and would not have affected the subjects to a great extent. We received no reports of local infections or other complications due to the needlestick.

9. Discussion

MRI is an important tool to diagnose and follow-up people with MS. The results of studies using standard MRI sequences are leading to the unravelling of the evolution and the effect of treatment in MS (38, 66). Advanced MRI sequences lead to the acquisition of important data, which in turn contribute to increase our understanding of the underlying mechanisms of MS.

Research targeted towards the brain often attracts large interest in the general population. From the beginning of life, we start an aging process. A robust imaging marker with the ability to measure the aging process and the inevitable deviations from chronological aging of the brain will be of great value in a research setting.

Visual MRI interpretation is still the gold standard when assessing MRI scans of MS subjects in a clinical setting. Quantitative MRI techniques in MS have been heavily investigated, with the aim of supporting clinical decisions in MS by implementation in MS guidelines. With the collection of big data and the development of machine learning methods in MS, several quantitative MRI methods are closer to clinical use than ever before.

During the PhD project, I had responsibility for re-examining a group of MS patients that had been investigated on two previous occasions. In parallel to the new follow-up of these patients, we used data from the first time point to investigate the interplay between depressive symptoms and fatigue, looking for associations with default mode network connectivity measured by resting state fMRI (Study 1). After the follow-up was completed, we used the whole longitudinal MS cohort and adapted state-of-the-art machine learning methods to estimate individual brain age based on structural T1-weighted MRI images (Study 2). Lastly, we used the acquired brain age marker to explore cognitive performance in comparison with other known structural MRI markers in our MS dataset (Study 3).

In taking an overview of the topic of my thesis and including the domains covered by the objectives for Studies 1–3, in the next section I discuss in broader terms the state of quantitative MRI in MS and its feasibility as possible biomarkers for clinical use.

9.1 MRI as an MS biomarker

The main objective for this thesis was to study the imaging data in the Oslo longitudinal MS cohort using state-of-the-art MRI post-processing methods to discover imaging markers for disease activity. The term 'biomarker' has been defined by a study group from the National Institute of Health as: '*a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to*

a therapeutic intervention' (248). Biomarkers are commonly perceived as biological markers, such as measuring infection by counting neutrophil leucocytes in the peripheral blood. Recently, studies investigating NfL from serum have provided results indicating that it might serve as an easily accessible biomarker of both treatment response and disease progression (152, 153, 249-251). However, neurofilament light chain lacks specificity as a biomarker and confounding factors need to be accounted for, such as BMI and blood volume (252).

Imaging markers in MS are needed to enable improved personalized treatment of multiple sclerosis and individual treatment selection (54, 253, 254). Clinically relevant imaging markers will need to fulfil required characteristics as relevant markers for disease activity, namely sensitivity to the imaging feature, reliability and specificity of the marker to other imaging features. An MRI marker for diagnostic purposes in MS will need to exclude all other possible differential diagnoses by having high levels of specificity, and ideally support early diagnosis. An imaging marker designed to evaluate treatment response should be reliable in a longitudinal setting and sensitive to subtle fluctuations in the imaging features. Advanced imaging markers that can be used to evaluate subtle and early microstructural damage, functional changes and metabolic differences will enable observation of treatment effects by reducing the consequences of neurodegeneration and neuroinflammation (54). New imaging tools will open up opportunities to study the true effects of brain plasticity, tissue repair and accumulation of damage (54).

9.1.1 Quantitative T1 and T2 MRI imaging markers

For decades, MS lesions and brain atrophy have been evaluated with increased sensitivity as image quality and processing methods have improved (54, 254, 255). The McDonald criteria for MS diagnosis include number, locations and presence of gadolinium-enhancing lesions as important imaging features, yet more advanced MRI markers have not been included (64). The 'clinico-radiological paradox' is understood as the lack of highly significant associations between observed disability and imaging markers in MS. The explanation for the 'clinico-radiological paradox' has shown to be either the lack of specificity in disability scores or lack of sensitivity concerning the imaging features (256).

Brain atrophy is the most commonly applied advanced imaging marker in MS (54, 254). Studies have found that improved methods for measuring brain atrophy have resulted in better correlations between MRI atrophy and clinical features (54). However, the term brain atrophy summarizes a variety of factors such as myelin loss, inflammation, reactive gliosis, and axonal loss, and a study has shown that confounding factors either overestimate or

underestimate many of the aforementioned factors (257). Other confounding factors are MRI acquisition and harmonization across scanners, image processing, pseudoatrophy, hydration status, and BMI. Especially for young adults, normal annual atrophy rates for total brain volume are very low (0.1–0.3%), hence the discrepancy between normal and pathological brain atrophy rates in MS (> 0.4–0.5%) requires robust and sensitive methods for brain atrophy measurement (54, 258). Several ongoing Phase II and Phase III MS trials and clinical MS studies are including brain atrophy as a secondary imaging outcome. Future research will aid in improving the normative values and cut-off scores for pathological brain atrophy in MS, potentially enabling the inclusion of brain atrophy in future MS guidelines (254).

Recently, regional structural or other specific imaging features in MS have been studied to reflect specific anatomical areas of interest, such as grey matter atrophy, cortical lesions, leptomeningeal inflammatory infiltrates, central vein sign, iron rim lesions, deep grey matter, magnetization transfer ratio, and myelin water imaging (54, 254). Interestingly, studies show that deep grey matter atrophy drives the disability progression in MS, also highlighting unique patterns of atrophy across MS phenotypes (259, 260). In a recently published overview of brain atrophy markers in MS, the authors state that the current state-of-art-methods for robust measuring of grey matter atrophy are not sufficiently reproducible for clinical use (261). A recent methodological study investigated automated hippocampal subfield segmentation in 11 healthy controls scanned on two different scanners, and highlighted the high sensitivity and reliability across MRI scanners (262). State-of-the-art deep grey matter segmentation allows for robust measures in a cross-sectional and longitudinal setting, enabling a sensitive method for measuring deep grey matter atrophy (262, 263). Convincing results for fast and robust unsupervised identification of MS and other lesions and longitudinal changes have been found using specific algorithms (264, 265).

Some regional MRI features, such as iron rim lesions, cortical lesions and central vein sign, are highly dependent on high-resolution images from ≥ 7T MRI scanners in order to be visualized, thus rendering these imaging markers unavailable for clinical use at present (54, 254). There is an unmet need for a robust imaging method to capture the dynamics of the chronic, low-grade inflammation in people with MS (266). Studies have revealed subtle differences in lesion formation, including important differences between MS phenotypes, where slowly expanding T2 hyperintense lesions and lesions with high susceptibility-weighted MRI signals in the rims were more frequent in progressive MS (PMS) than in RRMS (266).

Still, none of the above-mentioned imaging markers are in clinical use, and future studies will need to provide new insights into sensitivity, specificity, reproducibility, clinical utility, and feasibility of these candidate imaging markers (54, 254).

9.1.2 Other candidate imaging markers

A large body of evidence has highlighted alterations in brain perfusion in MS (261, 267). Specifically, perfusion MRI of the brain is a promising tool for detecting areas with focal inflammatory activity (267). When capturing brain perfusion using MRI, we indirectly quantify the delivery of blood to the arteries and capillaries in the brain (267). Perfusion MRI is usually measured either by using a paramagnetic contrast agent (dynamic susceptibility contrast and dynamic contrast-enhanced (DCE)) or by a labelling inflowing spin with a radiofrequency pulse, named arterial spin labelling (ASL) (267). Perfusion imaging in MS has been proposed as a supplement to boost early discovery of relapses, the formation of new lesions, and the effects of the different therapies in MS (267).

A research study showing increased innate immune activation in MS using translocator protein positron emission tomography (PET) is promising, although the current cost of PET scans is preventing the transition onto the clinical stage of MS care (266). Several studies of proton magnetic resonance spectroscopy (1H-MRS) have explored mitochondrial dysfunction and other metabolites responsible for neurodegeneration in MS (254). Recently, the possibility of sodium imaging has emerged as a possible technique to assess mitochondrial dysfunction, to serve as a marker for neuroaxonal dysfunction preceding neurodegeneration (254). MR elastography of the brain has demonstrated an association between brain stiffness and neurodegeneration in two studies that compared MS patients with healthy subjects (268). Furthermore, many studies have highlighted the utility of OCT as an imaging marker, where ganglion cell layer thickness and loss of retinal fibre thickness correlate with clinical parameters (269, 270). Performing OCT is quick, easy to perform, well-tolerated, and offers detailed imaging information on unmyelinated axons and retinal structures, as well as the possibility of acquiring oximetry data.

Enabling the identification of neurodegenerative processes preceding irreversible neurodegeneration is the optimal scenario for a biomarker in MS (271). Recent findings from neurochemistry and advances in imaging techniques are positive with respect to the progress made in metabolic and molecular imaging to identify energy dysfunction, cerebral hypoperfusion, and neurofilaments as a biochemical end-stage product of axonal damage (271).

9.2 Is brain age estimation a potential biomarker for MS?

Brain age is a candidate imaging marker for quantifying the pathological deviations from chronological age and normal aging of the brain (236). In the last decade (i.e. since 2010), a vast number of studies have been conducted on brain age, using machine learning methods and using mainly structural MRI sequences, for many diseases and groups of subjects, such as those with epilepsy, those with Alzheimer's disease, mothers, stroke patients, and those with psychiatric disorders (166, 225, 227-229, 272, 273). In Study 2 and Study 3, we investigated the possibility of estimating brain age by using machine learning methods based on structural MRI data, in our longitudinal MS cohort. At the time, only one abstract on reduced brain age following Natalizumab treatment in MS subjects had been published (274). In addition, only one previous study had explored longitudinal imaging data with brain age estimation using machine learning models (275). In line with the previously published results our results demonstrated good reliability of the brain age estimates across all time points, with intraclass correlation coefficient ranging from 0.78 to 0.92 using the residualized brain age estimates. We found that, in comparison with matched healthy controls, our MS subjects had significantly increased brain age. We also assessed the region-specific variety with eight unique brain age models, to show apparent patterns of brain aging in MS subjects. Lastly, our results suggested an accelerated brain aging for the MS subjects in comparison with chronological aging. Interestingly, when exploring the impact of lesion filling on the resulting brain age estimates, we did not find any significant differences. A recent study has highlighted that brain volumetric measurements in MS, provided by FreeSurfer, are not affected by lesion filling (276).

We have supplied overlapping MS data to the largest brain age study conducted to date, which is by Kaufmann et al., and similarly highlights the increased brain age gap observed for MS subjects in comparison with other disorders (166). Kaufmann et al. have also found a significant correlation between brain age gap and disability measured with EDSS ($r=0.23$), but we were not able to replicate the finding in our more modestly sized MS sample. Converging results from the two published MS studies (166, 277) on brain age estimation highlight the regional sensitivity and specificity for increased brain age gap. As reported in Paper II, the cerebellar/subcortical region showed the most evident sign of increased brain age gap, which was already present at baseline. However, the temporal region, and to some degree also the frontal and parietal brain regions, showed no signs of increased brain age. When the results of Study 2 and Study 3 were evaluated, we found a significant correlation with increased brain age and reduced processing speed. An earlier study investigating a large

cross-sectional MS sample found substantially weaker associations between structural brain damage and functional cognitive impairment in early MS patients with low disease burden, and these associations are supported by our findings (278). A systematic review of investigated associations between white matter lesion burden and cognitive function found a weak cumulative correlation (r=-0.30) (279). The authors highlight the need for simultaneous evaluations of multiple components of the complex pathology using selected cognitive and MRI features (279). Our cohort of early MS patients was modestly powered and exhibited low counts of disability progression over the course of five years (median EDSS 2.0 at all time points). We hypothesize that we need a longer follow-up period to be able to capture the dynamics of MS progression, and thus also uncover the individual trajectories.

Accelerated brain aging showed significant correlations with brain atrophy and change in white matter lesion load, which is understood as a proof-of-principle that our brain age method also encompasses the characteristics of two of the hallmark imaging markers in MS. In general, any change in the estimated brain morphometric variables between time points will inevitably influence the brain age estimates. Although we performed a thorough manual quality control (QC) of all images included in the brain age estimation, we cannot rule out that some of the changes in brain age between time points were partly driven by noise characteristics (subject motion, MRI artefacts, technical changes between acquisitions), in particular if the real biological changes were negligible, which is likely to have been the case for the relatively young and well-functioning MS subjects.

Recently, a preprint by Cole et al. found that MS and CIS subjects had older-appearing brains compared with healthy controls (brain-predicted age difference (brain-PAD) 11.8 years) (226). In total, 1204 MS/CIS subjects and 150 healthy controls were included in their preliminary study. There were significant differences between MS subtypes, with the most increased brain-PAD for the SPMS subjects. Increased brain-PAD was associated with increased EDSS, longer disease duration and younger age at diagnosis. The study by Cole et al. also found that lesion filling had no significant impact on the resulting brain age estimates ($r = 0.99$, $p < 0.0001$), and this result supports findings in our study. Cole et al.'s preliminary results validated our findings that brain-PAD in MS and CIS patients increased faster than chronological age. Moreover, their findings highlighted that the 'brain age paradigm' in MS has potential value for: (1) monitoring MS, (2) harmonizing an imaging marker across centres and different scanners, (3) conveying complex anatomical data in an intuitive and simple manner, (4) serving as an imaging marker of brain health with potential prognostic value, and (5) enabling the use of brain age as a surrogate outcome measure in MS (226).

Conventional brain age estimation requires standard T1-weighted structural MRI images, but recent publications have also explored the added value of other sequences for the model performance (280, 281).

Several possible machine learning algorithms are available. Due to the strong and often highly monotonic signal in the data, age prediction is a relatively simple task compared with several other applications (typically case-control classification), and in our experience the specific algorithm or implementation only marginally influences the results. In the analyses of our recent large-scale study of brain age (166), we evaluated different machine learning algorithms and found converging results with different models. We decided to choose XGBoost as the main method for our brain age studies, as it has been the lead solution in many machine learning competitions in the field (231). We also performed a shrinkage linear model (slm) and found converging results on brain age estimates by XGBoost, although our XGBoost model performed slightly better than the shrinkage linear model. In our work on the large-scale study we also investigated how sample size affected the performance of the brain age estimation model and found our training sample of 3208 subjects was acceptable (Fig. 15). We achieved a correlation of 0.91 between estimated brain age and chronological age in the training set, mean the absolute error was 6.4 years.

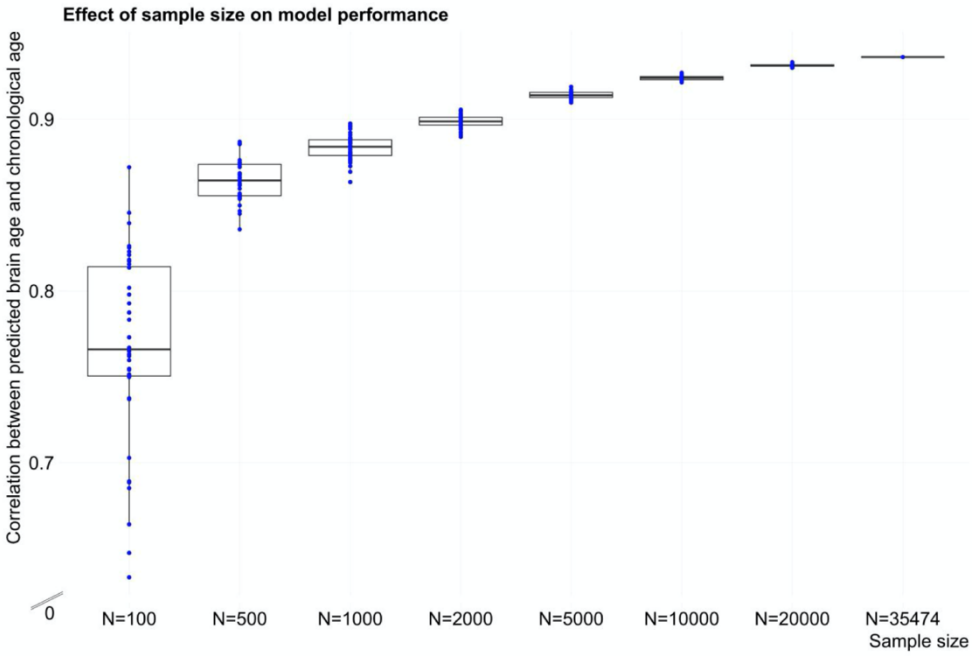

Fig. 15. The impact of sample size on brain age estimation model performance. Increasing the sample size will improve model performance and improve the variation. Reproduced from Kaufmann et al. (2019) (166), with permission from Springer Nature Ltd., obtained 4 February 2020.

Multimodal approaches in terms of imaging features in combination with high-dimensional statistical methods have been shown to outperform conventional unimodal approaches to imaging data, and are proving increasingly important in MS, for which sensitive methods for uncovering imaging responses to disease progression and treatment response are crucial (126). In addition to machine learning approaches, recent advances in computational methods have enabled deep learning methods for brain age estimation. In a recent study, a deep learning method based on structural T1-weighted MRI images was used and slightly outperformed the conventional machine learning models (282).

To summarize, our 'brain age paradigm' highlights the application of a sensitive tool that is able to capture individual trajectories for the neurodegenerative processes in the brain. Brain age estimation condenses all the complex structural MRI features into a single comprehensive estimate, enabling reliable longitudinal assessment, independent of lesion filling and sensitive to subtle changes, to address the effect of disease modifying treatments and predict future disability progression in MS. The article by Kaufmann et al. shows promising evidence of disease specificity, as the neurodegenerative and neurodevelopmental disorders showed different regional patterns, although we must acknowledge that many brain imaging features are common and not specific (166, 236).

9.3 Functional connectivity and clinical relevance

In Study 1 we tried to disentangle symptoms of depression and fatigue in subjects with early MS by analysing DMN connectivity acquired by an rs-fMRI sequence in a cross-sectional setting. Our findings with significant associations between DMN connectivity and both common and unique symptoms of fatigue and depressive symptoms further add to existing knowledge (109). By performing a multivariate decomposition analysis of all of the subscores for fatigue and depressive symptoms, we applied a novel approach that enabled the findings of unique DMN patterns in different MS subgroups to reflect the heterogeneity in the clinical presentation of the MS subjects. As expected, we found that high scores on both depressive symptoms and fatigue were associated with increased DMN connectivity. Interestingly, and clinically relevant, we found that MS subjects with high scores on fatigue and low scores on depressive symptoms were associated with DMN hyperconnectivity. We did not find any significant confounding effects when testing for associations with processing speed, disability, lesion volume, brain atrophy, and DMT. Other studies of MS have investigated the different regional components of the DMN and found distinct patterns of functional connectivity changes (109). Converging evidence that both fatigue and depression are

associated with increased activation of the DMN has been found in previous studies (96, 283). Our novel approach suggests that unique clinical patterns are addressed by performing a multivariate decomposition of the underlying clinical data, to enhance the disentanglement of the complex interplay of associated symptoms in MS.

A large body of evidence has demonstrated that the application of functional MRI in MS is able to capture the dynamic processes of functional reorganization (54). These changes in functional connectivity are present early in MS and are correlated with the extent of damage to the central nervous system (54). Studies have shown that functional brain plasticity, both by increased or decreased functional connectivity, can potentially commit to lessen the clinical consequences of structural damage to the brain (54). Conversely, findings from other studies suggest that the failure of these functional adaptations to compensate for neurological deficits is responsible for several neurological symptoms such as cognitive impairment and fatigue (54).

By capturing the connectivity of different networks in the brain by functional MRI acquisition, it is possible to measure indirectly neuronal activation and its pattern of activity in the brain (54). The fMRI sequence enables the location and measurement of the activation, as well as the connectivity of specific areas in the brain and identification of connectivity networks. Functional MRI acquisition can be captured either as resting state or task-based (visual, cognitive, sensorimotor), depending upon the topic of investigation. With task-based fMRI, MS patients have been found to have adaptive changes in which recruitment of additional motor areas and frontoparietal areas were thought to compensate for the loss of motor signal (284). Also, maladaptive changes observed on task-based fMRI show the abnormal recruitment and connectivity of both classical and non-classical sensorimotor and motor areas of the brain (284).

Interpreting the results of fMRI studies and cognition in MS has been one of the major challenges during the last decade (i.e. since 2010) (285, 286). Especially, since cognitive dysfunction have been linked to both decreased and increased DMN connectivity (285). The major point of discussion has been whether the increased functional connectivity in cognitively preserved MS patients is beneficial or whether it is damaging in a longitudinal setting (285). To address the issue of increased functional connectivity in MS, there is a need for longitudinal fMRI studies of MS (283, 285). A recent study with two-year follow-up data, including rs-fMRI from CIS patients, revealed early differences between CIS patients and healthy controls, suggestive of early protective mechanisms that minimized clinical disability (287). In terms of cognition in MS, rs-fMRI has been suggested as one of the important

neuroimaging modalities to enable increased understanding, preferably in large prospective longitudinal studies (88).

9.4 The challenge of investigating early MS patients

Early MS subjects exhibit a large variety of clinical presentations, most of them with low disease burden. Studies with a short follow-up period have not found robust associations between imaging markers and clinical data (256, 278). A previous study has found mainly low correlations ($r < 0.30$) between structural imaging markers and cognitive and clinical metrics (278). Previous publications relating to our MS cohort highlight the favourable clinical progression and the high cognitive performance of the MS subjects in comparison with healthy controls (159). A more modern treatment approach to MS, with early treatment and more treatment with highly effective DMTs is thought to delay the neurodegenerative processes, thereby slowing the rate of disease progression (48, 66). Our five-year observational period did not enable us to report any significant disease progression. This indicates the need for a longer observational period in our longitudinal MS cohort, to enable the establishment of individual MS trajectories. We have already collected new follow-up data for MS subjects almost seven years after diagnosis, and plan a 10-year follow-up study of the same cohort.

9.5 Methodological considerations

Several methodological considerations should be taken into account when interpreting the results of Studies 1–3, reported in Papers I–III respectively, and when assessing the generalizability of our results. These considerations are discussed in more detail in the following sections.

9.5.1 Imperfect healthy controls

The healthy controls used for both the training set and our matched controls for the brain age estimations were most likely affected by selection bias. This is a known issue when inviting healthy volunteers and is hard to avoid in general. In addition, as shown in Fig. 12 (Section 7.2.2), some age groups are harder to recruit than others. This was especially important in the case of our MS sample, since the majority of our subjects were in the age range 20–50 years, which is known as the most difficult age group to include as healthy controls (166).

Study 1 was a cross-sectional study to access changes in connectivity in the default mode network in association with symptoms of depression and fatigue. As always in most

scientific research questions, it is a strength to control for group differences between a study sample with a healthy control sample. In the presentation of our results in Paper I, we do not state that our findings are specific to MS patients, but rather that we built our analyses on within-group differences in DMN connectivity for the MS subjects. Changes in DMN connectivity within MS patients is an interesting topic, mainly due to the complex interactions and overlap in fatigue and depressive symptoms. MS patients are extremely heterogeneous in their clinical presentations, and this fact encouraged us to look into DMN connectivity on the basis of fatigue and depression. We consider our results representative of a typical MS cohort, and that our within-group analysis was worth the effort to explain some of the variance in symptoms burden. The lack of healthy controls was a concern, but in Paper I we argue that it is not a necessity to give novel insights.

For both Study 2 and Study 3, we lacked longitudinal healthy controls. Recruitment and follow-up of longitudinal controls has not been the focus of our MS research group to date, and for this reason, we have been dependent on healthy control sets from other collaborators. We are currently collecting a longitudinal cohort of healthy controls in our MS research group.

9.5.2 MRI harmonization and image processing

Software upgrades to the MRI scanners during the course of the study period might have affected the longitudinal MRI data. In addition, the magnet coil of an MRI scanner deteriorates over time and will inevitably affect longitudinal MRI data.

For the brain age estimation model, we had to address the scanner differences, as our MS subjects were scanned on a 1.5T MRI scanner longitudinally, but for time point 3 they were scanned both on a 1.5T scanner and a 3T scanner with two days between the scans. The matched healthy controls were scanned solely on the 3T MRI scanner, which induced some statistical considerations when we assessed the longitudinal MS data from the 1.5T scanner. Therefore, we included a scanner coefficient in the brain aging analyses to enable longitudinal assessment of brain aging, also inducing some confounding statistical implications.

Any change in the estimated brain morphometric variables between time points will inevitably influence the brain age estimates. As is the case for any clinical assessments, changes in morphometric variables based on MRI will be partly driven by biological differences in the brain and partly by differences in noise characteristics, as discussed previously (Section 9.2). Brain atrophy represents the final common result of different

pathophysiological substrates, including neuroaxonal and myelin loss, while reactive gliosis and inflammation have the potential to mask considerable tissue loss (54).

9.5.3 Clinical evaluations and statistical considerations

The fact that the investigators were not blinded for the MS subjects might have affected the resulting disability assessments. Also, different investigators for all time points could have affected the resulting clinical evaluations.

For the cognitive evaluations, we carried out a comprehensive investigation at time point 1 (Fig. 14). For the following two investigations, we selected a more specific cognitive battery, mainly so that the time needed to complete the cognitive follow-up was reduced in order to increase the likelihood of the MS subjects attending the follow-up. The selection in cognitive tests might have affected our results. In addition, the cognitive testing was mainly performed late evenings, which could have affected the resulting cognitive performance of the subjects. Most of the large practice-related learning effects observed in Study 3 are statistically significant, yet the lack of longitudinal healthy controls prevents any interpretation as to whether these longitudinal changes were as might be expected in the setting or if they were worse.

An understanding of the underlying mechanisms in the machine learning models of brain age estimation is highly relevant when interpreting the results (226). However, this has not been as straightforward as we had expected, and the results cannot be directly interpreted without considerable drawbacks (288). Machine learning models tend to be biased towards the mean, resulting in older people being underestimated and younger people being overestimated. To counter this confounding effect, we regressed out gender, age and age^2 from the raw brain age estimates in accordance with several previous studies (166, 229, 289). In our statistical models, we accounted for confounding effects and multiple testing to the best of our knowledge.

10. Conclusions and future research

This thesis is based on a longitudinal MS cohort of early MS subjects, which was studied using data from rs-fMRI and structural T1-weighted MRI sequences to evaluate the clinical utility of imaging markers. Fatigue, depressive symptoms, clinical tests and cognitive test performances were investigated to explore specific aspects of MS. The candidate imaging markers and our results indicated that our methods were promising for application in larger and future studies with longer follow-ups. Especially, brain age estimation using machine learning serves as a candidate imaging marker to incorporate the complexity of brain imaging features into a single comprehensive measure. It will become important to incorporate artificial intelligence methods in MS in order to achieve improved diagnostic precision and better estimation of individual trajectories (280).

MS is a complex disease with large heterogeneity, and therefore a well-defined MS population with extensive data collection will enable researchers to evaluate long-term trajectories of important domains such as fatigue, depressive symptoms, disability, and cognition. Quantitative MRI offers the opportunity to establish imaging markers with high sensitivity, specificity and reproducibility. Such markers have yet to be validated for clinical use in MS.

In Study 1, a correlation was identified between increased connectivity in the DMN and increased symptoms score for fatigue and depressive symptoms. Utilizing a statistical approach by PCA, attempts were made to disentangle the complex interplay of fatigue and depressive symptoms with some success. In Study 2 and Study 3, state-of-the-art machine learning methods were applied to complex structural imaging data from a training set of several thousand controls, to estimate individual brain age in MS subjects in a longitudinal setting. The candidate imaging marker were found reliable, sensitive and not affected by lesions. In Study 3, significant longitudinal learning effects were found in most cognitive domains, as well as significant correlations between imaging markers of brain age and thalamus volume with processing speed.

Overall, the findings support correlations of DMN connectivity with fatigue and depressive symptoms. Functional MRI will need to be addressed in larger longitudinal studies to enable the assessment of functional reorganization in brain plasticity in MS. Longitudinal fMRI studies will also enable studies to assess whether early increased connectivity is just a form of compensation to increase brain function or a detrimental process accelerating neurodegeneration in MS. Also, the concept of brain age as a high-dimensional marker of

brain health is promising and needs to be validated in future studies. To apply quantitative MRI markers in clinical use for MS, there is a need to harmonize MRI protocols and data processing, establish robust cut-off scores compared with healthy controls, apply imaging markers in multicentre studies, and include imaging markers in revised MS guidelines.

The next decades of MS research will hopefully include improved markers for early diagnosis, improved personalized treatment approaches and the possibility of better estimating future disability progression. Treatments will need to be approved for the progressive stages of the disease, in addition to new targeted treatments for remyelination in MS to reduce MS subjects' degree of disability in the future. Patient-reported outcome measures have to be incorporated in future studies to reflect the patient perspective in a better way. Collaborative efforts to acquire larger longitudinal datasets in MS will enable the application of sensitive and robust artificial intelligence methods to improve our understanding of MS (Fig. 16). With the increasing focus on patient-reported outcome measures, the implementation of the patient aspect in MS research will improve the quality of research and the quality of life for people with MS and their families.

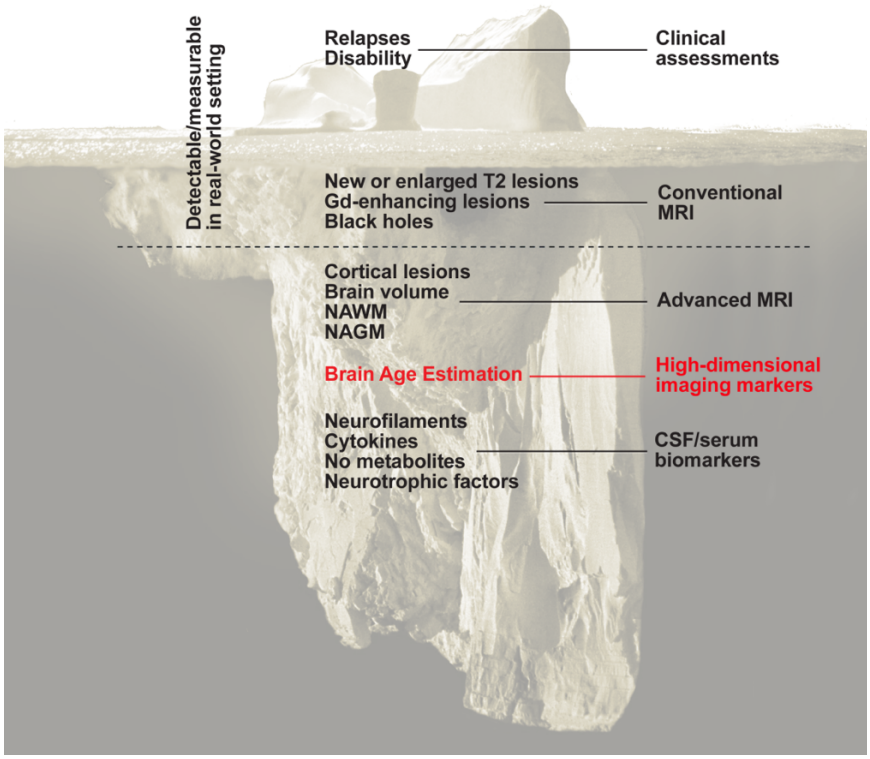

Fig. 16. The iceberg of discovering disease activity in MS. Gd – gadolinium, NAGM – normal-appearing grey matter, NAWM – normal-appearing white matter, NO – nitric oxide. Source: MS Research group, with assistance from Dag Bøgeberg.

11. References

1. Medaer R. Does the history of multiple sclerosis go back as far as the 14th century? Acta Neurol Scand. 1979;60(3):189-92.
2. Trust MS. History of MS 2018 [Available from: https://www.mstrust.org.uk/a-z/history-ms.
3. McAlpine D, Compston A. McAlpine's multiple sclerosis. 2006.
4. Rolak LA. MS: the basic facts. Clin Med Res. 2003;1(1):61-2.
5. Rivers TM, Sprunt DH, Berry GP. Observations on Attempts to Produce Acute Disseminated Encephalomyelitis in Monkeys. J Exp Med. 1933;58(1):39-53.
6. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. Lancet. 1972;1(7762):1240-1.
7. Naito S, Namerow N, Mickey MR, Terasaki PI. Multiple sclerosis: association with HL-A3. Tissue Antigens. 1972;2(1):1-4.
8. Paterson PY. Transfer of allergic encephalomyelitis in rats by means of lymph node cells. J Exp Med. 1960;111:119-36.
9. Karcher D, Van Sande M, Lowenthal A. Micro-electrophoresis in agar gel of proteins of the cerebrospinal fluid and central nervous system. J Neurochem. 1959;4(2):135-40.
10. Holmoy T. The discovery of oligoclonal bands: a 50-year anniversary. Eur Neurol. 2009;62(5):311-5.
11. Fletcher JM, Lalor SJ, Sweeney CM, Tubridy N, Mills KH. T cells in multiple sclerosis and experimental autoimmune encephalomyelitis. Clin Exp Immunol. 2010;162(1):1-11.
12. Sadovnick AD, Baird PA. The familial nature of multiple sclerosis: age-corrected empiric recurrence risks for children and siblings of patients. Neurology. 1988;38(6):990-1.
13. International Multiple Sclerosis Genetics C, Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med. 2007;357(9):851-62.
14. International Multiple Sclerosis Genetics Consortium I. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science. 2019;365(6460):eaav7188.
15. Society NMS. Research News and Progress 2019 [Available from: https://www.nationalmssociety.org/Research/Research-News-Progress.
16. MS International Federation. Atlas of MS 2013 [Available from: https://www.msif.org/about-us/who-we-are-and-what-we-do/advocacy/atlas/.
17. Visser EM, Wilde K, Wilson JF, Yong KK, Counsell CE. A new prevalence study of multiple sclerosis in Orkney, Shetland and Aberdeen city. J Neurol Neurosurg Psychiatry. 2012;83(7):719-24.
18. Berg-Hansen P, Moen SM, Harbo HF, Celius EG. High prevalence and no latitude gradient of multiple sclerosis in Norway. Multiple sclerosis (Houndmills, Basingstoke, England). 2014;20(13):1780-2.
19. Langer-Gould A, Zhang JL, Chung J, Yeung Y, Waubant E, Yao J. Incidence of acquired CNS demyelinating syndromes in a multiethnic cohort of children. Neurology. 2011;77(12):1143-8.
20. Wallin MT, Culpepper WJ, Coffman P, Pulaski S, Maloni H, Mahan CM, et al. The Gulf War era multiple sclerosis cohort: age and incidence rates by race, sex and service. Brain. 2012;135(Pt 6):1778-85.
21. Langer-Gould A, Brara SM, Beaber BE, Zhang JL. Incidence of multiple sclerosis in multiple racial and ethnic groups. Neurology. 2013;80(19):1734-9.

22. Alonso A, Hernan MA. Temporal trends in the incidence of multiple sclerosis: a systematic review. Neurology. 2008;71(2):129-35.
23. Koch-Henriksen N, Sorensen PS. The changing demographic pattern of multiple sclerosis epidemiology. The Lancet Neurology. 2010;9(5):520-32.
24. Simpson S, Jr., Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. J Neurol Neurosurg Psychiatry. 2011;82(10):1132-41.
25. Waubant E, Lucas R, Mowry E, Graves J, Olsson T, Alfredsson L, et al. Environmental and genetic risk factors for MS: an integrated review. Ann Clin Transl Neurol. 2019.
26. Ascherio A, Munger KL. Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. Ann Neurol. 2007;61(6):504-13.
27. Sintzel MB, Rametta M, Reder AT. Vitamin D and Multiple Sclerosis: A Comprehensive Review. Neurol Ther. 2017.
28. Salzer J, Hallmans G, Nystrom M, Stenlund H, Wadell G, Sundstrom P. Vitamin D as a protective factor in multiple sclerosis. Neurology. 2012;79(21):2140-5.
29. Orton SM, Wald L, Confavreux C, Vukusic S, Krohn JP, Ramagopalan SV, et al. Association of UV radiation with multiple sclerosis prevalence and sex ratio in France. Neurology. 2011;76(5):425-31.
30. van der Mei IA, Ponsonby AL, Dwyer T, Blizzard L, Simmons R, Taylor BV, et al. Past exposure to sun, skin phenotype, and risk of multiple sclerosis: case-control study. BMJ. 2003;327(7410):316.
31. Ramagopalan SV, Handel AE, Giovannoni G, Rutherford Siegel S, Ebers GC, Chaplin G. Relationship of UV exposure to prevalence of multiple sclerosis in England. Neurology. 2011;76(16):1410-4.
32. Sintzel MB, Rametta M, Reder AT. Vitamin D and Multiple Sclerosis: A Comprehensive Review. Neurol Ther. 2018;7(1):59-85.
33. Jagannath VA, Filippini G, Di Pietrantonj C, Asokan GV, Robak EW, Whamond L, et al. Vitamin D for the management of multiple sclerosis. The Cochrane database of systematic reviews. 2018;9:CD008422.
34. Berezowska M, Coe S, Dawes H. Effectiveness of Vitamin D Supplementation in the Management of Multiple Sclerosis: A Systematic Review. Int J Mol Sci. 2019;20(6).
35. Degelman ML, Herman KM. Smoking and multiple sclerosis: A systematic review and meta-analysis using the Bradford Hill criteria for causation. Mult Scler Relat Disord. 2017;17:207-16.
36. Riise T, Nortvedt MW, Ascherio A. Smoking is a risk factor for multiple sclerosis. Neurology. 2003;61(8):1122-4.
37. Ascherio A, Munger KL. Epidemiology of Multiple Sclerosis: From Risk Factors to Prevention-An Update. Semin Neurol. 2016;36(2):103-14.
38. Filippi M, Bar-Or A, Piehl F, Preziosa P, Solari A, Vukusic S, et al. Multiple sclerosis. Nat Rev Dis Primers. 2018;4(1):43.
39. Salzer J, Myhr K-M. Epstein–Barr virus is a necessary causative agent in the pathogenesis of multiple sclerosis: No. Multiple Sclerosis Journal. 2013;19(13):1692-3.
40. Goodin DS. The epidemiology of multiple sclerosis: insights to disease pathogenesis. Handbook of clinical neurology. 2014;122:231-66.
41. Wingerchuk DM. Environmental factors in multiple sclerosis: Epstein-Barr virus, vitamin D, and cigarette smoking. Mt Sinai J Med. 2011;78(2):221-30.
42. Ascherio A, Munger KL. Environmental risk factors for multiple sclerosis. Part I: the role of infection. Ann Neurol. 2007;61(4):288-99.

43. Marrie RA, Daumer M. A gestational dose of vitamin D per day keeps the MS doctor away. Neurology. 2017;88(1):13-4.
44. Horizon 2020 Framework Programme for Research and Innovation (2014-2020). MultipleMS. 2019.
45. ERACoSysMed. Sys4MS: Personalizing health care in Multiple Sclerosis using systems medicine tools. 2019.
46. International Multiple Sclerosis Genetics C. A systems biology approach uncovers cell-specific gene regulatory effects of genetic associations in multiple sclerosis. Nat Commun. 2019;10(1):2236.
47. Baecher-Allan C, Kaskow BJ, Weiner HL. Multiple Sclerosis: Mechanisms and Immunotherapy. Neuron. 2018;97(4):742-68.
48. Dobson R, Giovannoni G. Multiple sclerosis - a review. European journal of neurology. 2019;26(1):27-40.
49. Trapp BD, Nave KA. Multiple sclerosis: an immune or neurodegenerative disorder? Annu Rev Neurosci. 2008;31:247-69.
50. Grigoriadis N, van Pesch V, Paradig MSG. A basic overview of multiple sclerosis immunopathology. Eur J Neurol. 2015;22 Suppl 2:3-13.
51. Karussis D. The diagnosis of multiple sclerosis and the various related demyelinating syndromes: a critical review. J Autoimmun. 2014;48-49:134-42.
52. Kutzelnigg A, Lucchinetti CF, Stadelmann C, Bruck W, Rauschka H, Bergmann M, et al. Cortical demyelination and diffuse white matter injury in multiple sclerosis. Brain. 2005;128(Pt 11):2705-12.
53. Kutzelnigg A, Faber-Rod JC, Bauer J, Lucchinetti CF, Sorensen PS, Laursen H, et al. Widespread demyelination in the cerebellar cortex in multiple sclerosis. Brain Pathol. 2007;17(1):38-44.
54. Rocca MA, Preziosa P, Filippi M. Application of advanced MRI techniques to monitor pharmacologic and rehabilitative treatment in multiple sclerosis: current status and future perspectives. Expert Rev Neurother. 2018.
55. Peterson JW, Bo L, Mork S, Chang A, Trapp BD. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. Ann Neurol. 2001;50(3):389-400.
56. Bo L, Vedeler CA, Nyland HI, Trapp BD, Mork SJ. Subpial demyelination in the cerebral cortex of multiple sclerosis patients. J Neuropathol Exp Neurol. 2003;62(7):723-32.
57. Howell OW, Reeves CA, Nicholas R, Carassiti D, Radotra B, Gentleman SM, et al. Meningeal inflammation is widespread and linked to cortical pathology in multiple sclerosis. Brain. 2011;134(Pt 9):2755-71.
58. Lloyd AF, Davies CL, Holloway RK, Labrak Y, Ireland G, Carradori D, et al. Central nervous system regeneration is driven by microglia necroptosis and repopulation. Nat Neurosci. 2019;22(7):1046-52.
59. Schumacher GA, Beebe G, Kibler RF, Kurland LT, Kurtzke JF, McDowell F, et al. Problems of Experimental Trials of Therapy in Multiple Sclerosis: Report by the Panel on the Evaluation of Experimental Trials of Therapy in Multiple Sclerosis. Ann N Y Acad Sci. 1965;122:552-68.
60. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol. 1983;13(3):227-31.
61. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50(1):121-7.

62. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol. 2005;58(6):840-6.
63. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology. 2011;69(2):292-302.
64. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. The Lancet Neurology. 2018;17(2):162-73.
65. De Angelis F, Brownlee WJ, Chard DT, Trip SA. New MS diagnostic criteria in practice. Pract Neurol. 2019;19(1):64-7.
66. Giovannoni G, Butzkueven H, Dhib-Jalbut S, Hobart J, Kobelt G, Pepper G, et al. Brain health: time matters in multiple sclerosis. Mult Scler Relat Disord. 2016;9 Suppl 1:S5-S48.
67. Solomon AJ, Naismith RT, Cross AH. Misdiagnosis of multiple sclerosis: Impact of the 2017 McDonald criteria on clinical practice. Neurology. 2019;92(1):26-33.
68. MAGNIMS. Magnetic Resonance Imaging in MS. 2019.
69. Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. The Lancet Neurology. 2016;15(3):292-303.
70. Geraldes R, Ciccarelli O, Barkhof F, De Stefano N, Enzinger C, Filippi M, et al. The current role of MRI in differentiating multiple sclerosis from its imaging mimics. Nature reviews Neurology. 2018.
71. Borisow N, Mori M, Kuwabara S, Scheel M, Paul F. Diagnosis and Treatment of NMO Spectrum Disorder and MOG-Encephalomyelitis. Front Neurol. 2018;9:888.
72. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Neurology. 1996;46(4):907-11.
73. Tremlett H, Zhao Y, Rieckmann P, Hutchinson M. New perspectives in the natural history of multiple sclerosis. Neurology. 2010;74(24):2004-15.
74. Lunde HMB, Assmus J, Myhr KM, Bo L, Grytten N. Survival and cause of death in multiple sclerosis: a 60-year longitudinal population study. J Neurol Neurosurg Psychiatry. 2017;88(8):621-5.
75. Hobart J, Bowen A, Pepper G, Crofts H, Eberhard L, Berger T, et al. International consensus on quality standards for brain health-focused care in multiple sclerosis. Multiple sclerosis (Houndmills, Basingstoke, England). 2018:1352458518809326.
76. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014;83(3):278-86.
77. Dobson R, Giovannoni G. Multiple Sclerosis - a review. European journal of neurology. 2018.
78. Gass A, Costello FE. Optic neuritis in the diagnosis of MS: More than meets the eye. Neurology. 2018;91(12):545-6.
79. Osborne BJ, Volpe NJ. Optic neuritis and risk of MS: differential diagnosis and management. Cleve Clin J Med. 2009;76(3):181-90.
80. Weintraub MI. Optic neuritis and MS. Neurology. 1988;38(10):1660-1.
81. Filippi M, Preziosa P, Meani A, Ciccarelli O, Mesaros S, Rovira A, et al. Prediction of a multiple sclerosis diagnosis in patients with clinically isolated syndrome using the 2016 MAGNIMS and 2010 McDonald criteria: a retrospective study. Lancet Neurol. 2018;17(2):133-42.

82. MS International Federation. World MS Day. 2019.
83. Kurtzke JF. A new scale for evaluating disability in multiple sclerosis. Neurology. 1955;5(8):580-3.
84. Kurtzke FJ. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-52.
85. van Munster CE, Uitdehaag BM. Outcome Measures in Clinical Trials for Multiple Sclerosis. CNS Drugs. 2017;31(3):217-36.
86. Multiple Sclerosis Trust. Expanded Disability Status Scale (EDSS). 2019.
87. Rao SM, Leo GJ, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. Neurology. 1991;41(5):685-91.
88. Sumowski JF, Benedict R, Enzinger C, Filippi M, Geurts JJ, Hamalainen P, et al. Cognition in multiple sclerosis: State of the field and priorities for the future. Neurology. 2018;90(6):278-88.
89. Demaree HA, DeLuca J, Gaudino EA, Diamond BJ. Speed of information processing as a key deficit in multiple sclerosis: implications for rehabilitation. J Neurol Neurosurg Psychiatry. 1999;67(5):661-3.
90. Van Schependom J, D'Hooghe M B, Cleynhens K, D'Hooge M, Haelewyck MC, De Keyser J, et al. The Symbol Digit Modalities Test as sentinel test for cognitive impairment in multiple sclerosis. Eur J Neurol. 2014;21(9):1219-25, e71-2.
91. McNicholas N, O'Connell K, Yap SM, Killeen RP, Hutchinson M, McGuigan C. Cognitive dysfunction in early multiple sclerosis: a review. QJM: An International Journal of Medicine. 2018;111(6):359-64.
92. Landrø NI, Celius EG, Sletvold H. Depressive symptoms account for deficient information processing speed but not for impaired working memory in early phase multiple sclerosis (MS). J Neurol Sci. 2004;217(2):211-6.
93. Janardhan V, Bakshi R. Quality of life in patients with multiple sclerosis: the impact of fatigue and depression. Journal of the neurological sciences. 2002;205(1):51-8.
94. Feinstein A, Magalhaes S, Richard JF, Audet B, Moore C. The link between multiple sclerosis and depression. Nature reviews Neurology. 2014;10(9):507-17.
95. Patten SB, Marrie RA, Carta MG. Depression in multiple sclerosis. Int Rev Psychiatry. 2017;29(5):463-72.
96. Filippi M, Preziosa P, Rocca MA. Brain mapping in multiple sclerosis: Lessons learned about the human brain. NeuroImage. 2017.
97. Pujol J, Bello J, Deus J, Marti-Vilalta JL, Capdevila A. Lesions in the left arcuate fasciculus region and depressive symptoms in multiple sclerosis. Neurology. 1997;49(4):1105-10.
98. Bakshi R, Czarnecki D, Shaikh ZA, Priore RL, Janardhan V, Kaliszky Z, et al. Brain MRI lesions and atrophy are related to depression in multiple sclerosis. Neuroreport. 2000;11(6):1153-8.
99. Rocca MA, Pravata E, Valsasina P, Radaelli M, Colombo B, Vacchi L, et al. Hippocampal-DMN disconnectivity in MS is related to WM lesions and depression. Human brain mapping. 2015;36(12):5051-63.
100. Feinstein A, Roy P, Lobaugh N, Feinstein K, O'Connor P, Black S. Structural brain abnormalities in multiple sclerosis patients with major depression. Neurology. 2004;62(4):586-90.
101. Gobbi C, Rocca MA, Riccitelli G, Pagani E, Messina R, Preziosa P, et al. Influence of the topography of brain damage on depression and fatigue in patients with multiple sclerosis. Multiple sclerosis (Houndmills, Basingstoke, England). 2014;20(2):192-201.
102. Lerdal A, Celius EG, Krupp L, Dahl AA. A prospective study of patterns of fatigue in multiple sclerosis. European journal of neurology. 2007;14(12):1338-43.

103. Krupp LB, Serafin DJ, Christodoulou C. Multiple sclerosis-associated fatigue. Expert Rev Neurother. 2010;10(9):1437-47.
104. Bakshi R. Fatigue associated with multiple sclerosis: diagnosis, impact and management. Mult Scler. 2003;9(3):219-27.
105. Chaudhuri A, Behan PO. Fatigue in neurological disorders. Lancet. 2004;363(9413):978-88.
106. Palotai M, Guttmann CR. Brain anatomical correlates of fatigue in multiple sclerosis. Mult Scler. 2019:1352458519876032.
107. Rocca MA, Parisi L, Pagani E, Copetti M, Rodegher M, Colombo B, et al. Regional but not global brain damage contributes to fatigue in multiple sclerosis. Radiology. 2014;273(2):511-20.
108. Filippi M, Rocca MA, Colombo B, Falini A, Codella M, Scotti G, et al. Functional magnetic resonance imaging correlates of fatigue in multiple sclerosis. Neuroimage. 2002;15(3):559-67.
109. Bisecco A, Nardo FD, Docimo R, Caiazzo G, d'Ambrosio A, Bonavita S, et al. Fatigue in multiple sclerosis: The contribution of resting-state functional connectivity reorganization. Mult Scler. 2018;24(13):1696-705.
110. Young IR, Hall AS, Pallis CA, Legg NJ, Bydder GM, Steiner RE. Nuclear magnetic resonance imaging of the brain in multiple sclerosis. Lancet. 1981;2(8255):1063-6.
111. Wattjes MP, Rovira A, Miller D, Yousry TA, Sormani MP, de Stefano MP, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis--establishing disease prognosis and monitoring patients. Nature reviews Neurology. 2015;11(10):597-606.
112. Arevalo O, Riascos R, Rabiei P, Kamali A, Nelson F. Standardizing Magnetic Resonance Imaging Protocols, Requisitions, and Reports in Multiple Sclerosis: An Update for Radiologist Based on 2017 Magnetic Resonance Imaging in Multiple Sclerosis and 2018 Consortium of Multiple Sclerosis Centers Consensus Guidelines. J Comput Assist Tomogr. 2019;43(1):1-12.
113. Pasquier F, Leys D, Weerts JG, Mounier-Vehier F, Barkhof F, Scheltens P. Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. Eur Neurol. 1996;36(5):268-72.
114. Filippi M, Preziosa P, Banwell BL, Barkhof F, Ciccarelli O, De Stefano N, et al. Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. Brain. 2019.
115. Currie S, Hoggard N, Craven IJ, Hadjivassiliou M, Wilkinson ID. Understanding MRI: basic MR physics for physicians. Postgrad Med J. 2013;89(1050):209-23.
116. Glover GH. Overview of functional magnetic resonance imaging. Neurosurg Clin N Am. 2011;22(2):133-9, vii.
117. Gasperini C, Prosperini L, Tintore M, Sormani MP, Filippi M, Rio J, et al. Unraveling treatment response in multiple sclerosis: A clinical and MRI challenge. Neurology. 2019;92(4):180-92.
118. Sahraian MA, Radue E-W. Gadolinium Enhancing Lesions in Multiple Sclerosis. In: Radü E-W, Sahraian MA, editors. MRI Atlas of MS Lesions. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 45-74.
119. Kanda T, Ishii K, Kawaguchi H, Kitajima K, Takenaka D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. Radiology. 2014;270(3):834-41.

120. Zivadinov R, Bergsland N, Hagemeier J, Ramasamy DP, Dwyer MG, Schweser F, et al. Cumulative gadodiamide administration leads to brain gadolinium deposition in early MS. Neurology. 2019;93(6):e611-e23.
121. Traboulsee A, Simon JH, Stone L, Fisher E, Jones DE, Malhotra A, et al. Revised Recommendations of the Consortium of MS Centers Task Force for a Standardized MRI Protocol and Clinical Guidelines for the Diagnosis and Follow-Up of Multiple Sclerosis. AJNR Am J Neuroradiol. 2016;37(3):394-401.
122. Traboulsee A, Li D. Addressing Concerns Regarding the Use of Gadolinium in a Standardized MRI Protocol for the Diagnosis and Follow-Up of Multiple Sclerosis. AJNR Am J Neuroradiol. 2016;37(12):E82-E3.
123. Falk Delgado A, Van Westen D, Nilsson M, Knutsson L, Sundgren PC, Larsson EM, et al. Diagnostic value of alternative techniques to gadolinium-based contrast agents in MR neuroimaging-a comprehensive overview. Insights Imaging. 2019;10(1):84.
124. Schmierer K, Campion T, Sinclair A, van Hecke W, Matthews PM, Wattjes MP. Towards a standard MRI protocol for multiple sclerosis across the UK. Br J Radiol. 2019;92(1101):20180926.
125. Filippi M, Preziosa P, Rocca MA. MRI in multiple sclerosis: what is changing? Current opinion in neurology. 2018;31(4):386-95.
126. Kanber B, Nachev P, Barkhof F, Calvi A, Cardoso J, Cortese R, et al. High-dimensional detection of imaging response to treatment in multiple sclerosis. npj Digital Medicine. 2019;2(1):49.
127. Wang C, Beadnall HN, Hatton SN, Bader G, Tomic D, Silva DG, et al. Automated brain volumetrics in multiple sclerosis: a step closer to clinical application. Journal of neurology, neurosurgery, and psychiatry. 2016;87(7):754-7.
128. Pareto D, Sastre-Garriga J, Alberich M, Auger C, Tintore M, Montalban X, et al. Brain regional volume estimations with NeuroQuant and FIRST: a study in patients with a clinically isolated syndrome. Neuroradiology. 2019;61(6):667-74.
129. Beadnall HN, Wang C, Van Hecke W, Ribbens A, Billiet T, Barnett MH. Comparing longitudinal brain atrophy measurement techniques in a real-world multiple sclerosis clinical practice cohort: towards clinical integration? Ther Adv Neurol Disord. 2019;12:1756286418823462.
130. Brown JWL, Coles A, Horakova D, Havrdova E, Izquierdo G, Prat A, et al. Association of Initial Disease-Modifying Therapy With Later Conversion to Secondary Progressive Multiple Sclerosis. JAMA. 2019;321(2):175-87.
131. Paty DW, Li DK. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. II. MRI analysis results of a multicenter, randomized, double-blind, placebo-controlled trial. UBC MS/MRI Study Group and the IFNB Multiple Sclerosis Study Group. Neurology. 1993;43(4):662-7.
132. IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. 1993;43(4):655-.
133. Hauser SL, Bar-Or A, Comi G, Giovannoni G, Hartung HP, Hemmer B, et al. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. The New England journal of medicine. 2017;376(3):221-34.
134. Montalban X, Hauser SL, Kappos L, Arnold DL, Bar-Or A, Comi G, et al. Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. The New England journal of medicine. 2017;376(3):209-20.
135. Atkins H. Stem Cell Transplantation to Treat Multiple Sclerosis. JAMA. 2019;321(2):153-5.

136. Paolicelli D, Manni A, Iaffaldano A, Trojano M. Efficacy and Safety of Oral Therapies for Relapsing-Remitting Multiple Sclerosis. CNS Drugs. 2020.
137. Sykehusinnkjøp. LIS-avtale 1905 MS for immunmodulerende (sykdomsmodifiserende) medikamenter med indikasjonen multippel sklerose (MS). 2019.
138. McCall B. Alemtuzumab to be restricted pending review, says EMA. The Lancet. 2019;393(10182):1683.
139. European Medicines Agency. Use of multiple sclerosis medicine Lemtrada restricted while EMA review is ongoing. 2019.
140. Tjelle TE RC, Ohm IK, Pike E, Håheim, LL, Bidonde J, Fretheim A, Juvet LK. Disease-modifying treatments for relapsing remitting multiple sclerosis, including rituximab. A health technology assessment. 2019.
141. Whittam DH, Tallantyre EC, Jolles S, Huda S, Moots RJ, Kim HJ, et al. Rituximab in neurological disease: principles, evidence and practice. Pract Neurol. 2019;19(1):5-20.
142. Hauser SL. The Charcot Lecture | beating MS: a story of B cells, with twists and turns. Mult Scler. 2015;21(1):8-21.
143. Myhr KM, Torkildsen O, Lossius A, Bo L, Holmoy T. B cell depletion in the treatment of multiple sclerosis. Expert Opin Biol Ther. 2019;19(3):261-71.
144. Berntsson SG, Kristoffersson A, Bostrom I, Feresiadou A, Burman J, Landtblom AM. Rapidly increasing off-label use of rituximab in multiple sclerosis in Sweden - Outlier or predecessor? Acta Neurol Scand. 2018;138(4):327-31.
145. Ineichen BV, Moridi T, Granberg T, Piehl F. Rituximab treatment for multiple sclerosis. Mult Scler. 2019:1352458519858604.
146. Burt RK, Balabanov R, Burman J, Sharrack B, Snowden JA, Oliveira MC, et al. Effect of Nonmyeloablative Hematopoietic Stem Cell Transplantation vs Continued Disease-Modifying Therapy on Disease Progression in Patients With Relapsing-Remitting Multiple Sclerosis: A Randomized Clinical Trial. JAMA. 2019;321(2):165-74.
147. Burman J, Iacobaeus E, Svenningsson A, Lycke J, Gunnarsson M, Nilsson P, et al. Autologous haematopoietic stem cell transplantation for aggressive multiple sclerosis: the Swedish experience. J Neurol Neurosurg Psychiatry. 2014;85(10):1116-21.
148. Helse Bergen H. RAM-MS. 2019.
149. Baldassari LE, Feng J, Clayton BLL, Oh SH, Sakaie K, Tesar PJ, et al. Developing therapeutic strategies to promote myelin repair in multiple sclerosis. Expert Rev Neurother. 2019:1-17.
150. Göttle P, Förster M, Weyers V, Küry P, Rejdak K, Hartung H-P, et al. An unmet clinical need: roads to remyelination in MS. Neurological Research and Practice. 2019;1(1):21.
151. Clerico M, Artusi CA, Liberto AD, Rolla S, Bardina V, Barbero P, et al. Natalizumab in Multiple Sclerosis: Long-Term Management. Int J Mol Sci. 2017;18(5).
152. Domingues RB, Fernandes GBP, Leite F, Senne C. Neurofilament light chain in the assessment of patients with multiple sclerosis. Arq Neuropsiquiatr. 2019;77(6):436-41.
153. Disanto G, Barro C, Benkert P, Naegelin Y, Schadelin S, Giardiello A, et al. Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. Annals of neurology. 2017;81(6):857-70.
154. Khalil M, Pirpamer L, Hofer E, Voortman MM, Barro C, Leppert D, et al. Serum neurofilament light levels in normal aging and their association with morphologic brain changes. Nat Commun. 2020;11(1):812.
155. Bevan CJ, Cree BA. Disease activity free status: a new end point for a new era in multiple sclerosis clinical research? JAMA Neurol. 2014;71(3):269-70.
156. Laake P, Olsen BR, Benestad HB. Forskning i medisin og biofag. Oslo: Gyldendal; 2008.

157. Laake P, Hjartåker A, Thelle DS, Veierød MB. Epidemiologiske og kliniske forskningsmetoder. Oslo: Gyldendal Akademisk; 2007.
158. Nygaard GO, Walhovd KB, Sowa P, Chepkoech JL, Bjornerud A, Due-Tonnessen P, et al. Cortical thickness and surface area relate to specific symptoms in early relapsing-remitting multiple sclerosis. Multiple sclerosis (Houndmills, Basingstoke, England). 2015;21(4):402-14.
159. Nygaard GO, Celius EG, de Rodez Benavent SA, Sowa P, Gustavsen MW, Fjell AM, et al. A Longitudinal Study of Disability, Cognition and Gray Matter Atrophy in Early Multiple Sclerosis Patients According to Evidence of Disease Activity. PloS one. 2015;10(8):e0135974.
160. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Archives of neurology. 1989;46(10):1121-3.
161. Jason LA, Evans M, Brown M, Porter N, Brown A, Hunnell J, et al. Fatigue Scales and Chronic Fatigue Syndrome: Issues of Sensitivity and Specificity. Disabil Stud Q. 2011;31(1).
162. Beck AT SR, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: The Psychological Corporation; 1996.
163. Richard G, Kolskaar K, Sanders A-M, Kaufmann T, Petersen A, Doan NT, et al. Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. bioRxiv. 2018.
164. Moberget T, Doan NT, Alnaes D, Kaufmann T, Cordova-Palomera A, Lagerberg TV, et al. Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: a multisite mega-analysis of 983 patients and 1349 healthy controls. Molecular psychiatry. 2018;23(6):1512-20.
165. Doan NT, Engvig A, Zaske K, Persson K, Lund MJ, Kaufmann T, et al. Distinguishing early and late brain aging from the Alzheimer's disease spectrum: consistent morphological patterns across independent samples. NeuroImage. 2017;158:282-95.
166. Kaufmann T, van der Meer D, Doan NT, Schwarz E, Lund MJ, Agartz I, et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. Nat Neurosci. 2019;22(10):1617-23.
167. Neurostatus e-Test. 2016.
168. Roxburgh RH, Seaman SR, Masterman T, Hensiek AE, Sawcer SJ, Vukusic S, et al. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. Neurology. 2005;64(7):1144-51.
169. Manouchehrinia A, Westerlind H, Kingwell E, Zhu F, Carruthers R, Ramanujam R, et al. Age Related Multiple Sclerosis Severity Score: Disability ranked by age. Multiple sclerosis (Houndmills, Basingstoke, England). 2017:1352458517690618.
170. Rudick R, Antel J, Confavreux C, Cutter G, Ellison G, Fischer J, et al. Recommendations from the National Multiple Sclerosis Society Clinical Outcomes Assessment Task Force. Ann Neurol. 1997;42(3):379-82.
171. Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain. 1999;122 (Pt 5):871-82.
172. Weideman AM, Barbour C, Tapia-Maltos MA, Tran T, Jackson K, Kosa P, et al. New Multiple Sclerosis Disease Severity Scale Predicts Future Accumulation of Disability. Front Neurol. 2017;8:598.
173. Rao SM, Leo GJ, Haughton VM, St Aubin-Faubert P, Bernardin L. Correlation of magnetic resonance imaging with neuropsychological testing in multiple sclerosis. Neurology. 1989;39(2 Pt 1):161-6.

174. Rao SM, Leo GJ, Ellington L, Nauertz T, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning. Neurology. 1991;41(5):692-6.
175. Benedict RH, Fischer JS, Archibald CJ, Arnett PA, Beatty WW, Bobholz J, et al. Minimal neuropsychological assessment of MS patients: a consensus approach. Clin Neuropsychol. 2002;16(3):381-97.
176. Langdon DW, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al. Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Mult Scler. 2012;18(6):891-8.
177. Benedict RH, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al. Brief International Cognitive Assessment for MS (BICAMS): international standards for validation. BMC Neurol. 2012;12:55.
178. BICAMS. BICAMS. 2019.
179. Goretti B, Niccolai C, Hakiki B, Sturchio A, Falautano M, Minacapelli E, et al. The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS): normative values with gender, age and education corrections in the Italian population. BMC Neurol. 2014;14:171.
180. Walker LA, Osman L, Berard JA, Rees LM, Freedman MS, MacLean H, et al. Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS): Canadian contribution to the international validation project. J Neurol Sci. 2016;362:147-52.
181. Costers L, Gielen J, Eelen PL, Schependom JV, Laton J, Remoortel AV, et al. Does including the full CVLT-II and BVMT-R improve BICAMS? Evidence from a Belgian (Dutch) validation study. Mult Scler Relat Disord. 2017;18:33-40.
182. Niino M, Fukazawa T, Kira JI, Okuno T, Mori M, Sanjo N, et al. Validation of the Brief International Cognitive Assessment for Multiple Sclerosis in Japan. Mult Scler J Exp Transl Clin. 2017;3(4):2055217317748972.
183. Skorve E, Lundervold AJ, Torkildsen O, Myhr KM. The Norwegian translation of the brief international cognitive assessment for multiple sclerosis (BICAMS). Multiple sclerosis and related disorders. 2019;36:101408.
184. Smith A. Symbol digit modalities test: Manual. Los Angeles, CA: Western Psychological Services; 1982.
185. Delis DC, Kaplan E, Kramer JH. Delis–Kaplan executive function system: Examiner's manual. San Antonio, TX: The Psychological Corporation; 2001.
186. Delis DC, Kramer JH, Kaplan E, Ober BA. California verbal learning test – second edition. Adult version. Manual. San Antonio, TX: Psychological Corporation; 2000.
187. Delis D, Kramer J, Kaplan E, Ober B. California Verbal Learning Test (CVLT–II). Norwegian Manual Supplement. Stockholm: Pearson Assessment; 2004.
188. Benedict RHB. Brief visuospatial memory test - revised: Professional manual. Lutz, FL: Psychological Assessment Resources; 1997.
189. Spreen O, Strauss E. A compendium of neuropsychological tests : administration, norms, and commentary. 2nd ed. ed. New York: Oxford University Press; 1998.
190. Arnett PA, Smith MM, Barwick FH, Benedict RH, Ahlstrom BP. Oralmotor slowing in multiple sclerosis: relationship to neuropsychological tasks requiring an oral response. J Int Neuropsychol Soc. 2008;14(3):454-62.
191. Benedict RHB, DeLuca J, Phillips G, LaRocca N, Hudson LD, Rudick R. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Mult Scler. 2017;23(5):721-33.
192. Corfield F, Langdon D. A Systematic Review and Meta-Analysis of the Brief Cognitive Assessment for Multiple Sclerosis (BICAMS). Neurology and therapy. 2018;7(2):287-306.

193. Amato MP, Portaccio E, Goretti B, Zipoli V, Iudice A, Della Pina D, et al. Relevance of cognitive deterioration in early relapsing-remitting MS: a 3-year follow-up study. Mult Scler. 2010;16(12):1474-82.
194. Benedict RHB, Morrow SA, Guttman BW, Cookfair D, Schretlen DJ. Cognitive reserve moderates decline in information processing speed in multiple sclerosis patients. J Int Neuropsychol Soc. 2010;16(5):829-35.
195. Morrow SA, Drake A, Zivadinov R, Munschauer F, Weinstock-Guttman B, Benedict RH. Predicting loss of employment over three years in multiple sclerosis: clinically meaningful cognitive decline. Clin Neuropsychol. 2010;24(7):1131-45.
196. Strober L, DeLuca J, Benedict RH, Jacobs A, Cohen JA, Chiaravalloti N, et al. Symbol Digit Modalities Test: A valid clinical trial endpoint for measuring cognition in multiple sclerosis. 2018:1352458518808204.
197. Kalb R, Beier M, Benedict RH, Charvet L, Costello K, Feinstein A, et al. Recommendations for cognitive screening and management in multiple sclerosis care. Mult Scler. 2018;24(13):1665-80.
198. Rao SM, Martin AL, Huelin R, Wissinger E, Khankhel Z, Kim E, et al. Correlations between MRI and Information Processing Speed in MS: A Meta-Analysis. Mult Scler Int. 2014;2014:975803.
199. Lundervold AJ, Sundet KS. Norsk Versjon av CVLT-II. Manualsupplement. Oslo: Psykologiforlaget AB; 2004.
200. Bosnes O. California Verbal Learning Test-II (CVLT-II) utprøvd i et klinisk utvalg i Norge ; psykometriske egenskaper og krysskulturell validitet. Tidsskrift for Norsk psykologforening. 2007;44(7):887-92.
201. Siqveland J, Dalsbø TK, Harboe I, Leiknes KAJRfK. Måleegenskaper ved den norske versjonen av California Verbal Learning Test II (CVLT-II). 2014.
202. Smerbeck A, Benedict RHB, Eshaghi A, Vanotti S, Spedo C, Blahova Dusankova J, et al. Influence of nationality on the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Clin Neuropsychol. 2018;32(1):54-62.
203. Lerdal A, Wahl A, Rustoen T, Hanestad BR, Moum T. Fatigue in the general population: a translation and test of the psychometric properties of the Norwegian version of the fatigue severity scale. Scand J Public Health. 2005;33(2):123-30.
204. Penner IK, Raselli C, Stocklin M, Opwis K, Kappos L, Calabrese P. The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler. 2009;15(12):1509-17.
205. Palotai M, Cavallari M, Healy BC, Guttmann CR. A novel classification of fatigue in multiple sclerosis based on longitudinal assessments. Mult Scler. 2020:1352458519898112.
206. Krieger SC, Sumowski JF. Toward biologically meaningful categories of MS-related fatigue. Mult Scler. 2020:1352458520901705.
207. Amtmann D, Bamer AM, Noonan V, Lang N, Kim J, Cook KF. Comparison of the psychometric properties of two fatigue scales in multiple sclerosis. Rehabil Psychol. 2012;57(2):159-66.
208. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage. 1999;9(2):179-94.
209. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage. 1999;9(2):195-207.
210. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, et al. Bayesian analysis of neuroimaging data in FSL. Neuroimage. 2009;45(1 Suppl):S173-86.
211. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage. 2004;23 Suppl 1:S208-19.

212. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. Neuroimage. 2012;62(2):782-90.
213. Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage. 2002;17(1):479-89.
214. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage. 2012;61(4):1402-18.
215. Reuter M, Rosas HD, Fischl B. Highly accurate inverse consistent registration: a robust approach. NeuroImage. 2010;53(4):1181-96.
216. Damangir S, Manzouri A, Oppedal K, Carlsson S, Firbank MJ, Sonnesyn H, et al. Multispectral MRI segmentation of age related white matter changes using a cascade of support vector machines. Journal of the neurological sciences. 2012;322(1-2):211-6.
217. Battaglini M, Jenkinson M, De Stefano N. Evaluating and reducing the impact of white matter lesions on brain volume measurements. Hum Brain Mapp. 2012;33(9):2062-71.
218. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage. 2002;17(2):825-41.
219. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp. 2002;17(3):143-55.
220. Smith SM, Brady JM. SUSAN—A New Approach to Low Level Image Processing. International Journal of Computer Vision. 1997;23(1):45-78.
221. Beckmann CF, DeLuca M, Devlin JT, Smith SM. Investigations into resting-state connectivity using independent component analysis. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2005;360(1457):1001-13.
222. Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, et al. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. NeuroImage. 2014;95:232-47.
223. Nickerson LD, Smith SM, Ongur D, Beckmann CF. Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses. Front Neurosci. 2017;11:115.
224. Cole JH, Poudel RPK, Tsagkrasoulis D, Caan MWA, Steves C, Spector TD, et al. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. NeuroImage. 2017.
225. Franke K, Ziegler G, Klöppel S, Gaser C. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. NeuroImage. 2010;50(3):883-92.
226. Cole JH, Raffel J, Friede T, Eshaghi A, Brownlee W, Chard D, et al. Accelerated brain ageing and disability in multiple sclerosis. bioRxiv. 2019:584888.
227. Franke K, Gaser C. Ten Years of BrainAGE as a Neuroimaging Biomarker of Brain Aging: What Insights Have We Gained? Frontiers in Neurology. 2019;10(789).
228. Cole JH, Franke K. Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. Trends Neurosci. 2017;40(12):681-90.
229. Richard G, Kolskar K, Ulrichsen KM, Kaufmann T, Alnaes D, Sanders AM, et al. Brain age prediction in stroke patients: Highly reliable but limited sensitivity to cognitive performance and response to cognitive training. Neuroimage Clin. 2019;25:102159.
230. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; San Francisco, California, USA. 2939785: ACM; 2016. p. 785-94.
231. Chen TaG, Carlos, editor XGBoost: A Scalable Tree Boosting System. KDD ´16; 2016; San Francisco, California, US: ACM; 2016.

232. Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, et al. A multi-modal parcellation of human cerebral cortex. Nature. 2016;536(7615):171-8.
233. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002;33(3):341-55.
234. Madan CR. Advances in Studying Brain Morphology: The Benefits of Open-Access Data. Front Hum Neurosci. 2017;11:405.
235. Le TT, Kuplicki RT, McKinney BA, Yeh HW, Thompson WK, Paulus MP, et al. A Nonlinear Simulation Framework Supports Adjusting for Age When Analyzing BrainAGE. Front Aging Neurosci. 2018;10:317.
236. Bijsterbosch J. How old is your brain? Nature Neuroscience. 2019;22(10):1611-2.
237. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.
238. Ho D, Imai K, King G, Stuart E. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. Political Analysis. 2007;15:199–236.
239. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. 2019;567(7748):305-7.
240. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol Rev Camb Philos Soc. 2007;82(4):591-605.
241. Koerner TK, Zhang Y. Application of Linear Mixed-Effects Models in Human Neuroscience Research: A Comparison with Pearson Correlation in Two Auditory Electrophysiology Studies. Brain Sci. 2017;7(3).
242. Harrison XA, Donaldson L, Correa-Cano ME, Evans J, Fisher DN, Goodwin CED, et al. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ. 2018;6:e4794.
243. Douglas B MM, Bolker, B., Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, Articles. 2018;67.
244. Bernal-Rusiel JL, Greve DN, Reuter M, Fischl B, Sabuncu MR, Alzheimer's Disease Neuroimaging I. Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. NeuroImage. 2013;66:249-60.
245. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.
246. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb). 2005;95(3):221-7.
247. World Medical A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191-4.
248. Wagner JA, Atkinson AJ, Jr. Measuring Biomarker Progress. Clin Pharmacol Ther. 2015;98(1):2-5.
249. Kuhle J, Kropshofer H, Haering DA, Kundu U, Meinert R, Barro C, et al. Blood neurofilament light chain as a biomarker of MS disease activity and treatment response. Neurology. 2019;92(10):e1007-e15.
250. Novakova L, Zetterberg H, Sundstrom P, Axelsson M, Khademi M, Gunnarsson M, et al. Monitoring disease activity in multiple sclerosis using serum neurofilament light protein. Neurology. 2017;89(22):2230-7.
251. Varhaug KN, Barro C, Bjornevik K, Myhr KM, Torkildsen O, Wergeland S, et al. Neurofilament light chain predicts disease activity in relapsing-remitting MS. Neurol Neuroimmunol Neuroinflamm. 2018;5(1):e422.

252. Manouchehrinia A, Piehl F, Hillert J, Kuhle J, Alfredsson L, Olsson T, et al. Confounding effect of blood volume and body mass index on blood neurofilament light chain levels. Ann Clin Transl Neurol. 2020.
253. Rotstein D, Montalban X. Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. Nat Rev Neurol. 2019;15(5):287-300.
254. Cortese R, Collorone S, Ciccarelli O, Toosy AT. Advances in brain imaging in multiple sclerosis. Ther Adv Neurol Disord. 2019;12:1756286419859722.
255. Marciniewicz E, Podgorski P, Sasiadek M, Bladowska J. The role of MR volumetry in brain atrophy assessment in multiple sclerosis: A review of the literature. Adv Clin Exp Med. 2019.
256. Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. Curr Opin Neurol. 2002;15(3):239-45.
257. Rocca MA, Battaglini M, Benedict RH, De Stefano N, Geurts JJ, Henry RG, et al. Brain MRI atrophy quantification in MS: From methods to clinical application. Neurology. 2017;88(4):403-13.
258. De Stefano N, Stromillo ML, Giorgio A, Bartolozzi ML, Battaglini M, Baldini M, et al. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2016;87(1):93-9.
259. Eshaghi A, Marinescu RV, Young AL, Firth NC, Prados F, Jorge Cardoso M, et al. Progression of regional grey matter atrophy in multiple sclerosis. Brain. 2018;141(6):1665-77.
260. Eshaghi A, Prados F, Brownlee WJ, Altmann DR, Tur C, Cardoso MJ, et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. Ann Neurol. 2018;83(2):210-22.
261. Storelli L, Rocca MA, Pagani E, Van Hecke W, Horsfield MA, De Stefano N, et al. Measurement of Whole-Brain and Gray Matter Atrophy in Multiple Sclerosis: Assessment with MR Imaging. Radiology. 2018;288(2):554-64.
262. Brown EM, Pierce ME, Clark DC, Fischl BR, Iglesias JE, Milberg WP, et al. Test-retest reliability of FreeSurfer automated hippocampal subfield segmentation within and across scanners. NeuroImage. 2020:116563.
263. Pagnozzi AM, Fripp J, Rose SE. Quantifying deep grey matter atrophy using automated segmentation approaches: A systematic review of structural MRI studies. Neuroimage. 2019;201:116018.
264. Nguyen TD, Zhang S, Gupta A, Zhao Y, Gauthier SA, Wang Y. Fast and Robust Unsupervised Identification of MS Lesion Change Using the Statistical Detection of Changes Algorithm. AJNR Am J Neuroradiol. 2018;39(5):830-3.
265. Sundaresan V, Zamboni G, Le Heron C, Rothwell PM, Husain M, Battaglini M, et al. Automated lesion segmentation with BIANCA: Impact of population-level features, classification algorithm and locally adaptive thresholding. Neuroimage. 2019;202:116056.
266. Matthews PM. Chronic inflammation in multiple sclerosis - seeing what was always there. Nat Rev Neurol. 2019.
267. Lapointe E, Li DKB, Traboulsee AL, Rauscher A. What Have We Learned from MR Imaging in Multiple Sclerosis? AJNR American journal of neuroradiology. 2018.
268. Murphy MC, Huston J, 3rd, Ehman RL. MR elastography of the brain and its application in neurological diseases. Neuroimage. 2019;187:176-83.
269. Britze J, Frederiksen JL. Optical coherence tomography in multiple sclerosis. Eye (Lond). 2018.
270. Hanson JV, Lukas SC, Pless M, Schippling S. Optical Coherence Tomography in Multiple Sclerosis. Semin Neurol. 2016;36(2):177-84.

271. Van Schependom J, Guldolf K, D'Hooghe M B, Nagels G, D'Haeseleer M. Detecting neurodegenerative pathology in multiple sclerosis before irreversible brain tissue loss sets in. Transl Neurodegener. 2019;8:37.
272. Boyle R, Jollans L, Rueda-Delgado LM, Rizzo R, Yener GG, McMorrow JP, et al. Brain-predicted age difference score is related to specific cognitive functions: A multi-site replication analysis. bioRxiv. 2019:652867.
273. de Lange AG, Kaufmann T, van der Meer D, Maglanoc LA, Alnaes D, Moberget T, et al. Population-based neuroimaging reveals traces of childbirth in the maternal brain. Proc Natl Acad Sci U S A. 2019.
274. Raffel J, Cole J, Record C, Sridharan S, Sharp D, Nicholas R. Brain Age: A novel approach to quantify the impact of multiple sclerosis on the brain (P1.371). Neurology. 2017;88(16 Supplement).
275. Franke K, Gaser C. Longitudinal Changes in Individual BrainAGE in Healthy Aging, Mild Cognitive Impairment, and Alzheimer's Disease. GeroPsych. 2012;25(4):235-45.
276. Guo C, Ferreira D, Fink K, Westman E, Granberg T. Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. Eur Radiol. 2019;29(3):1355-64.
277. Høgestøl EA, Kaufmann T, Nygaard GO, Beyer MK, Sowa P, Nordvik JE, et al. Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis. Frontiers in Neurology. 2019;10(450).
278. Uher T, Krasensky J, Sobisek L, Blahova Dusankova J, Seidl Z, Kubala Havrdova E, et al. Cognitive clinico-radiological paradox in early stages of multiple sclerosis. Ann Clin Transl Neurol. 2018;5(1):81-91.
279. Mollison D, Sellar R, Bastin M, Mollison D, Chandran S, Wardlaw J, et al. The clinico-radiological paradox of cognitive function and MRI burden of white matter lesions in people with multiple sclerosis: A systematic review and meta-analysis. PLoS One. 2017;12(5):e0177727.
280. Shah NH, Milstein A, Bagley Ph DS. Making Machine Learning Models Clinically Useful. JAMA. 2019.
281. Richard G, Kolskar K, Sanders AM, Kaufmann T, Petersen A, Doan NT, et al. Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. PeerJ. 2018;6:e5908.
282. Jonsson BA, Bjornsdottir G, Thorgeirsson TE, Ellingsen LM, Walters GB, Gudbjartsson DF, et al. Brain age prediction using deep learning uncovers associated sequence variants. Nature Communications. 2019;10(1):5409.
283. Enzinger C, Pinter D, Rocca MA, De Luca J, Sastre-Garriga J, Audoin B, et al. Longitudinal fMRI studies: Exploring brain plasticity and repair in MS. Mult Scler. 2016;22(3):269-78.
284. Filippi M, Rocca MA. Present and future of fMRI in multiple sclerosis. Expert Rev Neurother. 2013;13(12 Suppl):27-31.
285. Schoonheim MM, Meijer KA, Geurts JJ. Network collapse and cognitive impairment in multiple sclerosis. Front Neurol. 2015;6:82.
286. Schoonheim MM, Geurts JJ, Barkhof F. The limits of functional reorganization in multiple sclerosis. Neurology. 2010;74(16):1246-7.
287. Rocca MA, Hidalgo de La Cruz M, Valsasina P, Mesaros S, Martinovic V, Ivanovic J, et al. Two-year dynamic functional network connectivity in clinically isolated syndrome. Mult Scler. 2019:1352458519837704.
288. Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M, et al. Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci. 2017;20(3):299-303.

289. de Lange AG, Kaufmann T, van der Meer D, Maglanoc LA, Alnaes D, Moberget T, et al. Population-based neuroimaging reveals traces of childbirth in the maternal brain. Proc Natl Acad Sci U S A. 2019;116(44):22341-6.

12. Errata

I

PLOS ONE

RESEARCH ARTICLE

Symptoms of fatigue and depression is reflected in altered default mode network connectivity in multiple sclerosis

Einar August Høgestøl[1]*, Gro Owren Nygaard[2], Dag Alnæs[3], Mona K. Beyer[4], Lars T. Westlye[3,5☯], Hanne Flinstad Harbo[1,2☯]

1 Department of Neurology, Institute of Clinical Medicine, University of Oslo, Oslo, Norway, **2** Department of Neurology, Oslo University Hospital, Oslo, Norway, **3** NORMENT, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway, **4** Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway, **5** Department of Psychology, University of Oslo, Oslo, Norway

☯ These authors contributed equally to this work.
* einar.august@gmail.com

OPEN ACCESS

Citation: Høgestøl EA, Nygaard GO, Alnæs D, Beyer MK, Westlye LT, Harbo HF (2019) Symptoms of fatigue and depression is reflected in altered default mode network connectivity in multiple sclerosis. PLoS ONE 14(4): e0210375. https://doi.org/10.1371/journal.pone.0210375

Editor: Marcello Moccia, Universita degli Studi di Napoli Federico II, ITALY

Received: December 19, 2018

Accepted: March 19, 2019

Published: April 1, 2019

Data Availability Statement: Data cannot be shared publicly because of patient privacy restrictions put upon the project by the Regional Ethical Committee. Summary data as published in this paper will however be available. We can apply to the Regional Ethical Committee of South Eastern Norway for permission to share data with other researchers or new collaborators. Any party who is interested may therefore contact us, and we will pursuit a permission to share the data, while still adhering to patient privacy requirements of the "Law of Health Research" in Norway. Readers or

Abstract

Background

Fatigue and depression are frequent and often co-occurring symptoms in multiple sclerosis (MS). Resting-state functional magnetic resonance imaging (rs-fMRI) represents a promising tool for disentangling differential associations between depression and fatigue and brain network function and connectivity. In this study we tested for associations between symptoms of fatigue and depression and DMN connectivity in patients with MS.

Materials and methods

Seventy-four MS patients were included on average 14 months after diagnosis. They underwent MRI scanning of the brain including rs-fMRI, and symptoms of fatigue and depression were assessed with Fatigue Severity Scale (FSS) and Beck Depression Inventory II (BDI). A principal component analysis (PCA) on FSS and BDI scores was performed, and the component scores were analysed using linear regression models to test for associations with default mode network (DMN) connectivity.

Results

We observed higher DMN connectivity with higher scores on the primary principal component reflecting common symptom burden for fatigue and depression (Cohen's $f^2 = 0.075$, $t = 2.17$, $p = 0.03$). The secondary principal component reflecting a pattern of low fatigue scores with high scores of depression was associated with lower DMN connectivity (Cohen's $f^2 = 0.067$, $t = -2.1$, $p = 0.04$). Using continuous mean scores of FSS we also observed higher DMN connectivity with higher symptom burden ($t = 3.1$, $p = 0.003$), but no significant associations between continuous sum scores of BDI and DMN connectivity ($t = 0.8$, $p = 0.4$).

researchers may contact the Head of research at the Neurological Department, Oslo University Hospital, Professor John-Anker Zwart (j.a.zwart@medisin.uio.no) to request the data and start such a process.

Funding: The project was supported by grants (Principal Investigator for both grants: HFH) from The Research Council of Norway (https://www.forskningsradet.no/en/Home_page/1177315753906) (grant number 240102) and the South-Eastern Regional Health Authorities of Norway (https://www.helse-sorost.no/south-eastern-norway-regional-health-authority) (grant number 2011059/ ES563338/Biotek 2021) (Principal Investigator for both grants: HFH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusion

Multivariate decomposition of FSS and BDI data supported both overlapping and unique manifestation of fatigue and depression in MS patients. Rs-fMRI analyses showed that symptoms of fatigue and depression were reflected in altered DMN connectivity, and that higher DMN activity was seen in MS patients with fatigue even with low depression scores.

Introduction

Multiple sclerosis (MS) is a heterogeneous disease of the central nervous system (CNS) with typical age of disease onset between 28 and 31 years [1]. One of the most common symptoms in multiple sclerosis (MS) is fatigue, affecting up to 90% of all MS patients [2–4]. Fatigue may have a large impact on the daily life of MS patients and may impair both quality of life and ability to work [2–4]. Depression is also a common symptom in MS with a lifetime prevalence of 40–60% [2, 3, 5]. The pathophysiology of these symptoms in MS is not fully understood [2–4, 6–8].

Structural MRI studies have shown different patterns of cortical thickness in MS patients who have either fatigue, depression or both depression and fatigue, but these cortical underpinnings only explain a proportion of the total variance of the neuropsychiatric symptoms [9]. Diverse results are reported concerning the presence and severity of fatigue in relation to structural MRI findings in MS (lesions, normal appearing white matter damage or grey matter damage) [3, 7, 10, 11]. Some have reported changes in regional cortico-subcortical pathways such as in the prefrontal cortex, thalamus and basal ganglia in patients with MS-related fatigue, while studies using utilizing whole-brain approaches have mostly been inconclusive [4, 7, 8, 11]. Both structural MRI and functional MRI (fMRI) have been applied in many studies with the aim to understand mechanisms responsible for clinical disability, depression, fatigue and cognitive impairment in MS [3, 7, 8, 10–13].

Functional connectivity (FC) can be conceptualized as the interaction between two different brain regions. Disconnection caused by white matter damage in MS leads to brain network dysfunction, named a disconnection syndrome [3]. Regional damage to the white and grey matter in MS patients is likely to disrupt brain network connectivity within cortical and subcortical networks [14]. fMRI has made it possible to assess the integration of activity across distant brain regions and has provided insight into functional brain networks.

Resting-state (rs) fMRI in MS has mainly been used to study the intrinsic functional architecture and connectivity of the brain and relation to disease progression and clinical impairment [14, 15]. In particular, rs-fMRI has highlighted the role of the default mode network (DMN) as a critical hub for both integration and flow of information [16]. The DMN comprises the precuneus, the posterior cingulate cortex (PCC), the angular gyrus, the medial prefrontal cortex (mPFC) and the inferior parietal regions [3, 14]. The DMN is most active when a person is not focused on a specific task, often referred to as wakeful rest [16]. Assuming a role of the DMN in introspection and rumination, DMN changes in MS patients have been proposed to be linked with cognitive dysfunction and depression [16–18].

Some fMRI studies have reported cortico-subcortical dysfunction in MS patients with fatigue, also specifically involving fronto-parietal regions and the basal ganglia [3, 4, 19, 20]. Another fMRI study reported that fatigue was mainly associated with rs-FC changes of the DMN, although with different components of the DMN uniquely involved [12]. A recent rs-fMRI study found that specific thalamo-cortical connections explained different components

of fatigue in MS patients [19]. Thus, there is evidence of altered DMN connectivity in MS patients with symptoms of both depression and fatigue. Although related, these symptoms do not always co-occur, and little is known about the different patterns of DMN alterations with different symptom burden [9]. On this background, we aimed to study the common and differential associations between symptoms of fatigue and depression and DMN connectivity using rs-fMRI in MS.

Materials and methods

Participants

We included in total 74 MS patients at Oslo University Hospital for a prospective longitudinal study. Some other data from this study have been published earlier [21, 22]. All participants were diagnosed between January 2009 and October 2012 with relapsing-remitting MS (RRMS) according to the revised McDonald Criteria [23] and were referred to brain MRI between January 2012 and January 2013. Seven participants did not perform the rs-fMRI sequence, and the remaining 67 participants were used in the current imaging analyses. The time intervals between MRI scans and clinical tests were for all except one patient within 11 days (mean 0.2, median 1, SD 3.0). One patient had to delay MRI and performed the scan two months after testing, but with no clinical relapse in that period. Exclusion criteria included age < 18 years or > 50 years, uncertain diagnosis, non-fluency in Norwegian, neurological or psychiatric disease, steroid intake or clinical relapse within the last six weeks, drug abuse, head trauma, pregnancy and previous adverse gadolinium reaction. Two patients were treated with the same selective serotonin reuptake inhibitor at the time of testing for their depressive symptoms. None of the patients received any medical treatment to improve their fatigue. The project was approved by the regional ethical committee of South Eastern Norway (REC ID:2011/1846), and all participants received oral and written information and gave their written informed consent.

All participants completed a comprehensive neurological examination, including expanded disability status scale (EDSS) by a Neurostatus certified medical doctor (http://www.neurostatus.net/) and symbol digits modalities test (SDMT) within the same week as their MRI examination. All participants also completed self-reported questionnaires concerning fatigue (Fatigue Severity Scale, FSS) [24], with 9 subscores covering the different dimensions of fatigue, and depressive symptoms (Beck Depressive Inventory II, BDI) [25] with a total of 21 subscores to encompass various features of depression. FSS mean score ≥ 4 was categorized as clinically significant fatigue, while BDI sum score ≥ 14 was categorized as clinically significant depressive symptoms [25].

MRI acquisition

The participants were scanned using the same 1.5 T scanner (Avanto, Siemens Medical Solutions; Erlangen, Germany) equipped with a 12-channel head coil. For rs-fMRI we used a T_2^* weighted echo-planar imaging (EPI) sequence (repetition time (TR) = 3000 milliseconds (ms), echo time (TE) = 70 ms, flip angle (FA) = 90˚, voxel size = 3.44 x 3.44 x 4 millimetre (mm), field-of-view (FOV) = 220, descending acquisition, GeneRalized Autocalibrating Partial Acquisition (GRAPPA) acceleration factor = 2), 28 transversally oriented slices, no gap, with a scan time of 7 minutes and 30 seconds, yielding 150 volumes. Three dummy volumes were collected to avoid T1 saturation effects. Structural MRI data were collected using a 3-D T1-weighted Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequence with the following parameters: TR / TE / time to inversion / FA = 2400 ms / 3.61 ms / 1000 ms / 8˚,

matrix 192 × 192, field of view = 240. Each scan lasted 7 minutes and 42 seconds and consisted of 160 sagittal slices with a voxel size of 1.20 × 1.25 × 1.25 mm.

FLAIR sequence parameters: TR / TE / time to inversion/ FA = 6000 ms / 3.33 ms / 2200 ms / variable T2, matrix 256 × 204, field of view = 260. Each scan lasted 7 min 02 sec and consisted of 176 sagital slices, with a slice thickness of 1 mm and a voxel size of 1.0 × 1.0 × 1.0 mm.

fMRI pre-processing and analysis

fMRI analysis was performed using FMRI Expert Analysis Tool (FEAT) Version 6.00, from FMRIB's Software Library [26, 27]. Head motion was corrected using MCFLIRT [28] before linear trends and low-frequency drifts were removed (high-pass filter of 0.01 Hertz). Image sequences were examined for excessive head motion causing image artefacts. FSL Brain extraction tool [29] was used to remove non-brain tissue. Spatial smoothing was performed using a Gaussian kernel filter with a full width at half maximum (FWHM) of 6 mm [30]. FMRIB's Nonlinear Image Registration tool (FNIRT) was used to register the participants fMRI volumes to Montreal Neurological Institute (MNI) 152 standard template using the T1-weighted scan as an intermediate, which had the non-brain tissue removed using procedures for automated volumetric segmentation in Freesurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/) [31].

Single-session independent component analysis (ICA) was performed for all runs using Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) [32]. The single-session ICA were submitted to FIX [33] for automatic classification into signal and noise components, in order to remove noise components from fMRI data. Data cleaning also included correction based on the estimated motion parameters for each run, using linear regression. FIX has been shown to effectively reduce motion induced variability, outperforming methods based on regression of motion parameters or spikes in the dataset [34].

The cleaned and MNI-conformed rs-fMRI datasets were submitted to temporal concatenation group independent component analysis (gICA) using MELODIC [32] with a model order of 30. These group level spatial components were then used as spatial repressors against the original rs-fMRI datasets to estimate subject-specific components and associated time series (dual regression [35]). The second group ICA component, encompassing the regions of the canonical DMN including the PCC, angular gyrus and mPFC, was thresholded at $z>4$ and used as a mask for extracting the mean DMN connectivity value from the subject specific dual-regression maps (Fig 1). The threshold $z>4$ ($p = 0.00006$) was pragmatically chosen based on previous experience.

Brain morphometry

Using the T1-weighted scans we performed cortical reconstruction and volumetric segmentation with FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/) [31]. Several processing steps, such as skull stripping, Talairach transforms, atlas registration as well as spherical surface maps and parcellations were then initialized to increase reliability and power [36].

Manual quality control of the MRI scans from patients was performed by trained research personnel to identify and edit segmentation errors where possible (n = 17 MRI scans). Lesion filling was performed utilizing automatically generated lesion masks from Cascade [37] with the lesion filling tool (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/lesion_filling) in FSL [27]. The lesion masks were assessed by a trained neuroradiologist and normalized to MNI space using FLIRT [28], with the corresponding T1 image as an intermediate. A probabilistic representation of the lesions across all patients is shown in S1 Fig. We estimated total brain volume from the Freesurfer output after lesion filling was performed (BrainSegVolNotVent) and extracted

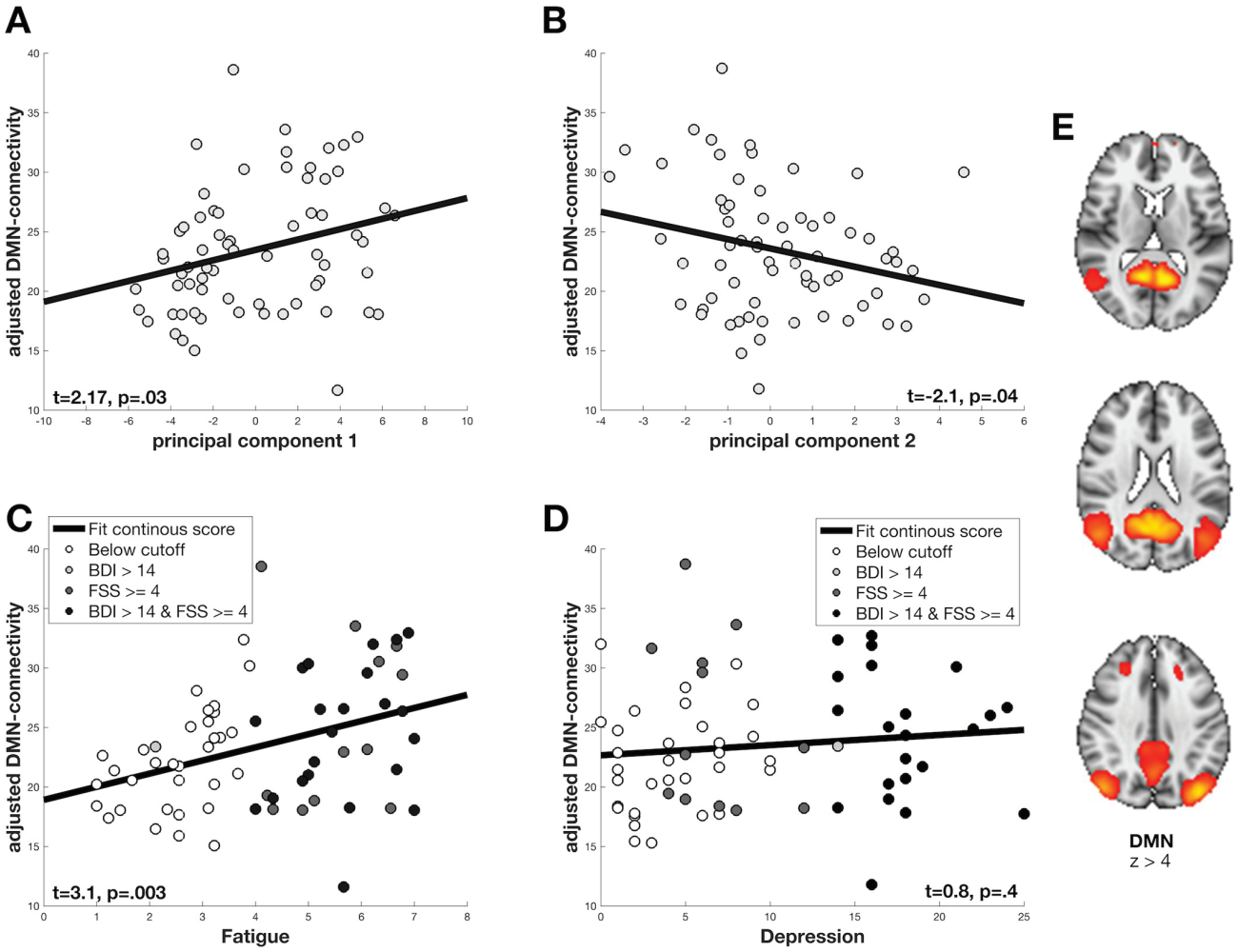

Fig 1. Associations between clinical symptoms and DMN connectivity. The correlation between adjusted DMN connectivity with the PCA components in A and B, and between adjusted DMN connectivity with FSS and BDI continuous scores in C and D. The grey tones for each subject represent clinical categories in C and D as described and shown in Table 1, and individual subject scores in A and B. (A) Increased PCA1 (high burden of both fatigue and depression) is positively correlated with DMN connectivity. (B) Decreased PCA2 (low burden of fatigue and high burden of depression) is negatively correlated with DMN connectivity. (C) Mean FSS correlated with DMN connectivity. (D) BDI sum scores correlated with DMN connectivity. Shown in E is the DMN component from the group independent component analysis (gICA). The component z-statistic map was thresholded at z>4. Depicted in three axial slices the posterior cingulate cortex (PCC) and the medial prefrontal cortex (mPFC) are masked out in red and yellow colours bilaterally.

https://doi.org/10.1371/journal.pone.0210375.g001

mean cortical thickness across the two hemispheres. Lesion volume was measured based the already mentioned lesion masks from Cascade, encompassing lesions in the whole brain.

Statistical analyses

We used MATLAB version 9.2 (The MathWorks Inc., Natick, MA, 2017) and R [38] (R Core Team, Vienna, 2018) for statistical analyses. BDI and FSS subscores for all participants were submitted to PCA, decomposing the data into orthogonal components. To increase the statistical power of the PCA, we kept the seven MS patients missing fMRI data. The PCA yielded component loading coefficients for each questionnaire as well as component subject scores, resulting in a ranked list of PCA components with their associations to each BDI and FSS

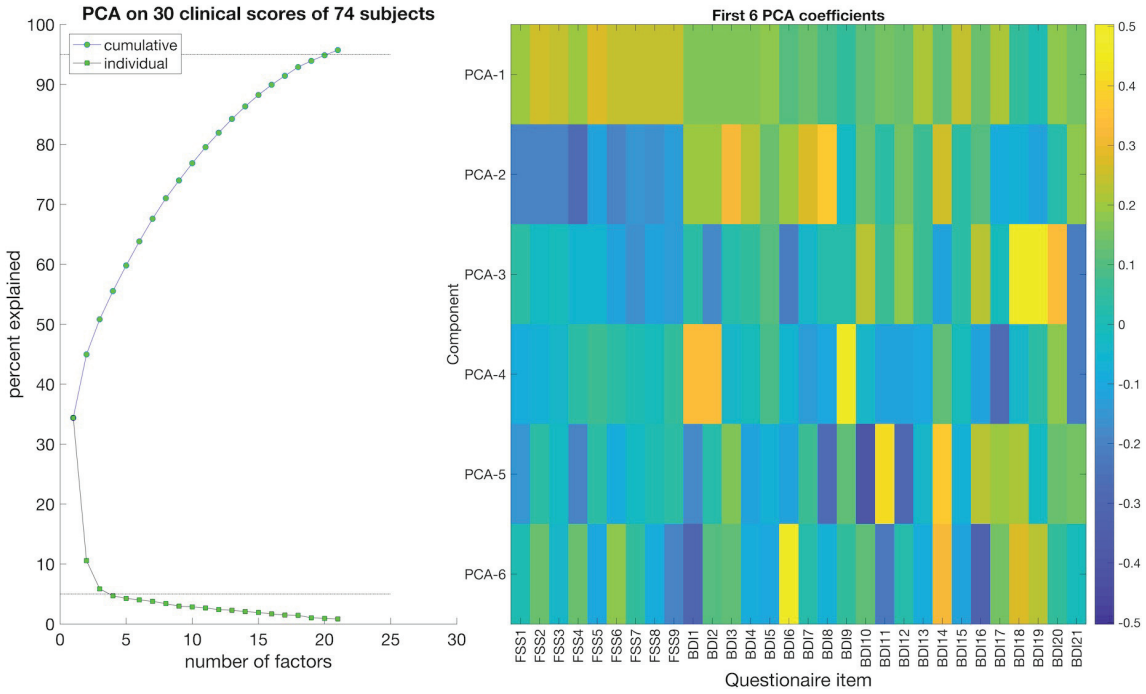

Fig 2. PCA from FSS and BDI subscores. PCA based on 30 clinical subscores (nine FSS and 21 BDI) for all participants. Left: The cumulative and individual explained variance of each PCA of the total variation in the clinical subscores. Right: A heatmap showing the first six PCA factors and their item loading on each component. Yellow and green boxes indicate association with high scores, while the blue boxes indicate association with low scores. The first PCA component (PCA1) captures common variance across BDI and FSS, while the second PCA component (PCA2) captures a pattern of covarying low FSS with high BDI scores.

https://doi.org/10.1371/journal.pone.0210375.g002

subscores (Fig 2). The subject scores for the two highest ranked PCA components were extracted for further analysis to test for associations with DMN connectivity.

Associations between DMN connectivity and clinical PCA scores were investigated using linear models, adjusting for age and sex. To evaluate effect sizes, we calculated Cohen´s f^2, also taking into account age and sex. For Cohen´s f^2 test, effect sizes are considered small (> 0.02), medium (> 0.15) and large (> 0.35). For clinical validation and comparison, we also estimated associations between DMN connectivity and the BDI and FSS continuous sum scores using multiple regression, adjusting for age and sex, and compared extreme groups based on conventional clinical thresholds (see above). To account for disability and cognitive impairment we also investigated the associations from the previously mentioned linear models with SDMT and EDSS scores.

Results

Participant demographics and characteristics

Table 1 summarizes demographic and clinical characteristics of the 74 included MS patients. The majority of the participants were women (70%), mean age was 35.0 years (range 21–49 years). The majority of the participants received disease modifying treatment (DMT), whereas 20% of the participants were never treated. The participants were included on average 14.1 months after the date of diagnosis and disease duration was on average 73.0 months (range 5–272 months).

Table 1. Demographic and clinical characteristics of the participants.

(a) Demographic characteristics	*Patients (n = 74)*
Female, n (%)	52 (70)
Age, mean years (range)	35.0 (21–49)
Education	
Years, mean (range)	14.9 (9–21)
≥ 15 years education n (%)	51 (69)
Working status	
Unemployed or 100% sick leave, n (%)	7 (9)
Working (part- og full-time), student or maternity leave, n (%)	67 (91)
(b) Clinical and MRI characteristics	*Patients (n = 74)*
Neurological disability	
EDSS, mean (range)	2.0 (0–6.0)
Number of total attacks, mean (range)	1.8 (0–5)
DMT	
No DMT, n (%)	15 (20)
Active DMTs, n (%)	48 (65)
Highly active DMTs, n (%)	11 (15)
Months on treatment before study, mean (range)	9.4 (0–34)
Months since diagnosis, mean (range)	14.1 (1–34)
Disease duration, mean months (range)	73.0 (5–272)
Cognitive disability	
SDMT, mean (range)	52.4 (30–80)
MRI	
Brain volume, mean cm^3 (SD, range)	1134.3 (98.2, 925.3–1356.6)
Lesion volume, mean cm^3 (SD, range)	8.58 (4.8, 2.5–26.1)
Lesion load, mean % (SD, range)	0.75 (0.39, 0.24–2.18)
Cortical thickness, mean mm (SD, range)	2.42 (0.09, 2.17–2.62)
(c) Self-reported questionnaires	*Patients (n = 74)*
FSS	
FSS, mean (standard deviation (SD))	4.2 (1.7)
Clinically significant fatigue (FSS mean ≥ 4), n (%)	41 (55)
BDI	
BDI sum, mean (SD)	9.1 (6.7)
Clinically significant depressive symptoms (BDI sum ≥ 14), n (%)	23 (31)
FSS and BDI status	
No fatigue (FSS mean < 4) and no depression (BDI sum < 14), n (%)	32 (43)
Fatigue (FSS mean ≥ 4) and no depression (BDI sum < 14), n (%)	19 (26)
No fatigue (FSS mean < 4) and depression (BDI sum ≥ 14), n (%)	1 (1)
Fatigue (FSS mean ≥ 4) and depression (BDI sum ≥ 14), n (%)	22 (30)

EDSS, Expanded Disability Status Scale; DMT, disease modifying treatment; SDMT, symbol digits modalities test; FSS, Fatigue Severity Scale; BDI, Beck Depression Inventory

https://doi.org/10.1371/journal.pone.0210375.t001

Fifty-five percent of all participants had clinically significant fatigue based on the FSS mean scores (FSS ≥ 4), and 31% of all participants had clinically significant depressive symptoms based on BDI sum scores (BDI > 14). There were no significant differences in FSS and BDI scores between patients with and without rs-fMRI. The first PCA component (PCA1), which reflected common variance across depression and fatigue (high FSS and BDI scores), explained

34% of the total variance in all FSS and BDI items (Fig 2). The second PCA component (PCA2), which reflected a characteristic pattern of low FSS with high BDI scores, explained 10% of the total variance in all FSS and BDI subscores (Fig 2).

Associations between clinical scores and DMN connectivity

Linear models revealed a significant positive correlation between PCA1 and DMN connectivity with small effect size (Cohen's $f^2 = 0.075$, t = 2.17, $p = 0.03$), indicating higher DMN connectivity with higher symptom burden. PCA2, which reflected a characteristic pattern of low FSS scores with high BDI scores, showed a significant negative correlation with DMN connectivity with small effect size (Cohen's $f^2 = 0.067$, t = -2.1, $p = 0.04$) (Fig 1). Linear models revealed a significant positive correlation between FSS continuous mean scores correlated with DMN connectivity (t = 3.1, $p = 0.003$), and a non-significant positive association for BDI continuous sum scores correlated with DMN connectivity (t = 0.8, $p = 0.39$).

Confounding effects in the cohort

Symbol digits modalities test. SDMT showed no significant association with DMN connectivity (t = 1.7, p = 0.09). The positive association between PCA1 and DMN connectivity remained significant (t = 3.0, p = 0.0045) when including SDMT in the model. The same model revealed a positive association between DMN connectivity and SDMT (t = 2.6, p = 0.011). The association between PCA2 and DMN became non-significant (t = -1.9, p = 0.061) when including SDMT in the model. The same model revealed a non-significant positive association between DMN connectivity and SDMT (t = 1.6, p = 0.12).

Expanded disability status scale. EDSS showed no significant association with DMN connectivity (t = 0.3, p = 0.77). The positive association between PCA1 and DMN connectivity remained significant (t = 2.2, p = 0.031) when including EDSS in the model. The same model showed a non-significant association between DMN connectivity and EDSS (t = -0.51, p = 0.61). The negative association between PCA2 and DMN connectivity remained significant (t = -2.0, p = 0.049) when including EDSS in the model. The same model revealed a non-significant positive association between DMN connectivity and EDSS (t = 0.25, p = 0.81).

Disease modifying treatment. DMT level showed a weak negative association with DMN connectivity (t = -1.8, p = 0.07). The positive association between PCA1 and DMN was reduced (t = 1.9, p = 0.06) when including DMT level in the model. The same model showed no association between DMN connectivity and DMT (t = -1.6, p = 0.12). The negative association between PCA2 and DMN connectivity remained significant (t = -2.1, p = 0.035) when including DMT level in the model. The same model revealed a weak negative association between DMN connectivity and DMT level (t = -2.0, p = 0.05).

Lesion volume. Lesion volume showed no significant association with DMN connectivity (t = -1.1, p = 0.27). The positive association between PCA1 and DMN connectivity remained significant (t = 2.1, p = 0.04) when including lesion volume in the model. The same model showed a non-significant association between DMN connectivity and lesion volume (t = -1.1, p = 0.30). The negative association between PCA2 and DMN connectivity was reduced (t = -1.8, p = 0.07) when including lesion volume in the model. The same model revealed no association between DMN connectivity and lesion volume (t = -0.65, p = 0.52).

Brain volume. Brain volume showed no significant association with DMN connectivity (t = -0.14, p = 0.89). The positive association between PCA1 and DMN connectivity remained significant (t = 2.2, p = 0.03) when including brain volume in the model. The same model showed no association between DMN connectivity and brain volume (t = 0.57, p = 0.57). The negative association between PCA2 and DMN connectivity remained significant (t = -2.0,

p = 0.05) when including brain volume in the model. The same model revealed a weak association between DMN connectivity and brain volume (t = -0.28, p = 0.78).

Cortical thickness. Mean cortical thickness across the two hemispheres showed no significant association with DMN connectivity (t = 1.1, p = 0.29). The positive association between PCA1 and DMN connectivity remained significant (t = 2.4, p = 0.02) when including cortical thickness in the model. The same model showed no association between DMN connectivity and cortical thickness (t = 1.5, p = 0.15). The negative association between PCA2 and DMN connectivity was reduced (t = -1.8, p = 0.07) when including cortical thickness in the model. The same model revealed no association between DMN connectivity and cortical thickness (t = 0.72, p = 0.47).

Associations with PCA1, PCA2, FSS and BDI sum scores. Linear models with PCA1 revealed significant associations between brain volume (t = -2.7, p = 0.01), EDSS (t = 3.1, p = 0.003) and SDMT (t = -2.5, p = 0.02), while not between lesion volume (t = -0.3, p = 0.74), DMT (t = -1.3, p = 0.21) and cortical thickness (t = -1.2, p = 0.25).

Linear models with PCA2 showed significant associations between lesion volume (t = 2.0, p = 0.05), but not between brain volume (t = -0.5, p = 0.61), DMT (t = -0.3, p = 0.75), EDSS (t = -0.2, p = 0.82), SDMT (t = -0.7, p = 0.49) or cortical thickness (t = -1.6, p = 0.12).

Linear models with FSS revealed significant associations with EDSS (t = 3.2, p = 0.002), SDMT (t = -2.1, p = 0.04) and brain volume (t = -2.2, p = 0.03), but not for DMT (t = -1.2, p = 0.25), lesion volume (t = -1.1, p = 0.27) and cortical thickness (t = -0.4, p = 0.72).

Linear models with BDI revealed significant associations with EDSS (t = 2.3, p = 0.02), SDMT (t = -2.2, p = 0.03) and brain volume (t = -2.7, p = 0.009), but not for DMT (t = -1.2, p = 0.22), lesion volume (t = 0.79, p = 0.44) and cortical thickness (t = -1.8, p = 0.07).

Discussion

To understand the variability and mechanisms of fatigue and depression is a key clinical question in MS. This study is to our knowledge among the first to study the complex interaction of fatigue and depression in patients with MS by multivariate decomposition analyses of these symptoms in relation to DMN connectivity measured by rs-fMRI.

Fatigue and depression represent common and strong predictors for quality of life in patients with MS, yet the pathophysiological mechanisms of fatigue and depression in MS patients are poorly understood. Converging lines of evidence have suggested associations between different symptoms (such as fatigue, cognitive impairment, depression) and the organization and synchronization of large-scale brain networks as measured by fMRI [3]. Here, using multivariate decomposition of symptoms scores and rs-fMRI data we report significant associations between DMN connectivity and both common and unique symptoms of depression and fatigue in patients with MS. The symptoms presenting in MS patients vary between individuals and is assumed to result primarily from demyelination and microscopic CNS tissue damage [3]. Structural MRI studies have found diverse regional correlates with different MS symptoms [9–11]. Our results show correlation between DMN FC and FSS and BDI scores in MS, which support and further adds to previous knowledge.

One third of the participants in our study had both fatigue and depression, in line with other studies of MS patients [9]. It is important to underline, that in this study, as in most MS papers, depressive symptoms are evaluated by self-reported psychometric scales, and no formal diagnosis of depressive mood disorder has been made [5]. Some previous studies have excluded MS patients with depressive symptoms when investigating the associations between symptoms of fatigue and FC changes [7, 11, 12], while a diffusion tensor imaging study analysed MS patients in subgroups based on the presence of depressive symptoms and fatigue [8].

Here, we wanted to disentangle the complex interaction between symptoms of depression and fatigue by multivariate decomposition analyses, enabling a novel approach in the study of fatigue and depression in MS.

We found a significant positive correlation between DMN connectivity and the burden of fatigue and depression (PCA1 in Fig 1). DMN hyperconnectivity has been demonstrated in depression [39]. A recent study investigated FC changes in MS patients with depression and suggested a functional link between depression and cognitive impairment [18]. A functional link between depression and Alzheimer's disease has also been reported [40]. The same study proposed that depression in MS patients is a result of the demyelination and microscopic CNS tissue damage itself, and not a secondary symptom [18]. A study on primary and secondary progressive MS patients found associations between cognitive impairment and reduction in resting state connectivity [41]. Our findings support the hypothesis that symptoms of depression and fatigue are associated with altered DMN connectivity in MS, possibly influencing the normal function of the DMN as a critical hub of integration and flow of information.

We found that the second PCA component (PCA2) reflecting low burden of fatigue and a high burden of depressive symptoms was negatively correlated with DMN connectivity, indicating that the clinical presentation of fatigue with no depression was associated with DMN hyperconnectivity. DMN hyperconnectivity in fatigue has been demonstrated in a group of breast cancer survivors, where enhanced intrinsic DMN connectivity with the frontal gyrus was associated with persistent fatigue after completed treatment [42]. Our results indicate hyperconnectivity in fatigued MS patients unrelated to depression, possibly caused by the inflammation or structural damage in the brain. Our findings of different DMN patterns depending on the symptom burden of fatigue and depression, may reflect the heterogeneity of symptoms in MS patients, as also reported in a recent review [4]. It has also been reported that fatigue in MS patients, in the absence of depressive symptoms, may be driven by rs-FC changes in the DMN [12]. This study also uncovered that unique components of the DMN was associated with different FC changes. Such regional DMN analyses were beyond the scope of our study.

When adjusting our findings for cognitive impairment, the positive correlation of the first PCA component with DMN connectivity increased while the negative correlation with the second PCA component were slightly decreased. Disability did not have a confounding effect on the correlation between the PCA components and DMN connectivity. Yet we found a significant positive correlation between both BDI and FSS and EDSS, indicating higher disability with higher symptoms of fatigue and depression. Patients with more effective DMTs showed a trend towards decreased symptom burden of fatigue and depression. Adjusting our findings for DMT level weakened our results with PCA1 and DMN connectivity, while the results between PCA2 and DMN connectivity remained significant. Furthermore, adjusting for cognitive impairment seemed to only strengthen our results, while when adjusting for disability our results remained the same. Including whole-brain volume and cortical thickness, which are sensitive indices of brain morphometry, in the analyses did not affect the correlations between the PCA components and DMN connectivity. Lower brain volume was associated with higher scores of both FSS and BDI. When we controlled for lesion volume in our analyses between the PCA components and DMN connectivity, the second PCA component was reduced, while the first PCA component remained significant. Lesion volume was not associated with neither FSS or BDI scores.

Our sample size is modest, but the participants were very thoroughly characterized and comprise a relatively homogenous group in terms of age, cognitive and physical disability, disease duration, education and clinical course. Concerning fatigue, the participants in our study scored a mean of 4.2 for FSS, which is lower than reported in some larger studies [43].

However, the FSS scores for the participants included in this study were in line with a recent Norwegian MS study [6]. Fatigue may impair the quality of life and contribute to the establishment and maintenance of depressive symptoms [4]. The mean BDI sum score in our dataset was 9.1, which is lower than reported in some studies [5], but comparable with a Swedish study [44]. Possible reasons for the relatively low BDI sum score in our sample include the low age, newly diagnosed RRMS, short disease duration and few brain lesions in our MS patients [21].

Adjusting our results for whole-brain volume, cortical thickness, DMT level, lesion volume, SDMT or EDSS did not alter our observed associations significantly. A more detailed analysis of structural MRI and rs-fMRI data could give further insights into the pathophysiology of depression and fatigue in MS. The associations between the two most prominent PCA components and DMN connectivity identified by rs-fMRI in our study suggest separate underlying alterations in the functional connectome. Previous studies assessing cortical morphometry in an overlapping patient sample reported regional associations between cortical surface areas and several clinical manifestations, where the most prominent structural association were smaller cortical surface area and volume significantly associated with depressive symptoms [21].

In addition to our modest sample size, other limitations should be considered when interpreting our results. We did not include lesion filling as part of the fMRI analysis pipeline, but have included both lesion volume and brain volume (after lesion filling) in our analyses to account for confounding effects. In MS patients, permanent damage affects the white matter of the CNS and can cause disconnection syndromes [3]. The FC and large-scale networks depend on structural connections, and inter-individual variability in DMN connectivity, and its association with clinical traits, might be mediated by degree of demyelination, atrophy of both the grey and white matter and microscopic CNS damage [17]. The lack of healthy controls in our study does not allow us to test for specificity, i.e. to which degree any associations between brain connectivity and clinical symptoms generalize to other groups. Yet, our results only focus on the DMN connectivity changes in relation to neuropsychiatric symptoms within the MS group. Future studies are needed to test if our results can be generalized to other populations.

Conclusion

In conclusion, multivariate decomposition of FSS and BDI symptom data supported that the clinical manifestations of fatigue and depression in patients with MS reflect both overlapping and unique variability in the FSS and BDI subscores. The observed differential correlations between symptoms of fatigue and depression and DMN connectivity underline the heterogeneity and complexity of fatigue and depression in MS. Our analyses revealed that high burden of both fatigue and depression was associated with DMN hyperconnectivity, while we also found hyperconnectivity in DMN to be associated with high burden of fatigue in absence of depression. Effect sizes were in general relatively small, and further investigations into the mechanisms of fatigue and depression in MS are warranted. Multivariate decomposition analyses of MS symptoms in relation to default mode network (DMN) connectivity measured by resting-state-fMRI (rs-fMRI) is a promising method to pursue these questions.

Supporting information

S1 File.
(DOCX)

Acknowledgments

We thank all the patients who participated in our study.

Author Contributions

Conceptualization: Einar August Høgestøl, Gro Owren Nygaard, Dag Alnæs, Lars T. Westlye, Hanne Flinstad Harbo.

Data curation: Gro Owren Nygaard, Mona K. Beyer.

Formal analysis: Einar August Høgestøl, Dag Alnæs.

Funding acquisition: Hanne Flinstad Harbo.

Investigation: Einar August Høgestøl, Gro Owren Nygaard.

Methodology: Einar August Høgestøl, Dag Alnæs, Lars T. Westlye.

Project administration: Einar August Høgestøl, Mona K. Beyer, Hanne Flinstad Harbo.

Supervision: Gro Owren Nygaard, Mona K. Beyer, Lars T. Westlye, Hanne Flinstad Harbo.

Validation: Einar August Høgestøl.

Visualization: Einar August Høgestøl, Dag Alnæs.

Writing – original draft: Einar August Høgestøl.

Writing – review & editing: Einar August Høgestøl, Gro Owren Nygaard, Dag Alnæs, Mona K. Beyer, Lars T. Westlye, Hanne Flinstad Harbo.

References

1. Goodin DS. The epidemiology of multiple sclerosis: insights to disease pathogenesis. Handb Clin Neurol. 2014; 122:231–66. https://doi.org/10.1016/B978-0-444-52001-2.00010-8 PMID: 24507521.
2. Janardhan V, Bakshi R. Quality of life in patients with multiple sclerosis: the impact of fatigue and depression. J Neurol Sci. 2002; 205(1):51–8. Epub 2002/11/01. PMID: 12409184.
3. Filippi M, Preziosa P, Rocca MA. Brain mapping in multiple sclerosis: Lessons learned about the human brain. Neuroimage. 2017. https://doi.org/10.1016/j.neuroimage.2017.09.021 PMID: 28917696.
4. Penner IK, Paul F. Fatigue as a symptom or comorbidity of neurological diseases. Nat Rev Neurol. 2017. https://doi.org/10.1038/nrneurol.2017.117 PMID: 29027539.
5. Feinstein A, Magalhaes S, Richard JF, Audet B, Moore C. The link between multiple sclerosis and depression. Nat Rev Neurol. 2014; 10(9):507–17. https://doi.org/10.1038/nrneurol.2014.139 PMID: 25112509.
6. Lerdal A, Celius EG, Krupp L, Dahl AA. A prospective study of patterns of fatigue in multiple sclerosis. Eur J Neurol. 2007; 14(12):1338–43. Epub 2007/10/02. https://doi.org/10.1111/j.1468-1331.2007.01974.x PMID: 17903208.
7. Bisecco A, Caiazzo G, d'Ambrosio A, Sacco R, Bonavita S, Docimo R, et al. Fatigue in multiple sclerosis: The contribution of occult white matter damage. Mult Scler. 2016; 22(13):1676–84. https://doi.org/10.1177/1352458516628331 PMID: 26846989.
8. Gobbi C, Rocca MA, Pagani E, Riccitelli GC, Pravata E, Radaelli M, et al. Forceps minor damage and co-occurrence of depression and fatigue in multiple sclerosis. Mult Scler. 2014; 20(12):1633–40. Epub 2014/04/18. https://doi.org/10.1177/1352458514530022 PMID: 24740370.
9. Hanken K, Eling P, Klein J, Klaene E, Hildebrandt H. Different cortical underpinnings for fatigue and depression in MS? Mult Scler Relat Disord. 2016; 6:81–6. https://doi.org/10.1016/j.msard.2016.02.005 PMID: 27063629.
10. Cruz Gomez AJ, Ventura Campos N, Belenguer A, Avila C, Forn C. Regional brain atrophy and functional connectivity changes related to fatigue in multiple sclerosis. PLoS One. 2013; 8(10):e77914. Epub 2013/10/30. https://doi.org/10.1371/journal.pone.0077914 PMID: 24167590; PubMed Central PMCID: PMCPMC3805520.

11. Rocca MA, Parisi L, Pagani E, Copetti M, Rodegher M, Colombo B, et al. Regional but not global brain damage contributes to fatigue in multiple sclerosis. Radiology. 2014; 273(2):511–20. Epub 2014/06/14. https://doi.org/10.1148/radiol.14140417 PMID: 24927473.

12. Bisecco A, Nardo FD, Docimo R, Caiazzo G, d'Ambrosio A, Bonavita S, et al. Fatigue in multiple sclerosis: The contribution of resting-state functional connectivity reorganization. Mult Scler. 2018; 24 (13):1696–705. Epub 2017/09/16. https://doi.org/10.1177/1352458517730932 PMID: 28911257.

13. Cruz-Gomez AJ, Ventura-Campos N, Belenguer A, Avila C, Forn C. The link between resting-state functional connectivity and cognition in MS patients. Mult Scler. 2014; 20(3):338–48. Epub 2013/07/06. https://doi.org/10.1177/1352458513495584 PMID: 23828871.

14. Basile B, Castelli M, Monteleone F, Nocentini U, Caltagirone C, Centonze D, et al. Functional connectivity changes within specific networks parallel the clinical evolution of multiple sclerosis. Mult Scler. 2014; 20(8):1050–7. Epub 2013/12/12. https://doi.org/10.1177/1352458513515082 PMID: 24326671.

15. Sbardella E, Petsas N, Tona F, Pantano P. Resting-State fMRI in MS: General Concepts and Brief Overview of Its Application. Biomed Res Int. 2015; 2015:212693. https://doi.org/10.1155/2015/212693 PMID: 26413509; PubMed Central PMCID: PMCPMC4564590.

16. Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci. 2008; 1124:1–38. Epub 2008/04/11. https://doi.org/10.1196/annals.1440.011 PMID: 18400922.

17. Rocca MA, Valsasina P, Martinelli V, Misci P, Falini A, Comi G, et al. Large-scale neuronal network dysfunction in relapsing-remitting multiple sclerosis. Neurology. 2012; 79(14):1449–57. Epub 2012/09/08. https://doi.org/10.1212/WNL.0b013e31826d5f10 PMID: 22955126.

18. Bonavita S, Sacco R, Esposito S, d'Ambrosio A, Della Corte M, Corbo D, et al. Default mode network changes in multiple sclerosis: a link between depression and cognitive impairment? Eur J Neurol. 2017; 24(1):27–36. Epub 2016/09/17. https://doi.org/10.1111/ene.13112 PMID: 27633185.

19. Hidalgo de la Cruz M, d'Ambrosio A, Valsasina P, Pagani E, Colombo B, Rodegher M, et al. Abnormal functional connectivity of thalamic sub-regions contributes to fatigue in multiple sclerosis. Mult Scler. 2017:1352458517717807. https://doi.org/10.1177/1352458517717807 PMID: 28657428.

20. Finke C, Schlichting J, Papazoglou S, Scheel M, Freing A, Soemmer C, et al. Altered basal ganglia functional connectivity in multiple sclerosis patients with fatigue. Mult Scler. 2015; 21(7):925–34. Epub 2014/11/14. https://doi.org/10.1177/1352458514555784 PMID: 25392321.

21. Nygaard GO, Celius EG, de Rodez Benavent SA, Sowa P, Gustavsen MW, Fjell AM, et al. A Longitudinal Study of Disability, Cognition and Gray Matter Atrophy in Early Multiple Sclerosis Patients According to Evidence of Disease Activity. PLoS One. 2015; 10(8):e0135974. Epub 2015/08/19. https://doi.org/10.1371/journal.pone.0135974 PMID: 26280173; PubMed Central PMCID: PMCPMC4539191.

22. Nygaard GO, Walhovd KB, Sowa P, Chepkoech JL, Bjornerud A, Due-Tonnessen P, et al. Cortical thickness and surface area relate to specific symptoms in early relapsing-remitting multiple sclerosis. Mult Scler. 2015; 21(4):402–14. Epub 2014/08/21. https://doi.org/10.1177/1352458514543811 PMID: 25139946.

23. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol. 2011; 69(2):292–302. Epub 2011/03/10. https://doi.org/10.1002/ana.22366 PMID: 21387374; PubMed Central PMCID: PMCPMC3084507.

24. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol. 1989; 46(10):1121–3. Epub 1989/10/01. PMID: 2803071.

25. Beck AT SR, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: The Psychological Corporation, 1996.

26. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004; 23 Suppl 1:S208–19. Epub 2004/10/27. https://doi.org/10.1016/j.neuroimage.2004.07.051 PMID: 15501092.

27. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. Neuroimage. 2012; 62 (2):782–90. Epub 2011/10/08. https://doi.org/10.1016/j.neuroimage.2011.09.015 PMID: 21979382.

28. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17(2):825–41. Epub 2002/10/16. PMID: 12377157.

29. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp. 2002; 17(3):143–55. Epub 2002/10/23. https://doi.org/10.1002/hbm.10062 PMID: 12391568.

30. Smith SM, Brady JM. SUSAN—A New Approach to Low Level Image Processing. International Journal of Computer Vision. 1997; 23(1):45–78. https://doi.org/10.1023/a:1007963824710

31. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage. 1999; 9(2):179–94. Epub 1999/02/05. https://doi.org/10.1006/nimg.1998.0395 PMID: 9931268.

32. Beckmann CF, DeLuca M, Devlin JT, Smith SM. Investigations into resting-state connectivity using independent component analysis. Philos Trans R Soc Lond B Biol Sci. 2005; 360(1457):1001–13. Epub 2005/08/10. https://doi.org/10.1098/rstb.2005.1634 PMID: 16087444; PubMed Central PMCID: PMCPMC1854918.

33. Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, et al. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage. 2014; 95:232–47. Epub 2014/03/25. https://doi.org/10.1016/j.neuroimage.2014.03.034 PMID: 24657355; PubMed Central PMCID: PMCPMC4154346.

34. Pruim RH, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage. 2015; 112:267–77. Epub 2015/03/17. https://doi.org/10.1016/j.neuroimage.2015.02.064 PMID: 25770991.

35. Nickerson LD, Smith SM, Ongur D, Beckmann CF. Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses. Front Neurosci. 2017; 11:115. https://doi.org/10.3389/fnins.2017.00115 PMID: 28348512; PubMed Central PMCID: PMCPMC5346569.

36. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage. 2012; 61(4):1402–18. Epub 2012/03/21. https://doi.org/10.1016/j.neuroimage.2012.02.084 PMID: 22430496; PubMed Central PMCID: PMCPMC3389460.

37. Damangir S, Manzouri A, Oppedal K, Carlsson S, Firbank MJ, Sonnesyn H, et al. Multispectral MRI segmentation of age related white matter changes using a cascade of support vector machines. J Neurol Sci. 2012; 322(1–2):211–6. https://doi.org/10.1016/j.jns.2012.07.064 PMID: 22921728.

38. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.

39. Kaiser RH, Andrews-Hanna JR, Wager TD, Pizzagalli DA. Large-Scale Network Dysfunction in Major Depressive Disorder: A Meta-analysis of Resting-State Functional Connectivity. JAMA Psychiatry. 2015; 72(6):603–11. Epub 2015/03/19. https://doi.org/10.1001/jamapsychiatry.2015.0071 PMID: 25785575; PubMed Central PMCID: PMCPMC4456260.

40. Damoiseaux JS, Prater KE, Miller BL, Greicius MD. Functional connectivity tracks clinical deterioration in Alzheimer's disease. Neurobiol Aging. 2012; 33(4):828 e19–30. https://doi.org/10.1016/j.neurobiolaging.2011.06.024 PMID: 21840627; PubMed Central PMCID: PMCPMC3218226.

41. Rocca MA, Valsasina P, Absinta M, Riccitelli G, Rodegher ME, Misci P, et al. Default-mode network dysfunction and cognitive impairment in progressive MS. Neurology. 2010; 74(16):1252–9. Epub 2010/04/21. https://doi.org/10.1212/WNL.0b013e3181d9ed91 PMID: 20404306.

42. Hampson JP, Zick SM, Khabir T, Wright BD, Harris RE. Altered resting brain connectivity in persistent cancer related fatigue. Neuroimage Clin. 2015; 8:305–13. https://doi.org/10.1016/j.nicl.2015.04.022 PMID: 26106555; PubMed Central PMCID: PMCPMC4474178.

43. The Goldman Consensus statement on depression in multiple sclerosis. Mult Scler. 2005; 11(3):328–37. Epub 2005/06/17. https://doi.org/10.1191/1352458505ms1162oa PMID: 15957516.

44. Sundgren M, Maurex L, Wahlin A, Piehl F, Brismar T. Cognitive impairment has a strong relation to non-somatic symptoms of depression in relapsing-remitting multiple sclerosis. Arch Clin Neuropsychol. 2013; 28(2):144–55. https://doi.org/10.1093/arclin/acs113 PMID: 23291310.

II

frontiers in Neurology

ORIGINAL RESEARCH
published: 30 April 2019
doi: 10.3389/fneur.2019.00450

Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis

Einar A. Høgestøl[1], Tobias Kaufmann[2], Gro O. Nygaard[3], Mona K. Beyer[1,4], Piotr Sowa[4], Jan E. Nordvik[5], Knut Kolskår[2,6,7], Geneviève Richard[2,6,7], Ole A. Andreassen[2], Hanne F. Harbo[1,3] and Lars T. Westlye[2,7]*

[1] Institute of Clinical Medicine, University of Oslo, Oslo, Norway, [2] NORMENT, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway, [3] Department of Neurology, Oslo University Hospital, Oslo, Norway, [4] Division of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway, [5] Catosenteret Rehabilitation Centre, Son, Norway, [6] Sunnaas Rehabilitation Hospital HT, Nesodden, Norway, [7] Department of Psychology, University of Oslo, Oslo, Norway

OPEN ACCESS

Edited by:
Jorge Matias-Guiu, Complutense University of Madrid, Spain

Reviewed by:
Robert Günther Marschallinger, University of Salzburg, Austria
Franca Wagner, Bern University Hospital, Switzerland
Ambrosio Miralles, Universidad Europea de Madrid, Spain

***Correspondence:**
Einar A. Høgestøl
einar.august@gmail.com

Specialty section:
This article was submitted to Multiple Sclerosis and Neuroimmunology, a section of the journal Frontiers in Neurology

Received: *20 December 2018*
Accepted: *12 April 2019*
Published: *30 April 2019*

Citation:
Høgestøl EA, Kaufmann T, Nygaard GO, Beyer MK, Sowa P, Nordvik JE, Kolskår K, Richard G, Andreassen OA, Harbo HF and Westlye LT (2019) Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis. Front. Neurol. 10:450. doi: 10.3389/fneur.2019.00450

Multiple sclerosis (MS) is an inflammatory disorder of the central nervous system. By combining longitudinal MRI-based brain morphometry and brain age estimation using machine learning, we tested the hypothesis that MS patients have higher brain age relative to chronological age than healthy controls (HC) and that longitudinal rate of brain aging in MS patients is associated with clinical course and severity. Seventy-six MS patients [71% females, mean age 34.8 years (range 21–49) at inclusion] were examined with brain MRI at three time points with a mean total follow up period of 4.4 years (±0.4 years). We used additional cross-sectional MRI data from 235 HC for case-control comparison. We applied a machine learning model trained on an independent set of 3,208 HC to estimate individual brain age and to calculate the difference between estimated and chronological age, termed brain age gap (BAG). We also assessed the longitudinal change rate in BAG in individuals with MS. MS patients showed significantly higher BAG (4.4 ± 6.6 years) compared to HC (Cohen's D = 0.69, $p = 4.0 \times 10^{-6}$). Longitudinal estimates of BAG in MS patients showed high reliability and suggested an accelerated rate of brain aging corresponding to an annual increase of 0.41 (SE = 0.15) years compared to chronological aging ($p = 0.008$). Multiple regression analyses revealed higher rate of brain aging in patients with more brain atrophy (Cohen's D = 0.86, $p = 4.3 \times 10^{-15}$) and increased white matter lesion load (WMLL) (Cohen's D = 0.55, $p = 0.015$). On average, patients with MS had significantly higher BAG compared to HC. Progressive brain aging in patients with MS was related to brain atrophy and increased WMLL. No significant clinical associations were found in our sample, future studies are warranted on this matter. Brain age estimation is a promising method for evaluation of subtle brain changes in MS, which is important for predicting clinical outcome and guide choice of intervention.

Keywords: multiple sclerosis, brain age, magnetic resonance imaging, machine learning, longitudinal

INTRODUCTION

Multiple sclerosis (MS) is an inflammatory, demyelinating disease of the CNS. The pathophysiology of MS can be divided into acute inflammation during a relapse and chronic inflammation thought to continuously perturb neuroaxonal homeostasis and drive neurodegeneration (1). Development of robust brain imaging markers that can parse between-subject heterogeneity of the clinical trajectories, predict future progression of disability, and monitor the effects of treatment for MS patients, is a major aim with important clinical implications (2, 3). Current imaging markers with relevance for MS are associated with disease activity and progression, and include, among other features, number or volume of hyperintense brain lesions visible on T2-weighted MRI images, contrast-enhancing T1 lesions, increased annual brain volume loss and T1-hypointense "black holes" (2, 4, 5). Increased rate of total brain volume loss, which is best captured using longitudinal designs (6), reflects accelerated neurodegeneration (7), and regional analyses may boost the correlations between estimated brain atrophy and disability (2).

However, identifying robust associations between clinical outcomes and MRI measures has been challenging (8). This clinico-radiological paradox in MS is likely explained by a combination of lack of sensitivity and specificity both in the clinical and imaging domain. Brain age estimation uses machine learning to train a model that can accurately predict the individual age from brain imaging data (9–11). Utilizing sensitive measures of MRI-based brain morphometry, brain age estimation provides a robust imaging-based biomarker with potential to yield novel insights into similarities and differences of disease pathophysiology across brain disorders (11, 12). Such imaging-based brain age has been shown to be reliable both within and between MRI scanners, and is a candidate biomarker of an individual's brain health and integrity (10–12). Different approaches to brain age estimation exploit information from a variety of brain regions (e.g., hippocampus, subcortical, gray matter, and white matter) or MRI sequences (e.g., T1, T2, diffusion tensor imaging and functional MRI) to inform the model (12). An older appearing brain, which is related to advanced physiological and cognitive aging and mortality (12, 13), has been found across several brain disorders, and region specific brain age patterns in patient cohorts have shown potential differential genetic effects, including genetic pleiotropy between global brain age and MS (11). To our knowledge, only two preprint manuscripts (11, 14) and one abstract (15) have reported brain age estimations in MS, and all reported older appearing brains in patients with MS compared to HC.

Here, combining cross-sectional and sensitive measures of MRI-based regional and global brain morphometry in MS and HC (cross-sectional only), we tested the hypothesis that MS patients have higher brain age than HC. Next, using longitudinal MRI data in MS patients we tested the hypothesis that brain aging accelerates in MS and that the rate of acceleration is associated with a more severe clinical outcome.

Abbreviations: BAG, Brain Age Gap; Cereb., Cerebellar; DMT, Disease Modifying Treatment; EDA, Evidence of Disease Activity; HC, Healthy Controls; ICC, Intraclass Correlation Coefficient; MS, Multiple Sclerosis; MSSS, Multiple Sclerosis Severity Scale; NEDA, No Evidence of Disease Activity; Subcort., Subcortical; WMLL, White Matter Lesion Load.

MATERIALS AND METHODS

Participants

We recruited 76 MS patients at Oslo University Hospital (16, 17). All patients were diagnosed with MS between January 2009 and December 2012 according to the revised McDonald Criteria (18) and were enrolled in the study on average 14 months (±11.8) after the date of diagnosis (time point 1). Exclusion criteria included age $<$ 18 years or $>$ 50 years, uncertain diagnosis, non-fluency in Norwegian, neurological or psychiatric disease, drug abuse, head trauma, pregnancy, and previous adverse gadolinium reaction. Most patients also participated in two follow-up examinations on average 26 months (±11.7, time point 2, $n = 60$) and 66 months (±13.3, time point 3, $n = 62$) after the date of diagnosis. At each visit, all patients completed a neurological examination by a Neurostatus certified medical doctor (http://www.neurostatus.com) within the same week as their MRI scan. Disease-modifying treatments were categorized into the following groups; 0: no treatment; 1: glatiramer acetate, interferons, teriflunomide, or dimetylfumarate; and 2: fingolimod, natalizumab, or alemtuzumab. Many patients ($n = 58$) were also included in a partly overlapping study with a larger cross-sectional MS group ($n = 254$) (11).

The HC group was recruited through newspaper ads or after a stratified random selection drawn from the Norwegian National Population Registry to two parallel studies (13, 19). Exclusion criteria included estimated IQ (intelligence quotient) $<$70, history of neurologic or psychiatric disease and current medication significantly affecting the nervous system (20).

This study was carried out in accordance with the recommendations of the Regional Committee for Medical and Health Research Ethics with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the South East Regional Committee for Medical and Health Research Ethics.

MRI Acquisition

All MS patients were scanned at up to three time points between January 2012 and August 2017 in a study setting, using the same 1.5 T scanner (Avanto, Siemens Medical Solutions; Erlangen, Germany) equipped with a 12-channel head coil. Structural MRI data were collected using a 3D T1-weighted MPRAGE (Magnetization Prepared Rapid Gradient Echo) sequence, with the following parameters: TR (repetition time)/TE (echo time)/flip angle/voxel size/FOV (field of view)/slices/scan time/matrix/time to inversion = 2,400 ms/3.61 ms/8°/1.20 × 1.25 × 1.25 mm/240/160 sagittal slices/7:42 min/192 × 192/1,000 ms. The MRI sequence was kept identical during the scanning period. FLAIR (Fluid attenuation inversion recovery), T2 and pre- and post-gadolinium 3D T1 sequences were attained and used for neuroradiological evaluation (17).

Fifty-eight of the MS patients were also scanned at Oslo University Hospital on a 3 T GE 750 Discovery MRI scanner with a 32-channel head coil at time point 3 between August 2016 and June 2017 during the same week they were scanned at the 1.5 T scanner for time point 3. HCs were scanned solely on the 3 T scanner at one time point to provide cross-sectional data. Structural MRI data were collected using a 3D high-resolution IR-prepared FSPGR (fast spoiled gradient echo) T1-weighted sequence (3D BRAVO) with the following parameters: TR (repetition time)/TE (echo time)/flip angle/voxel size/FOV (field of view)/slices/scan time = 8.16 ms/3.18 ms/12°/1 × 1 × 1 mm/256 × 256 mm/188 sagittal slices/4:42 min.

MRI Pre- and Post-processing

Using the T1-weighted scans we performed cortical reconstruction and volumetric segmentation with FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/) (21). To extract reliable volume and thickness estimates, images included in the longitudinal 1.5 T MRI dataset were processed with the longitudinal stream in FreeSurfer (22). Specifically an unbiased within-subject template space and image was created using robust, inverse consistent registration (23). Several processing steps, such as skull stripping, Talairach transforms, atlas registration as well as spherical surface maps and parcellations were then initialized with common information from the within-subject template, increasing reliability and power (22).

Manual quality control of the MRI scans from patients was performed by trained research personnel to identify and edit segmentation errors where possible ($n = 43$ MRI scans) and exclude data of insufficient quality ($n = 6$ MRI scans). In addition, eight brain scans were removed due to missing sequences of the 263 MRI scans from MS patients. Lesion filling was performed utilizing automatically generated lesion masks from Cascade (24) with the lesion filling tool (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/lesion_filling) in FSL (25). The lesion masks were assessed by a trained neuroradiologist and normalized to MNI space using FLIRT (26), with the corresponding T1 image as an intermediate. A probabilistic representation of the lesions across all patients is shown in **Supplementary Figure 1**.

Brain Age Estimation Model

The training set for brain age estimation included MRI scans from 3,208 HC >12 years (54% women, mean age 47.5 (±19.8), age range 12–95) obtained from several publicly available datasets (**Supplementary Figure 2**) and processed in the same MRI pipeline.

We trained one machine learning model for each sex to predict brain age following a recent implementation (11). The features were derived from the Human Connectome Project parcellation of the cortex (27), comprising 180 regions of interest per hemisphere for thickness, area, and volume, respectively. In addition, we used subcortical and cerebellar parcellations from Freesurfer. The full set comprised 1,118 features in total. We used extreme gradient boosting, "xgboost" package in R (28), as the main method for our brain age studies as it has been the lead solution on many machine learning competitions in the field and due to our data being highly monotonic. We compared xgboost to shrinkage linear models (https://cran.r-project.org/package=care) and found converging results, although xgboost performed slightly better in our data (**Supplementary Table 1**). We trained one extreme gradient boosting tree machine learning model per sex on the training set to predict age using the 28 brain imaging features (learning rate eta = 0.01, optimal number of rounds determined in a nested cross-validation loop within the training set, other parameters as default). A 10-fold cross-validation confirmed good performance and generalizability in the combined model for females and males (**Supplementary Figure 3**, $r = 0.91$).

Next, for all patients and HC in the test set, we estimated brain age and calculated the brain age gap (BAG, defined as the difference between chronological age and imaging-based brain age). Using linear regressions, we removed any common variance with age, age^2 and sex to account for confounding factors before submitting the residualized version of BAG to further analyses (29). When pooling estimates of BAG from the 1.5 T and 3 T scanners, we adjusted BAG for scanner effect on BAG estimates by extracting the scanner coefficient from a LME (linear mixed effects) model. When comparing BAG between patients and matched HCs we report the actual adjusted difference in BAG between these two groups.

In addition to the estimation of brain age based on features from the whole brain, we also performed brain age estimation of regional subsets of features (11, 13). We used the lobe parcellation labels from Freesurfer (21) to identify features that overlapped with a given lobe and performed similar machine learning procedures sets as described for the whole brain using occipital, frontal, temporal, cingulate, insula, and subcortical/cerebellar features alone, respectively.

Statistical Analyses

We used R (R Core Team, Vienna, 2018) for statistical analyses. All LME models accounted for age, age^2, sex, and scanner (30). We estimated annual change in BAG by dividing the total change in BAG by the relevant time interval. We utilized the longest time interval between time points and excluded MS patients lacking longitudinal data ($n = 8$). A score of 0 indicates that the rate of brain aging corresponds to chronological aging, and positive and negative values correspond to accelerated and decelerated brain aging compared to chronological aging, respectively. For each brain region we tested the relative rate of brain aging on a group level by performing one-sample t-tests on BAG with 0 as test value. We estimated the annual global brain atrophy by comparing estimated total brain volume from the Freesurfer output (BrainSegVolNotVent) between time points. Based on Freesurfer volumetric output, we also compared the volumetric and normalized measurements (divided by estimated total intracranial volume) between MS patients and HC (**Supplementary Table 2**).

To assess reliability of brain age across time we computed the intraclass correlation coefficient (ICC) using the R package "irr" (https://CRAN.R-project.org/package=irr). Figures were made using "ggplot2" (31) and "cowplot" (https://CRAN.R-project.org/package=cowplot) in R. To control for multiple testing we adjusted the p-values using false discovery rate (FDR) (32)

procedures implemented in the R package "p.adjust" (http://stat.ethz.ch/R-manual/R-devel/library/stats/html/p.adjust.html). The LME models were performed using the R package "nlme" (https://CRAN.R-project.org/package=nlme).

RESULTS

Participant Demographics and Characteristics

Table 1 summarizes the demographic and clinical characteristics of all MS patients. Key demographic variables regarding HC are summarized in **Supplementary Table 2**. The majority of the MS patients were women (71%), 96% had relapsing-remitting MS and mean age at inclusion was 34.8 years (±7.2). On average they were examined 1.2, 2.2, and 5.5 years after diagnosis. Most patients used first line treatment; 65, 48, and 37% at time point 1, 2, and 3, respectively. Second line treatments were used by 13, 23, and 32% of the MS patients at time point 1, 2, and 3, respectively. At time point 2 and 3, 53 and 44% of the patients were categorized as having NEDA (No Evidence of Disease Activity)−3 (no clinical progression, no new lesions observable in MRI and no new attacks). At time point 2, 43% of the patients with EDA (Evidence of Disease Activity) had changed their disease modifying treatment (DMT). At time point 3, 77% of the patients with EDA had changed their DMT.

TABLE 1 | Demographic and clinical characteristics of the multiple sclerosis patients.

	Time point 1	Time point 2	Time point 3
(a) Demographic characteristics	*n* = 76	*n* = 75	*n* = 62
Female (%)	54 (71)	54 (72)	44 (71)
Age, mean years (SD)	34.8 (7.2)	35.8 (7.2)	40.0 (7.3)
≥15 years education (%)	53 (70)	NA	50 (81)
Disease duration, mean months (SD)	71.7 (63.0)	79.7 (57.1)	125.1 (60.2)
Age at first symptom, mean years (SD)	29.3 (6.7)		
Months since MS diagnosis, mean (SD)	14.0 (11.8)	26.3 (11.7)	66.2 (13.3)
Positive OCB status (%)	69 (91)		
Disease modifying treatment			
None (%)	17 (22)	22 (29)	19 (31)
First line (%)	49 (65)	36 (48)	23 (37)
Second line (%)	10 (13)	17 (23)	20 (32)
(b) Clinical evaluation			
Multiple sclerosis classification			
RRMS (%)	73 (96)	72 (96)	60 (95)
PPMS (%)	2 (3)	2 (3)	1 (2)
SPMS (%)	1 (1)	1 (1)	2 (3)
Neurological disability			
EDSS, median (SD, range)	2.0 (0.9, 0-6)	2.0 (0.9, 0-4)	2.0 (1.3, 0-6)
MSSS (SD)	4.9 (1.9)	4.5 (2.0)	2.6 (1.8)
Number of total attacks, mean (SD)	1.8 (1.0)	2.0 (1.0)	2.6 (1.3)
Nine hole peg test			
Dominant hand, mean seconds (SD)	20.0 (3.1)	NA	20.6 (8.4)
Non-dominant hand, mean seconds (SD)	20.8 (2.8)	NA	21.1 (5.9)
Timed 25 feet walk test, mean seconds (SD)	4.0 (0.7)	3.9 (0.8)	4.0 (1.1)
(c) NEDA assessment			
NEDA-3 (%)		40 (53)	27 (44)
NEDA-4 (%)		17 (30)	18 (32)

OCB, oligoclonal bands; RRMS, relapsing-remitting multiple sclerosis; PPMS, primary progressive multiple sclerosis; SPMS, secondary progressive multiple sclerosis; EDSS, expanded disability status scale; MSSS, multiple sclerosis severity scale; NEDA, no evidence of disease activity.

Cross-Sectional Case-Control Analyses (3 T)

At time point 3 (3T data) we found significantly higher BAG for the MS group compared to matched HC for all brain regions except the temporal region (**Figure 1**; **Supplementary Table 4**). The most prominent differences in BAG were 4.4 years for global BAG (Cohen's D = 0.69) and 6.2 years for subcortical and cerebellar brain regions (Cohen's D = 0.72).

At time point 3, 58 MS patients underwent one MRI scanning in the 1.5 T and one in the 3 T scanner with 2 days apart. Whereas, absolute estimates of brain age varied between scanners for all brain regions except insula (BAG scanner difference −6.08 to 10.60 years, see **Supplementary Table 3**; **Supplementary Figures 4**), brain age estimates from the two scanners were highly correlated for global BAG and all brain regions (r = 0.67–0.86, $p < 0.001$), supporting the reproducibility.

Volumetric data showed no significant differences in measures of whole brain, gray matter and white matter. When using normalized measurements (divided by estimated total intracranial volume), we found significant differences between normalized whole brain (Cohen's D = 0.45) and gray matter (Cohen's D = 0.46) volumes (**Supplementary Table 2**).

Longitudinal MS Sample (1.5 T)

The correlations between chronological age and global brain age were r = 0.71 for time point 1, r = 0.70 for time point 2, and r = 0.69 for time point 3. After adjusting for scanner effects mean global BAG was 2.8 (±9.0) for time point 1, 3.3 (±9.4) for time point 2, and 4.6 (±9.8) for time point 3 in the longitudinal MS sample (**Supplementary Figures 5A,B**). Some patients exhibited reduced estimates of brain age over time, likely partly explained by an effect of MRI noise characteristics (subject motion, MRI artifacts or any other technical changes between acquisitions), while in the same period the biological changes were negligible.

We found a significant annual increase in global BAG of 0.41 (SE = 0.15) years (p = 0.008) in patients with MS (**Figure 2**; **Supplementary Tables 6, 7**). No regional measures showed significantly decreasing or increasing BAG at the group level (**Supplementary Tables 7, 8**). Our dataset included a low number of other MS phenotypes than RRMS, and we did not find significant correlations between brain atrophy rates or annual change in global BAG among these.

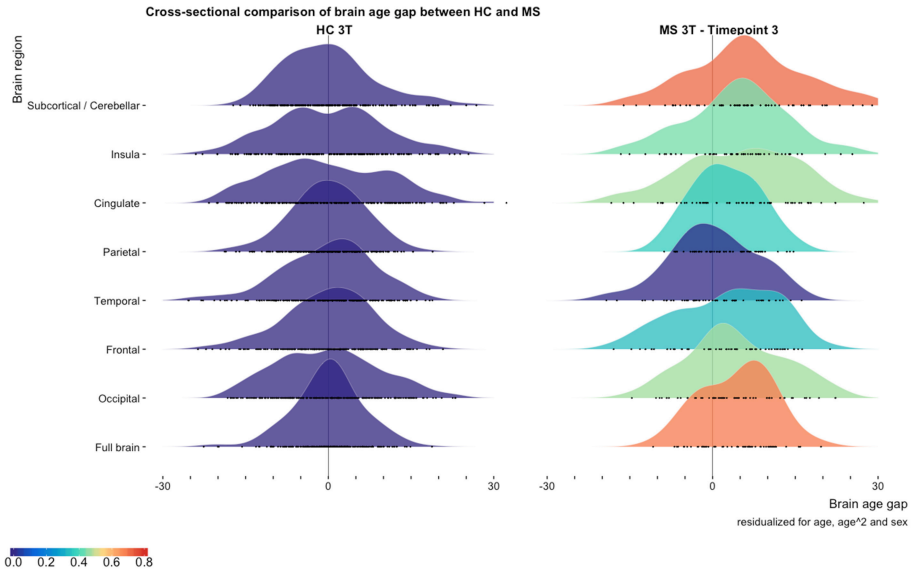

FIGURE 1 | Cross-sectional comparison of brain age gap between multiple sclerosis patients and healthy controls. The distribution of brain age gaps across brain regions based on the cross-sectional 3 T MRI data from matched HC and multiple sclerosis patients at time point 3. We found increased brain age gaps for all brain regions except from the temporal brain region. Brain age gaps are residualized for age, age^2, and sex. Cohen's D effect sizes for the brain age gap between HC and multiple sclerosis patients are depicted using the color bar. All BAG estimates are depicted as black circles on the x-axes.

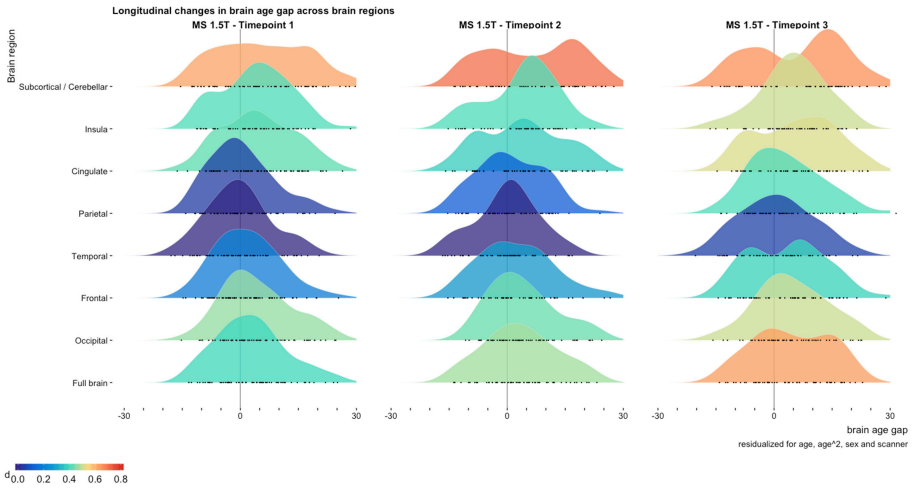

FIGURE 2 | Longitudinal changes in brain age gap across brain regions. The distribution of brain age gaps across brain regions based on the longitudinal 1.5 T MRI sample. Brain age gaps from the MS sample are compared with the cross-sectional 3 T HC sample and residualized for age, age^2, sex, and scanner. The full brain estimates showed a significant accelerated rate of brain aging compared to chronological aging [annual increase in brain age gap 0.41 ($p = 0.008$)]. Cohen's D effect sizes for the brain age gap between MS and HC are depicted using the color bar. All BAG estimates are depicted as black circles on the x-axes.

We found no significant difference in BAG between the raw and the lesion filled MRI scans, and the BAG scores from the two versions were highly correlated BAG ($r = 0.98$). Data processed with the longitudinal stream in FreeSurfer had significantly lower BAG than the cross-sectionally processed MRI scans (mean difference in BAG 4.9 years, $p < 0.001$) and lesion filled MRI scans (difference in BAG 5.1 years, $p \leq 0.001$) (**Supplementary Figures 6, 7**; **Supplementary Tables 4, 5**).

ICCs for all brain regions across all time points varied from 0.79 to 0.94 for residualized BAG and 0.78–0.95 for predicted age. Cerebellar and subcortical brain regions showed highest reliability with an ICC of 0.94 for BAG and 0.95 for predicted age (**Supplementary Table 9**).

Mean annualized estimated change in global brain volume from all three time points. from Freesurfer was −0.30% (SD = 0.53%). ICC for global brain volume was 0.97–0.99. Mean annualized change in WMLL was 504

mm^3 (± 28 mm^3). ICC for WMLL at time point three was 0.93–0.99.

Associations Between Global Brain Age and Clinical Outcomes

Table 2 (BAG) and **Table 3** (annual rate of brain aging) show summary statistics from the multiple regressions testing for associations with demographic, clinical, and MRI variables in the longitudinal MS group (**Supplementary Tables 10, 11**). After accounting for multiple testing, significant associations were found between BAG and WMLL (Cohen's D = −1.23, $p = 3.0 \times 10^{-4}$) and global brain atrophy (Cohen's D = −0.07, $p = 0.01$), respectively, indicating higher BAG at baseline with higher WMLL at time point three and increased brain atrophy over time. Further, changes in BAG over time was significantly associated with brain atrophy over time (Cohen's D = 0.86, $p = 4.3 \times 10^{-15}$) and change in WMLL (Cohen's D = 0.55, $p = 0.015$), indicating higher rates of brain aging in patients with higher levels of brain atrophy and more progressive changes in WMLL. WMLL also showed a significant correlation with BAG for cerebellar and subcortical regions (Cohen's D = −1.23, $p = 3.2 \times 10^{-3}$).

DISCUSSION

Using cross-sectional and longitudinal MRI data as basis for brain age estimation based on machine learning, we tested the hypotheses that patients with MS on average show higher brain age than healthy controls, and that the rate of brain aging is associated with clinical trajectories. Cross-sectional analysis revealed higher brain age gap in patients with MS compared to healthy controls, and longitudinal analysis showed increased rates of brain aging in patients with higher rates of brain atrophy and increasing WMLL.

MS patients had on average 4.4 years higher BAG compared to HC (Cohen's D = 0.68), in line with preliminary findings in a partly overlapping cross-sectional sample (11). To our knowledge, other studies comparable to ours are not yet available, and further studies are warranted. Global brain age differences may disguise relevant regional effects. Indeed, for subcortical and cerebellar brain regions we found a higher BAG in MS compared with HCs (BAG 6.2 years, Cohen's D = 0.72), which was already evident at time point 1 (BAG 5.7 years, Cohen's D = 0.63). The regional variability may reflect differential affinity of MS pathology across the brain, which is also supported by lesion probability maps in MS (2, 33).

In our longitudinal patient sample, the average annual rate of brain aging for global BAG exceeded that of chronological aging by 0.41 years per year ($p = 0.008$). Although further studies are needed, this apparent accelerated aging of the brain may partly be explained by chronic inflammatory processes that drive neurodegeneration in MS (1).

As expected, we found relatively robust associations between brain atrophy and brain aging ($r = 0.79$, $p = 4.3 \times 10^{-15}$). Of notice, regional brain aging and BAG is sensitive to subtle brain changes that may not necessarily be picked up in the global brain atrophy measures. Indeed, the associations between change in WMLLs and annual rate of brain aging were significant for occipital, temporal, and parietal brain regions in addition to the global estimate (**Supplementary Table 10**). For BAG we did indeed only see significant associations with brain volume and BAG for occipital, frontal, parietal, and cingulate regions (**Supplementary Table 11**). This shows that regional brain age estimation may capture regional specificity of MS pathology (10–12, 33).

Multiple regression analyses revealed only nominally significant ($p < 0.05$, uncorrected) associations between some clinical, cognitive, and imaging variables and BAG as well as brain aging for specific brain regions. However, these associations did not survive correction for multiple testing, and further studies are needed to assess the robustness of these observations. A previous study in healthy individuals reported significant associations between BAG and performance on specific cognitive tests, including spatial Stroop and symbol coding, with poorer performance in individuals with an over-estimated age (13). Preliminary results from a partly overlapping cross-sectional sample revealed a significant association between BAG and Expanded Disability Status Scale (Fisher $z = 0.23$) (11), indicating that patients with higher clinical disease burden have older appearing brains. Further studies are needed to test the generalizability and robustness of these findings, both in clinical and healthy samples.

Brain age estimation is a useful framework that allows us to leverage large scale brain imaging databases for training robust machine learning models and apply automated prediction on the individual level. Further, whereas the approach builds on the vast amount of previous atrophy and lesion research, it contributes beyond that by downsampling a lot of information from the entire brain into a single holistic score in an automated fashion. As an example, in our data, we found no association between brain atrophy and change in Multiple Sclerosis Severity Scale (MSSS) ($r = 0.03$, $p = 0.80$), yet our brain age estimation approach revealed associations with change in MSSS for brain aging of the cerebellar & subcortical regions ($r = 0.36$, $p = 5.1 \times 10^{-3}$, not significant after correcting for multiple testing) (**Supplementary Table 12**).

Some limitations should be considered when interpreting the results. First, although the cross-sectional case-control comparison and the within-patient longitudinal analysis jointly suggest accelerated brain aging in patients with MS, a longitudinal sample of HCs would have enabled us to directly compare the rate of brain aging between patients and controls. Next, the current brain age model was exclusively based on gross morphometric features, and extending the range of brain imaging features, including indices of white matter microstructural properties and myelin integrity, may increase sensitivity to

TABLE 2 | Pearson's correlations between brain age gap and relevant clinical and MRI variables.

	Fullbrain		Frontal		Parietal		Cereb. / Subcort.	
Clinical variables	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***
9HPT Non-dominant	**0.36**	**5.8×10^{-3}**	0.03	0.80	0.16	0.22	**0.28**	**0.030**
Change in 9HPT Non-dominant	**0.28**	**0.035**	0.05	0.68	0.14	0.31	0.21	0.12
DMT Level	0.01	0.93	0.03	0.80	−0.05	0.70	**0.26**	**0.046**
Gender	**−0.28**	**0.031**	0.05	0.68	−0.18	0.17	−0.04	0.78
MRI variables	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***
WMLL	**0.46**	**3.0×10^{-4}**	0.19	0.16	0.24	0.07	**0.38**	**3.2×10^{-3}**
Change in WMLL	**0.30**	**0.022**	0.12	0.34	0.20	0.13	**0.34**	**9.6×10^{-3}**
Brain volume	−0.25	0.06	**−0.43**	**8.8×10^{-4}**	**−0.35**	**7.3×10^{-3}**	−0.24	0.07
Brain atrophy	**−0.33**	**0.011**	**−0.31**	**0.017**	**−0.37**	**4.7×10^{-3}**	−0.13	0.32
ICV	−0.01	0.94	**−0.29**	**0.027**	−0.20	0.13	−0.02	0.87

Significant associations are highlighted with bold ($p < 0.05$). Associations which were still significant after adjusting for false discovery rate are underlined. Cereb., cerebellar; Subcort., subcortical; 9HPT, nine hole peg test; Cor., correlation; DMT, disease-modifying therapies; WMLL, white matter lesion load; ICV, intracranial volume.

TABLE 3 | Pearson's correlations between annual rate of brain aging and relevant clinical and MRI variables on time point 3.

	Fullbrain		Frontal		Parietal		Cereb. / Subcort.	
Clinical variables	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***
EDSS	0.09	0.49	−0.01	0.95	−0.15	0.25	0.22	0.08
Change in EDSS	0.16	0.23	0.09	0.50	−0.03	0.83	**0.29**	**0.026**
MSSS	−0.03	0.84	−0.09	0.47	−0.21	0.11	0.17	0.20
Change in MSSS	0.17	0.21	0.10	0.46	0.05	0.68	**0.36**	**5.1×10^{-3}**
9HPT Non-dominant	**0.29**	**0.028**	0.15	0.27	0.01	0.92	**0.30**	**0.021**
Change in 9HPT Non-dominant	**0.31**	**0.017**	0.20	0.14	0.08	0.53	**0.32**	**0.014**
DMT Level	**−0.28**	**0.031**	−0.22	0.09	−0.17	0.21	−0.08	0.54
MRI variables	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***	**cor.**	***p***
WMLL	**0.29**	**0.026**	0.21	0.11	0.19	0.16	0.01	0.96
Change in WMLL	**0.30**	**0.015**	0.19	0.12	**0.35**	**4.3×10^{-3}**	0.00	0.98
Brain volume	−0.01	0.93	−0.08	0.54	−0.03	0.83	0.10	0.44
Brain atrophy	**−0.79**	**4.3×10^{-15}**	**−0.79**	**1.6×10^{-15}**	**−0.72**	**1.1×10^{-11}**	−0.07	0.57

Significant associations are highlighted with bold ($p<0.05$). Associations which were still significant after adjusting for false discovery rate are underlined. Cereb., cerebellar; Subcort., subcortical; Cor., correlation; EDSS, expanded disability status scale; MSSS, multiple sclerosis severity score; 9HPT, nine hole peg test; DMT, disease-modifying therapies; WMLL, white matter lesion load.

clinical trajectories in MS. When analyzing clinical associations with estimates of brain age gap we include clinical tests which relies heavily on the spine, although morphometric data from the spine are not included in our brain age estimation model. Finally, although prospective data is a substantial strength of our study, our design does not allow for causal inference (e.g., related to treatment status). Data from analyses of brain age compared to disease modifying treatments are provided in **Supplementary Figure 8,** and **Supplementary Tables 10, 11, 13**. Our current brain age estimation model aimed at identifying deviations from healthy aging trajectories, future studies could potentially benefit from establishing unique models based on disease specific training sets.

In conclusion, using advanced cross-sectional imaging data and machine learning we report higher brain age in patients with MS compared to healthy controls. Longitudinal analysis suggested accelerated brain aging in MS patients with higher levels of brain atrophy and longitudinal progression of changes in WMLL. Brain age estimation is a framework that allows us to downsample the complex brain imaging features into a single individual "score" using automated machine learning, enabling us to gain new insights into the complex brain structure. Jointly, these results corroborate that brain age estimation is a promising and intuitive tool with potential to establish a comprehensive measure of brain health which may guide a personalized treatment approach in MS.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Regional Committee for Medical

and Health Research Ethics with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the South East Regional Committee for Medical and Health Research Ethics.

AUTHOR'S NOTE

A preprint of this manuscript was published at bioRxiv (https://www.biorxiv.org/content/10.1101/440412v1) on October 10th 28 (34).

AUTHOR CONTRIBUTIONS

EH, TK, GN, HH, and LW contributed to the conception and design of the study. EH, TK, GN, KK, GR, HH, and LW contributed to the acquisition and analysis of data. EH, TK, GN, KK, HH, and LW drafted the text and figures. All authors contributed to the review and editing.

FUNDING

The project was supported by grants from The Research Council of Norway (NFR, grant number 240102, 249795, and 223273) and the South-Eastern Health Authorities of Norway (grant number 257955, 2014097).

ACKNOWLEDGMENTS

We thank all the patients who participated in our study. We acknowledge the collaboration with members of the Multiple Sclerosis Research Group at the University of Oslo and Oslo University hospital, especially Professor Elisabeth G. Celius.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2019.00450/full#supplementary-material

REFERENCES

1. Friese MA, Schattling B, Fugger L. Mechanisms of neurodegeneration and axonal dysfunction in multiple sclerosis. *Nat Rev Neurol.* (2014) 10:225–38. doi: 10.1038/nrneurol.2014.37
2. Rocca MA, Comi G, Filippi M. The role of T1-weighted derived measures of neurodegeneration for assessing disability progression in multiple sclerosis. *Front Neurol.* (2017) 8:433. doi: 10.3389/fneur.2017.00433
3. Filippi M, Preziosa P, Rocca MA. MRI in multiple sclerosis: what is changing? *Curr Opin Neurol.* (2018) 31:386–95. doi: 10.1097/WCO.0000000000000572
4. Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol.* (2016) 15:292–303. doi: 10.1016/S1474-4422(15)00393-2
5. Filippi M, Preziosa P, Rocca MA. Brain mapping in multiple sclerosis: Lessons learned about the human brain. *Neuroimage.* (2017) 190:32–45. doi: 10.1016/j.neuroimage.2017.09.021
6. Altmann DR, Jasperse B, Barkhof F, Beckmann K, Filippi M, Kappos LD, et al. Sample sizes for brain atrophy outcomes in trials for secondary progressive multiple sclerosis. *Neurology.* (2009) 72:595–601. doi: 10.1212/01.wnl.0000335765.55346.fc
7. Popescu V, Klaver R, Versteeg A, Voorn P, Twisk JW, Barkhof F, et al. Postmortem validation of MRI cortical volume measurements in MS. *Hum Brain Mapp.* (2016) 37:2223–33. doi: 10.1002/hbm.23168
8. Chard D, Trip SA. Resolving the clinico-radiological paradox in multiple sclerosis. *F1000Res.* (2017) 6:1828. doi: 10.12688/f1000research.11932.1
9. Franke K, Ziegler G, Klöppel S, Gaser C. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *Neuroimage.* (2010) 50:883–92. doi: 10.1016/j.neuroimage.2010.01.005
10. Cole JH, Poudel RPK, Tsagkrasoulis D, Caan MWA, Steves C, Spector TD, et al. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *Neuroimage.* (2017) 163:115–24. doi: 10.1016/j.neuroimage.2017.07.059
11. Kaufmann T, van der Meer D, Doan NT, Schwarz E, Lund MJ, Agartz I, et al. Genetics of brain age suggest an overlap with common brain disorders. *bioRxiv.* [Preprint] (2018) doi: 10.1101/303164
12. Cole JH, Franke K. Predicting age using neuroimaging: innovative brain ageing biomarkers. *Trends Neurosci.* (2017) 40:681–90. doi: 10.1016/j.tins.2017.10.001
13. Richard G, Kolskar K, Sanders AM, Kaufmann T, Petersen A, Doan NT, et al. Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. *PeerJ.* (2018) 6:e5908. doi: 10.7717/peerj.5908
14. Cole JH, Raffel J, Friede T, Eshaghi A, Brownlee W, Chard D, et al. Accelerated brain ageing and disability in multiple sclerosis. *bioRxiv.* [Preprint] (2019) doi: 10.1101/584888
15. Raffel J, Cole J, Record, C, Sridharan, S, Sharp, D, Nicholas, R. Brain age: a novel approach to quantify the impact of multiple sclerosis on the brain (P1.371). *Neurology.* (2017) 88(16 Suppl.):P1.371.
16. Nygaard GO, Celius EG, de Rodez Benavent SA, Sowa P, Gustavsen MW, Fjell AM, et al. A longitudinal study of disability, cognition and gray matter atrophy in early multiple sclerosis patients according to evidence of disease activity. *PLoS ONE.* (2015a) 10:e0135974. doi: 10.1371/journal.pone.0135974
17. Nygaard GO, Walhovd KB, Sowa P, Chepkoech JL, Bjornerud A, Due-Tonnessen P, et al. Cortical thickness and surface area relate to specific symptoms in early relapsing-remitting multiple sclerosis. *Mult Scler.* (2015b) 21:402–14. doi: 10.1177/1352458514543811
18. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol.* (2011) 69:292–302. doi: 10.1002/ana.22366
19. Moberget T, Doan NT, Alnaes D, Kaufmann T, Cordova-Palomera A, Lagerberg TV, et al. Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: a multisite mega-analysis of 983 patients and 1349 healthy controls. *Mol Psychiatry.* (2018) 23:1512–20. doi: 10.1038/mp.2017.106
20. Doan NT, Engvig A, Zaske K, Persson K, Lund MJ, Kaufmann T, et al. Distinguishing early and late brain aging from the Alzheimer's disease spectrum: consistent morphological patterns across independent samples. *Neuroimage.* (2017) 158:282–95. doi: 10.1016/j.neuroimage.2017.06.070
21. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. Segmentation, I, and surface reconstruction. *Neuroimage.* (1999) 9:179–94. doi: 10.1006/nimg.1998.0395
22. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage.* (2012) 61:1402–18. doi: 10.1016/j.neuroimage.2012.02.084
23. Reuter M, Rosas HD, Fischl B. Highly accurate inverse consistent registration: a robust approach. *Neuroimage.* (2010) 53:1181–96. doi: 10.1016/j.neuroimage.2010.07.020
24. Damangir S, Manzouri A, Oppedal K, Carlsson S, Firbank MJ, Sonnesyn H, et al. Multispectral MRI segmentation of age related white matter changes using a cascade of support vector machines. *J Neurol Sci.* (2012) 322:211–6. doi: 10.1016/j.jns.2012.07.064

25. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage.* (2012) 62:782–90. doi: 10.1016/j.neuroimage.2011.09.015
26. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage.* (2002) 17:825–41. doi: 10.1006/nimg.2002.1132
27. Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, et al. A multi-modal parcellation of human cerebral cortex. *Nature.* (2016) 536:171–8. doi: 10.1038/nature18933
28. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: *KDD'16.* San Francisco, CA: ACM (2016). p. 785–94.
29. Le TT, Kuplicki RT, McKinney BA, Yeh HW, Thompson WK, Paulus MP, et al. A nonlinear simulation framework supports adjusting for age when analyzing BrainAGE. *Front Aging Neurosci.* (2018) 10:317. doi: 10.3389/fnagi.2018.00317
30. Bernal-Rusiel JL, Greve DN, Reuter M, Fischl B, Sabuncu MR, Alzheimer's Disease Neuroimaging I. Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. *Neuroimage.* (2013) 66:249–260. doi: 10.1016/j.neuroimage.2012.10.065
31. Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* New York, NY: Springer-Verlag (2016).
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B.* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
33. Gourraud PA, Sdika M, Khankhanian P, Henry RG, Beheshtian A, Matthews PM, et al. A genome-wide association study of brain lesion distribution in multiple sclerosis. *Brain.* (2013) 136(Pt 4):1012–24. doi: 10.1093/brain/aws363
34. Hogestol EA, Kaufmann T, Nygaard GO, Beyer MK, Sowa P, Nordvik JE, et al. Cross-sectional and longitudinal brain scans reveal accelerated brain aging in multiple sclerosis. *bioRxiv.* [Preprint] (2018) doi: 10.1101/440412

Conflict of Interest Statement: EH has received honoraria for lecturing from Merck and Sanofi-Genzyme. MB has received honoraria for lecturing from Novartis and Biogen Idec. OA has received honoraria for lecturing from Lundbeck. HH has received travel support, honoraria for advice or lecturing from Biogen Idec, Sanofi-Genzyme, Merck, Novartis, Roche, and Teva and an unrestricted research grant from Novartis.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.