Magnetic resonance imaging study of brain atrophy in multiple sclerosis patients treated with autologous hematopoietic stem cell transplantation

Hyunwoo Lee

Integrated Program in Neuroscience

McGill University

Final Submission: September 2017

A thesis submitted to McGill University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

Abstract

Multiple sclerosis (MS) is a chronic autoimmune, inflammatory, and demyelinating disease of the central nervous system. Patients with MS suffer from inflammatory lesions in the brain, neuroaxonal degeneration, and brain atrophy; these pathologies manifest as episodes of relapses, physical and cognitive impairment, and accumulation of disability. Autologous hematopoietic stem cell transplantation (aHSCT) is a treatment approach that is based on the idea that complete abrogation of the old, autoreactive immune system and reconstitution of a new, self-tolerant immune system would result in prolonged remission of MS disease activity. aHSCT can induce substantial long-term reduction or even stoppage of relapses or new focal white matter lesion formation in MS, but its effect on brain atrophy is uncertain. The main objective of this thesis was to determine the effect of aHSCT on brain volume as measured with magnetic resonance imaging, and related degenerative processes in MS.

First, I modeled the time course of whole-brain (WB) atrophy in the Canadian "Multiple sclerosis – bone marrow transplantation" (MS-BMT) cohort. In this cohort, aHSCT was immediately followed by an early acceleration in the rate of WB atrophy. The dose of busulfan used for immunoablation (an index of chemotherapy-related neurotoxicity) and the baseline volume of T1-weighted white matter lesions (a marker of the amount of focally injured tissue that may be committed to degeneration prior to aHSCT) were significant factors associated with the accelerated atrophy. The accelerated atrophy slowed continuously over approximately 2.5 years, after which the average rate of WB atrophy was consistent with the rate observed in normal aging.

Second, I modeled the time courses of grey- and white-matter (GM, WM) atrophy in the Canadian MS-BMT cohort. Both the busulfan dose and the baseline volume of T1-weighted WM lesions were significant predictors of the early, accelerated WM atrophy following aHSCT, whereas only the busulfan dose was a significant predictor of the GM atrophy. The rates of both GM and WM atrophy subsequently slowed to levels seen in normal-aging, although in different timeframes: atrophy in the GM slowed over the first 1 to 2 years, whereas atrophy in the WM slowed over the first 2 to 3 years.

Third, I modeled the time courses of WB, GM, and WM atrophy in the "High-dose immunosuppressive therapy and autologous hematopoietic cell transplantation for relapsing-remitting multiple sclerosis" (HALT-MS) trial. The dose of the lower intensity chemotherapy regimen with BiCNU, etoposide, ara-C and melphalan (BEAM), and the baseline volume of T1-weighted WM lesions were significant predictors of the early, accelerated WB and WM atrophy. For the GM, only the BEAM dose was a significant predictor of atrophy. Compared to the Canadian MS-BMT cohort, the HALT-MS cohort experienced a shorter and milder course of the early, accelerated brain volume loss. The accelerated WB atrophy slowed continuously over approximately the first year of follow-up. Atrophy in the grey matter also slowed over the first year, whereas atrophy in the white matter slowed over a longer period, approximately the first 1 to 2 years.

Fourth, I estimated the size of effects of different MRI scanner upgrade or change combinations, and of T1-weighted sequence changes, on whole-brain volume change measurements, using a large cohort from the Alzheimer's Disease Neuroimaging Initiative (ADNI). A linear mixed-effects model was applied to estimate the effects of the scanner and the sequence changes. Inter-vendor scanner changes generally led to greater effects compared to intra-vendor scanner upgrades. Change in the T1-weighted sequence, within the same scanning platform, also led to a significant effect, comparable to that from inter-vendor scanner changes. Inclusion of the corrective terms led to better model goodness-of-fits, and thus, provided more reliable estimates of WB atrophy rates.

In summary, I showed that the early acceleration of brain atrophy in MS patients treated with aHSCT is likely due to chemotherapy-related toxicity as well as loss of WM tissues already in the process of degenerating due to MS-related injury prior to the treatment. Stopping focal white matter inflammation with aHSCT can eventually lead to slowing of brain atrophy in MS to, usually to rates consistent with normal aging. In addition, I showed that the effects of MRI scanner changes on brain atrophy measurement can be significant, and it would be beneficial to account for these effects during statistical analysis.

Résumé

La sclérose en plaques (SEP) est une maladie chronique auto-immune, inflammatoire et démyélinisante du système nerveux central. Les patients atteints de SEP souffrent de lésions inflammatoires dans le cerveau, de dégénérescence neuro-axiale et d'atrophie cérébrale; Ces pathologies se manifestent comme des épisodes de rechutes, de déficiences cognitives et de progression du handicap. l'immunosuppression suivie d'une greffe de cellules souches hématopoïétiques autologues (aHSCT) est une approche de traitement qui repose sur l'idée qu'une abrogation complète de l'ancien système immunitaire autorréactif et la reconstitution d'un nouveau système immunitaire autonome entraînera une rémission prolongée de l'activité de la maladie de SEP. L'aHSCT peut induire une réduction significative à long terme ou même un arrêt des rechutes ou une nouvelle formation de lésion focale dans la SEP, mais son effet sur l'atrophie cérébrale est incertain. L'objectif principal de cette thèse est de déterminer l'effet de l'aHSCT sur le volume du cerveau tel que mesuré avec l'imagerie par résonance magnétique et les processus dégénératifs apparentés dans la SEP.

Tout d'abord, j'ai modélisé la trajectoire temporelle de l'atrophie du cerveau entier (CE) dans la cohorte « L'essai canadien sur la greffe de moelle osseuse » (MS-BMT). L'aHSCT a été immédiatement suivie d'une accélération précoce du taux d'atrophie du CE. La dose de traitement de chimiothérapie de busulfan à haute intensité (un indice de neurotoxicité liée à la chimiothérapie) et le volume de base des lésions de la matières blanche pondérée par T1 (un marqueur de la quantité de tissu lésé focalement qui pourrait être compromis à la dégénérescence avant l'aHSCT) étaient des facteurs significatifs associés à l'atrophie accélérée. L'atrophie accélérée a ralenti de façon continue pendant environ 2,5 ans, après quoi le taux moyen d'atrophie du CE était compatible avec le taux observé dans le vieillissement normal.

Deuxièmement, j'ai modélisé les trajectoires temporelles de l'atrophie des matières grasses et des matières blanches (MG, MB) dans la cohorte canadienne MS-BMT. La dose de busulfan et le volume de base des lésions MB pondérées par T1 étaient des prédicteurs significatifs de l'atrophie MB accélérée précoce après l'aHSCT, alors que seule la dose de busulfan était un prédicteur significatif de l'atrophie MG. Les taux d'atrophie MG et MB ont ensuite ralenti aux niveaux observés dans le vieillissement normal, bien que dans des délais différents: l'atrophie dans le MG a ralenti au cours des 1 à 2 premiers ans, alors que dans le MB a ralenti au cours des

2 à 3 premiers ans. En outre, il y a eu une réduction significative des taux à long terme de l'atrophie MG par rapport aux taux de référence.

Troisièmement, j'ai modélisé les trajectoires temporelles de l'atrophie CE, MG et MB dans la cohorte « High-dose immunosuppressive therapy and autologous hematopoietic cell transplantation for relapsing-remitting multiple sclerosis » (HALT-MS). La dose de régime de chimiothérapie BEAM à intensité intermédiaire et le volume de base des lésions MB pondérées par T1 étaient des prédicteurs significatifs de l'atrophie précoce et accélérée de la CE et de la MB. Pour le MG, seule la dose BEAM était un prédicteur significatif d'atrophie. Par rapport à la cohorte canadienne de MS-BMT, la cohorte HALT-MS a connu des trajectoires plus courtes et plus légeres de la perte de volume cérébral précoce et accélérée. L'atrophie accélérée de la CE a ralenti de façon continue au cours de la première année de suivi. L'atrophie dans la matière grise a également ralenti au cours de la première année, et que dans la matière blanche a ralenti sur une période plus longue, environ les 1 à 2 premières années.

Quatrièmement, j'ai estimé l'impact de différentes combinaisons de modification ou de changement de scanners IRM et de changements de séquence pondérés T1, sur des mesures de changement de volume de cerveau entier, en utilisant une grande cohorte de l'étude de « Alzheimer's Disease Neuroimaging Initiative » (ADNI). Un modèle linéaire à effets mixtes a été appliqué pour estimer les effets du scanner et les changements de séquences. J'ai également évalué si l'inclusion des termes correctifs dans le modèle a conduit à des améliorations dans les mesures de qualité. Les changements de scanners inter-fournisseur ont généralement entraîné des effets plus importants par rapport aux mises à niveau de scanner intra-fournisseur. Le changement de la séquence pondérée T1, dans la même plate-forme de scanner, a également entraîné un effet significatif, comparable à celui des changements de scanner entre fournisseurs. L'inclusion des termes correctifs a conduit à une meilleure qualité de modélisation des modèles, et donc, a fourni des estimations plus fiables des taux d'atrophie de la CE.

En résumé, j'ai montré que l'accélération précoce de l'atrophie cérébrale chez les patients atteints de SEP traités par un l'aHSCT est probablement due à une toxicité liée à la chimiothérapie ainsi qu'à la perte de tissus MB déjà en cours de dégénérescence due à une lésion liée à la SEP avant le traitement. L'arrêt de l'inflammation focale de la matières blanche peut éventuellement conduire à un ralentissement considérable de l'atrophie cérébrale dans la SEP. En outre, j'ai montré que les

effets de modifications ou changement des scanners IRM peuvent être significatifs sur la mesure de l'atrophie cérébrale et qu'il serait avantageux de tenir compte de ces effets lors de l'analyse statistique.

Acknowledgements

This work would not have been possible without the guidance and support of wonderful people around me.

First and foremost, I would like to acknowledge Dr. Douglas L. Arnold for supervising my work with great kindness, extraordinary dedication, extreme patience, immense knowledge, enlightening advice, and unwavering support. I wish to thank you from the bottom of my heart.

The advisory committee members: Dr. Sridar Narayanan, Dr. D. Louis Collins and Dr. G. Bruce Pike for insightful ideas and inputs not only on the direction of the projects, but also on any other matters.

Members of the MRS lab and the BIC:

Dr. Robert A. Brown for answering my endless questions, for years, on every topic, no matter how minor or major the issue, with unending patience and sincerity.

Dr. Kunio Nakamura for countless hours of thought-provoking discussions that allowed me to better understand the theories and steps of image processing and measurement of brain atrophy.

Dr. Jacqueline T. Chen for getting me started when I first joined the lab and knew very little about aHSCT.

Dr. Mishkin Derakhshan, Mr. Simon Francis and Dr. Vladimir Fonov for informative discussions on image processing. Dr. Haz-Edine Assemlal for helpful discussions, as well as for helping with the French translation of the thesis abstract. Ms. Ellie Tobman for helping with administrative matters. Dr. Sergey Kuznetsov for helping with computer-related matters. Mr. Zografos Caramanos for being very supportive and helping with very many things. Dr. Dumitru Fetco, Dr. Josefina Maranzano, Dr. Rezwan Ghassemi, Dr. David Rudko and Dr. Stanley Hum for their kind advices and encouragements throughout my studies. Mr. Xu Fei, the man of great wisdom.

The Canadian MS-BMT study: Dr. Harold L. Atkins, Dr. Mark S. Freedman and Ms. Marjorie Bowman for valuable discussions with regards to the treatment procedures and clinical aspect of immunoablation and aHSCT.

HALT-MS Study: Dr. Richard A. Nash for valuable discussions with regards to the treatment procedures.

Thesis committee: Dr. Roland G. Henry (external examiner), Dr. Alain Dagher (internal examiner), Dr. Mallar Chakravarty (external member) and Dr. Christine L. Tardif (internal member) for the time and expertise they have invested during the thesis examination and the oral defense.

IPN mentor: Dr. Pierre Lachapelle for guiding me through the Ph.D. program milestones.

The EndMS program, Ms. Anne-Marie Bismuth and Ms. Anik Schoenfeldt: for their dedication to organizing and providing unique educational and networking opportunities through events, conferences, summer schools, and the SPRINT program.

Throughout my studies, I received invaluable financial support from the McGill Integrated Program in Neuroscience, the CIHR Neuroinflammation Training Program, the EndMS SPRINT Program and the Fonds de la recherche en santé du Québec. The Canadian MS-BMT study was funded by the Multiple Sclerosis Scientific Research Foundation. The HALT-MS study was funded by the National Institute of Health. The ADNI study was funded by a public-private partnership (http://adni.loni.usc.edu/about/funding/).

Importantly, I would like to thank the patients that contributed to the data analyzed in this work. This work would not have been possible without their determination and courage.

Lastly, I would like to acknowledge my family for their constant support and encouragement throughout my studies.

General Preface

This thesis is the result of my original research, conducted under the supervision of Dr. Douglas Arnold at McGill University. The thesis begins with the review of the relevant literature, followed by research objectives and rationale (chapter 1). The main body comprises four manuscripts, of which I am the primary author, that were prepared for publication in peer-reviewed journals (chapters 2 to 5). This is followed by the final summary and conclusions (chapter 6), and the reference list.

aHSCT has been shown to be effective in reducing inflammatory activity in MS related to relapses and focal WM lesions shown on MRI. However, its effect on brain atrophy is less well understood. Although it has been shown that the rates of WB atrophy accelerate following aHSCT, the cause(s) of this acceleration are uncertain. Moreover, the long-term effect of aHSCT on brain atrophy is unknown. Chapters 2 to 4 (manuscripts 1 to 3) of this thesis extend the current understanding of the aforementioned issues by studying two independent cohorts of MS patients treated with aHSCT.

A related issue that frequently interferes with reliable measurement of brain atrophy in longitudinal MRI studies is changes or upgrades to MRI hardware or pulse sequences during the study. Chapter 5 (manuscript 4) sought to assess the effect of MRI scanner changes on measurement of brain atrophy by analyzing a dataset from the Alzheimer's Disease Neuroimaging Initiative study.

Contributions to knowledge and contributions of authors

Manuscript 1: Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis

Published in *Multiple Sclerosis Journal*. 2017 Mar;23(3):420-431.

Contributions to knowledge:

Characterized the long-term evolution of brain atrophy in the Canadian MS-BMT patients treated with busulfan-based immunoablation and aHSCT, by statistically modeling the time course of whole-brain atrophy for up to 10 years.

Demonstrated that treatment-related toxicity (represented by the dose of busulfan) and degeneration of tissue injured prior to aHSCT (represented by the baseline volume of T1-weighted white matter lesions) contributed to the early, acceleration of whole-brain atrophy observed following immunoablation and aHSCT.

Demonstrated that the early, accelerated whole-brain atrophy slowed continuously over approximately 2.5 years in this cohort.

Showed that there was no evidence of significant changes in the whole-brain tissue water content after aHSCT.

Demonstrated that after the early acceleration, the average rate of whole-brain atrophy in this cohort was consistent with the rate seen in normal aging.

Contributions of authors:

Authors: Hyunwoo Lee, Sridar Narayanan, Robert A. Brown, Jacqueline T. Chen, Harold L. Atkins, Mark S. Freedman, Douglas L. Arnold

Hyunwoo Lee: Study conception and design, acquisition of data, analysis and interpretation, drafting of manuscript

Sridar Narayanan: Study conception and design, acquisition of data, critical revision

Robert A. Brown: Study conception and design, critical revision

Jacqueline T. Chen: Acquisition of data, critical revision

Harold L. Atkins: Acquisition of data, critical revision

Mark S. Freedman: Acquisition of data, critical revision

Douglas L. Arnold: Study conception and design, acquisition of data, analysis and interpretation, critical revision

Manuscript 2: Impact of immunoablation and autologous hematopoietic stem cell transplantation on grey matter and white matter atrophy in multiple sclerosis

Accepted for publication in *Multiple Sclerosis Journal*

Contributions to knowledge:

Statistically modeled the long-term time courses of grey matter and white matter atrophy in the Canadian MS-BMT patients.

Demonstrated that both the busulfan dose and the baseline T1-weighted white matter lesion volume were significantly associated with the early, accelerated white matter atrophy following aHSCT; only the busulfan dose was significantly associated with the early, accelerated grey matter atrophy.

Demonstrated that the treatment-related atrophy of each tissue compartment occurred with different half-times: the atrophy in the grey matter slowed over the first 1 to 2 years of follow-up, and the atrophy in the white matter slowed over approximately the first 2 to 3 years of follow-up.

Showed that there was no evidence of significant changes in the grey matter and white matter tissue water content after aHSCT. This suggested that the early acceleration of volume loss did not result from "pseudoatrophy" due to resolution of edema in this cohort.

Demonstrated that there was a significant reduction in the long-term rates of grey matter atrophy compared to the baseline rates.

Demonstrated that after the early, accelerated periods, the average rates of grey matter and white matter atrophy were comparable to the rates seen in normal aging.

Contributions of authors:

Authors: Hyunwoo Lee, Kunio Nakamura, Sridar Narayanan, Robert A. Brown, Jacqueline T. Chen, Harold L. Atkins, Mark S. Freedman, Douglas L. Arnold

Hyunwoo Lee: Study conception and design, acquisition of data, analysis and interpretation, drafting of manuscript

Kunio Nakamura: Acquisition of data, critical revision

Sridar Narayanan: Study conception and design, acquisition of data, critical revision

Robert A. Brown: Study conception and design, critical revision

Jacqueline T. Chen: Acquisition of data, critical revision

Harold L. Atkins: Acquisition of data, critical revision

Mark S. Freedman: Acquisition of data, critical revision

Douglas L. Arnold: Study conception and design, acquisition of data, analysis and interpretation, critical revision

Manuscript 3: Brain atrophy in relapsing remitting multiple sclerosis following high-dose immunosuppressive therapy and autologous hematopoietic cell transplantation in the HALT-MS trial

Manuscript in preparation

Contributions to knowledge:

Characterized the long-term evolution of brain atrophy in the HALT-MS patients treated with BEAM-based immunosuppression and aHSCT, by statistically modeling the time course of whole-brain, grey matter and white matter atrophy for up to 5 years.

Demonstrated that both the BEAM dose and the baseline T1-weighted white matter lesion volume were significantly associated with the early, accelerated whole-brain and white matter atrophy following treatment; only the BEAM dose was significantly associated with the early, accelerated grey matter atrophy.

Showed that timeframes of the early, accelerated atrophy in each tissue compartment were shorter than that observed after the more intensive immunoablation used in the Canadian MS-BMT trial. In the HALT-MS trial, the atrophy in the grey matter slowed over the first year of follow-up; the atrophy in the white matter slowed over approximately the first 1 to 2 years of follow-up; altogether, the atrophy in the whole-brain slowed over the first year of follow-up.

Confirmed that over the long-term, adequate immunosuppression or immunoablation and aHSCT can reduce brain atrophy rates to that consistent with normal aging.

Contributions of authors:

Authors: Hyunwoo Lee, Kunio Nakamura, Robert A. Brown, Sridar Narayanan, Richard A. Nash, Douglas L. Arnold

Hyunwoo Lee: Study conception and design, acquisition of data, analysis and interpretation, drafting of manuscript

Kunio Nakamura: Acquisition of data, critical revision

Robert A. Brown: Study conception and design, critical revision

Sridar Narayanan: Study conception and design, critical revision

Richard A. Nash: Acquisition of data, critical revision

Douglas L. Arnold: Study conception and design, acquisition of data, analysis and interpretation, critical revision

Manuscript 4: Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain atrophy measurements

Manuscript in preparation

Contributions to knowledge:

Using a large cohort of subjects from the Alzheimer's disease neuroimaging initiative (ADNI), estimated the effects of 1) intra-vendor MRI scanner upgrades, 2) inter-vendor scanner changes and 3) standardized T1-weighted sequence change on longitudinal percentage brain volume changes estimated using FSL-SIENA.

Demonstrated a large effect of inter-vendor scanner changes, equivalent to a year's worth of atrophy in AD and MCI.

Demonstrated a much smaller effect of intra-vendor scanner upgrades, comparable to half a year's worth of normal aging in this elderly group, or about 20% of the annual change in the AD subjects.

Demonstrated a large effect of T1-weighted sequence change, comparable to that from inter-vendor scanner changes.

Showed that modeling brain volume loss with a linear mixed-effects model that includes corrective terms for scanner and sequence changes led to better model goodness-of-fits.

Contributions of authors:

Authors: Hyunwoo Lee, Kunio Nakamura, Robert A. Brown, Sridar Narayanan, Douglas L. Arnold

Hyunwoo Lee: Study conception and design, acquisition of data, analysis and interpretation, drafting of manuscript

Kunio Nakamura: Study conception and design, acquisition of data, critical revision

Robert A. Brown: Study conception and design, critical revision

Sridar Narayanan: Study conception and design, critical revision

Douglas L. Arnold: Study conception and design, acquisition of data, analysis and interpretation, critical revision

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AD: Alzheimer's Disease

ADNI: Alzheimer's Disease Neuroimaging Initiative

aHSCT: Autologous Hematopoietic Stem Cell Transplantation

ANOVA: Analysis of Variance

ATG: Antithymocyte Globulin

BEAM: A conditioning regimen consisting of BiCNU, Etoposide, Ara-C, and Melphalan

BBB: Blood Brain Barrier

BMT: Bone Marrow Transplantation

BPF: Brain Parenchymal Fraction

BSI: Boundary Shift Integral

CD: Cluster of Differentiation

CIS: Clinically Isolated Syndrome

CNS: Central Nervous System

CSF: Cerebrospinal Fluid

DIR: Double Inversion Recovery

DMT: Disease Modifying Therapy

EDSS: The Kurtzke Expanded Disability Status Scale

FLAIR: Fluid Attenuated Inversion Recovery

FU: Follow-up

GA: Glatiramer Acetate

Gd-enhancing: Gadolinium-Enhancing

GM: Grey Matter

GMF: Grey Matter Fraction

GMV: Grey Matter Volume

IA: Immunoablation

IFN: Interferon

HLA: Human Leukocyte Antigen

HSC: Hematopoietic Stem Cell

MCI: Mild Cognitive Impairment

MRI: Magnetic Resonance Imaging

MRS: Magnetic Resonance Spectroscopy

MS: Multiple Sclerosis

MSFC: The Multiple Sclerosis Functional Composite

MTR: Magnetization Transfer Ratio

NABT: Normal Appearing Brain Tissue

NAGM: Normal Appearing Grey Matter

NAWM: Normal Appearing White Matter

NBV: Normalized Brain Volume

NGMV: Normalized Grey Matter Volume

n.s.: Not Significantly Different

NWMV: Normalized White Matter Volume

PBVC: Percentage Brain Volume Change

PJI: Pairwise Jacobian Integration

PPMS: Primary Progressive Multiple Sclerosis

PRMS: Progressive Relapsing Multiple Sclerosis

pT2: Pseudo-T2

RRMS: Relapsing Remitting Multiple Sclerosis

SIENA: Structural Image Evaluation, using Normalisation, of Atrophy

SPGR: Spoiled Gradient Recalled Echo

SPM: Statistical Parametric Mapping

SPMS: Secondary Progressive Multiple Sclerosis

T1LV: T1-weighted Lesion Volume

T2LV: T2-weighted Lesion Volume

TBI: Total Body Irradiation

TBM: Tensor Based Morphometry

TE: Echo Time

TR: Repetition Time

VBM: Voxel Based Morphometry

WB: Whole Brain

WBV: Whole Brain Volume

WM: White Matter

WMF: White Matter Fraction

WMV: White Matter Volume

Chapter 1 Introduction

1.1 Overview of Multiple Sclerosis

Multiple sclerosis (MS) is a chronic autoimmune, inflammatory, and demyelinating disease of the central nervous system (CNS) that leads to neuroaxonal damage and accumulation of disability. MS affects more than 2.3 million people worldwide, including nearly 100,000 Canadians.[1] The prevalence of MS in Canada is one of the highest in the world, ranging between 55-240 per 100,000.[2] There is a strong female preponderance in the incidence of MS, but males who develop MS may undergo a more rapid decline in cognition and disability.[3] With an average age at onset of around 30,[4] MS is one the most common causes of non-traumatic neurological disability in young adults.[5]

As yet, the cause of MS is unknown but is thought to be multifactorial, involving multiple genetic and environmental factors. While human leukocyte antigen (HLA) genes are a well-known genetic susceptibility factor, there may be non-HLA regions associated with MS susceptibility.[6] First-degree relatives of a patient have greater risks of also developing MS.[7,8] Potential environmental risk factors include smoking, latitude, and infectious agents such as Epstein-Barr virus.[8,9]

1.1.1 Symptoms and Clinical Courses of Multiple Sclerosis

MS is associated with a wide range of symptoms. During relapse, the most common include sensory disturbances (numbness, tingling, temperature sensitivity, Lhermitte's sign, pain), visual impairments (blurred vision from internuclear opthalmoplegia, optic neuritis), motor dysfunction (muscle weakness, stiffness and spasms, gait ataxia, impaired balance, tremor), bladder and bowel dysfunction, sexual dysfunction, cognitive impairment (attention, reasoning, executive function), speech impairment, depression, and fatigue.[7,10]

The course of MS is unpredictable, not only because it varies among patients but also because it changes with time within the same patient. Four types of disease course have been defined, based on clinical characteristics: relapsing remitting multiple sclerosis (RRMS), secondary progressive

multiple sclerosis (SPMS), primary progressive multiple sclerosis (PPMS), and progressive relapsing multiple sclerosis (PRMS).[11] According to Lublin et al., RRMS is defined as "clearly defined disease relapses with full recovery or with sequelae and residual deficit upon recovery; periods between disease relapses characterized by a lack of disease progression."[11] SPMS is "initial RR disease course followed by progression with or without occasional relapses, minor remissions, and plateaus."[11] PPMS is "disease progression from onset with occasional plateaus and temporary minor improvements allowed."[11] Finally, PRMS is "progressive disease from onset, with clear acute relapses, with or without full recovery; periods between relapses characterized by continuing progression."[11]

Clinical onset in about 85% of patients who later develop MS begins with clinically isolated syndrome (CIS), the first episode of neurological disturbance associated with inflammatory demyelination.[12–14] CIS is always isolated in time by definition, and is usually, but not always, isolated in space.[13] A CIS episode usually resolves without intervention.[13] After a second clinical attack (relapse) or magnetic resonance imaging (MRI) lesions that demonstrate dissemination in space and time, a CIS patient is diagnosed with RRMS.[15] Subsequently, RRMS patients experience a series of relapses followed by variable degrees of recovery.[16] This phase usually lasts for about two decades,[17] then this RRMS phase may gradually transition to SPMS.[12] SPMS is characterized by progressive worsening of disability in the absence of relapses.[12]

About 10-15% of MS patients have PPMS, which is characterized by a progressive worsening of disability from the onset, without distinct relapses, at a rate of disability progression similar to that of SPMS.[11,12] The course characterized by the progressive worsening of disability from the onset with distinct relapses, formerly known as PRMS[11], is now considered a part of PPMS, i.e. PP-active.[12]

1.1.2 Evaluation of Disability in Multiple Sclerosis

Currently, the Kurtzke Expanded Disability Status Scale (EDSS) is the de facto standard for the measurement of disability and progression used both clinically and in clinical trials for MS.[18,19] The EDSS grades disability across eight functional system subscales, and its score ranges from 0 (normal neurologic examination) to 10 (death due to MS).[18] EDSS scores from 1.0-3.0 and 3.0-5.0 indicate minimal and moderate disabilities, respectively.[20] A score above 6.0, however,

indicates a more severe disability that requires constant assistance in walking.[20] A score above 8.5 indicates that the patient is mostly restricted to bed.[20] Weaknesses of the EDSS include: 1) high inter-observer variability around the lower end of the scale, 2) heavy bias towards ambulatory disability especially in the mid-upper regions of the scale, and 3) nonlinearity of the scale.[20]

The multiple sclerosis functional composite (MSFC) is another method of evaluating the disabilities associated with MS.[20] The MSFC comprises three tests: the 9-hole peg test, the timed 25-foot walk, and the paced auditory serial addition test.[20] Strengths of the MSFC include: 1) linearity of the scale, 2) inclusion of a cognitive assessment, and 3) high sensitivity and reproducibility.[20] Weaknesses of the MSFC include: 1) significant practice effects, 2) lack of a measure of vision, 3) difficulty in the interpretation of the MSFC Z-scores and 4) potential limitation in the comparability of results from different studies.[20]

1.2 Pathology of Multiple Sclerosis

MS has been traditionally thought of as a disease that predominantly affects the white matter (WM), hallmarked by an autoimmune response against myelin resulting in multifocal demyelinating WM plaques (lesions) visible on conventional MRI and histopathological studies of the post-mortem brain. However, recent evidence has highlighted the fact that the focal WM lesions are only one part of the whole picture; MS pathology is also found in normal-appearing tissues as well as the grey matter (GM).

1.2.1 Pathology in Focal White Matter Lesions

Focal WM lesions are areas of focal demyelination with varying degrees of inflammation, gliosis, and axonal injury. They are found throughout the CNS, but occur most commonly around the ventricles or deep WM. Specific characteristics of the focal WM lesions include: 1) distinguishable demarcation and color/texture depending on the stage of activity and repair,[21,22] 2) formation around central veins,[23] and 3) the presence of various combinations of myelin-laden macrophages, lymphocytes, microglia, and reactive astrocytes.[21,22]

Focal WM lesions are heterogeneous in terms of associated degrees of inflammation, de/remyelination, as well as oligodendrocyte and axonal injury.[24] One way of staging demyelinating activity in WM lesions is in terms of the sequence of myelin degradation products in macrophages, i.e. early active, late active, inactive demyelinated, early remyelinated, and late remyelinated (shadow plaque) stages.[25] Active lesions are classified as early or late depending on immunoreactivity for specific myelin proteins, i.e. early active lesions contain minor myelin proteins and late active lesions contain major myelin proteins.[25] Inactive demyelinated lesions are still infiltrated by macrophages, but they are no longer myelin-laden.[25] However, remyelinated lesions, especially the shadow plaques, are characterized by variable extents of remyelination with uniformly thin myelin sheaths.[25,26]

Also, types of active WM lesion can be classified based on the topographical distribution of the macrophages: 1) acute active types are hypercellular and have macrophages throughout the lesion. The macrophages contain both early and late myelin degradation products; 2) chronic active types have hypercellular rims and hypocellular center, with active macrophages (also containing early and late myelin degradation products) clustered at the border. Here, the center of the lesion is inactive; 3) smoldering rim types have few active macrophages, containing early and late myelin degradation products, restricted to the border, while the center is also inactive; 4) chronic inactive types are hypocellular, demyelinated throughout, and contain no early or late myelin degradation products.[24,27]

Axonal injury and loss is an important component of MS pathology that may be associated with progression of the disease. While demyelination is potentially reversible, axonal transection is not.[28] Axonal injury starts from the early stages of MS, as indicated by 1) the accumulation of amyloid precursor protein (a marker of axonal transport failure, and thus axonal damage) within acute active lesions as well as in the border of chronic active lesions,[29] and 2) the prevalence of axonal transection, especially in active lesions from patients with short disease duration.[30] These findings suggest that the degree of axonal transection is correlated with the degree of inflammation.[29,30] Therefore, toxic inflammatory mediators such as protease, cytokines, oxidative products, and free radicals may be associated with axonal injury.[31] Yet, ongoing axonal injury is also evident at a low, but significant level in demyelinated chronic inactive lesions.[32]

1.2.2 Pathology in Normal-Appearing White Matter

Normal-appearing white matter (NAWM) refer to WM areas that appear normal (i.e. are not lesional). The term "normal-appearing" is used because the non-lesional WM is not, in fact, normal, but shows evidence of chronic injury, including blood-brain-barrier (BBB) disruption, mild inflammation, microglial activation, gliosis, axonal swellings and injury, and increased protease expression.[33] Diffuse inflammation in the NAWM is significantly more pronounced in progressive MS (e.g. SPMS and PPMS) compared to acute and relapsing MS.[34] Accordingly, NAWM in SPMS and PPMS patients exhibits significantly more diffuse axonal injury, as indicated by increased numbers of axonal swellings and axonal spheroids.[34] For example, normal-appearing corpora callosa of MS patients have significantly reduced axonal density as well as significant atrophy compared to those of normal controls.[35] One potential mechanism for the axonal injury in NAWM is Wallerian degeneration, a process originating from axonal transections within focal WM lesions.[36] However, there is also a finding that the focal WM lesion load is not significantly correlated with microglia activation or axonal injury in the NAWM. This indicates that diffuse WM injury may develop independently of focal WM lesions.[34]

1.2.3 Pathology in Grey Matter

While GM pathology in MS has long been recognized,[37] its extent has been underestimated until recently. This is largely due to the fact that conventional analysis techniques, such as conventional MRI or histochemical stains (e.g. Luxol Fast Blue), have low sensitivity to GM lesions.[38] An early study found that about 5%, 4%, and 17% of the total number of plaques were positioned in cortex, central GM, and cortex/WM junction, respectively.[37]

However, recent immunohistochemical findings suggest that GM pathology in MS is as extensive, if not more so, than that of WM and is found in GM regions including cingulate gyrus, frontal cortex, temporal cortex, motor cortex, hippocampus, thalamus, hypothalamus, basal ganglia, cerebellum and spinal cord, among others. For example, Huitinga et al. found hypothalamic lesions in 16 out of 17 (2 PP, 15 RR) MS cases.[39] Bo et al. showed that in 20 (10 SP, 7 PP, 3 RR) MS cases, the mean percentage of demyelinated area in the cerebral cortex is significantly higher than that in WM (26.5% vs. 6.5%, p=0.001), with particular prominence in

the cingulate gyrus.[38] Vercellino et al. studied 6 cases of RR and SPMS and found a higher percentage of demyelinated cortex and a significant reduction of neuronal density in 2 SPMS cases (48% and 25.5%, vs. the mean of 2.75% in the other 3 RR and 1 SP cases); overall, the mean percentages of demyelinated cortex and WM were 14.8% and 21.75%, respectively.[40] Again, cortical lesions were more frequent in the cingulate gyrus.[40] Geurts et al. found extensive hippocampal demyelination in 15 out of 19 (9 SP, 7 PP, 1 RR) MS cases; none were found in healthy controls.[41] Gilmore et al. showed that in 14 cases (11 SP, 2 PP, 1 RR), the mean proportion of demyelinated area in GM was significantly higher than that in WM (28.8% vs. 15.6%, $p<0.001$).[42] While demyelination was found in cortical (motor cortex and cingulate) and deep GM (thalamus) areas, it was especially pronounced in the spinal cord and cerebellum.[42] In particular, the proportion of demyelinated GM was significantly greater than that of WM in the spinal cord, cerebellum, and motor cortex.[42] Kutzelnigg et al. suggested that cortical demyelination is overall more prominent in progressive MS (SP and PP).[34]

Furthermore, demyelinated areas in the cortex have significantly increased levels of axonal transection, dendrite transection, and importantly, neuronal apoptosis, compared to healthy cortex or myelinated MS cortex.[43] Indeed, cortical lesions have significantly reduced neuronal density compared to adjacent normal cortex.[40] Wegner et al. reported significantly reduced neocortical thickness in MS patients compared to controls, although they found there was no significant correlation between the mean cortical thickness and the mean extent of cortical demyelination.[44] Significant reductions in the neuronal density, neuronal cell size, glial density, and synaptic density were also found in leukocortical lesions compared to normal-appearing MS neocortex.[44]

Four types of cortical lesions have been defined according to their locations: Type 1 (leucocortical) lesions extend across both WM and GM; Type 2 (intracortical) lesions reside entirely within the cerebral cortex, and are perivascular; Type 3 (subpial) lesions extend from the pial surface into the cortex; Type 4 lesions span the cortex.[38,43] In an analysis of 112 cortical lesions from 50 MS patients, Peterson et al. reported that the types 1, 2, and 3 accounted for 34%, 16%, and 50% of the cortical lesions, respectively.[43] In another analysis of 109 cortical lesions from 20 MS patients (3 RR, 10 SP, 7 PP), Bo et al. reported that the types 1, 2, 3, and 4 accounted for 14.4%, 17%, 60%, and 8% of the cortical lesions, respectively.[38] Types 1 and 2

represent a small proportion of the cortical demyelination, but are found in all stages of MS including acute, RR, SP, and PPMS; particularly, they are the dominant types in acute and RRMS.[45]

The relationship between GM demyelination and meningeal inflammation is currently being recognized. In terms of histopathological analyses, Magliozzi et al. has showed that B-cell follicles are present in the cerebral meninges of SPMS cases, but not PPMS cases.[46] In particular, the follicles were found adjacent to large subpial lesions; the authors suggested that the follicles may be associated with cortical injury through release of soluble factors, such as pathogenic antibodies, pro-inflammatory cytokines, or proteolytic enzymes, that diffuse into the cortex through the pial membrane.[46] Howell et al. showed that the B-cell follicle-like structures are predominantly found in deep cerebral sulci, and that the follicle-positive SPMS cases have significantly greater area of demyelination compared to follicle-negative SPMS cases.[47] Notably, in the follicle-positive SPMS cases, the study found a significant correlation between the relative incidence of follicle aggregates and the level of total meningeal infiltrates (T- and B-lymphocytes). Further, the level of total meningeal infiltrates was significantly correlated with the degree of microglial activation as well as the extent of GM demyelination in the forebrain.[47] Additionally, the follicle-positive MS cases had significantly younger age at disease onset, age at conversion to SPMS, and age of death.[46,47] Choi et al. showed that in 26 PPMS cases there were significant aggregates of T- and B-lymphocytes in the meninges with absence of the tertiary lymphoid-like structures.[48] This study found a significant correlation between the total number of meningeal lymphocytes and the extent of cortical demyelination.[48] Microglia activation was significantly higher in PPMS cases compared to controls.[48] Also, greater meningeal inflammation was associated with greater axon/dendrite loss in GM lesions.[48] Note that the above studies mostly investigated the chronic, progressive stages of MS. For early MS cases, Lucchinetti et al. has shown evidence of active demyelination (myelin-laden macrophages in cortical lesions), cortical inflammation especially in leukocortical lesions (perivascular T- and B-cell infiltrates), and meningeal inflammation (both diffuse and focal perivascular).[49] Cortical demyelination was significantly correlated with both the diffuse and the focal perivascular meningeal inflammation.[49] Also, cortical demyelination was topographically related to meningeal inflammation.[49] While cortical inflammation is prominent in early MS, it may resolve rapidly; this may explain the lack of parenchymal lymphocytes, macrophages, and tertiary

lymphoid-like follicle structures in the cortical lesions of progressive MS.[43,48] In an MRI study of 229 MS patients (10 CIS, 171 RRMS, 44 SPMS, 74 PPMS), Absinta et al. reported 1.7-fold higher prevalence of leptomeningeal gadolinium enhancement (a potential marker of focal leptomeningeal inflammation) in SP and PP MS (33%) than in RRMS (19%); in particular, the PPMS cases had the highest prevalence at 38%.[50]

1.3 MRI in Multiple Sclerosis

MRI non-invasively produces in-vivo images of soft tissue with high resolution and contrast that allow for quantitative measurements of changes due to disease pathology, disease progression, or therapeutic intervention.

Due to its sensitivity and relative specificity to MS pathology, MRI has become an important component in the diagnosis of MS. With the aim of facilitating earlier and more sensitive diagnoses of MS, the 2010 revisions to the McDonald Criteria focused on refining and establishing the definition of the dissemination of MRI lesions in both time and space.[15] This section reviews the conventional MRI modalities that were used in the work presented in this thesis.

1.3.1 T1-weighted Imaging

There are two types of spin relaxation properties: spin-spin relaxation and spin-lattice relaxation. Spin-spin relaxation refers to exponential decay of the transverse spin magnetization towards its equilibrium. Spin-lattice relaxation refers to recovery of the longitudinal spin magnetization towards its equilibrium. Accordingly, T1 (spin-lattice relaxation time) is defined as the time when 63% of the longitudinal magnetization has been recovered.[51] T1-weighted imaging uses a short repetition time (TR) and short echo time (TE) to achieve tissue contrast based on T1 values.[51] In general, fat has a short T1 and a long T2 while in the case of water, both are long. Therefore, tissues with high fat content, such as WM, appear bright. Regions of high water content, such as cerebrospinal fluid (CSF), however, appear dark. Correspondingly, GM is darker than WM (Figure 1-1).

T1-weighted hypointense WM lesions refer to areas on T1-weighted images with reduced signal intensities relative to surrounding WM. An MRI-histopathology comparison study showed that the degree of hypointensity (i.e. isointense, mildly hypointense, severely hypointense) in chronic T1-weighted lesions was strongly correlated with decreasing axonal density, indicating that these "black holes" represent areas of severe tissue destruction.[52] In particular, severely hypointense lesions were associated with a complete loss of axons.[52] Indeed, hypointense lesion loads are significantly correlated with higher EDSS scores,[53,54] more rapid EDSS progression rate in SPMS patients,[54] and longer disease duration.[55] Also, patients with SPMS or PPMS were shown to have significantly greater T1-weighted lesion volume (T1LV) and higher T1/T2 lesion ratios (i.e. how much of T2-weighted lesions are also T1-weighted lesions).[55] T1-weighted lesion load was shown to be better correlated to disability than was T2-weighted lesion load.[53,54] However, not all hypointense lesions are chronic; acute focal inflammation can also produce T1 hypointensity. One study showed that over 6 months, less than half of initially hypointense lesions stayed hypointense while the rest resolved and became isointense.[56]

T1-weighted lesions that enhance upon injection of gadolinium (Gd) contrast agent (i.e. Gd-enhancing lesions) represent areas of blood-brain barrier breakdown.[57] The enhancement mainly occurs in early, active lesions with intense inflammatory activity but not in chronic active/inactive lesions.[57,58] The vast majority of Gd-enhancing lesions usually resolve within 6 months.[59] Therefore, Gd-enhancing lesions are a sensitive marker of active inflammation.

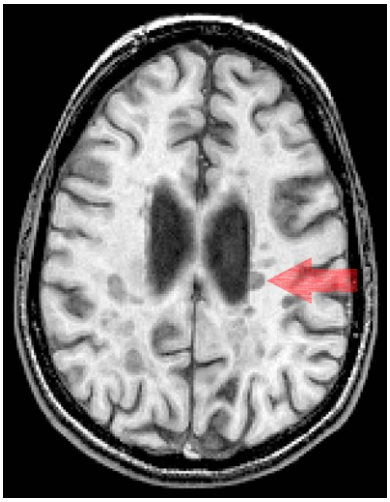

Figure 1-1: Example of a T1-weighted image

The arrow indicates a T1-weighted lesion.

1.3.2 T2-weighted Imaging

T2 (spin-spin relaxation time) is defined as the time when 63% of the transverse magnetization has decayed.[51] T2-weighted imaging uses a long TR and long TE to achieve tissue contrast based on T2 values.[51] As opposed to T1-weighted imaging, tissues with high fat content, such as WM, appear dark. Regions of high water content, such as CSF, appear bright. Here, WM is darker than GM (Figure 1-2).

T2-weighted imaging is sensitive to WM abnormalities caused by a variety of processes, including those in normal aging, migraine, ischaemic changes, chemotherapy, and multiple sclerosis.[60] Therefore, T2-weighted hyperintense WM lesions are sensitive, but not specific markers of pathology. For example, T2-weighted imaging can detect early reactive MS lesions, which are not usually as evident on T1-weighted imaging. However, T2-weighted lesions can contain various degrees of axonal and myelin loss, indicating the lack of specificity.[61] Fisher et al. showed, for instance, in a post-mortem MRI-histopathology comparison study that only 55% of T2-weighted hyperintense regions were demyelinated (as opposed to 83% in the regions abnormal on T2-weighted, T1-weighted, and magnetization transfer ratio (MTR) images). On the other hand, 0% of the T2-hyperintense only regions were likely to be chronic inactive lesions (as opposed to 68% of the regions abnormal on T2-weighted, T1-weighted, and MTR images).[62] This could contribute to the reasons why T2 lesion burden is relatively weakly correlated to disability, compared to T1-weighted lesions or MTR.[53,61,63]

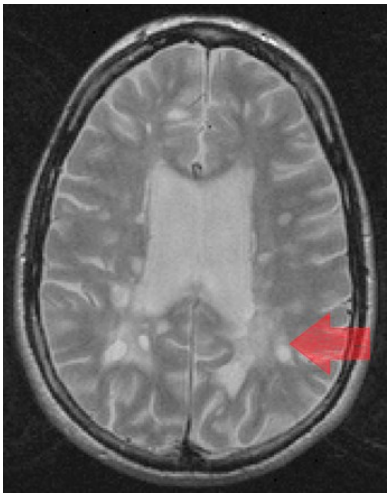

Figure 1-2: Example of a T2-weighted image

The arrow indicates a T2-weighted lesion.

1.3.3 Proton Density-weighted Imaging and Fluid Attenuated Inversion Recovery

Proton density-weighted imaging uses a long TR and short TE to minimize the effects of both T1 and T2.[51] Here, the signal intensity depends on the density of protons in the tissue. Due to the use of a long TR, proton density-weighted images can be acquired at the same time as T2-weighted images, using a dual-echo sequence.

One issue with T2-weighted imaging is that identification of lesions that are near the ventricles or around the cortex can be confounded by high signal from the CSF. This problem can in part be overcome by using the fluid attenuated inversion recovery (FLAIR) sequence, which is a type of T2-weighted imaging that uses a long TR and long TE, plus an extra radiofrequency pulse that suppresses signal from CSF. Therefore, tissue with high water content appears bright but the CSF appears dark. This approach is especially useful for detecting periventricular, cortical, or juxtacortical lesions that are often masked by the adjacent CSF.[51,64]

1.4 Therapies for Multiple Sclerosis

At the present time, MS cannot be cured. The management of MS involves a comprehensive approach that includes symptom management, relapse management, quality of life intervention, and disease modifying therapies (DMT)s.

Numerous drugs are available to treat MS symptoms, such as bladder and bowel dysfunction, depression, fatigue, or pain. Acute relapses are treated with high-dose corticosteroids, which may be effective in shortening the duration of a relapse and limiting the residual neurological deficits.[65–67] However, corticosteroids are associated with serious side effects when administered chronically, and their long-term efficacy on prevention of relapse recurrence and reduction in disability is unclear. As a result, corticosteroids are no longer used as a DMT.[65,66] DMTs are immunomodulatory agents that play a key role in the treatment strategy as they are, to a certain extent, capable of modifying and slowing the natural course of MS. According to the Multiple Sclerosis Society of Canada, there are 11 DMTs approved in Canada as of April 2016: Avonex (Interferon (IFN) B-1a, intramuscular, for RRMS, SPMS with relapses, CIS), Betaseon (IFN B-1b, intramuscular, for RRMS, SPMS with relapses, CIS), Extavia (IFN B-1b, subcutaneous, for RRMS, SPMS with relapses, CIS), Rebif (IFN B-1a, subcutaneous, for RRMS, SPMS with

relapses, CIS), Plegridy (pegIFN B-1a, subcutaneous, for RRMS), Copaxone (glatiramer acetate, subcutaneous, for RRMS, CIS), Tecfidera (dimethyl fumarate, oral, for RRMS), Aubagio (teriflunomide, oral, for RRMS), Gilenya (fingolimod, oral, for RRMS), Tysabri (natalizumab, intravenous, for RRMS), and Lemtrada (alemtuzumab, intravenous, for RRMS).

All of these drugs are intended for relapsing forms of MS. Current DMTs have been approved based on evidence of effectiveness on endpoints related to relapses and inflammation; these include reduction in relapse rate, reduction in sustained accumulation of disability, and reduction in MRI activity (e.g. volume and number of T2-weighted lesions, number of gadolinium enhancing lesions). Patients with PPMS or SPMS without relapses generally do not respond to these agents, and currently there is no DMT approved for treatment of PPMS (although a novel B cell therapy (ocrelizumab) has recently shown modest efficacy in PPMS, and is likely to be approved for this indication).[68]

Based on their risk-benefit profile, current DMTs are categorized into first-, second-, and third-line treatments. This classification system differs slightly between countries, but the first-line drugs generally include IFN B-1a and -1b, glatiramer acetate, dimethyl fumarate, and teriflunomide.[69] A patient typically starts with a first-line drug, unless he has a highly active MS. If the patient has poor response to, or has safety/tolerability issues with the drug, he may be prescribed another first-line drug, or even a later-line drug, which can be more effective but potentially less safe.[69]

There are several limitations with the current generation DMTs. They need to be administered continuously, and are very expensive, with the cost ranging from $15,000-$34,000 per year in Canada.[70] Yet, they are only partially effective and many patients continue to experience relapses, MRI activity, and progression.[71,72] Moreover, they are not effective once a patient enters the progressive stage of MS, as they do not have neuroprotective or regenerative effects. While short-term effects have been studied, long-term effects of the DMTs are not established. On the whole, DMTs are associated with serious side effects. For example, natalizumab is associated with activation of JC virus, which causes potentially fatal progressive multifocal leukoencephalopathy.[71,72]

1.5 Autologous Hematopoietic Stem Cell Transplantation for treatment of Multiple Sclerosis

Autologous hematopoietic stem cell transplantation (aHSCT) as a promising alternative approach to the treatment of MS that is based on the theory that a one-time "resetting" of a patient's immune system can produce prolonged remission of disease activity. In principle, ablation of the immune system can permanently eliminate autoreactive pathogenic clones. Subsequently, reinfusion of purified hematopoietic stem cells (HSC)s can regenerate an antigen-naïve immune system without the autoimmune memory cells. Since the first trial in 1995[73] there have been more than 20 aHSCT trials worldwide that treated patients with different MS phenotypes.[74,75] While the trials shared the general goal of immune-resetting, each used a different approach in terms of the inclusion criteria, different immunoablative conditioning regimens, and different methods of HSC purification. The studies have reported generally good outcomes in terms of reducing relapse rates and focal WM lesion activity on MRI. Still, there exist debates on the factors that can potentially influence the effectiveness and toxicity of aHSCT.

The administration of aHSCT has been mainly reserved for patients with 1) significant disability, reflected in the median baseline EDSS of 6.0 or higher in most of the treated cohorts, and 2) poor prognosis, as represented by ongoing relapses, increasing EDSS score, and MRI activity despite the use of conventional DMTs.[75] This is due to the fact that aHSCT is associated with high toxicity and transplant-related mortality. The mortality rate from 1995 to 2000 was 7.3%, although more recently it has fallen to 1.3%.[76] Therefore, an important issue is the selection of patients who can benefit most from the procedure while withstanding the risks. Saccardi et al. reported that younger ages (e.g. less than 40) may be associated with lower mortality and higher progression-free survival rates.[77]

It is difficult to directly compare results across different trials due to the differences in the inclusion criteria, treatment procedures, follow-up durations, and definitions of endpoints. Still, aHSCT appears to be effective for reducing disability progression and inflammatory activity. Reported three year progression-free survival rates include 36%,[78] 73%,[79] and 91%[80]. Furthermore, 5 year progression-free survival rates include 66%,[81] 77%,[82] and 82%.[83] Relapse-free survival rates included 86.3% at 3 years,[80] and 85% at 5 years[81]; some studies even reported

zero relapses by 2 years[84] and up to 10 years follow-up (FU).[85] Some studies have also reported improvements in the EDSS or MSFC scores in some patients.[80,81,85–88]

Significant reductions in MRI lesion activity have also been reported. Mancardi et al. showed that the number of Gd+ lesions decreased from 341 prior to aHSCT to 5 at 3 months FU, and none subsequently until the end of FU.[89] 5 year MRI-free survival rate was 75% in the HALT-MS study[80], and 85% in the Swedish study.[82] Notably, the Canadian MS – Bone Marrow Transplantation (MS-BMT) study reported zero new Gd+ or T2-weighted lesions for up to 10 years FU.[85] Reduction of the total T2-weighted lesion volume has also been reported by multiple studies.[80,86] Taken together, these results suggest that younger patients with active inflammation and relatively less disability are more likely to benefit from aHSCT.[90]

The choice of the immunoablative conditioning regimen is a matter of active discussion, as it may have a direct impact on the effectiveness and toxicity of aHSCT. On one hand, a more intensive, myeloablative regimen can better suppress episodic inflammation. On the other hand, it means the toxicity of the regimen is proportionately higher. Similarly, a less intense regimen may have less side effects but at the expense of its ability to control inflammation.[75] The most widely used regimen in aHSCT for MS has been BiCNU, Etoposide, Ara-C and Melphalan (BEAM), an intermediate intensity chemotherapy regimen. High intensity approaches like busulfan chemoregimen and total body irradiation (TBI) have also been used, although it is now being recognized that TBI generally leads to poorer outcomes and high toxicity/mortality rate.[75] Again, it is difficult to directly compare the effectiveness and toxicity of these regimens due to the lack of a comparison study. However, it is clear that it will be important to tailor a regimen that can maximize the benefits while minimizing toxicity.

All immunoablative regimens are inherently cytotoxic. Side effects are unavoidable and include neutropenic fever, infectious complications, urinary tract and respiratory infections, and liver toxicity. Chemotherapeutic agents also induce CNS toxicity and are lethal to progenitor cells and oligodendrocytes.[91] Accordingly, aHSCT may be associated with treatment-related brain atrophy. This is further discussed in section 1.6.6.

With the exception of the ASTIMS trial, all aHSCT trials for MS have been observational cohort studies without control groups.[92,93] This poses some limitations on comparison of the degree of disease control between aHSCT and DMTs. Notably, the ASTIMS phase-2 randomized trial

showed that the patients treated with a BEAM regimen-based aHSCT had significantly lower cumulative numbers of new T2-weighted lesions, Gd+ lesions, and relapse rates compared to those treated with mitoxantrone.[92] This suggests that aHSCT may be a promising approach for treatment of MS, if its risks can be ameliorated. A randomized controlled trial of aHSCT is necessary to show this definitively, and has been proposed.[94]

The work presented in this thesis uses data from two aHSCT trials for MS: the Canadian MS-BMT study, and the HALT-MS study.[80,95] Both studies had similar aims: the Canadian MS-BMT aimed to eliminate inflammatory activity so as to preserve remaining neurological function and prevent further irreversible damage; the HALT-MS study "hypothesized that control of inflammation in earlier RRMS may provide prolonged remission with the potential to reverse neurologic dysfunction."[80] The Canadian MS-BMT study treated 23 patients (11 SP, 12 RR) with a high-intensity regimen comprising busulfan, cyclophosphamide, and rabbit antithymocyte globulin (ATG); the HALT-MS study treated 24 RRMS patients with an intermediate-intensity regimen comprising BEAM and rabbit ATG. Both cohorts had failed conventional DMTs for MS. The treatment procedure is summarized as follows: 1) baseline evaluation, 2) mobilization of HSCs with granulocyte colony-stimulating factor to increase the number of HSCs in the blood stream, 3) HSC collection followed by cluster of differentiation (CD) 34+ HSC selection to leave out any remaining effector cells, 4) immune-ablation through chemotherapy using chemotherapeutic conditioning regimen, 5) aHSCT, and 6) clinical and MRI follow-up (Figure 1-3). Subject demographics for these two cohorts are shown in Table 1-1.

Figure 1-3: Outline of the procedures for autologous hematopoietic stem cell

Table 1-1: Subject demographics for the Canadian MS-BMT and the HALT-MS cohorts

	Canadian MS-BMT	**HALT-MS**
MS Subtype	24 SP or RRMS	25 RRMS (24 underwent treatment)
Mean baseline Expanded Disability Status Scale score (SD)	5.0 (1.1)	4.4 (0.6)
% of subjects with Gd+ lesions, during baseline prior to aHSCT	63 (SP: 86, RR: 50)	46
Mean baseline T1-w lesion volume (SD), ml	7.9 (7.8)	1.2 (2.6)
Mean baseline T2-w lesion volume (SD), ml	20.0 (18.0)	10.9 (13.4)
Conditioning regimen	Busulfan, Cyclophosphamide, Rabbit ATG	BEAM, Rabbit ATG
# of subjects with clinical relapses during follow-up	0	3 by year 5
# of subjects with two or more Gd+ and/or new T2 lesions during follow-up	0	2 by year 5
# of subjects with EDSS increase > 0.5	7 (by year 1.5 to 6)	2 by year 5
Death	1 (due to treatment related chemotoxicity)	2 (1 MS progression, 1 asthma)

1.6 Brain Atrophy in Multiple Sclerosis

Brain atrophy, defined as the loss of brain tissue, represents the net effect of the destructive pathological processes that occur in MS (Figure 1-4). Here, brain atrophy is defined as a loss of brain volume. Changes in whole-brain (WB) volume can be measured using MRI with high sensitivity and reproducibility. MRI studies have shown that MS patients suffer significantly higher rates of WB atrophy compared to healthy control subjects.[96] Brain atrophy is correlated with disability progression,[97] and its rate differs between MS subtypes.[98] Accordingly, WB atrophy is often used as an outcome measure in the evaluation of DMTs.

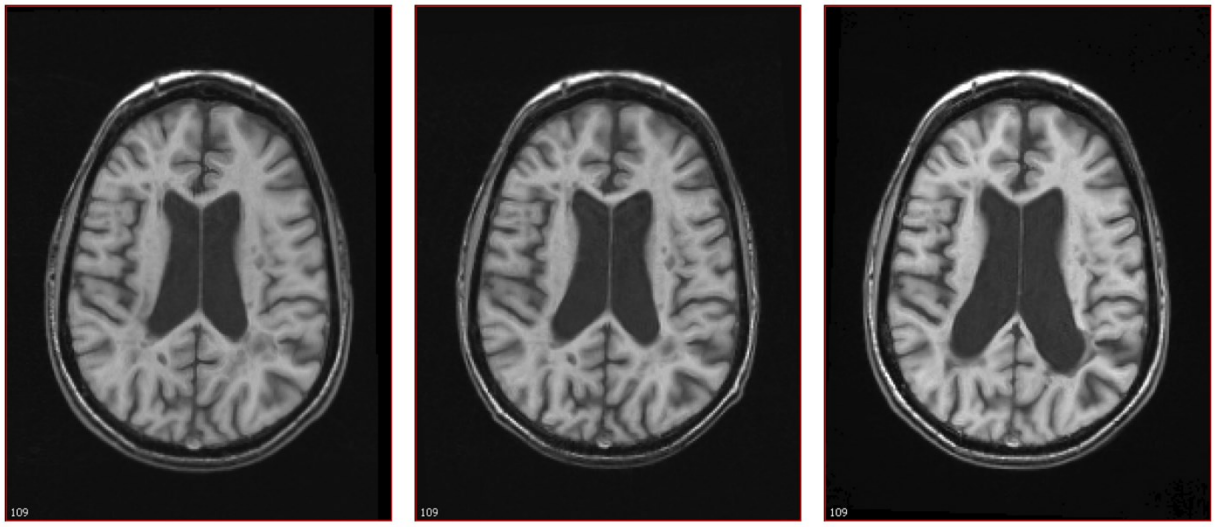

Figure 1-4: Example of brain atrophy in a patient over 10 years

(Left: Baseline, Center: 1y follow-up, Right: 10y follow-up)

Brain atrophy in MS affects both WM and GM.[99] Focal WM lesions represent sites of extensive axonal, oligodendrocyte, and myelin injury and their volumes can shrink as they mature.[30] From early phases of MS, significant Wallerian degeneration takes place not only within lesions but also in peri-plaque NAWM.[100] Even in their subacute/chronic phases though, WM lesions may be associated with atrophy of their surrounding WM regions.[101]

Although it has been suggested that GM atrophy may be, at least in part, independent of WM disease, this hypothesis is still the subject of research. It has been shown that WM lesions in the optic pathway can explain deep GM (lateral geniculate nucleus) atrophy, suggesting potential retrograde degeneration.[102] Another study showed that WM lesion probability within the ipsilateral region was significantly correlated with regions of deep GM atrophy.[103] Several studies have shown furthermore that increasing T2-weighted lesion volume (T2LV) is associated with greater cortical and subcortical GM volume loss in RRMS patients.[104,105] Indeed, significant correlations between WM lesion loads and GM atrophy were found.[106,107] Additionally, it is likely that the effect of WM lesions on cortical atrophy depends on location.[108]

As described in previous sections, GM in MS patients can be affected by extensive demyelination and neuroaxonal injury. Therefore, it is possible that pathological processes within GM itself also contribute to GM atrophy. Indeed, cortical lesions are a significant correlate of EDSS progression and cognitive impairment.[109] While some studies found a significant correlation between GM lesions and GM atrophy,[110] others did not.[111] This discrepancy is potentially due to the fact these MRI studies have used GM lesions visible on double-inversion recovery MRI. Double inversion recovery (DIR) is relatively more sensitive to GM lesions compared to T2-weighted or FLAIR sequences but still misses most of the cortical and deep GM lesions, especially the intracortical and the subpial types. For example, a double-inversion recovery study using a 1.5 T scanner missed 92% and 93% of intracortical and subpial lesions, respectively.[112] Using a 3 T scanner led to increased sensitivity compared to using a 1.5 T scanner, but still missed most of the true cortical lesions.[113] Using a 7 T scanner led to a further increase in sensitivity to intracortical and subpial lesions, but still missed over 40% of the histopathologically-confirmed lesions.[114] Therefore, it may take a long time before GM lesions can be reliably detected using MRI.

Given the extensive GM pathology in MS, measurement of GM atrophy is becoming an area of active research. GM atrophy is significantly correlated with cognitive impairment and disability,[99,106] and some studies have found that it is less likely to be affected by pseudoatrophy.[115,116] However, GM atrophy is relatively more difficult to measure because the cortex is thin and convoluted, and the GM-WM boundary is often ill-defined.[117]

This section first outlines several atrophy measurement techniques that are commonly used in the MS field. Structural Image Evaluation, using Normalisation, of Atrophy (SIENA) and Pairwise Jacobian Integration (PJI), the two atrophy measurement techniques used in this thesis, are discussed in more detail. Then, it reviews important findings from the previous studies on brain atrophy in normal aging, brain atrophy in MS, pseudoatrophy, and brain atrophy after aHSCT for treatment of MS. Effects of MRI scanner changes/updates, potential factors that confound the interpretation of atrophy findings, are also discussed. Finally, it discusses the mixed-effects model, the main statistical analysis approach used in this thesis.

1.6.1 Techniques for Measurement of Brain Atrophy

Various methods have been developed for quantitative measurement of brain volume change. An early approach to measure cerebral volume was to segment 20mm central slices of the brain (i.e. "four contiguous slices from each scan with the most caudal at the level of the velum interpositum cerebri"), as opposed to the WB.[118] This method has the scan-rescan coefficient of variation of 0.56%.[118] Currently however, WB segmentation approaches are generally used. The most widely used segmentation-based method in MS is brain parenchymal fraction (BPF).[119] BPF is defined as the ratio of segmented brain parenchymal volume over the total volume contained within the brain surface contour. Therefore it is intrinsically normalized and also relatively insensitive to gradient distortion.[119] The scan-rescan coefficient of variation for BPF is 0.19%.[119] A related method for the measurement of GM and WM volume change is grey matter fraction (GMF) and white matter fraction (WMF). Here, the denominator of the ratio stays the same but the numerators are grey matter volume (GMV) and white matter volume (WMV), respectively. The scan-rescan coefficient of variation for GMF is 1.1%.[99] Note that BPF, GMF, and WMF use only one image to calculate the brain volume; they are cross-sectional methods.

A widely used longitudinal registration-based method in the MS field is SIENA.[120] In short, SIENA estimates the brain-CSF edge displacements between two images and calculates the percentage brain volume change between them. This method is further reviewed in the next section. The median absolute error of SIENA is 0.15%.[120] A cross-sectional variation of SIENA is SIENAX, which measures volumes of WB, GM, and WM normalized (to the skull); the median absolute error of SIENAX is 0.5-1%.[120] Boundary shift integral (BSI) is another

registration-based method that is widely used in the Alzheimer's disease (AD) field. BSI integrates the intensities within the regions of brain-CSF boundary shift between two images, normalizes the intensities, and estimates the volume change between them.[121] A comparison of SIENA and BSI reported that both methods have similar error of 0.2%, and are highly correlated ($r=0.87$, $p<0.0001$).[122] However, it was also shown that SIENA systemically reports larger volume changes by about 20% compared to BSI.[122]

Non-linear deformation-based methods include voxel-based morphometry (VBM) and tensor-based morphometry (TBM). Briefly, VBM non-linearly registers all input images into a symmetric common space, segments them into GM, WM, and CSF, and then performs group comparisons on a voxel-by-voxel basis.[123] TBM is a general term for the methods that identify local structural differences between brain images by utilizing the Jacobian determinant of the deformation fields that nonlinearly register one image to another.[124] The work presented in this thesis uses PJI, which is a type of TBM that calculates percentage WB, GM, and WM volume changes; the scan-rescan absolute error is 0.32±0.24% for the GM volume.[117] This method is further discussed in the next section.

1.6.2 Description of SIENA and PJI

This section describes the longitudinal atrophy measurement techniques used in the work presented in this thesis: SIENA and PJI.[117,120] Both use two T1-weighted scans of a same subject (timepoint 1 referred to as "first", timepoint 2 referred to as "second") as inputs. In this thesis, the first timepoint generally corresponds to the baseline evaluation MRI scan, whereas the later timepoints correspond to subsequent follow-up scans.

SIENA calculates a percentage brain volume change (PBVC) between two scans for the whole-brain.[120] First, brain extraction is done to output a binary brain mask, the segmented brain image, and an external skull surface image for each scan.[125] A feature of SIENA is the use of the exterior skull surface during registration steps; this is to reduce the potential confounding effect of imaging geometry changes that may confound the true brain volume changes.[125,126] In estimating a constant scale factor, the exterior skull surface, which is assumed to be constant in size and shape, is used as a scaling constraint in the registration.[125] Second is the registration step where the second brain is registered to the first brain via a set of linear registrations that

incorporates the scaling and skew constraints produced from skull registrations. An issue with this approach is that the second brain image has been through several image processing steps, resulting in slight blurriness compared to the first image. To address this interpolation bias,[127] SIENA decomposes the total (i.e. from the second to the first) transformation into two, and transforms both the first and the second images to a halfway position between the two; this results in both images having gone through similar degree of processing.[125] The output of this step is pairs of the registered head images and brain masks. Third, the two brain masks are combined to prevent possible mis-segmentations from affecting the change analysis, and then applied to the registered head images to produce two registered brain images.[125] Fourth step is the edge motion detection. Initially, SIENA finds all brain surface edge points, including those in the brain-ventricle boundary.[120] Next, for each edge point in the first brain, the image gradient direction is found and used to find the surface normal unit vector.[120] This information is also used to find the direction (i.e. "atrophy" or "growth") of the motion.[120] Next is a 1-D example of the edge displacement calculation: for a point in the first scan, an intensity profile perpendicular to the edge (represented as a "1-D array") is created and filled with the image values.[120] Then, a second intensity profile is filled with image values of exactly the same positions, but this time from the second scan.[120] The relative shift between the two intensity profiles is used to estimate the edge motion.[120] Notably, this step involves additional preprocessing in a way that the derivatives of the profiles, rather than the raw profiles, are compared to measure the edge motion; this is done to reduce the effects of intensity or contrast differences between the two scans.[120] The final step is the PBVC calculation. Initially, the mean perpendicular brain surface motion is calculated by the following formula: $\mathrm{l} = \frac{v \sum m}{aN}$, where "l" is the mean surface motion, "$\sum m$" is the summation of edge motion over all edge points, "v" is the voxel volume, "N" is the number of edge points, and "a" is voxel cross-sectional area.[120] It should be noted that the quantity "v$\sum m$" represents the total change in volume between the two scans; however this value is influenced by the number of edge points, which is inversely proportional to the slice thickness.[120] Therefore, dividing "v$\sum m$" by "aN" is done to reduce the influence of slice thickness by normalizing the total volume change for the number of edge points found.[120] Finally, the PBVC is calculated by the following formula: $\mathrm{PBVC} = \frac{l*A}{V} * 100$, where A is the actual brain surface area, and V is the actual brain volume.[120] The error associated with this method is about 0.2%.[120]

SIENAX is the cross-sectional variation of SIENA; it also uses the skull as a scale constraint in the process of spatially normalizing a single image with respect to a standard space.[120] The output is the normalized brain volume (NBV) as well as normalized grey matter volume (NGMV) and normalized white matter volume (NWMV).[120]

This paragraph outlines a longitudinal implementation of PJI by Nakamura et al., which can be used to calculate a percentage change in volume between two scans from a same subject for the whole-brain, GM, and WM.[117] Similarly to SIENA, the method starts by registering the first and the second scans to each other, using the skull as a scale constraint.[117] Then, the images are transformed into a halfway space.[117] Nonlinear registration of the two transformed images outputs a series of deformation fields, which contain information about the difference in shape of the brain between the two images.[123] For instance, deformation fields can be thought of as vector fields that map points in one image to corresponding points in another image; the Jacobian of the deformation field describes local shear, stretch, and rotation associated with the mapping.[124] TBM utilizes the property that the determinant of the Jacobian matrix is the ratio of the second to the first volume elements, thereby giving information about local volumetric expansion (i.e. $\det J > 1$) or shrinkage (i.e. $\det J < 1$).[123] PJI estimates percent changes in brain volume by averaging the volumetric changes within the brain mask; furthermore, GM and WM volume changes can be calculated by averaging the volumetric changes within the GM and the WM masks, respectively.[117]

An important issue in the measurement of GM and WM atrophy in MS is the presence of WM lesion, which causes bias in segmentation in a way that increasing lesion volume leads to underestimation of GMV and hence overestimation of WMV.[128] One way to reduce the influence of WM lesion on segmentation is to fill the lesion areas with intensities derived from their surrounding NAWM.[129] This approach is implemented within the preprocessing steps of both PJI and SIENA/SIENAX.[117,129]

1.6.3 Brain Atrophy in Normal Aging

It is well known that a human brain grows and then shrinks with increasing age. The effect of normal aging on brain volume is a topic of active research. Several studies have proposed various regression models to explain the age effect on GM and WM volumes during adulthood

(Table 1-2 lists examples of these findings, focusing on those that have included the rates during young/middle adulthood, which is the range most relevant to the patients treated with aHSCT). It appears that the numerical results somewhat vary between studies, probably due to the differences in subjects and analysis methods. However, the common pattern is that the GM volume declines linearly throughout the adulthood years whereas the WM volume increases until late 30's/early 40's, plateaus until about 50's, and then declines thereafter.

Longitudinal studies generally provide greater statistical power for detecting small changes. Several studies have used longitudinal designs to directly measure whole volume changes in healthy subjects (Table 1-3 lists examples of these findings, focusing on those that have included the rates during young/middle adulthood). Again, the numerical results vary among studies due to the differences in subjects, imaging protocols, and analysis methods. However, it can be approximated that from the ages 30 to 60, the average WB atrophy rate falls somewhere between -0.2% to -0.3% per year.

Table 1-2: Effect of normal aging on grey matter and white matter volume during young/middle adulthood

Study	Analysis Method	Subjects	GMV and WMV change findings
Jernigan 2001[130]	Semi-automated brain extraction and tissue segmentation (GM, WM, CSF)	78 healthy volunteers Sex: 41 F, 37 M Age range: 30 to 99, (mean: 64, SD: 17.4) Cross-sectional	14% of the cerebral cortex, 35% of the hippocampus, and 26% of the cerebral WM volume lost between the ages of 30 to 90. The rates of volume loss started to accelerate around age 50.
Ge 2002[131]	Semi-automated brain extraction and tissue segmentation (GM, WM, CSF)	54 healthy volunteers Sex: 32 F, 22 M Age range: 20 to 86, (mean: 46.8, SD: 19.3) Cross-sectional	**GMV**: A linear model estimated a steady change of -0.09%/y between the ages of 20 to 86. **WMV**: A quadratic model estimated volume growth until age 40, and decline thereafter. Once the WMV started to decline, its rate was faster than that of GMV.
Sowell 2003[132]	Brain extraction, tissue segmentation (GM, WM, CSF), tissue parcellation	176 healthy volunteers Sex: 86 F, 90 M Age range: 7 to 87, (mean: 31, SD: 21.3) Cross-sectional	**GMV**: A nonlinear decline in GMV that was most rapid between the ages of 7 to 60. **WMV**: A quadratic model estimated volume growth until age 43, and declined thereafter.
Liu 2003[133]	Semi-automated brain extraction, segmentation, and then percentage volume change over 3.5y	90 healthy volunteers Sex: 41 F, 49 M Age range: 14 to 77 Longitudinal, with average interval between scans: 3.5y	**GMV**: Age group ≤ 34, mean: -0.11%/y Age group 35-54, mean: -0.02%/y Age group ≥ 55, mean: -0.27%/y **WMV**: Age group ≤ 34, mean: 0.07%/y Age group 35-54, mean: -0.43%/y Age group ≥ 55, mean: -0.58%/y
Allen 2005[134]	Automated brain extraction and tissue segmentation (GM, WM, CSF). Then, manual parcellation (frontal, temporal, parietal, and occipital lobes)	87 healthy volunteers Sex: 44 F, 43 M Age range: 22 to 88, (mean: ~48, SD: ~18) Cross-sectional	**GMV**: A linear model estimated about 12% cortical volume loss between the ages of 30 to 80 (i.e. 0.24%/y) **WMV**: A quadratic model estimated volume growth until about age 50, and then decline; the decline started to accelerate around age 65.
Van Haren 2008[135]	Semi-automated brain extraction and tissue segmentation (GM, WM, CSF)	113 HC Sex: 37 F, 76 M Age range: 17 to 56, (mean: 35) Longitudinal, with mean FU duration: 5y	**GMV**: -0.5%/y (-2.5% over 5y) **WMV**: 0.32%/y (1.6% over 5y)

Table 1-3: Longitudinal studies of the effect of normal aging on whole brain volume during young/middle adulthood

Study	Analysis method	Subjects	Whole-brain volume (WBV) change findings
Fox 2000[136]	BSI	26 HC Sex: 9 F, 17 M Age range: 30 to 59, (mean: 47.1) Average interval between scans: 1y	Median: -0.3%/y (IQR: -0.1 to -0.6)
Liu 2003[133]	Semi-automated brain extraction, and then percentage volume change over 3.5y	90 healthy volunteers Sex: 41 F, 49 M Age range: 14 to 77 Average interval between scans: 3.5y	Age group ≤ 34, mean: -0.06%/y Age group 35-54, mean: -0.18%/y Age group ≥ 55, mean: -0.39%/y
Scahill 2003[137]	BSI	39 healthy volunteers Sex: 21 F, 18 M Age range: 31 to 84 Average interval between scans: 1.7y	Whole age range, mean: -0.32%/y (95% CI: 0.10 to 0.54) Ages 30 to 39, N=8, mean: -0.25%/y Ages 40 to 49, N=10, mean: -0.26%/y Ages 50 to 59, N=10, mean: -0.24%/y Ages 60 to 69, N=6, mean: -0.3%/y Ages 70 to 84, N=5, mean: -0.4%/y
Henley 2006[138]	BSI	7 HC Sex: 4 F, 3 M Age range: n/a (mean: 40.7, SD: 10.5) Average interval between scans: 0.5y	Mean: -0.26%/y
Van Haren 2008[135]	Semi-automated brain extraction and tissue segmentation (GM, WM, CSF)	113 HC Sex: 37 F, 76 M Age range: 17 to 56, (mean: 35) Mean FU duration: 5y	Mean: -0.16%/y (-0.8% over 5y)
De Stefano 2015[96]	SIENA PBVC	35 HC Sex: 20 F, 15 M Age range: 21 to 60, (mean: 37) Mean FU duration: 6.3y	Mean: -0.27%/y ± 0.15%

1.6.4 Brain Atrophy in patients with MS

There are a huge number of studies that have reported brain atrophy outcomes of some sort. For instance, a simple PubMed search for "brain atrophy multiple sclerosis" finds 1600 articles. However, it is difficult to directly compare numerical results from different studies that have used different protocols and atrophy measurement techniques.

Still, cross-sectional studies have consistently shown that patients with MS have significantly lower brain volume compared to age-matched normal controls (Table 1-4 lists examples of these findings). These findings suggest that in general, SP patients have significantly more WB, GM, and WM atrophy compared to RR patients. SP and PP patients appear to have similar degrees of atrophy.

Longitudinal studies have shown that the rates of atrophy may differ between MS subtypes (Table 1-5 lists examples of these findings). Again, heterogeneities in the study design (e.g. differences in patient characteristics, DMT status) and analysis methods make it difficult to directly compare results from different studies. Overall, average WB atrophy rates in different cohorts of MS patients appear to range from -0.5 to -1.0%/y, which are about two to three times higher than that of normal aging. In general, atrophy appears to accelerate with disease progression.

Table 1-4: Cross-sectional findings of significant brain atrophy in patients with MS

Study	Analysis Method	Subjects	Findings
Rudick 1999[119]	BPF	16 HC vs. 68 RR (IFN B-1a) vs. 72 RR (placebo); Part of a longitudinal study	**HC vs. MS (IFN B-1a or placebo)**: Mean BPF at baseline was significantly lower in MS patients (0.831 [IFN B-1a] or 0.830 [placebo] vs. 0.871 [HC], p=0.0001)
Ge 2000[139]	Percentage brain parenchyma volume: Brain extraction, then normalized to the volume of the intracranial contents	20 HC vs. 36 MS (27 RR, 9 SP)	**MS vs. HC**: PBV was significantly lower in MS patients (mean 85.5 vs. 88.2%, p=0.007) **RR vs. SP**: No significant difference in PBV between the RR and the SP patients
Lin 2003[140]	Normalized brain volume: Brain extraction, then normalized to the total intracranial volume	31 HC vs. 97 MS (49 RR, 48 SP)	**RR vs. HC**: RR patients had significantly lower supratentorial (976 vs. 1033ml, p<0.05) and greater lateral ventricle (18.3 vs. 13.5ml, p<0.05) volumes **SP vs. HC**: SP patients had significantly lower supratentorial (941 vs. 1033ml, p<0.001), greater lateral ventricle (24.9 vs. 13.5ml, p<0.001), lower cerebellar (114.9 vs. 122.7ml, p<0.05), lower brainstem (17.2 vs. 18.7ml, p<0.001), and lower upper cervical cord (2.7 vs. 3.2ml, p<0.001) volumes
Tedeschi 2005[141]	WM and GM fractions, normalized to the intracranial volume	104 HC vs. 629 MS (427 RR, 140 SP, 30 PP)	**MS vs. HC**: MS patients had significantly lower WM-f (33.8 vs. 35.1, p<0.001) and GM-f (51.0 vs. 53.1, p<0.001) **SP vs. RR**: SP patients had significantly lower WM-f (32.8 vs. 34.2, p<0.001) and GM-f (50.0 vs. 51.4, p<0.001) **RR vs. PP**: No significant differences **SP vs. PP**: No significant differences
Pagani 2005[142]	Normalized brain volume (SIENAX)	10 HC vs. 70 MS (20 RR, 19 SP, 31 PP)	**Baseline NBV (ml, (SD))**: HC=1602, (80); RR=1517, (70); SP=1464, (90); PP=1473, (70)
Horakova 2008[143]	Normalized brain, GM, peripheral GM, WM volume (SIENAX)	27 HC vs. 147 RRMS; HC significantly older; Part of a longitudinal study	**Baseline NBV, median:** **NBV, peripheral NGMV, NWMV**: MS patients had significantly lower volumes. **Total NGMV**: No significant differences
Fisniku 2008[106]	WM and GM fractions, normalized to the intracranial volume	25 HC vs. 29 CIS vs. 44 MS (33 RR, 11 SP)	**MS vs. HC**: MS patients had significantly lower GM-f (0.49 vs. 0.51, p<0.001) and WM-f (0.28 vs. 0.29, p<0.017) **MS vs. CIS**: MS patients had significantly lower GM-f (0.47 vs. 0.50, p<0.001); WM-f difference was not significantly different (n.s). **CIS vs. HC**: GM-f and WM-f differences were n.s. **SP vs. RR**: SP patients had significantly lower GM-f (0.45 vs. 0.48, p<0.003); WM-f difference was n.s.

Fisher 2008[99]	BPF, GMF, WMF	17 HC vs. 7 CIS vs. 47 MS (28 RR, 19 SP); Part of a longitudinal study	**Mean BPF at baseline** **HC or CIS vs. SP**: SP patients had significantly lower BPF (0.801 vs. 0.862 [HC] or 0.861[CIS], $p<0.05$)
			Mean GMF at baseline **HC or CIS vs. SP**: SP patients had significantly lower GMF (0.528 vs. 0.554 [HC] or 0.551 [CIS], $p<0.05$) **SP vs. RR**: SP patients had significantly lower GMF (0.528 vs. 0.537, $p<0.05$)
			Mean WMF at baseline **HC, CIS, RR vs. SP**: SP patients had significantly lower WMF (0.280 vs. 0.308 [HC] or 0.309 [CIS], 0.304 [RR], $p<0.05$)
Chu 2015[144]	Normalized brain volume (SIENAX)	9 HC vs. 26 MS (22 CIS or RR, 4 SP or PP)	**MS vs. HC, on 1.5T scanner**: MS had significantly lower NBV (1465.7 vs. 1513.8 ml, $p=0.04$) **MS vs. HC, on 3.0T scanner**: MS patients had significantly lower NBV (1414.1 vs. 1485.1 ml, $p=0.006$)

Table 1-5: Longitudinal findings of brain atrophy rates in MS (treated or untreated)

Study	Analysis Method	Subjects	MS Treatment Status	Findings
Losseff 1996[118]	Central slices	**29 MS** (13 RR, 16 SP); FU duration: 1.5y	Anti-CD4 antibody	**Mean mlΔ/y** **MS**: -3.4ml/y (i.e. -1.1%/y)
Rudick 1999[119]	BPF	16 HC vs. **68 RR (IFN B-1a) vs.** **72 RR (placebo)**; FU: 2y	IFN B-1a or Placebo	**Mean BPF %Δ (SD)** **RR Placebo**: -1.22% (1.30) over 2y **RR IFN B-1a**: -0.996 (1.22) over 2y; 1st year rate: -0.763%/y, 2nd year rate: -0.233%/y
Fox 2000[136]	BSI	**26 MS (9 PP, 6 SP, 6 RR, 5 benign)** vs. 26 HC; FU: 1y	Untreated	**Median BSI %Δ/y (IQR)** **All MS subjects**: -0.8%/y (-0.2 to -1.0) **Benign**: -0.7%/y (-0.3 to -0.9) **RR**: -0.8%/y (0.2 to -0.9) **SP**: -0.6%/y (0.3 to -1.3) **PP**: -0.9%/y (-0.3 to -1.4)
Sormani 2004[145]	SIENA PBVC	**207 RRMS (102 GA, 105 placebo)**; FU: 1.5y	Glatiramer acetate (GA) or placebo	**Mean PBVC %Δ (SD)** **RR Placebo**: -2.0% (1.8) over 1.5y **RR GA treated**: -1.5% (1.6) over 1.5y
Jasperse 2007[146]	SIENA PBVC	**89 MS (74 relapse onset, 15 progressive onset)**; FU: 2.2y	Relapse onset: 28% on DMT; Progressive onset: 0% on DMT	**Mean PBVC %Δ/y (SD)** **All patients**: -0.9%/y (0.8) **Relapse onset**: -1.0%/y (0.8) **Progressive onset**: -0.9%/y (0.6)
Anderson 2007[147]	SIENA PBVC	16 HC vs. **33 RRMS**; FU: 3y	Untreated, but some went on DMT later on	**Mean PBVC %Δ (SD)** **RR**: -2.34% (1.33) over 3y
Anderson 2007[147]	BSI	16 HC vs. **33 RRMS**; FU: 3y	Untreated, but some went on DMT later on	**Mean BSI %Δ (SD)** **RR**: -1.88% (1.13) over 3y
Horakova 2008[143]	SIENA PBVC and SIENAX	27 HC vs. **147 RRMS**; FU: 5y	Avonex, Steroid, Azathioprine	**Median %Δ over 5y in RRMS patients who had a serial MRI at all timepoints, N=36**
				RR, PBVC: -3.65% over 5y, i.e. -0.73%/y
				RR, GMV: -5.9% over 5y, i.e. -1.18%/y
				RR, peripheral GMV: -5.5% over 5y, i.e. -1.1%/y
				RR, WMV: 0.5% over 5y, i.e. 0.1%/y

Fisher 2008[99]	BPF, GMF, WMF	**17 HC vs. 7 CIS vs. 47 MS (28 RR, 19 SP) vs. 8 CIS-MS converts vs. 7 RR-SP converts**; FU duration: 4y	IFN-B, glatiramer acetate, methotrexate, azathioprine	**Mean BPF %Δ/y (SD)** **HC**: -0.066 (0.22) **CIS**: -0.003 (0.15) **CIS-MS converts**: -0.15 (0.14) **RRMS**: -0.23 (0.32) **RR-SPMS converts**: -0.35 (0.18) **SPMS**: -0.39 (0.31)	Increasing BPF rate associated with increasing disease severity
				Mean GMF %Δ/y (SD) **HC**: -0.028 (0.24) **CIS**: -0.028 (0.25) **CIS-MS converts**: -0.096 (0.23) **RRMS**: -0.23 (0.34) **RR-SPMS converts**: -0.35 (0.37) **SPMS**: -0.39 (0.50)	Compared to HC, rate was 3.4x in CIS-MS, 8.1x in RR, 12.4x in RR-SP converts, 14x in SP
				Mean WMF %Δ/y (SD) **HC**: -0.076 (0.35) **CIS**: 0.11 (0.25) **CIS-MS converts**: -0.24 (0.29) **RRMS**: -0.24 (0.72) **RR-SPMS converts**: -0.33 (0.53) **SPMS**: -0.25 (0.49)	WMF rate was similar in all disease categories, all at 3x compared to HC
Kappos 2010[148]	SIENA PBVC	**429 RR (1.25mg fingolimod) vs. 425 RR (0.5mg fingolimod) vs. 418 placebo**; FU duration: 2y	Fingolimod 1.25mg or 0.5mg	**Fingolimod 1.25, mean PBVC %Δ (SD)**: -0.89% (1.30) over 2y	
				Fingolimod 0.5mg, mean PBVC %Δ (SD): -0.84% (1.31) over 2y	
				Placebo, mean PBVC %Δ (SD): -1.31% (1.50) over 2y	
Di Filippo 2010[149]	SIENA PBVC	**92 CIS**; FU duration: 1y	All untreated at baseline, but 2 patients on DMT at 1y FU	**Mean PBVC %Δ/y (SD)** **CIS**: -0.38%/y (0.55)	
Barkhof 2010[150]	SIENA PBVC	**90 Placebo (92% RR, rest SP) vs. 76 Ibudilast 30mg/d (94% RR, rest SP) vs. 82 Ibudilast 60mg/d (93% RR, rest SP)**; FU duration: 2y	Ibudilast 30mg/d or 60mg/d	**Ibudilast 60 mg/d, mean PBVC %Δ (SD)**: By 1st year: -0.79%/y (1.02) By 2nd year: -1.64% (1.67), i.e. -0.82%/y	
				Ibudilast 30 mg/d, mean PBVC %Δ (SD): By 1st year: -1.05%/y (1.03) By 2nd year: -1.97% (1.83), i.e. -0.99%/y	
				Placebo, mean PBVC %Δ (SD): By 1st year: -1.20%/y (1.15)	

De Stefano 2010[98]	SIENA PBVC	**963 MS (579 RR, 139 SP, 88 PP)**; FU: minimum 1y	Untreated	**Mean PBVC %Δ/y (SD)** **CIS**: -0.40%/y (0.47) **RR**: -0.49%/y (0.65) **SP**: -0.64%/y (0.68) **PP**: -0.56%/y (0.55)
Popescu 2013[151]	SIENA PBVC	**261 MS (18 CIS, 97 RR, 69 SP, 77 PP)**; FU: 8y	124 patients had DMT	**Median PBVC %Δ/y (IQR)** **All MS subjects**: -0.69%/y (-1.17 to -0.19) **CIS**: -0.31%/y (-0.49 to 0.15) **RR**: -0.69%/y (-1.2 to -0.19) **SP**: -0.81%/y (-1.2 to -0.23) **PP**: -0.64%/y (-1.2 to -0.19)
De Stefano 2015[96]	SIENA PBVC	**206 MS (180 RR, 14 SP, 12 PP)**; FU: 7.5y	85% on DMT during the study period	**Mean PBVC %Δ/y (SD)** **All MS subjects**: -0.51%/y (0.27) **RR**: -0.52%/y (0.29) **SP and PP**: -0.45%/y (0.18)

1.6.5 Pseudoatrophy

Several studies noted that some anti-inflammatory DMTs may have a delayed effect on reducing the rate of brain volume loss in MS patients (Table 1-6 lists examples of these findings).[119,152] Perplexingly, the reductions in these cases were preceded by periods of accelerated volume loss, especially during the first few months after initiation of the therapy. This phenomenon has been termed "pseudoatrophy" under the assumption that the accelerated atrophy was not due to actual tissue loss but was associated with the resolution of inflammation. Several studies suggested that the pseudoatrophy effect was more prominent in WM,[115,116,153] which is generally more inflamed than GM.[154] Two proposed mechanisms of pseudoatrophy are 1) resolution of edema (i.e. tissue water shifts) and 2) reduction in the volume of inflammatory cells.

Duning et al. demonstrated that fluctuations in hydration status were associated with reversible but significant whole-brain volume (WBV) changes.[155] This finding implies that if resolution of edema leads to relative dehydration of the brain, pseudoatrophy would presumably be paralleled by a decrease in the brain tissue water content. Changes in brain water content can be detected through changes in the overall T2 relaxation time of whole brain. The latter can be estimated from dual spin echo images by calculating the "pseudo-T2". The pseudo-T2 (pT2) MRI metric is sensitive to bulk brain tissue water content and is also significantly correlated with WBV change.[156] Accordingly, the brain volume loss due to tissue water shift would be accompanied by a decrease in the pT2.

Inflammation in MS is mediated by factors including T-cells, B-cells, macrophages, and microglial activation.[157] Returning of the microglial cells back to the resting state might also explain pseudoatrophy, but such process is indistinguishable from actual tissue loss using conventional MRI. Recently, Dwyer et al. showed that short-term (6 months) WB, GM, and WM volume losses in RRMS patients after IFN B-1a treatment were significantly correlated with decreased percentage in the blood of CD4+ T-cells expressing interleukin-17F, a marker of pro-inflammatory cytokines.[158]

Pseudoatrophy may also have affected MS patients treated with aHSCT, a powerful anti-inflammatory therapy. Indeed, all of the WB atrophy findings so far have suggested acceleration of the atrophy rates immediately following aHSCT. However, it must be noted that these patients were also subjected to immunoablative regimen-related CNS toxicity, which could have caused a

significant amount of actual tissue loss. For example, Petzold et al. showed that the level of neurofilament heavy chain NfH-SMI35 (a protein biomarker for neuroaxonal degeneration) significantly increased in MS patients following BMT using TBI conditioning regimen.[159] Therefore, pseudoatrophy is likely only a part of the processes behind brain atrophy in MS patients treated with aHSCT.

Table 1-6: Examples of acceleration of volume loss upon initiation of anti-inflammatory disease-modifying treatments

Study	DMT	Analysis Method	Subjects	Findings pertaining to the early accelerated volume loss. The numbers indicate average values (mean or median)
Rudick 1999[119]	IFN B-1a	BPF	RRMS, 68 Treated vs. 72 Placebo	**Treated vs. Placebo:** No significant difference in the % BPF reduction during the 1st year of FU (-0.996 vs. -1.22). During the 2nd year of FU, % BPF reduction was significantly lesser in the treated group (-0.233 vs. -0.521, p=0.03).
Molyneux 2000[160]	IFN B-1b	Central slices	SPMS, 360 Treated vs. 358 Placebo	**Treated vs. Placebo:** No significant difference in the % volume loss over 3y FU (-2.91 vs. -3.86). At 6m of FU, a trend towards greater volume reduction in the treated group (-1.39 vs. -0.89).
Rovaris 2001[161]	GA	Brain Parenchymal Segmentation	RRMS, 113 Treated vs. 114 Placebo	**Treated vs. Placebo:** No significant difference in the % volume reduction over 18m of FU (1.2 vs. 1.4). In the treated group, % change during the first 9m of FU (-0.8) was higher than that from 9m to 18m (-0.4).
Frank 2004[162]	IFN B-1b	Semi-automated Segmentation (DeCarli 1992)	RRMS, 30 Treated	The mean % brain volume changes from baseline to 1y, 2y, and 3y FU were -1.35, -1.48, and -1.68, respectively. This implies that the majority of the volume reduction occurred during the first year of FU.
Hommes 2004[163]	IV Immuno globulin	BPF	SPMS, 156 Treated vs. 158 Placebo	**Treated vs. Placebo:** Significantly less % BPF reduction in the treated group (-0.62 vs. -0.88) over 2y of FU. During the 1st year of FU, the median % BPF reduction was -0.4 for the treated group and -0.6 for the placebo.
Hardmeier 2005[164]	IFN B-1a	BPF	RRMS, 386 Treated	**Entire group:** % BPF reduction during the 1st year of FU (-0.686) was greater than those during the 2nd (-0.377) and the 3rd (-0.378) years of FU.
				Subgroup with frequent scans (N=138): % BPF reduction during the first 4m of FU was significantly greater than that from 4m to 12m (-0.482 vs. -0.228, p<0.05).
Miller 2007[152]	Natalizumab	BPF	RRMS, 627 Treated vs. 315 Placebo	**Treated vs. Placebo:** No significant difference in the % BPF reduction over 2y of FU (-0.80 vs. -0.82). During the 1st year of FU, % BPF reduction was significantly greater in the treated group (-0.56 vs. -0.40, p=0.002). During the 2nd year of FU, % BPF reduction was significantly lesser in the treated group (-0.24 vs. -0.43, p=0.004).

Mikol 2008[165]	IFN B-1a, GA	SIENA	RRMS, 230 IFN B-1a Treated vs. 230 GA Treated	**IFN B-1a group:** -1.240% PBVC over 96w of FU. Mean PBVC during weeks 0 to 48 was -0.82%, which was double the PBVC over weeks 48 to 96 (-0.41%).
				GA group: -1.073% PBVC over 96w of FU. Mean PBVC during weeks 0 to 48 was -0.76%, which was double the PBVC over weeks 48 to 96 (-0.32%).
O'Connor 2009[166]	IFN B-1b, GA	SIENA	RRMS, 899 IFN B-1b Treated (500µg) vs. 897 IFN B-1b Treated (250µg) vs. 448 GA Treated	**IFN B-1b (500µg) group:** Mean PBVC over 3y FU was -0.64%. Mean PBVC from screening-y1, y2-y2, and y2-y3 were approximately -0.9%, +0.25%, and +0.1%, respectively.
				IFN B-1b (250µg) group: Mean PBVC over 3y FU was -0.65%. Mean PBVC from screening-y1, y2-y2, and y2-y3 were approximately -0.85%, +0.2%, and +0.1%, respectively.
				GA group: Mean PBVC over 3y FU was -0.61%. Mean PBVC from screening-y1, y2-y2, and y2-y3 were approximately -0.75%, +0.15%, and +0.1%, respectively.
Montalban 2009[167]	IFN B-1b	BPF	PPMS, 36 IFN B-1b treated vs. 37 Placebo	**IFN B-1b group:** Mean BPF % change over 2y of FU was -1.013. Changes during the 1st year and the 2nd year of FU were -0.659 and -0.350, respectively.
				Placebo group: Mean BPF % change over 2y of FU was -0.461. Changes during the 1st year and the 2nd year of FU were -0.203 and -0.288, respectively.
Radue 2010[168]	Natalizumab, IFN-B1a	BPF	RRMS, 582 Placebo + IFN-B1a vs. 589 Natalizumab + IFN-B1a	**Placebo + IFN-B1a group:** Mean BPF % change from baseline to 2y FU was -0.82. The reduction during the 2nd year (-0.40) was similar to that during the 1st year (-0.42).
				Natalizumab + IFN-B1a group: Mean BPF % change from baseline to 2y FU was -0.81. The reduction during the 2nd year (-0.31) was less than that during the 1st year (-0.50).
Vidal-Jordana 2013[115]	Natalizumab	SIENA (WB), SPM5 (GM, WM)	RRMS, 45 Treated	**WB, entire cohort:** The mean PBVCs during the 1st and the 2nd years of FU were -1.10% and -0.51% respectively, p=0.037. **WB, sub-cohort with Gd+ at baseline:** The mean PBVCs during the 1st and the 2nd years of FU were -1.55% and -0.41% respectively, p=0.026. **WB, sub-cohort with Gd- at baseline:** The mean PBVCs during the 1st and the 2nd years of FU were -0.53% and -0.64%, n.s..
				GM and WM: During the 1st year of FU, GMF increased by 1.15% and WMF decreased by -0.72%. Sub-cohort with Gd+ at baseline had a trend towards a greater WMF decrease (-2.81 vs. 0.31, p=0.071). GMF change was not significantly different between the Gd+ and the Gd- sub-cohorts (1.56 vs. 0.62).

Dwyer 2015[158]	IFN B-1a	SIENA (WB), SIENAX (GM, WM)	RRMS, 23 Treated vs. 15 Healthy Controls	**WB treated:** The mean PBVCs from baseline to 3m, 3m to 6m, and baseline to 6m FU were -0.95%, -0.08%, and -0.91%, respectively. **WB healthy controls:** The mean PBVCs from baseline to 3m, 3m to 6m, and baseline to 6m FU were +0.24%, -0.32%, and +0.01%, respectively.
				GM treated: The mean changes from baseline to 3m, 3m to 6m, and baseline to 6m FU were -1.52%, -0.46%, and -1.66%, respectively. **GM healthy controls:** The mean changes from baseline to 3m, 3m to 6m, and baseline to 6m FU were +0.01%, -0.60%, and -0.51%, respectively.
				WM treated: The mean changes from baseline to 3m, 3m to 6m, and baseline to 6m FU were -0.41%, +0.30%, and -0.21%, respectively. **WM healthy controls:** The mean changes from baseline to 3m, 3m to 6m, and baseline to 6m FU were +0.51%, -0.03%, and +0.58%, respectively.
Fisher 2015[153]	Intramuscular IFN B-1a	BPF, GMF, WMF	RRMS, 62 Treated vs. 69 Placebo	**WB:** The mean BPF % change over 2y of FU was -1.12% in the treated group and -1.30% in the placebo group, (n.s.). The changes during the 1st year FU were -0.80% (treated) and -0.70% (placebo), (n.s.). The changes during the 2nd year FU were -0.33% (treated) and -0.61% (placebo), (p=0.04). For the treated group, the rate during 1st year FU was significantly greater than that during 2nd year FU, p=0.005. No significant difference for the placebo.
				GM: The mean GMF % change over 2y of FU was -1.02% in the treated group and -1.42% in the placebo group, (n.s.). The changes during the 1st year FU were -0.70% (treated) and -0.73% (placebo), (n.s.). The changes during the 2nd year FU were -0.31% (treated) and -0.69% (placebo), (p=0.03). For the treated group, the rate during 1st year FU was marginally greater than that during 2nd year FU, p=0.07. No significant difference for the placebo.
				WM: The mean WMF % change over 2y of FU was -1.29% in the treated group and -1.06% in the placebo group, (n.s.). The changes during the 1st year FU were -0.93% (treated) and -0.62% (placebo), (n.s.). The changes during the 2nd year FU were -0.36% (treated) and -0.44% (placebo), (n.s.). For the treated group, the rate during 1st year FU was marginally greater than that during 2nd year FU, p=0.08. No significant difference for the placebo.

Vidal-Jordana 2016[116]	IFN B	SPM8 (WB, GM, WM)	RRMS, 123 Treated	During the 1st year of FU, the mean WBV, GMV, and WMV changes were -0.52%, -0.79%, and -0.11%, respectively. In all regions, the patients with Gd+ at baseline showed higher atrophy rates, although not statistically significant.

1.6.6 Brain Atrophy after aHSCT for Treatment of MS

A small number of observational cohort studies have measured WB atrophy in MS patients after treatment with aHSCT (Table 1-7 lists these findings). Despite the small numbers of studied patients and the lack of control groups, they have consistently shown a clear pattern of atrophy; during the first year of follow-up, the patients suffered significant brain volume loss that was well beyond those seen in natural history of MS, and even beyond those observed with, so-called, pseudoatrophy. Two studies with more frequent scans (Chen 2006 and Nash 2015) revealed that the majority of the volume loss occurred during the first 6 months of follow-up. This volume loss was paralleled by profound suppression of relapses and MRI WM lesion activity. Although this pattern appears similar to that observed with pseudoatrophy, the mechanism may be more complex due to the added insult from immunoablative regimens. A range of CNS neurotoxicity, including seizure, myelopathy, encephalopathy, leukoencephalopathy, cerebellar dysfunction and cerebrovascular complications, has been reported depending on the type and strength of the chemotherapeutic agent.[169] Conditions described as "chemobrain" exemplify cases of the treatment-related cognitive impairments that occur during or shortly after receiving chemotherapy.[170] In addition, brain atrophy is being recognized in patients without neurological comorbidities who were treated with cytotoxic chemotherapy. For example, VBM studies of breast cancer patients treated with adjuvant chemotherapy have reported that the transient cognitive decline may be associated with GM and WM atrophy.[170–173]

The effect of immunoablative regimens, plus the effect of degenerative processes related to MS (e.g. degeneration of tissues that were irreparably injured prior to aHSCT) may have contributed to the pronounced atrophy in the patients. However, no previous study has evaluated the association between these factors and atrophy in MS patients treated with aHSCT. Furthermore, the long-term effect of aHSCT on atrophy is unclear.

Table 1-7: Examples of whole brain atrophy rates in MS patients treated with aHSCT

Study	Conditioning Regimen	Analysis Method	Subjects	PBVC Outcome Average Value, (SD if available)	Note
Inglese 2004[174]	BEAM / Rabbit ATG	SIENA PBVC	10 SPMS	Mean Baseline to m12: -1.87% (2.19) M12 to m24: -1.88% (0.78)	
Chen 2006[175]	Busulfan / Cyclophosphamide / Rabbit ATG	SIENA PBVC	9 SPMS	Median Baseline to m1: -15.1%/y, N=9 M1 to m12: -1.6%/y, N=9 M12 to m24: -0.9%/y, N=5 M24 to m36: -0.8%/y, N=4	Annualized rates were reported
Rocca 2007[176]	TBI / Cyclophosphamide / ATG	SIENA PBVC	14 SPMS	Median Baseline to m12: -2.33%, N=14 M12 to m24: -1.35%, N=11 For the 9 patients with 3y FU Baseline to m12: -1.92 M12 to m24: -1.24% M24 to m36: -0.69%	Avoided pseudoatrophy by setting m1 FU scan as the baseline
Roccata gliata 2007[177]	BEAM / Rabbit ATG	SIENA PBVC	9 SPMS	Mean Baseline to m12: -1.10% (1.71) M12 to m24: -1.55% (1.11) Baseline to m24: -2.72 (2.0) M24 to last available FU (m48): -1.17% (2.96), or -0.45%/y (0.90)	Avoided pseudoatrophy by setting m1 FU scan as the baseline
Petzold 2010[159]	TBI / Cyclophosphamide / Horse ATG	SIENA PBVC	14 SPMS	Mean Baseline to m12: -2.1% Baseline to m24: -3.6% Baseline to m36: -3.9%	Avoided pseudoatrophy by setting m1 FU scan as the baseline
Nash 2015[80]	BEAM / Rabbit ATG	SIENA PBVC	24 RRMS	Median Screening to m6: -1.1% Screening to y1: -1.2% Screening to y2: -1.1% Screening to y3: -2.1%	Unequal number of subjects per timepoint N=23 N=19 N=21 N=17

1.6.7 Effect of MRI Scanner Change/Upgrade

In order to minimize the confounding effects of non-physiological/technical factors on atrophy measurements, all subjects enrolled in a study should ideally be scanned throughout using a consistent scanning protocol. However, this is not always the case. Many studies employ multiple sites to increase the number of participants and thereby to increase the power of the study. But this poses a new issue: each site likely has its unique combination of scanning protocol including scanner hardware, software, routine procedures, and acquisition parameters. Furthermore, longitudinal studies are sometimes subjected to protocol changes, including scanner hardware upgrades that affect image characteristics like tissue contrast, signal-to-noise ratio, geometric distortion, and intensity non-uniformity. Although a researcher can impose standardized sets of acquisition parameters and scanning procedures throughout the study, it is practically difficult for the researcher to control for the changes in the scanner hardware and software. These differences/changes may increase variability in the atrophy measure and mask the effect of interest that the researcher is looking for. Unfortunately, there were major MRI hardware upgrades during the follow-up phases of both the Canadian MS-BMT and the HALT-MS studies.

Several studies have looked at the potential impact of scanning protocol changes on brain atrophy measures, but the results seem to be mixed. A scan-rescan study by Preboske et al. showed that changing the image contrast of a conventional spoiled gradient recalled echo (SPGR) scan, simulated by 1) re-scanning the same subject with different flip angles and 2) re-scanning the same subject with fast SPGR sequence, resulted in large volume changes that exceed the disease effects in AD.[178] The same study also showed that increasing levels of head motion (simulated by adding motion blurring to the baseline images) and image noise (simulated by adding noise to the baseline images) resulted in volume changes in a level-dependent manner.[178] Han et al. showed that the average cortical thickness variability did not change significantly across an intra-manufacturer scanner upgrade from Siemens Sonata 1.5T to Siemens Avanto 1.5T.[179] The within-scanner measurement variability, however, was significantly reduced after the upgrade.[179] Similarly, Jovicich et al. showed that the scan-rescan reproducibility of the volumes of subcortical, ventricular, and intracranial volumes in the intra-manufacturer scanner upgrade condition (i.e. scan using Sonata and rescan using Avanto) was not significantly different from those of the pre- and post-upgrade conditions (i.e. scan-rescan

using only Sonata or Avanto).[180] The inter-manufacturer scan-rescan condition (i.e. scan using Sonata and rescan using GE Signa) also did not show significantly different reproducibility, although the mean volume difference showed some bias.[180] Stonnington et al. performed VBM to compare AD patients and healthy controls, who were scanned on 6 different scanners that had the same platform (GE Signa 1.5T) and similar hardware elements but slightly varying TR, TE, and flip angle. They also did not find a significant effect of the scanner differences.[181] Note that the above studies were all single-site studies. Kruggel et al. analyzed the multi-site Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset that included scans from Siemens, GE, and Philips scanners.[182] Despite the use of a standardized protocol and quality control procedure, repeated scans on different scanners gave intra-subject variabilities of the WB, GM, and WM volumes that were roughly 10 times higher than that of the repeated scans on a same scanner.[182]

Several studies have recommended mixed-effects model as a powerful method to analyze both cross-sectional and longitudinal data collected from multiple scanners.[183–186] Fennema-Notestine et al. showed that a linear mixed-effects model with site as a random effect yielded a better fit for a multi-site hippocampal atrophy data in AD and normal aging, compared to a regression model without the site term or a model with the site term as a fixed effect.[183] Bernal-Rusiel et al. demonstrated that linear mixed-effects model provided higher sensitivity and specificity compared to repeated-measures analysis of variance (ANOVA) or cross-sectional analysis of the slope, in modelling longitudinal hippocampal and entorhinal cortex atrophy in AD, mild cognitive impairment (MCI), and healthy controls.[184] Jones et al. reported that adding a categorical fixed-effect site term to a linear mixed-effects model led to substantially lower residual variance in modelling of longitudinal WB, GM, WM, and ventricular CSF atrophy in MS.[185] Chua et al. tested several forms of linear mixed-effects models with a fixed-effect protocol term, on a longitudinal MS brain atrophy data acquired on a single scanner but using 19 distinct protocols. They reported that the model with random intercept and slope with protocol specific residual variance provided the best fit in terms of Akaike Information Criterion.[186] These results suggest that accounting for the effect of scanner change is necessary in the analysis of scans acquired using more than one protocol. Mixed-effects models are one way to do so.

1.7 Methodology: Analyzing Longitudinal Atrophy Data using a Mixed-Effects Model

A longitudinal study involves observing each subject on multiple occasions over a set period of time, thereby allowing researchers to characterize changes in the responses of interest and to explore factors potentially associated with the change.[187,188] Longitudinal design generally has more statistical power compared to cross-sectional design due to the fact that each subject serves as his own control and the confounding effect of between-subject variability is reduced.[187] As such, the longitudinal approach is increasingly common in biomedical research, including studies of brain atrophy using neuroimaging data.[184]

Still, there are certain properties of longitudinal data that must be taken into account during the analysis. First is that repeated measurements within subjects are correlated, which invalidates the assumption of independence.[188] Further, the variability of the response at the end of the study is not necessarily the same as that at the beginning of the study; such heterogeneous variability invalidates the assumption of homogeneity of variance.[188] Note that these assumptions are precisely what many of the conventional statistical methods, such as standard linear regression, t-test, or ANOVA, are based upon.[188] Therefore these conventional techniques may not be appropriate for the analysis of longitudinal data.

Another issue that frequently rises in longitudinal studies is unbalanced data, in which the observations are not necessarily made at the same set of common scheduled times and/or the subjects do not necessarily have the same number of repeated observations due to missing data that commonly result from missing visits or study dropouts.[188] Conventional methods that require balanced data, such as repeated-measures ANOVA, may not be appropriate for this situation.[188]

The work presented in this thesis makes extensive use of a statistical analysis method called the mixed-effects model, which is a flexible way to analyze unbalanced repeated measures data. A mixed-effects model contains both fixed and random effects. In the context of a brain atrophy study, the fixed effects estimate parameters affecting the population-average brain volume estimate, which are assumed to be shared by all patients. The model further assumes that each patient has his own mean brain atrophy trajectory over time and the subject-specific deviations

from the population average are estimated with the random effects (for example, slopes estimate the subject-specific rates of atrophy and intercepts estimate the subject-specific starting point of the atrophy).[188] This way, the model accommodates both between-subject variability and within-subject correlation among unbalanced repeated measures data.[188]

In particular, some of the work presented in this thesis makes use of nonlinear mixed-effects models. Compared to the linear model, the nonlinear model has a theoretical advantage of being able to accommodate mechanistic models, i.e. the models whose form is determined by past experiences or theoretical considerations about the underlying mechanisms involved in generating the raw data.[189] This is done by allowing nonlinearity in the parameters, i.e. the derivative of the model function with respect to the parameters can still depend on at least one other parameter. This feature allows one to use certain functions such as the exponential function, and to account for nonlinear characteristics of the data such as asymptotes. Nonlinear models can often fit nonlinear data with fewer parameters than linear models, resulting in a more parsimonious model.[189] In this sense, the parameters estimated using a nonlinear model generally have more physically interpretable meanings.[189] For example, they may directly correspond to the amount of brain volume loss associated with some baseline predictors tested in the model.

The potential drawback of using nonlinear models is the complexity in building and fitting of the model. A nonlinear model requires specification of its equations and the starting values to be used during iterative computation of the parameter estimates.[189] The model may fail to converge if the main function is not well-suited to describe the data or inadequate starting values are supplied.[189] Therefore, it is often critical to derive good initial estimates for these parameters from previous knowledge or a theoretical consideration.

1.8 Research Objectives and Rationale

Previous MRI studies of MS patients treated with aHSCT have reported early, accelerated whole-brain volume loss immediately following aHSCT. Understanding the mechanisms behind this volume loss remains an important question because 1) the treatment can be associated with the equivalent of several years of brain atrophy over a short time, and 2) the consequences of this one-time treatment on long-term rates of brain atrophy is unknown. Several questions remain to

be answered in order to further determine the effect of chemotherapy and aHSCT on brain atrophy and related degenerative processes in patients with MS.

First, an important issue to consider is whether the accelerated brain volume loss is due to pseudoatrophy (e.g. due to resolution of inflammation-related edema). Second, it is necessary to assess whether toxicity of the chemotherapy could be associated with the atrophy. Third, it is unknown whether MS patients, who have already suffered damage to brain tissue, might have a predisposition to chemotherapy-related atrophy. Fourth, it remains to be answered whether the prolonged suppression of new MS-related focal inflammatory activity by aHSCT is paralleled by long-term reduction of MS-related brain atrophy. Fifth, it is important to better understand to what extent the effects of scanner changes can confound brain volume measurements.

The main objective of this thesis is to determine the effect of autologous hematopoietic stem cell transplantation on brain volume and related degenerative processes in multiple sclerosis. To do so, this thesis attempts to answer the following specific questions:

1) Does brain atrophy accelerate following aHSCT? If so, what is the timeframe and to what extent does the acceleration affect grey matter and white matter?
2) Do MS patients who receive higher doses or more intensive chemotherapy suffer greater brain atrophy because of more pronounced neurotoxicity?
3) Are MS patients with higher baseline white matter lesion load predisposed to greater brain atrophy after aHSCT because they have relatively more sub-lethally injured tissues that are more vulnerable to chemotoxicity?
4) Do brain atrophy rates in MS patients treated with aHSCT return to that of normal aging because the therapy results in complete or almost complete suppression of MS-related inflammatory activity?
5) Do MRI scanner changes make atrophy measurements unreliable? If so, can scanner changes be corrected for to obtain more reliable estimates of atrophy rates?

Chapter 2 Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis

2.1 Preface

Multiple sclerosis is currently understood as an inflammatory autoimmune disease that leads to demyelination and neurodegeneration in the central nervous system. Inflammation plays an important role in all stages of MS, exemplified by a finding that inflammatory infiltrates are present in tissue blocks from both relapsing and progressive MS cases, and that the number of inflammatory cells significantly correlate with markers of axonal injury.[190] Accordingly, most of the currently available DMTs are anti-inflammatory with either immunomodulatory or immunosuppressive characteristics; they have been approved based on evidence of effectiveness on endpoints related to focal white matter inflammation, such as relapse rates and MRI lesions.[191] However, these DMTs do not benefit all patients. Some patients continue to experience relapses, develop new MRI lesions, and have disability worsening despite being on therapy. Moreover, current DMTs that benefit patients with relapsing MS have failed for patients with progressive MS (with exception of ocrelizumab).[192] This led to debates as to whether stopping focal white matter inflammation can effectively halt neurodegeneration in MS. The consensus is that inflammation (both focal and diffuse) are strongly linked to neurodegeneration in MS, but further evidence is necessary before a conclusion can be drawn about their potential causative relationship.[193,194] At least, a complete control of MS-related inflammation is required, and its long-term impact needs to be assessed.

As discussed in chapter 1, the primary goal of aHSCT for MS is to produce prolonged remission of MS-related inflammatory activity by eliminating autoreactive pathogenic clones and repopulating an antigen-naïve system free of the autoimmune memory cells. Indeed, different trials have reported reductions in inflammatory activity and disability progression in the treated patients.[74] Also, a small number of observational cohort studies have reported whole-brain atrophy measurement data over short-term follow-up. However, two questions remain unanswered up to now. First, which factors can explain the significantly accelerated short-term whole-brain atrophy that is observed across different aHSCT for MS trials? Second, what is the

long-term effect of aHSCT on whole-brain atrophy in MS? Does the stoppage of measurable MS-related inflammatory activity result in reduction of whole-brain atrophy rates in MS patients? How do the long-term rates of whole-brain atrophy compare to those already reported in comparable age-matched MS or normal control population?

The following manuscript entitled "Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis" describes the short- and long-term courses of whole-brain atrophy in the patients enrolled in the Canadian MS-BMT trial, which used a particularly intensive immunoablative regimen and was the first treatment to fully stop all measurable CNS inflammatory activity over up to 10 years of follow-up.[195] A non-linear mixed-effects model is used to model the time courses of whole-brain atrophy with two hypothesized factors that could have contributed to short-term changes: the dose of chemotherapy drug (an index of treatment-related neurotoxicity) and baseline volume of T1-weighted white matter lesions (a marker of the amount of lethally injured tissues committed to degeneration prior to the treatment). An important aspect of this study is that the complete absence of new focal white matter lesion formation after aHSCT allowed for assessment of brain atrophy without the confound of new MRI-visible, MS-related inflammatory activity. In addition, long-term rates of atrophy, under this condition, could be estimated to explore the potential effect of aHSCT on brain atrophy.

2.2 Manuscript: Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis

Hyunwoo Lee, B.Sc. [1], Sridar Narayanan, Ph.D. [1], Robert A. Brown, Ph.D. [1], Jacqueline T. Chen, Ph.D. [2], Harold L. Atkins, M.D. [3], Mark S. Freedman, M.D. [4], Douglas L. Arnold, M.D. [1]

[1] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada

[2] Department of Neurosciences, Lerner Research Institute, Cleveland Clinic, Cleveland, USA

[3] Ottawa Hospital Blood and Marrow Transplant Program, The Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Canada

[4] University of Ottawa and the Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Canada

Published in: ***Multiple Sclerosis Journal*****. 2017 Mar;23(3):420-431.**

2.2.1 Abstract

Background: A cohort of patients with poor-prognosis multiple sclerosis underwent chemotherapy-based immune-ablation followed by immune-reconstitution with an autologous hematopoietic stem cell transplant (IA/aHSCT). This eliminated new focal inflammatory activity, but resulted in early acceleration of brain atrophy.

Objective: We modeled the time course of whole-brain volume in 19 patients to identify the baseline predictors of atrophy and to estimate the average rate of atrophy after IA/aHSCT.

Methods: Percentage whole-brain volume changes were calculated between the baseline and follow-up MRIs (mean duration: 5 years). A mixed-effects model was applied using two predictors: total busulfan dose and baseline volume of T1-weighted white-matter lesions.

Results: Treatment was followed by accelerated whole-brain volume loss averaging 3.3%. Both the busulfan dose and the baseline lesion volume were significant predictors. The atrophy slowed progressively over approximately 2.5 years. There was no evidence that resolution of edema contributed to volume loss. The mean rate of long-term atrophy was -0.23%/year, consistent with the rate expected from normal-aging.

Conclusion: Following IA/aHSCT, MS patients showed accelerated whole-brain atrophy that was likely associated with treatment-related toxicity and degeneration of 'committed' tissues. Atrophy eventually slowed to that expected from normal-aging, suggesting that stopping inflammatory activity in MS can reduce secondary-degeneration and atrophy.

2.2.2 Introduction

The hypothesis behind immunoablation and autologous hematopoietic stem cell transplantation (IA/aHSCT, a type of bone-marrow transplantation (BMT)) for MS is to target active inflammation by eradicating the autoreactive immune system and then reconstituting a new, self-tolerant system.[95] The Canadian MS-BMT study recruited highly active, poor-prognosis relapsing-remitting (RR) and secondary-progressive (SP) MS patients who had progressive disability, and had failed conventional therapy for MS.[95] The aim of IA/aHSCT therapy was to eliminate ongoing MS-related inflammatory activity in order to prevent further irreversible damage and preserve remaining neurological function.[95]

Following IA/aHSCT, no subject had evidence of acute inflammatory disease activity, i.e., no relapses or new focal white matter (WM) lesions on MRI.[85] However, rates of whole-brain (WB) atrophy did not show an immediate benefit; a previous analysis by Chen et al. of 9 MS-BMT patients with SPMS and a median follow-up of 18 months, showed accelerated WB atrophy that was ten times faster immediately after treatment than before treatment.[175] Similar outcomes have been reported by others after BMT.[174,176,177] Potential explanations for the short-term acceleration have included: 1) neurotoxicity of the conditioning regimen used for IA, 2) the consequences of

extensive pre-treatment inflammatory demyelination, and 3) "pseudoatrophy" due to resolution of edema.[174–177] The effects of all of these are expected to be temporary, although the timeframe remains unexplored.

A better understanding of the extent and time course of brain atrophy after IA/aHSCT is an essential component in the evaluation of the efficacy of this therapy, and will provide valuable insights on the mechanisms involved in brain atrophy in MS. We applied a mixed-effects model to the long-term follow-up MRI WB volume data of the MS-BMT patients with the goal of identifying potential predictors of the early, accelerated atrophy and to estimate the group rate of atrophy after IA/aHSCT. Mixed-effects models include both fixed and random effects. For example, a model explaining WB volume might contain fixed effects for slope and intercept, which, like the explanatory variables in regression models, represent the population-average estimate that is assumed to be shared by all patients.[188] The model further assumes that each patient has his/her own mean WB volume trajectory over time; subject-specific random effects estimate the departure from the group mean for each subject.[188] This way, the model is used both to explain inter-subject variation and also to account for correlation among multiple measurements within the same patient.

We hypothesized that the rate of atrophy immediately following IA/aHSCT is due to the combined effect of i) the chemotherapeutic regimen used for IA and, ii) the amount of tissue lethally injured by neuroinflammation and committed to degeneration prior to IA/aHSCT, and estimated the duration of the accelerated atrophy. Furthermore, we hypothesized that MS-related WB atrophy loses its "driving force" after IA/aHSCT. The absence of new focal WM lesions implies cessation of new lesion-related processes that may lead to neuroaxonal degeneration, specifically irreversible axonal transection and subsequent axonal degeneration in normal-appearing WM (NAWM). Thus, if IA/aHSCT halts inflammatory activity, it could prevent neurodegeneration with the rates of atrophy eventually slowing to that expected for normal-aging.[190]

2.2.3 Methods

Patients and Treatment Regimen

The Canadian MS-BMT study was approved by the Ottawa Health Science Network REB, McGill Faculty of Medicine Institutional REB, and St. Michael's REB (#2000374-O1H). Written informed consent was obtained from all participants. Twenty-four patients were enrolled.[95] The treatment procedure is shown in Figure 2-1.[95] One patient died shortly after aHSCT due to chemotoxicity, and only 23 patients were followed-up.[85] Due to the lengthy follow-up, all but two subjects were scanned on multiple MRI scanners over the course of the study. We selected 19 subjects with a maximum of one scanner change: 17 subjects were initially scanned with a Siemens Symphony, which was upgraded to a Symphony TIM, and two subjects were scanned entirely on the Siemens Symphony TIM. These 19 subjects had varying durations of MRI follow-up (mean: 5y, range: 1.5-10y). Subject demographics are shown in Table 2-1. The MRI protocol is shown in Table 2-2.

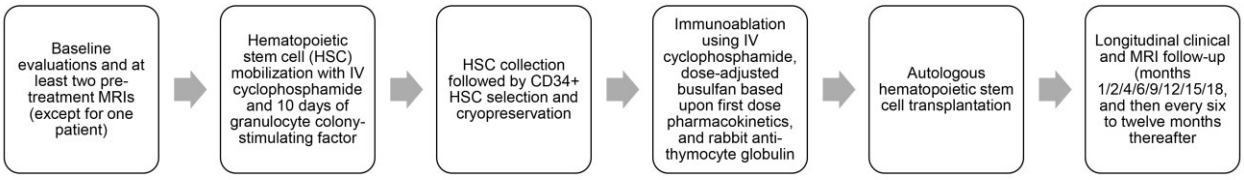

Figure 2-1: Outline of the procedure for immunoablation and autologous hematopoietic stem cell transplantation for treatment of multiple sclerosis

Table 2-1: Basic subject demographics – Canadian MS-BMT

There were no significant differences in the characteristics between the included and the excluded subjects.

Characteristics	Whole-cohort (N=19)	SPMS-cohort (N=7)	RRMS-cohort (N=12)	Excluded subjects (N=4 SPMS)
Mean age (SD) [range], yr	32.7 (6.2) [23-45]	30.9 (4.7) [26-40]	33.8 (6.9) [23-45]	31.8 (4.6) [26-37]
Sex, Female:Male	12:7	5:2	7:5	2:2
Mean disease duration (SD), onset to IA/aHSCT, yr	6.9 (2.7)	7.4 (2.5)	6.5 (2.8)	10.1 (7.5)
Mean Expanded Disability Status Scale score (SD)	5.0 (1.1)	5.7 (0.8)	4.6 (1.2)	5.0 (1.2)
Mean annualized relapse rates for 2-years before IA/aHSCT	1.0 (0.5)	1.2 (0.5)	1.0 (0.5)	0.5 (0.4)
% of subjects with new or enlarging T2-weighted lesions, during baseline evaluation prior to IA/aHSCT	74	100	58	50
% of subjects with gadolinium-enhancing lesions, during baseline evaluation prior to IA/aHSCT	63	86	50	0
Mean baseline gadolinium-enhancing lesion volume (SD), ml	0.5 (1.1)	1.1 (1.6)	0.1 (0.2)	0
Mean baseline T1-weighted lesion volume (SD), ml	7.9 (7.8)	12.2 (10.0)	5.3 (5.2)	10.2 (10.2)
Mean baseline T2-weighted lesion volume (SD), ml	20.0 (18.0)	31.3 (21.6)	13.5 (12.1)	26.9 (18.1)
Mean total busulfan dose (SD), mg	723.6 (146.0)	802.4 (151.9)	677.7 (126.7)	850 (268.7)
Numbers of subjects that have reached each MRI follow-up time point, years from aHSCT, (N)	1.5y (19); 2y (17); 2.5y (16); 3y (13); 4y (12); 4.5y (11); 5y (9); 5.5y (8); 7y (7); 9.5y (5); 10y (3)			3y (4); 8y (2); 9y (1)
Mean number of MRI scans available per subject, (SD) [range]	15 (5) [9-23]			n/a
Total number of T1-weighted scans used for the analysis	288			n/a

Table 2-2: MRI protocol for the images used in this analysis

	T1-weighted Pre-Gadolinium	Axial Proton density-weighted / T2-weighted
Field strength	1.5T	1.5T
Sequence	3D FLASH	2D multi-slice fast/turbo spin-echo
TR (ms)	28	2070
TE (ms)	11	12/86
Field of view (mm)	256	256
Slices	60	60
Thickness (mm)	3	3
Matrix	256x192	256x192
Flip angle	30	90 (excitation), 180 (refocusing)
ETL/Turbo factor	n/a	5

Whole-brain atrophy measurement

Longitudinal WB volume changes were calculated with FSL-SIENA using pairs of pre-contrast T1-weighted images.[120] For each subject, percentage brain volume change (PBVC) was calculated between the last baseline MRI (reference point, defined as 100%) and the follow-up scans. Normalized brain volumes (NBV) at baseline were calculated with FSL-SIENAX.[120]

T2w- and T1w-lesion volume measurement

For all timepoints, hyper-intense WM lesions were identified on T2-weighted scans with an in-house Bayesian classifier and then manually corrected by a reader. Hypo-intense WM lesions were identified on pre-contrast T1-weighted scans as follows: The voxels within the areas labeled as T2-weighted lesions were identified. Then, those with T1-weighted image intensity of 85% or lower than surrounding NAWM voxels were identified as lesions.

Pseudo-T2 relaxation time measurement

Acceleration of brain volume loss upon initiation of anti-inflammatory treatment is often ascribed to "pseudoatrophy" due to resolution of inflammation, i.e., volume reduction associated with factors like decreased tissue water or decreased volume of inflammatory cells but not neural tissue loss. Two-point "pseudo" T2-relaxation time (pT2) is a metric sensitive to changes in tissue water content that is derived from dual-echo turbo spin-echo MRI with the formula $pT2=(E_2-E_1)/\ln(S_1/S_2)$ where S_N are the signal intensities at each echo time E_N.[156] For each patient, we estimated pT2 changes in normal-appearing brain tissue (NABT: tissue defined as

cortical and subcortical grey matter (GM) or WM using FSL-SIENAX,[120] and not identified as WM lesion) from baseline to early post-treatment follow-up intervals, when the effect of resolution of edema, if present, is expected to be the largest. Decreased pT2 relative to baseline would be consistent with resolution of edema.

Baseline whole-brain atrophy rate

For each subject, a linear regression line was fitted between the baseline measurements (PBVC from the first baseline to the second, and if available, the third baseline). Then, the slope was multiplied by the annualization factor to calculate the subject-specific baseline atrophy rate. 14 subjects had two baseline MRIs; 4 subjects had three. The mean baseline rate among these 4 subjects was calculated separately.

Model of atrophy: entire follow-up

We modeled the time course of post-treatment WB atrophy as the summation of three processes: accelerated treatment-related change in WB volume after IA/aHSCT, normal-aging and potentially persisting MS-related atrophy (Figure 2-2), and apparent volume change due to a change in MRI scanner. We employed a non-linear mixed-effects model (an extension of linear mixed-effects model that allows for nonlinearity in parameters). The nonlinear approach allows for a mechanistic model, whose form is determined by past experiences or theoretical considerations about the underlying mechanisms involved in generating the data.[189] Accordingly, the estimated model parameters generally have more physically interpretable meanings.[189]

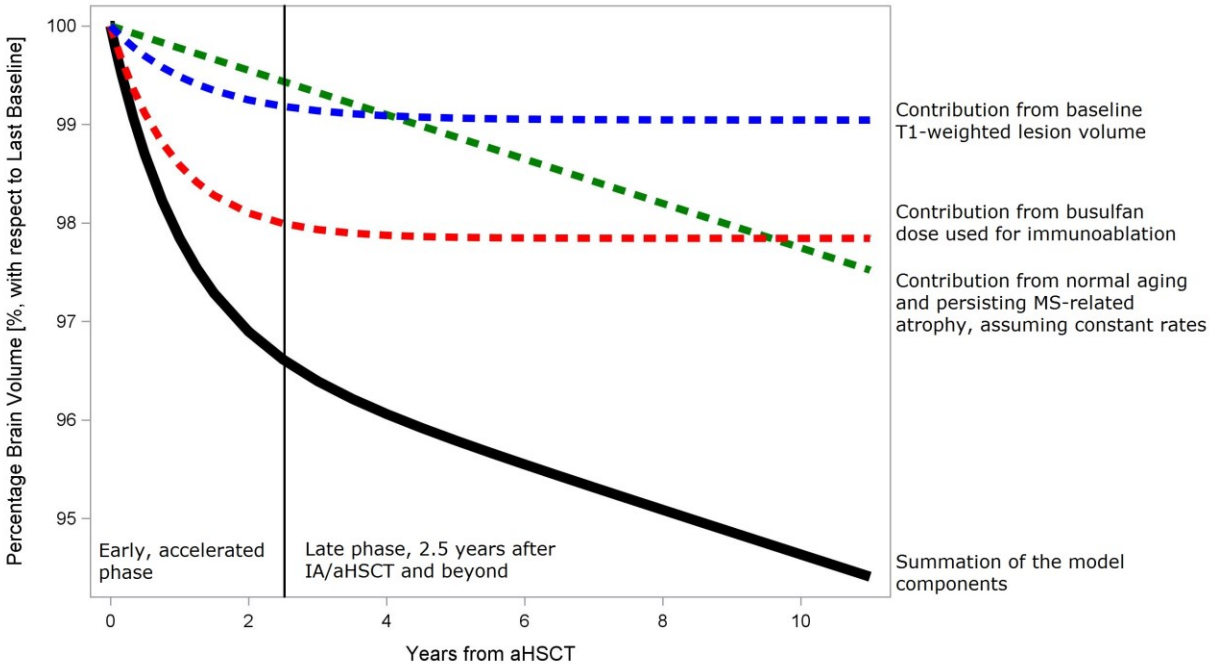

Figure 2-2: Illustration of the model of post-treatment whole-brain atrophy

We modeled the time course of post-treatment whole-brain atrophy as the summation of several processes. The accelerated whole-brain atrophy after IA/aHSCT was modeled as exponential-decay curves; their contributions are expected to be the largest at the beginning and then reduced with time. We used two curves, one related to busulfan dose and another related to T1 lesion volume, a marker of pre-existing focal white matter injury. The rate of whole-brain atrophy due to normal-aging is approximately -0.27%/year with the analysis methods we used for the age range of our patients. We included a linear term to account for this and any persisting atrophy related to the MS disease process.

The following nonlinear mixed-effects model was fitted to the group WB volume data using the NLMIXED procedure (SASv9.3):

$$Volume_{WB}(time\ from\ aHSCT) =$$

$$(a_1 * BusDose) * exp^{-time*(b_1*BusDose)} \quad \textit{[Component\_1]}$$

$$+ (a_2 * T1LV) * exp^{-time*(b_2*T1LV)} \quad \textit{[Component\_2]}$$

$$+ (InteractionCoef * BusDose * T1LV) * exp^{-time*(Exponent*BusDose*T1LV)} \quad \textit{[Interaction]}$$

$$+ (LINRATE + RANDOM_{SLOPE}) * time + RANDOM_{INTERCEPT} \quad \textit{[Component\_3]}$$

$$+ \left(EFFECT_{ScannerChg}\right) * CategoricalVar_{Scanner} \quad \textit{[Component\_4]}$$

$$+ AsymptoteVol + error$$

[Components 1 and 2]: Accelerated treatment-related change in WB volume

Based on our earlier observations,[175] we modeled the WB volume change as an exponential decay. We hypothesized that two subject-specific baseline predictors were associated with the decay: total dose of busulfan given for immunoablation (*BusDose*) and T1-weighted WM lesion volumes (*T1LV*). [Component_1] estimated two parameters describing the decay curve associated with *BusDose*: a_1 was the exponential coefficient (to determine the amount of volume loss explained by the function) and b_1 was the exponential decay constant (to determine how

quickly the volume loss occurred). [Component_2] also estimated the exponential coefficient and decay constant parameters (a_2 and b_2, respectively) describing the curve associated with *T1LV*.

For two reasons, we used *BusDose* as an index of the treatment-related neurotoxicity due to the combined effect of the entire conditioning regimen. First, the busulfan doses were somewhat more variable because of adjustment for first dose pharmacokinetics. Second, busulfan easily crosses the blood-brain-barrier and it is likely that CNS penetration is necessary for lymphocytotoxicity in sites of active MS inflammation.[196]

We chose *T1LV* as a marker of pre-existing focal tissue injury over T2-weighted lesion volume because the T1-weighted lesion volume is more strongly correlated with significant injury and axonal loss.[61,62] The baseline *T1LV* associated with gadolinium-enhancement was small relative to the total *T1LV*; subtracting the volume of enhancing lesions from the volume of *T1LV* and using it as a predictor variable gave similar results.

[Interaction]: We also tested an interaction between *BusDose* and *T1LV* to explore whether the effect of *BusDose* on atrophy is altered in patients with different *T1LV*.

AsymptoteVol represented the volume remaining after the exponential function had levelled-off.

[Component_3]: Normal-aging and persisting MS-related atrophy

These exponential functions may not be sufficient to fully account for the course of atrophy following aHSCT. Apart from the treatment-related atrophy, a smaller-scale atrophy is expected to be present throughout the follow-up due to i) normal-aging and ii) any persisting MS-related atrophy. The temporal evolution of WB atrophy in MS patients may be reasonably modeled as a linear function.[98] WB atrophy related to normal-aging within the age range of our patients may also be approximated as a linear function.[137] Therefore, atrophy due to the combination of these two factors may be modeled as a linear function. We estimated the yearly rate of this atrophy with the slope parameter *LINRATE*. Each subject was expected to show different rates, and this variability was modeled with a random subject-specific slope $RANDOM_{SLOPE}$. Similarly, $RANDOM_{INTERCEPT}$ modeled the inter-subject variability in the intercepts.

[Component_4]: Bias in volume change measurements due to scanner changes

A scanner change may affect estimation of atrophy due to altered MR image characteristics. This effect was estimated by using a categorical variable representing the scanner used for the acquisition, e.g. Symphony or Symphony TIM, assigned to each timepoint.

Model of atrophy: late follow-up

The early, accelerated brain atrophy that occurs after IA/aHSCT is not maintained in the long-term, and appears to attenuate over time.[175] Although new, peripherally-mediated, focal WM lesions were absent after IA/aHSCT, we cannot affirm whether the treatment halted diffuse/CNS-compartmentalized inflammation.[34] To specifically explore whether IA/aHSCT improves MS-related atrophy over the long term, we estimated the long-term rates with a linear mixed-effects model using only the data from 2.5 years after IA/aHSCT and beyond (NLMIXED, SASv9.3).

$$Volume_{WB}(time \geq 2.5y\ after\ aHSCT) =$$

$$(LINRATE + RANDOM_{SLOPE}) * time$$

$$+ \left(EFFECT_{ScannerChg}\right) * CategoricalVar_{Scanner}$$

$$+ Intercept + RANDOM_{INTERCEPT}$$

$$+ error$$

2.2.4 Results

Baseline whole-brain atrophy rate

Pre-treatment scan intervals were short (median:0.17y) and varied between patients (SD:0.64y). The mean annualized baseline rate of WB atrophy was: 0.25%/y (N=18, SE:0.98) with one outlier.

Without that outlier, the rate was: -0.58%/y (N=17, SE:0.55). For the patients with enhancing lesions at baseline, the rate was: -1.6%/y (N=10, SE:0.7); for those without, the rate was: 0.9%/y (N=7, SE:0.6). Four patients had three baseline MRIs over an average of 1.3y. For these subjects,

the rate was: -1.5%/y (N=4, SE:0.7). There was no significant correlation between the baseline Expanded Disability Status Scale (EDSS) scores and the atrophy rates.

pT2 changes from baseline

Average pT2 changes from baseline to selected early follow-up months did not change significantly according to paired t-tests (Table 2-3) and were not correlated with baseline volumes of enhancing lesions.

Table 2-3: Average pT2 differences from baseline to selected early follow-up timepoints

	Average change in pT2 of normal-appearing brain tissue, Mean (SD) [range], paired t-test outcome
Baseline to month 1 (N=19)	-0.17 (1.54) [-2.8 to 2.7], n.s.
Baseline to month 6 (N=18)	-0.1 (1.71) [-3.4 to 3.1], n.s.
Baseline to month 12 (N=17)	1.22 (2.57) [-2.4 to 6.2], n.s.

Model of atrophy: entire follow-up

Estimates of the parameters for the model describing the entire follow-up period are shown in Table 2-4. Overall, the exponential decay predicted about 3.3% volume loss. Both BusDose and T1LV were significant predictors of the exponential coefficients and the decay constants. Inputting the mean BusDose, the exponential model associated with BusDose effect predicted about 2.2% loss with an approximate half-life of 0.65y. Similarly, the exponential model associated with T1LV effect predicted about 0.96% loss with an approximate half-life of 0.90y. The mean rate of the smaller-scale, linear atrophy was -0.23%/y. The interaction effect was not significant and thus removed from the model. The effect of scanner upgrade was not statistically significant. Figure 2-3 plots the actual PBVC measurements for the individual subjects and the fitted individual and group atrophy model trajectories for the entire follow-up. The amount of WB atrophy following aHSCT was not significantly different between the subjects who had an increase in the EDSS (N=4) and the subjects who had stability or a decrease in the EDSS (N=15).

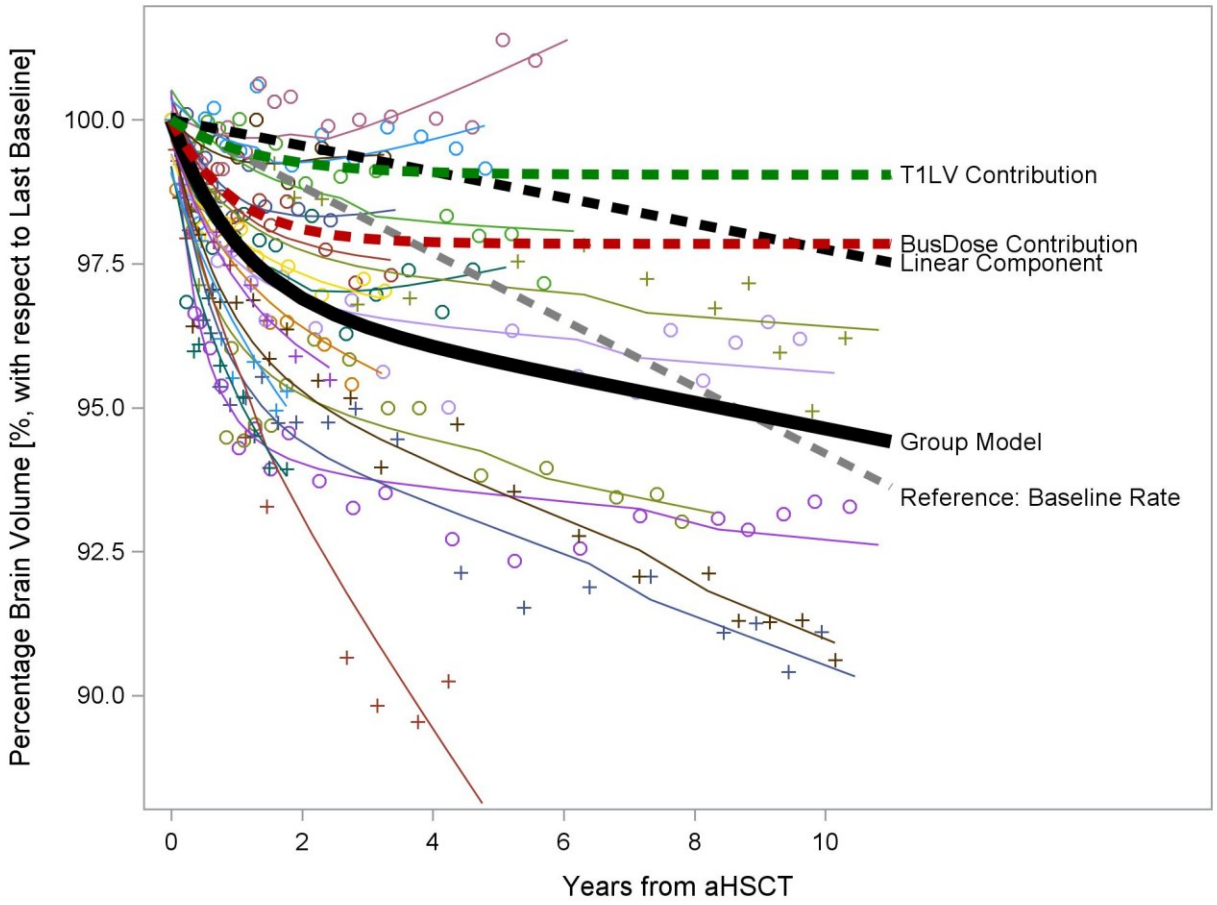

Figure 2-3: Evolution of post-treatment atrophy: Entire follow-up – Canadian MS-BMT

Overall, the exponential models indicated about 3.3% whole-brain volume loss immediately following IA/aHSCT. The exponential decay associated with the total busulfan dose (a marker of the treatment-related neurotoxicity) was relatively more pronounced and shorter-lasting compared to that of the baseline T1-weighted lesion volume (an index of the amount of tissue lethally injured before treatment). Apart from the exponential decay, a smaller-scale atrophy (assuming a constant rate) was present, consistent with normal-aging. The thick solid and dashed lines represent the group model, obtained by inputting the mean values of BusDose and T1LV into the fixed-effects.

Legend:

Reference (dashed grey line) = Reference line for the mean baseline whole-brain atrophy rate (-0.58%/y)

BusDose Contribution (dashed red line) = Exponential decay model associated with total busulfan dose

T1LV Contribution (dashed green line) = Exponential decay model associated with baseline T1LV

Linear Component (dashed black line) = Linear model of a small-scale atrophy, potentially due to normal-aging

Group Model (solid black line) = Net model for the whole cohort, i.e. the sum of T1LV and BusDose curves, and the Linear component

Colored dots (round or cross) = Actual PBVC measurements from the last baseline for the individual subjects

Colored unlabelled fitted lines = Fitted model trajectories for the individual subjects

Table 2-4: Parameter estimates for the model of post-treatment atrophy: entire follow-up

	Parameter, N=19, df=17	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Whole Cohort; Entire follow-up	**a_1: contribution of *BusDose* on exponential coefficient**	**0.0030**	**0.00059**	**5.08**	**<.0001**	**0.0017**	**0.0042**
	b_1: contribution of *BusDose* on decay constant	**0.0015**	**0.00068**	**2.16**	**0.045**	**0.000035**	**0.0029**
	a_2: contribution of *T1LV* on exponential coefficient	**0.12**	**0.021**	**5.74**	**<.0001**	**0.077**	**0.17**
	b_2: contribution of *T1LV* on decay constant	**0.098**	**0.055**	**1.78**	**0.093**	**-0.018**	**0.21**
	LINRATE: % change/year	**-0.23**	**0.17**	**-1.36**	**0.19**	**-0.57**	**0.12**
	$EFFECT_{ScannerChg}$	-0.23	0.20	-1.11	0.28	-0.65	0.20
	AsymptoteVol	**96.71**	**0.47**	**205.96**	**<.0001**	**95.72**	**97.70**
	Error variance	0.44	0.040	11.03	<.0001	0.35	0.52
	$RANDOM_{SLOPE}$	0.31	0.12	2.67	0.016	0.065	0.55
	Random $Effect_{Covariance}$	0.29	0.18	1.62	0.12	-0.088	0.66
	$RANDOM_{INTERCEPT}$	1.56	0.60	2.6	0.019	0.30	2.82

Model of atrophy: late follow-up

Estimates of the parameters for atrophy after 2.5 years are shown in Table 2-5. 15 subjects with follow-up duration of 2.5 years or longer contributed to this analysis. According to the half-lives calculated above, approximately 7% of the *BusDose* and 15% of the *T1LV* model processes were remaining at this point. The mean atrophy rates estimated by the model were: entire-cohort: -0.23%/y, SPMS-cohort: -0.30%/y, RRMS-cohort: -0.14%/y. The effect of scanner hardware upgrade was not statistically significant. There was no significant correlation between the baseline NBVs and the atrophy rates during the late period. Figure 2-4 plots the actual PBVC measurements for the individual subjects and the fitted individual and group atrophy model trajectories for the late follow-up.

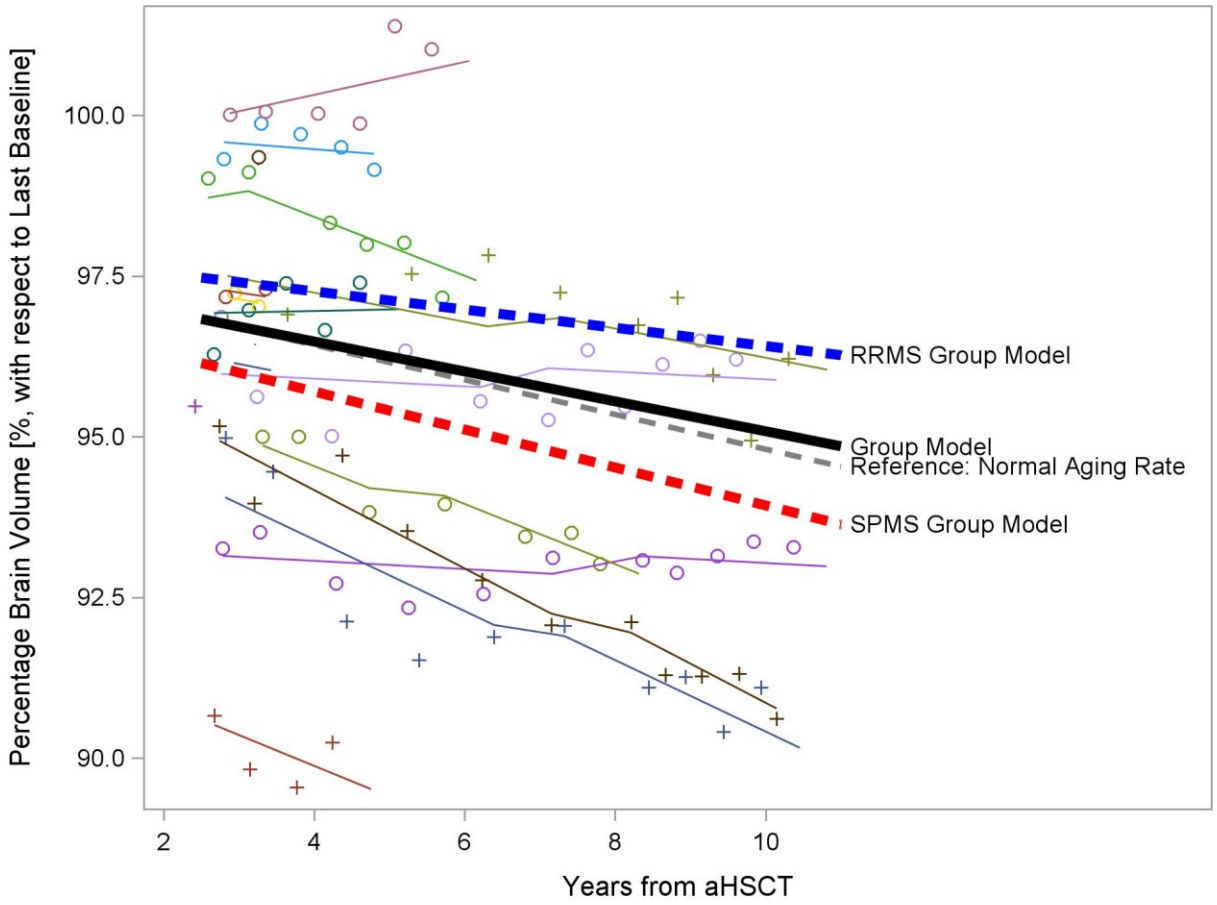

Figure 2-4: Evolution of post-treatment atrophy: Late follow-up – Canadian MS-BMT

Assuming a constant rate of atrophy throughout the period, the average whole-brain atrophy rate during the late follow-up period (i.e. 2.5y after treatment and beyond) was close to those observed in healthy normal-aging. The thick solid and dashed lines represent the fixed-effects model of long-term atrophy progression, with the average intercepts being the starting points. The lower starting point of the SPMS-cohort indicates that the accelerated atrophy was relatively more pronounced during the first 2.5y after treatment. A potential reason is that on average they received more doses of busulfan and had >2x the baseline WM lesion volumes compared to the RRMS-cohort. 15 subjects with follow-up duration of 2.5 years or longer were included in this analysis.

Legend:

Reference (dashed grey line) = Reference line for the normal-aging related whole-brain atrophy rate (-0.27%/y, from De Stefano et al., 2015)

Group Model (solid black line) = Model for the whole cohort (-0.23%/y)

SPMS Group Model (dashed red line) = Model for the SPMS cohort (-0.30%/y)

RRMS Group Model (dashed blue line) = Model for the RRMS cohort (-0.14%/y)

Colored dots (round or cross) = Actual PBVC measurements from the last baseline for the individual subjects

Colored unlabelled fitted lines = Fitted model trajectories for the individual subjects

Table 2-5: Parameter estimates for the model of post-treatment atrophy: late follow-up

	Parameter, N=15, df=13	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Whole Cohort	**LINRATE: % change/year**	**-0.23**	**0.11**	**-2.07**	**0.059**	**-0.48**	**0.0098**
	$EFFECT_{ScannerChg}$	0.35	0.26	1.34	0.20	-0.21	0.91
	Intercept	96.83	0.65	148.09	<.0001	95.42	98.24
	Error variance	0.28	0.048	5.75	<.0001	0.17	0.38
	$RANDOM_{SLOPE}$	0.090	0.053	1.7	0.11	-0.024	0.20
	Random Effect Covariance	0.031	0.26	0.12	0.91	-0.53	0.59
	$RANDOM_{INTERCEPT}$	5.41	2.39	2.27	0.041	0.25	10.58
	Parameter for SPMS, N=7, df=5	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**95% Confidence Limits**	
SPMS Cohort	**LINRATE: % change/year**	**-0.30**	**0.12**	**-2.45**	**0.058**	**-0.60**	**0.014**
	$EFFECT_{ScannerChg}$	0.41	0.30	1.36	0.23	-0.37	1.19
	Intercept	96.14	0.60	161.16	<.0001	94.61	97.68
	Error variance	0.32	0.063	5.05	0.0039	0.16	0.48
	$RANDOM_{SLOPE}$	0.064	0.044	1.46	0.20	-0.049	0.18
	Random Effect Covariance	-0.096	0.16	-0.6	0.58	-0.51	0.31
	$RANDOM_{INTERCEPT}$	1.88	1.15	1.63	0.16	-1.09	4.85
	Parameter for RRMS, N=8, df=6	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**95% Confidence Limits**	
RRMS Cohort	**LINRATE: % change/year**	**-0.14**	**0.21**	**-0.68**	**0.52**	**-0.65**	**0.37**
	$EFFECT_{ScannerChg}$	0.19	0.57	0.34	0.75	-1.19	1.57
	Intercept	97.47	1.24	78.52	<.0001	94.44	100.51
	Error variance	0.14	0.049	2.77	0.033	0.016	0.25
	$RANDOM_{SLOPE}$	0.14	0.11	1.21	0.27	-0.14	0.42
	Random Effect Covariance	-0.12	0.58	-0.22	0.84	-1.53	1.28
	$RANDOM_{INTERCEPT}$	8.64	4.98	1.73	0.13	-3.55	20.83

2.2.5 Discussion

We modeled brain atrophy using data obtained from serial MRI scans of patients participating in the Canadian MS-BMT trial. These patients received high dose busulfan/cyclophosphamide/rATG followed by aHSCT which completely eliminated focal inflammatory activity on MRI for up to 10 years of follow-up.[85] Treatment was associated with an early and short-term acceleration of atrophy. Our results in the present study suggest that early, short-term acceleration of atrophy is dependent on the toxicity of the chemotherapy (as measured by the dose of busulfan) and the volume of pre-existing focally injured tissue (as measured by T1-weighted lesion volume), but not resolution of edema (as measured by change in pT2). Prior to IA/aHSCT, our patients, refractory to standard MS therapies, exhibited baseline rates of WB atrophy similar to those of untreated MS patients.[98] Over several years, the mean rate of atrophy in these MS patients gradually slowed to rates that have been observed in normal-aging cohorts (-0.2 to -0.3%/y in comparable age groups).[96,136,137]

We hypothesized that the evolution of WB atrophy after IA/aHSCT is a complex process that includes treatment-related WB volume loss and the net-effect of normal-aging and potentially remaining MS-related atrophy. The results from the entire follow-up model suggest that chemotherapy toxicity and the volume of pre-existing irreversible focal tissue injury are significant predictors of the accelerated atrophy after IA/aHSCT. Comparing the two predictors, atrophy related to *BusDose* was more pronounced and relatively short-lasting whereas that related to *T1LV* was smaller but longer-lasting. This implies that a chemotherapy-related process largely drives the acutely-accelerated atrophy during the first year or so, while residual lesion-related processes predominate the later stages.

Neurotoxicity is intrinsic to all categories of chemotherapeutic drugs, although its relationship to brain atrophy is less understood.[75,197] Several voxel-based morphometry studies of cancer patients have reported atrophy of GM and WM as early as one month after chemotherapy.[170–173,198] Atrophy was still observed at one year after chemotherapy,[170,198] and the long-term outcomes were mixed with evidence of both recovery and non-recovery.[171,172] The basis of this atrophy is unclear, but several cytotoxic chemotherapeutic agents (e.g. BCNU, cytarabine) have been shown to be toxic in a concentration-dependent manner for normal CNS neural progenitor cells and non-dividing oligodendrocytes.[91,199] Busulfan is more intense than regimens containing

BCNU/cytarabine, and its neurotoxicity may be stronger.[95] One question worth noting is whether chemotherapy-related atrophy is more pronounced in MS patients, where pre-existing injury may make the CNS more susceptible to additional insults. Unfortunately, the present study did not include non-MS controls undergoing HSCT following an analogous high-dose chemotherapy regimen. Alternatively, we tested for an interaction between *T1LV* and *BusDose*. Within the ranges of the tested predictors, the effect of additional *BusDose* on atrophy was insignificantly altered in presence of a higher *T1LV*. This suggests that greater focal inflammatory damage in the MS brain does not increase sensitivity to the damaging effects of chemotherapy.

We also considered the effect of pre-treatment focal tissue injury per se on brain atrophy after treatment. In MS patients, significant axonal damage and loss is present in focal demyelinated lesions as well as in NAWM.[29,35,200] Demyelinated MS lesions contain significantly more transected axons than normal-appearing tissues,[30] and one possible mechanism for tissue loss is Wallerian degeneration.[100] This process, which persists for months to years in CNS,[201] could be one explanation for the significance of the *T1LV*-related exponential model, which had a similar temporal profile. Additionally, retrograde degeneration could have indirectly contributed to atrophy of GM regions. The average long-term atrophy rate in our SPMS-cohort was higher than that of patients classified as RRMS. The potential reason is that the SPMS-cohort had on average 2.3x more T1w- and T2w-lesion volumes and 11x more gadolinium-enhancing lesion volume at baseline. The larger volume of focally injured tissue in the SPMS patients may imply a larger volume of tissue undergoing degeneration, thus resulting in more brain volume loss per unit time.

"Pseudoatrophy" due to resolution of edema, cannot explain the early increase in atrophy rates. This is supported by several findings. First, the average change over time in pT2 in NABT did not show the decreases that would be expected if the change in brain volume were partially related to a decrease in brain water content. A previous study estimated that a unit decrease in pT2 could produce 0.116% brain volume loss.[156] In the present study, the largest average decrease in NABT pT2 was observed at month 1 (-0.17ms); this would be expected to produce less than 0.02% volume loss. Second, decreases in T2-weighted lesion volume might explain the increase in atrophy rates, however, during the first year following IA/aHSCT, the changes in the volume of T2-weighted WM lesions were less than the brain volume changes by an order of magnitude. These findings supplement previous studies of MS patients treated with BMT using

different immunoablative regimens. In these studies, the mean PBVC during the first year of follow-up were as follows: Inglese et al., -1.87% (BEAM)[174]; Rocca et al., -2.33% (total body irradiation)[176]; Roccatagliata et al., -1.10% (BEAM)[177]; Petzold et al., -2% (cyclophosphamide and total body irradiation).[159] Notably, these studies have used the follow-up scan 1 month after BMT as the 'baseline' to avoid the potential confounding effect of pseudoatrophy. Our results indicate that even during the first month of follow-up, the accelerated atrophy following IA/aHSCT does not appear to have been associated with resolution of edema or T2-weighted lesions. The study by Petzold et al. showed that SPMS patients had significantly increased neuroaxonal damage 1 month after aHSCT compared to baseline or placebo-treated SPMS patients.[159] This suggests that neurotoxicity and neurodegeneration account for the volume loss after BMT. Resolution of inflammation, such as microglia transitioning from an activated state to a resting state, may also have contributed to atrophy; such a process is indistinguishable from neurodegeneration using conventional MRI.

We conducted an independent analysis to investigate the late atrophy rates from 2.5 years after treatment and onwards, when the effects associated with treatment-related neurotoxicity have largely resolved. The patients had marked baseline inflammatory activity; after treatment, they were free of new focal MRI activity and relapses throughout the follow-up. Theoretically, this means no new WM lesion-related secondary degeneration developed in the patients. This does not inform on the stoppage of other processes that potentially contribute to atrophy, such as diffuse/compartmentalized inflammation or GM demyelination, both undetectable with the imaging protocol used. The impact of this immunoablative therapy on any primary degenerative process, if it were present in our patients, is unclear.

The average rate of late atrophy in our patients was -0.23%/y. Since we lacked an age-matched healthy control group with which to make a direct comparison, we compared this rate to values in the literature obtained using a similar imaging protocol and the same measurement technique. De Stefano et al. reported the mean atrophy rate of -0.27%/y (SD:±0.15%/y) in healthy controls with the mean age of 37, which was comparable to that of our patients.[96] While the average rates were similar, the small sample size led to greater variance of rates in our patients. Nine out of the 15 subjects had atrophy rates lower than -0.27%/y, while 10 subjects had atrophy rates lower than -0.42%/y (-0.27%/y minus 1SD), and 14 subjects had atrophy rates lower than -0.57% (-

0.27% minus 2SD). This suggests that the long-term atrophy rate was within the normal-aging range in the majority of patients.

If, as our data suggests, the long-term atrophy rates in our patients have returned to near-normal levels, it could serve as an indirect evidence that IA/aHSCT halts inflammatory activity, which subsequently reduces the accelerated degeneration seen in patients with MS, and that neurodegeneration is, in fact, all secondary to ongoing inflammatory processes in the nervous system.

2.2.6 Acknowledgements

Funding acknowledgements

This work was supported by a Multiple Sclerosis Scientific Research Foundation grant from the Multiple Sclerosis Society of Canada [grant number: OHRI REB # 2000374-O1H]. Hyunwoo Lee was supported by the Doctoral Training Award from Fonds de recherche du Québec and a Neuroinflammation Training Program Award from the Canadian Institutes of Health Research.

Chapter 3 Impact of immunoablation and autologous hematopoietic stem cell transplantation on grey and white matter atrophy in multiple sclerosis

3.1 Preface

The previous chapter measured and modeled the time courses of whole-brain atrophy in the patients enrolled in the Canadian MS-BMT study. Whole-brain atrophy is an important metric for assessing neurodegeneration in MS, as it is a representation of irreversible structural tissue damage and loss.[202] However, studies have suggested that these destructive pathological processes, including neurodegeneration, may be better represented by atrophy in the grey matter. Grey matter atrophy starts early and accelerates during the course of MS and is better correlated with disability than whole-brain or white matter atrophy.[99] Furthermore, grey matter atrophy may be less susceptible to the so-called pseudoatrophy effect that can confound interpretation of the whole-atrophy findings.[115,116]

This chapter investigates the impact of aHSCT on grey and white matter atrophy in MS. In doing so, it addresses a series of questions aimed at further understanding of the effects seen on whole-brain atrophy in the Canadian MS-BMT patients. Specifically, how do grey matter and white matter individually contribute to the accelerated atrophy observed immediately following aHSCT?

In the following manuscript entitled "Impact of immunoablation and autologous hematopoietic stem cell transplantation on grey and white matter atrophy in multiple sclerosis", grey and white matter atrophy are separately measured in the patients enrolled in the Canadian MS-BMT study. Then, a non-linear mixed-effects model is used to model the time courses of grey and white matter atrophy. As in the previous chapter on whole-brain atrophy in these patients, two hypothesized predictors are used in the model: the dose of the chemotherapy drug busulfan (as a marker of the neurotoxicity of the conditioning regimen used for immunoablation) and the baseline volume of T1-weighted white matter lesions (a marker of the amount of tissue damage from pre-treatment inflammatory demyelination). Furthermore, long-term rates of grey and white matter atrophy are estimated.

3.2 Manuscript: Impact of immunoablation and autologous hematopoietic stem cell transplantation on grey and white matter atrophy in multiple sclerosis

Hyunwoo Lee, B.Sc. [1], Kunio Nakamura, Ph.D. [1,2], Sridar Narayanan, Ph.D. [1], Robert Brown, Ph.D. [1], Jacqueline Chen, Ph.D. [3], Harold L. Atkins, M.D. [4], Mark S. Freedman, M.D. [5], Douglas L. Arnold, M.D. [1]

[1] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[2] Department of Biomedical Engineering, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA

[3] Department of Neurosciences, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA

[4] Ottawa Hospital Blood and Marrow Transplant Program, The Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada

[5] Department of Medicine (Neurology), The Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada

Accepted for publication in: ***Multiple Sclerosis Journal***

3.2.1 Abstract

Background: Immunoablation and autologous hematopoietic stem-cell transplantation (IA/aHSCT) halts relapses, white-matter (WM) lesion formation, and pathological whole-brain atrophy in MS patients. Whether the latter was due to effects on grey-matter (GM) or WM warranted further exploration.

Objective: To model GM and WM volume changes after IA/aHSCT to further understand the effects seen on whole-brain atrophy.

Methods: GM and WM volume changes were calculated from serial baseline and follow-up MRIs ranging from 1.5 to 10.5 years in 19 MS patients treated with IA/aHSCT. A mixed-effects model with two predictors (total busulfan dose and baseline T1-weighted WM lesion volume "T1LV") characterized the time-courses after IA/aHSCT.

Results: Accelerated short-term atrophy of 2.1% and 3.2% occurred in GM and WM respectively, on average. Both busulfan dose and T1LV were significant predictors of WM atrophy, whereas only busulfan was a significant predictor of GM atrophy. Compared to baseline, a significant reduction in GM atrophy, not WM atrophy, was found. The average rates of long-term GM and WM atrophy were -0.18%/y (SE:0.083) and -0.07%/y (SE:0.14), respectively.

Conclusion: Chemotherapy-related toxicity affected both GM and WM. WM was further affected by focal T1-weighted lesion-related pathologies. Long-term rates of GM and WM atrophy were comparable to those of normal-aging.

3.2.2 Introduction

Brain atrophy associated with volume loss is a pathologic feature of MS, occurring at rates about two-fold higher than those of age-matched normal-aging.[96] Both the grey matter (GM) and white matter (WM) are affected, as indicated by significantly reduced volumes in both relapsing-remitting (RR) and progressive MS patients compared to healthy controls.[106,203] In particular, GM atrophy occurs at greater rates and is better correlated with disability and cognitive impairment than WM lesion load or whole-brain (WB) atrophy.[99,106,107,203] Furthermore, the rate of GM atrophy appears to accelerate with disease progression, whereas this is not the case for WM.[99] These findings suggest GM atrophy is a potentially relevant marker of disease progression in MS.

The mechanisms underlying GM atrophy in MS are multifactorial and poorly understood. Axonal transection in focal WM lesions, which causes Wallerian degeneration and axonal loss in WM,[100] may also result in retrograde degeneration and GM loss.[102] However, GM damage may also be independent of WM pathology. Cortical demyelination is common and extensive,[38] and is

associated with significant neuritic, oligodendrocytic, and neuronal injury in both progressive and early MS.[34,43,49] Cortical demyelinating lesions generally show relatively little inflammation, although they can be inflammatory early on.[43,49] Meningeal inflammation also contributes to cortical pathology.[47,49] Furthermore, extra-lesional cortex in MS contains abnormal neuronal morphology,[44] mitochondrial dysfunction,[204] smaller neurons and reduced axonal density.[205] These could also contribute to overall cortical atrophy.

Immunoablation and purified autologous hematopoietic stem cell transplantation (IA/aHSCT) has been shown to reduce relapses, expanded disability status scale (EDSS) progression, and new MRI WM lesion (gadolinium-enhancing or T2-weighted) formation in groups of MS patients. In particular, the Canadian MS-BMT trial reported complete cessation of relapses and new WM lesion formations in RR or secondary-progressive (SP) patients who were refractory to conventional disease-modifying treatments.[195] Moreover, the average rate of WB atrophy in the same patients declined to a level comparable to normal-aging over approximately 2-3 years, following an early acceleration associated with the immunoablative chemotherapy.[206]

It is unknown how GM and WM individually contribute to the accelerated atrophy observed immediately following initiation of IA/aHSCT. Understanding their individual contribution would help to clarify the mechanism of treatment related atrophy. Acceleration of GM volume loss would support primary neurotoxicity of the treatment regimen accompanied by degeneration of GM tissues that were irreversibly injured prior to treatment. Acceleration of WM volume loss would support degeneration of WM components such as axons or myelin due to the treatment regimen, further accompanied by degeneration of WM tissues that were irreversibly injured prior to treatment. Significant atrophy due to chemotherapy-related toxicity may be an unavoidable side-effect of IA/aHSCT, and may be associated with disability outcomes after treatment. Furthermore, it remains to be confirmed whether the subsequent slowing of WB atrophy to a normal-aging rate is reflected in both GM and WM levels in order to better understand the mechanism of atrophy in these patients.

We aimed to determine the short- and the long-term effect of IA/aHSCT on the courses of GM and WM atrophy in the Canadian MS-BMT patients. First, we used a longitudinal registration-based method that had a higher power for measurement of GM atrophy than other commonly used methods like SIENAX, Statistical Parametric Mapping (SPM), or FreeSurfer.[117] Then, we

used a mixed-effects model, which can handle unbalanced repeated measures data that frequently occur in longitudinal trials. This approach was similar to what we have used in our previous WB atrophy study.[206] Based on our previous knowledge, we hypothesized that for the first 2-3 years, the atrophy rates in both GM and WM are accelerated due to neurotoxicity of the conditioning regimen used for IA, plus ongoing secondary degeneration of tissues damaged by pre-treatment inflammatory demyelination. This, however, would eventually slow to the rates seen in normal-aging, in keeping with what we have reported in WB atrophy.[206]

3.2.3 Methods

Patients

The Canadian MS-BMT study was approved by the research ethics boards of the participating institutions. All participants provided written informed consent. This study comprised 24 patients that received immunoablative-dose busulfan, cyclophosphamide, and rabbit anti-thymocyte globulin chemotherapy followed by purified aHSCT. One death occurred due to chemotoxicity. The detailed procedure was previously reported.[195]

The follow-up duration ranged from 1.5 to 10.5 years (mean: 5.5y, N=23). To minimize the impact of MRI hardware upgrades, we selected 19 subjects. 17 subjects were initially scanned with a single 1.5T Siemens Symphony scanner. This scanner was upgraded to a Siemens Symphony total imaging matrix (TIM) system about 3.5 years after IA/aHSCT, on average, well after the period of accelerated atrophy. Scanning for the remaining two patients started only after the upgrade was completed. The four excluded subjects had two or more scanner changes. No significant baseline differences were found between the included and the excluded subjects (Table 3-1). The MRI protocol was identical to that shown in chapter 2. Examples of the T1-weighted images used in the study are shown in Figure 3-1.

Table 3-1: Baseline subject characteristics – Canadian MS-BMT

Characteristics	Included subjects (N=19; 12 RR + 7 SP)	Excluded subjects (N=4 SP)
Mean age (SD) [range], yr	32.7 (6.2) [23-45]	31.8 (4.6) [26-37]
Sex, Female:Male	12:7	2:2
Mean disease duration (SD), onset to IA/aHSCT, yr	6.9 (2.7)	10.1 (7.5)
Mean Expanded Disability Status Scale score (SD)	5.0 (1.1)	5.0 (1.2)
Mean annualized relapse rates for 2-years before IA/aHSCT	1.0 (0.5)	0.5 (0.4)
% of subjects with new or enlarging T2-weighted lesions, during baseline evaluation prior to IA/aHSCT	74	50
% of subjects with gadolinium-enhancing lesions, during baseline evaluation prior to IA/aHSCT	63	0
Mean baseline gadolinium-enhancing lesion volume (SD), ml	0.5 (1.1)	0
Mean baseline T1-weighted lesion volume (SD), ml	7.9 (7.8)	10.2 (10.2)
Mean baseline T2-weighted lesion volume (SD), ml	20.0 (18.0)	26.9 (18.1)
Mean total busulfan dose (SD), mg	723.6 (146.0)	850 (268.7)
Numbers of subjects that have reached each MRI follow-up time point, years from aHSCT, (N)	1.5y (19); 2.5y (17); 3y (15); 4.5y (11); 5.5y (9); 6y (7); 7.5y (6); 10y (5); 10.5y (2)	3y (4); 8y (2); 9y (1)
Median number of MRI scans available per subject, [range]	18 [11-24]	n/a

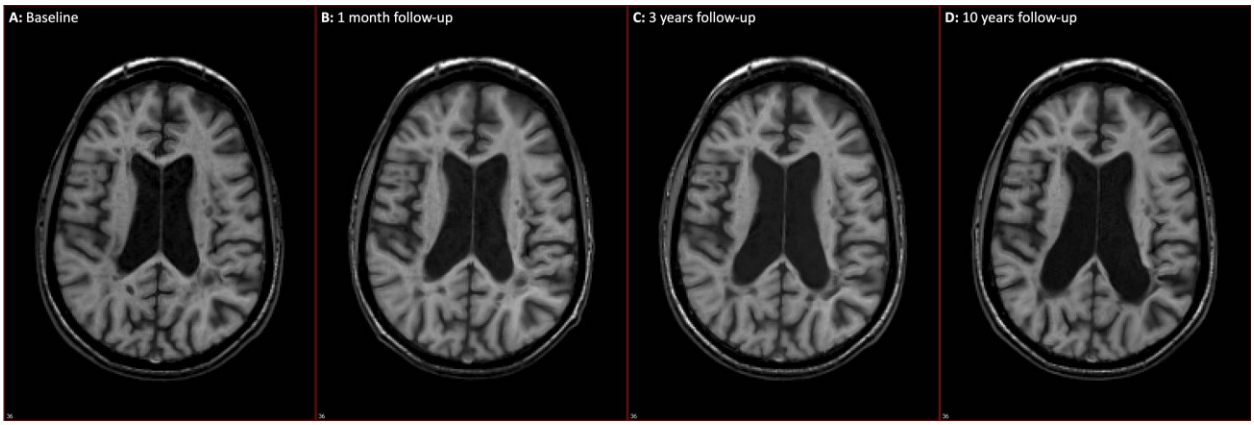

Figure 3-1: Example of T1-weighted images from a patient with 10 years of follow-up

A: Baseline scan; B: 1 month follow-up after IA/aHSCT (percentage volume changes from baseline: WB: -3.0%, GM: -2.4%, WM: -4.1%); C: 3 years follow-up (WB: -5.3%, GM: -3.7%, WM: -7.9% from baseline); D: 10 years follow-up (WB: -6.5%, GM: -5.0%, WM: -9.0% from baseline)

T2-weighted and T1-weighted lesion volume measurement

Hyper-intense WM lesions were identified on T2-weighted scans with an in-house Bayesian classifier and then manually corrected. Hypo-intense WM lesions were identified on pre-contrast T1-weighted scans as clusters of voxels within T2-weighted lesions having an intensity of 85% or lower than neighboring normal-appearing WM (NAWM).

Volume change measurement

Total (cortical and deep) GM and total WM volume changes from the baseline were calculated from pre-contrast T1-weighted images using Jacobian integration (JI), a longitudinal registration-based method that outputs a percent change in volume between two timepoints.[117] JI calls a series of pre- and post-processing steps that include GM and WM segmentation and lesion-filling to reduce the bias on segmentation of GM and WM due to the presence of WM lesion.[117] Examples of the segmented images are shown in Figure 3-2. As the absolute volume of cortical GM is far greater than that of deep GM,[207] it is likely that the cortical GM constitute a dominant proportion within the total GM volume. For a subject *K*, the volume at a follow-up timepoint *N* (in percent, with respect to the baseline) was calculated as follows: $Vol_{K,N} = 100\% + \%Change_{K,from\_Baseline\text{-}to\text{-}N}$. This produced GM and WM trajectories for each subject.

14 subjects had two baseline MRIs, and 4 subjects had three. For each subject, the baseline atrophy rate was calculated by fitting a linear regression line and multiplying the slope by the annualization factor.

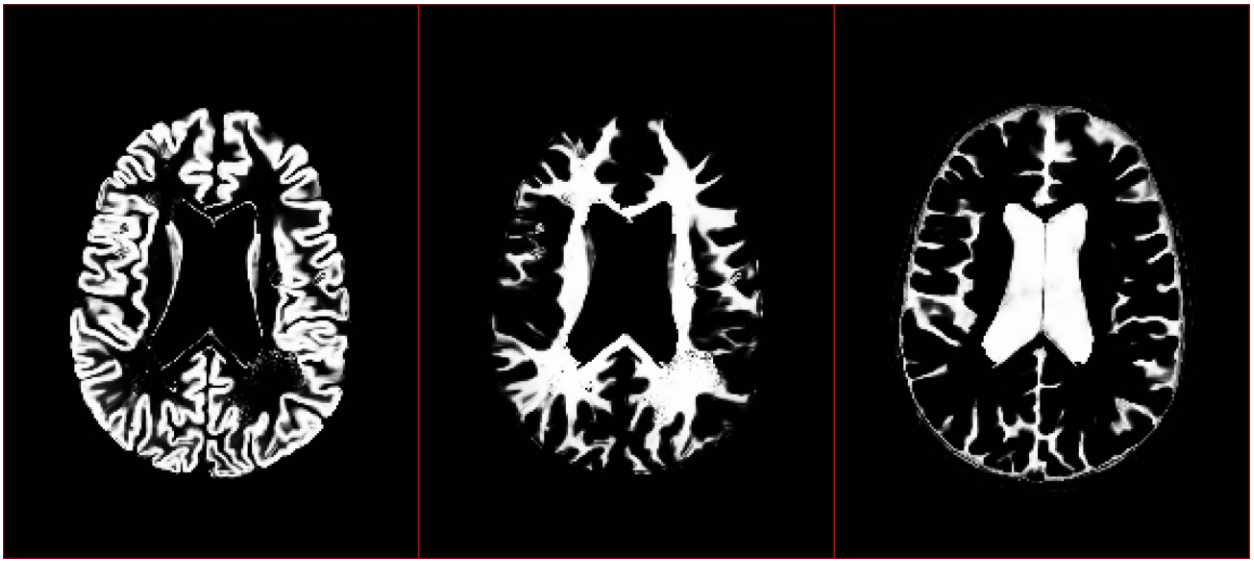

Figure 3-2: Example of grey matter, white matter, and cerebrospinal fluid segmentations

Left: grey matter; Middle: white matter; Right: CSF

Pseudo-T2-relaxation time calculation

Brain volume may decrease after anti-inflammatory therapy independently of actual tissue loss due to loss of tissue water that would accompany the resolution of inflammatory activity.[156] To explore the effect of tissue water fluctuation on tissue volume change, we used the change in a two-point estimated "pseudo" T2-relaxation time (pT2) as a marker of change in tissue water content.[156] pT2 was derived from dual-echo turbo spin-echo MRI with the formula $pT2=(E_2-E_1)/\ln(S_1/S_2)$ where S_N are the signal intensities at each echo time E_N. Average pT2 was calculated separately in normal-appearing GM and WM masks (excluding T2-weighted lesions) obtained with FSL-SIENAX.[208] Then, differences in the average pT2 were calculated between the baseline and the follow-up points at months 1, 6, 12, and 24.

Statistical analysis

We modeled the courses of GM and WM atrophy as a net effect of four processes, each represented by a sub-component of a time-dependent non-linear mixed-effects model. This approach was previously applied to model WB atrophy in the same patients.[206] We assumed that the GM and WM atrophy progressed in a comparable pattern: that is, temporary acceleration and eventual stabilization. First, exponential decay functions were used to model the early, accelerated atrophy observed after aHSCT. Second, a straight-line function was used to model the atrophy due to normal-aging and, if present, ongoing MS-related atrophy. Third, a categorical variable was used to estimate non-physiological volume fluctuation due to major MRI hardware upgrade. The analysis was performed using the SASv9.4 NLMIXED procedure.

$$Volume_{GM\ or\ WM}(time\ from\ aHSCT) =$$

$$(a_1 * BusDose) * exp^{-time*(b_1*BusDose)} \quad \textit{[Component\_1]}$$

$$+ (a_2 * T1LV) * exp^{-time*(b_2*T1LV)} \quad \textit{[Component\_2]}$$

$$+ (InteractionCoef * BusDose * T1LV) * exp^{-time*(Exponent*BusDose*T1LV)} \quad \textit{[Interaction]}$$

$$+ (LINRATE + RANDOM_{SLOPE}) * time + RANDOM_{INTERCEPT} \quad \textit{[Component\_3]}$$

$$+ \left(EFFECT_{ScannerChg}\right) * CategoricalVar_{Scanner} \quad \textit{[Component\_4]}$$

$$+ AsymptoteVol + error$$

[Component_1] was an exponential function whose coefficient (a_1*BusDose; representing the quantity of volume decay) and decay constant (b_1*BusDose; controlling the rate of decay) depended on the total dose of busulfan given for immunoablation (BusDose). Busulfan is a cytotoxic agent that easily crosses the blood-brain-barrier and penetrates the CNS.[196] Therefore, we used BusDose alone as an index of the treatment-related toxicity due to the combined effect of the entire conditioning regimen. There was variation in BusDose due to adjustment for first dose pharmacokinetics, and we modeled for an acute and dose-dependent effect on atrophy.

[Component_2] was defined similarly, except that it depended on the baseline T1-weighted WM lesion volumes (T1LV). Baseline T1-weighted and T2-weighted lesion volumes were highly correlated, and we decided to use T1LV in the model because T1-weighted lesions are characterized by extensive axonal transection, which leads to subsequent Wallerian degeneration and axonal loss extending into NAWM.[35,100] We assumed that T1LV denoted the tissues that were irreparably injured and committed to atrophy before treatment. BusDose and T1LV were uncorrelated. The AsymptoteVol parameter represented the remaining volume after the exponential function had levelled-off. We also explored the potential **[Interaction]** between BusDose and T1LV, i.e. whether the patients with higher T1LV were more susceptible to atrophy due to the effect of varying BusDose. Beyond BusDose and baseline WM lesion volumes, there were no significant correlations between atrophy and other baseline clinical characteristics.

Treatment-related atrophy is temporary.[206] However, the combined effect of normal-aging and potential MS-related atrophy that remains after treatment may yield a small-scale atrophy throughout the follow-up. We approximated the average rate of this atrophy with the slope parameter LINRATE, assuming a straight-line function, **[Component_3]**, that was independent of the temporarily-accelerated, treatment-related atrophy. Random effects ($RANDOM_{SLOPE}$, $RANDOM_{INTERCEPT}$) estimated the departures from the group average slope and intercept for each subject. We did not find significant effects of age and sex on atrophy rates, so these were not included in the model.

A hardware upgrade can alter various MR image characteristics, including contrast, signal-to-noise ratio, geometric distortion and intensity non-uniformity. This may distort atrophy measurements by introducing non-physiological volume fluctuation across the point of upgrade. **[Component_4]** used a categorical variable ($CategoricalVar_{Scanner}$) to estimate the fluctuation ($EFFECT_{ScannerChg}$) due to the TIM upgrade.

Individual rates of atrophy after IA/aHSCT were estimated by adding the values of LINRATE and subject-specific random slopes. These rates were compared against the baseline rates using Wilcoxon signed-rank test as the baseline rates were not normally distributed. Additionally, we tested whether the atrophy rates differed between those who had an increase in the Expanded Disability Status Scale (EDSS) and those who did not.

3.2.4 Results

Baseline atrophy rates:

The median baseline atrophy rates were (median, IQR): GM, -1.54%/y (-3.10, -0.43); WM, 0.36%/y (-0.57, 3.30). Due to the short (median: 0.17y) and varied (SD: 0.64y) baseline scan intervals, any noise in the volume measurements, such as the effects of new inflammatory lesions during the baseline evaluation, would have been further amplified by the annualization factors. To alleviate this issue, we separately calculated the baseline rates in the four patients who had three baseline MRIs over an average of 1.3y (median, IQR): GM, -1.03%/y (-2.31, -0.45); WM, -0.89%/y (-1.77, 0.58).

Longitudinal pT2 change from baseline:

Average differences in the pT2 between the baseline and the follow-up points at months 1, 6, 12, and 24 are shown in Table 3-2. No significant differences in average pT2 in GM or WM were observed after Bonferroni correction.

Table 3-2: Average pT2 differences from baseline to selected early follow-up timepoints

Mean baseline pT2 in GM: 116.65ms Mean baseline pT2 in WM: 101.46ms	**Average change in pT2 of normal-appearing GM, Mean (SD) [range], paired t-test outcome**	**Average change in pT2 of normal-appearing WM, Mean (SD) [range], paired t-test outcome**
Baseline to month 1 (N=19)	0.77ms (2.13) [-2.9, 5.4], n.s.	-0.55ms (1.67) [-3.9, 3.2], n.s.
Baseline to month 6 (N=19)	1.18ms (2.64) [-4.4, 6.8], n.s.	-0.64ms (2.03) [-4.8, 2.6], n.s.
Baseline to month 12 (N=18)	2.07ms (2.92) [-1.5, 8.4], n.s. after Bonferroni correction	0.57ms (3.06) [-5.7, 5.5], n.s.
Baseline to month 24 (N=18)	1.72ms (3.68) [-2.9, 9.2], n.s.	0.13ms (3.11) [-3.6, 6.6], n.s.

Model of the GM and WM atrophy:

Estimates of the model parameters are shown in Table 3-3.

GM: Overall, the exponential functions predicted about 2.1% GM volume loss with an approximate half-life of 0.19y. The exponential parameters indicated that only busulfan dose was a significant predictor. T1LV was not significant in this model (p=0.59). The interaction term was removed from the model due to failed convergence. The slope parameter LINRATE modeled GM atrophy as a straight-line function with the annualized rate of -0.18%/y, p=0.041. The overall effect of scanner upgrade was not statistically significant (-0.030%, p=0.81). Figure 3-3 shows the summary of the model.

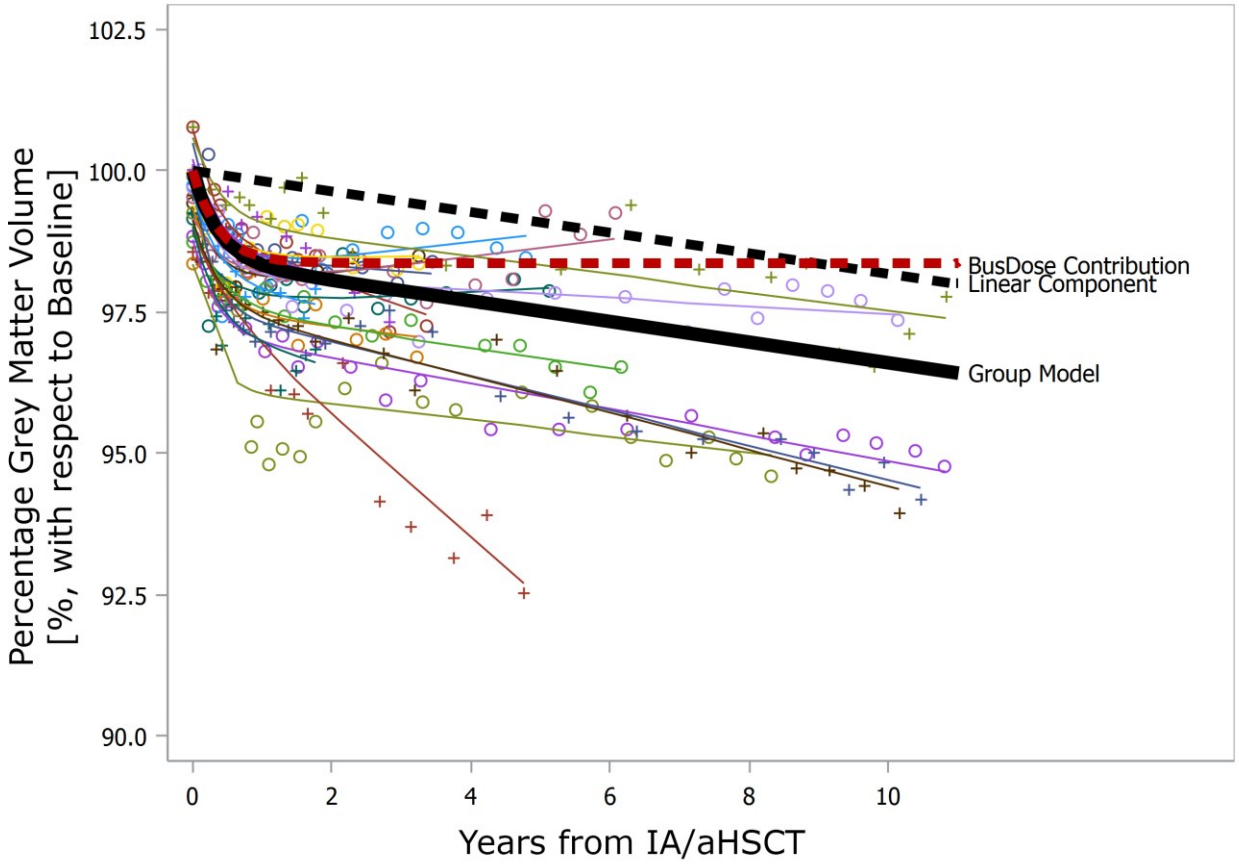

Figure 3-3: Evolution of grey matter atrophy after IA/aHSCT – Canadian MS-BMT

On average, the model predicted about 2.1% of GM volume loss immediately following treatment. Total busulfan dose was a significant predictor of GM atrophy. Baseline WM lesion volume was not a significant predictor. The thick solid and dashed lines represent the group average model for all subjects, obtained by inputting the average value of total busulfan dose into the fixed-effects.

Legend:

Group Model (solid black line) = Average group model for all subjects

Linear Component (dashed black line) = Linear model of a small-scale atrophy, potentially due to normal-aging

Busulfan Dose Contribution (dashed red line) = Exponential model associated with total busulfan dose

Colored Dots = Actual percent volume change measurements from the baseline for the individual subjects

Colored Unlabeled Fitted Lines = Fitted model trajectories for the individual subjects

WM: Overall, the exponential functions predicted about 3.2% WM volume loss with an approximate half-life of 1.18y. Both busulfan dose and T1LV were significant predictors of atrophy with a significant negative interaction. The slope parameter LINRATE indicated insignificant rate of average WM atrophy (-0.07%/y, p=0.60). The effect of scanner upgrade was significant with an average volume fluctuation of -0.50% (p=0.0091); the effect on the WM compartment was significantly larger than that on the GM compartment (p=0.03, N=17, two-tailed t-test). Figure 3-4 shows the summary of the model.

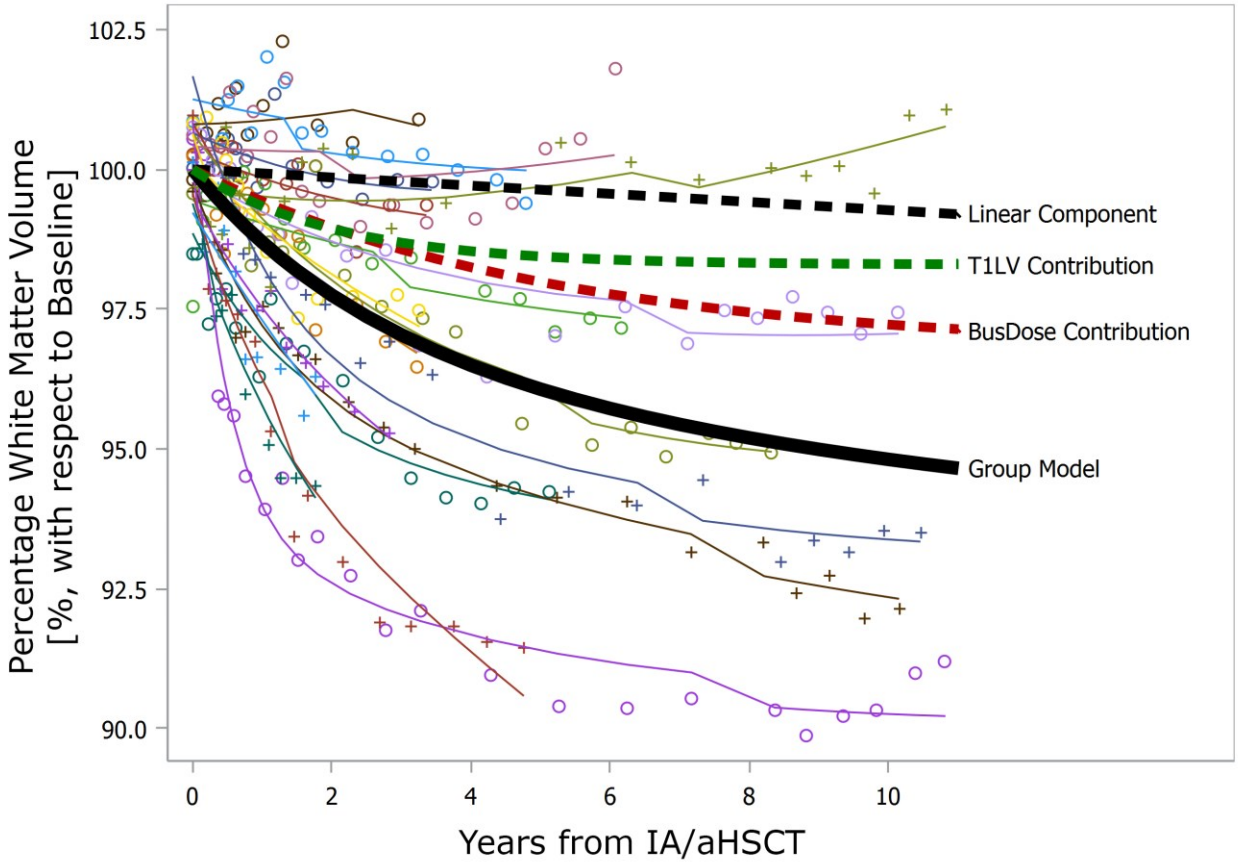

Figure 3-4: Evolution of white matter atrophy after IA/aHSCT – Canadian MS-BMT

On average, the model predicted about 3.2% of WM volume loss immediately following treatment. Both the total busulfan dose and the baseline WM lesion volume were significant predictors of WM atrophy. The thick solid and dashed lines represent the group average models for all subjects, obtained by inputting the average values of total busulfan dose and T1LV into the fixed-effects.

Legend:

Group Model (solid black line) = Average group model for all subjects

Linear Component (dashed black line) = Linear model of a small-scale atrophy, potentially due to normal-aging

Busulfan Dose Contribution (dashed red line) = Exponential model associated with total busulfan dose

T1LV Contribution (dashed green line) = Exponential model associated with baseline T1-weighted WM lesion volume

Colored Dots = Actual percent volume change measurements from the baseline for the individual subjects

Colored Unlabeled Fitted Lines = Fitted model trajectories for the individual subjects

Table 3-3: Parameter estimates for the models of GM and WM atrophy

Region	Parameter, df=17	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Grey Matter	**a_1: contribution of *BusDose* on decay quantity**	**0.0023**	**0.00024**	**9.3**	**<.0001**	**0.0017**	**0.0028**
	b_1: contribution of *BusDose* on decay rate	**0.0037**	**0.00093**	**4.03**	**0.0009**	**0.0018**	**0.0057**
	a_2: contribution of *T1LV* on decay quantity	**-0.0080**	**0.015**	**-0.55**	**0.59**	**-0.039**	**0.023**
	b_2: contribution of *T1LV* on decay rate	**0.14**	**0.17**	**0.82**	**0.42**	**-0.22**	**0.50**
	LINRATE: % change/year	**-0.18**	**0.083**	**-2.21**	**0.041**	**-0.36**	**-0.0082**
	$EFFECT_{ScannerChg}$	-0.030	0.12	-0.24	0.81	-0.29	0.23
	AsymptoteVol	**97.91**	**0.19**	**503.36**	**<.0001**	**97.50**	**98.32**
	Model Variance	0.23	0.020	11.58	<.0001	0.19	0.27
	$RANDOM_{SLOPE}$	0.085	0.031	2.73	0.014	0.019	0.15
	Random $Effect_{Covariance}$	-0.0075	0.051	-0.15	0.88	-0.11	0.10
	$RANDOM_{INTERCEPT}$	0.48	0.17	2.82	0.012	0.12	0.83
White Matter	**a_1: contribution of *BusDose* on decay quantity**	**0.0045**	**0.0012**	**3.69**	**0.0018**	**0.0019**	**0.0070**
	b_1: contribution of *BusDose* on decay rate	**0.00027**	**0.000069**	**3.92**	**0.0011**	**0.00012**	**0.00041**
	a_2: contribution of *T1LV* on decay quantity	**0.22**	**0.020**	**11.05**	**<.0001**	**0.18**	**0.26**
	b_2: contribution of *T1LV* on decay rate	**0.062**	**0.011**	**5.93**	**<.0001**	**0.040**	**0.084**
	InteractionCoef: contribution of interaction between *BusDose* and *T1LV* on decay quantity	**-0.00026**	**0.000037**	**-7.04**	**<.0001**	**-0.00034**	**-0.00018**
	InteractionExponent: contribution of interaction between *BusDose* and *T1LV* on decay rate	**-1.61E-07**	**1.53E-06**	**-0.11**	**0.92**	**-3.39E-06**	**3.07E-06**
	LINRATE: % change/year	**-0.073**	**0.14**	**-0.53**	**0.60**	**-0.36**	**0.22**
	$EFFECT_{ScannerChg}$	-0.50	0.17	-2.94	0.0091	-0.85	-0.14
	AsymptoteVol	**96.81**	**0.85**	**114.28**	**<.0001**	**95.02**	**98.60**
	Model Variance	0.39	0.035	11.41	<.0001	0.32	0.47
	$RANDOM_{SLOPE}$	0.20	0.084	2.37	0.030	0.022	0.38
	Random $Effect_{Covariance}$	0.16	0.11	1.51	0.15	-0.063	0.38
	$RANDOM_{INTERCEPT}$	0.75	0.27	2.77	0.013	0.18	1.33

Comparisons between the baseline and the post-treatment atrophy rates:

GM atrophy rates after aHSCT (average: -0.18%/y) were significantly lower than the baseline rates, p=0.0237, N=18, Wilcoxon signed-rank test. There was no significant difference between baseline and post-treatment atrophy rates in WM.

Comparisons between the patients with EDSS increase versus stable/decrease:

There was a trend towards lower atrophy rates in the patients who had stable (N=7) or improved (N=8) EDSS scores after IA/aHSCT versus those who progressed (N=4), although the differences were not statistically significant.

3.2.5 Discussion

GM atrophy may better reflect neurodegeneration and is potentially less susceptible to pseudoatrophy compared to WB atrophy. We assessed the effect of IA/aHSCT on GM and WM atrophy in MS patients that showed no evidence of relapses or new MRI lesions during up to 10.5 years of follow-up.[195] Immediately following treatment, short-term atrophy rates were accelerated in both GM and WM. Total dose of busulfan was a significant predictor of atrophy in GM. In WM, both the busulfan dose and the baseline volume of WM lesion significantly predicted atrophy. Following the accelerated periods, the average rates of both GM and WM atrophy were consistent with normal-aging level.

We observed substantial GM and WM atrophy following IA/aHSCT, as shown by the average (N=19) change of -2.12% (SD=0.97) in GM and -0.96% (SD=1.42) in WM over the first month of follow-up. GM and WM atrophy were correlated, suggesting the possibility of a common mechanism. First, we considered the potential contribution of so-called pseudoatrophy. However, the lack of measured decreases in pT2 did not support the presence of volume loss due to reduction in tissue water. In fact, the average volume loss during the first month of follow-up was significantly greater in the GM than in the WM, which is contrary to the typical findings that pseudoatrophy mainly occurs in the WM.[153] These imply the presence of mechanisms other than pseudoatrophy that can affect both GM and WM tissues. Nevertheless, the lack of evidence for changes in tissue water does not completely exclude the possibility of atrophy related to resolution of inflammation; there could also have been a decrease in the volume of inflammatory

cells, such as microglia, due to change in activation state, which would be indistinguishable from true tissue loss on MRI.

Still, it is likely that true tissue loss predominated the accelerated volume loss during the early follow-up period after IA/aHSCT, considering that the model provided evidence of the significant roles of the baseline predictors, busulfan dose and pre-treatment T1LV. In GM, higher busulfan doses were associated with greater atrophy. In WM, both busulfan dose and T1LV significantly predicted greater atrophy. These results show the immediate, dose-dependent effect of busulfan on both GM and WM. CNS cells exposed to chemotherapeutic drugs are subject to toxic effects, including direct toxicity, oxidative stress, metabolic alteration, and inflammation, which can induce atrophy and cognitive dysfunction.[209] For example, both cortical and deep GM volumes were shown to be decreased in groups of breast cancer survivors treated with chemotherapy.[210] This indicates that GM atrophy in MS patients after IA/aHSCT is attributable to at least two different processes, one due to potentially remaining MS-related pathology and another due to the immunoablative chemotherapy. Compared to neurologically-healthy patients, the overall degree of treatment-related atrophy might be greater in MS patients due to extensive neuroaxonal injury that might render the brain more vulnerable to chemotoxicity. We tested whether the atrophy associated with busulfan dose could be more severe in patients with presumably more injured brain (higher T1LV). The negative interaction in the model for WM tissue indicated this was not the case within our patients.

Contrary to previous studies,[104,203] we did not find a significant relationship between T1LV and GM atrophy. However, a direct comparison is difficult because in our patients the high rate of chemotherapy-related atrophy could have masked the effect of retrograde degeneration, which is slow,[211] is less prominent in progressive phases,[212] and is only a part of the GM pathology.[213] Our imaging protocol was insensitive to GM lesions, another potentially relevant predictor, although their association with GM atrophy is unclear.[111] We did not measure cortical and deep GM volumes separately, but it is known that atrophy occurs in both GM regions.[214,215] The high rate of atrophy immediately following the immunoablative treatment suggests that the effects of the treatment predominate in this period. Yet, the observation that long-term rates of GM and WM atrophy in our patients are comparable to those of normal aging implies that atrophy rates in both cortical and deep GM regions have stabilized. In the present study, we focused on the

separate predictors of WM and GM atrophy. Future analyses will explore the predictors of the separate components of GM, including deep GM and cortex.

It is well-established that axonal transection within WM lesions results in Wallerian degeneration impacting NAWM.[100] Figure 3-5 shows the group model of GM and WM atrophy for average busulfan dose and T1LV values. While the overall percentage volume loss was greater in WM, the decay progressed over a longer period. The overall time-course of WM atrophy was comparable to that found in CNS Wallerian degeneration, which takes months to years.[201]

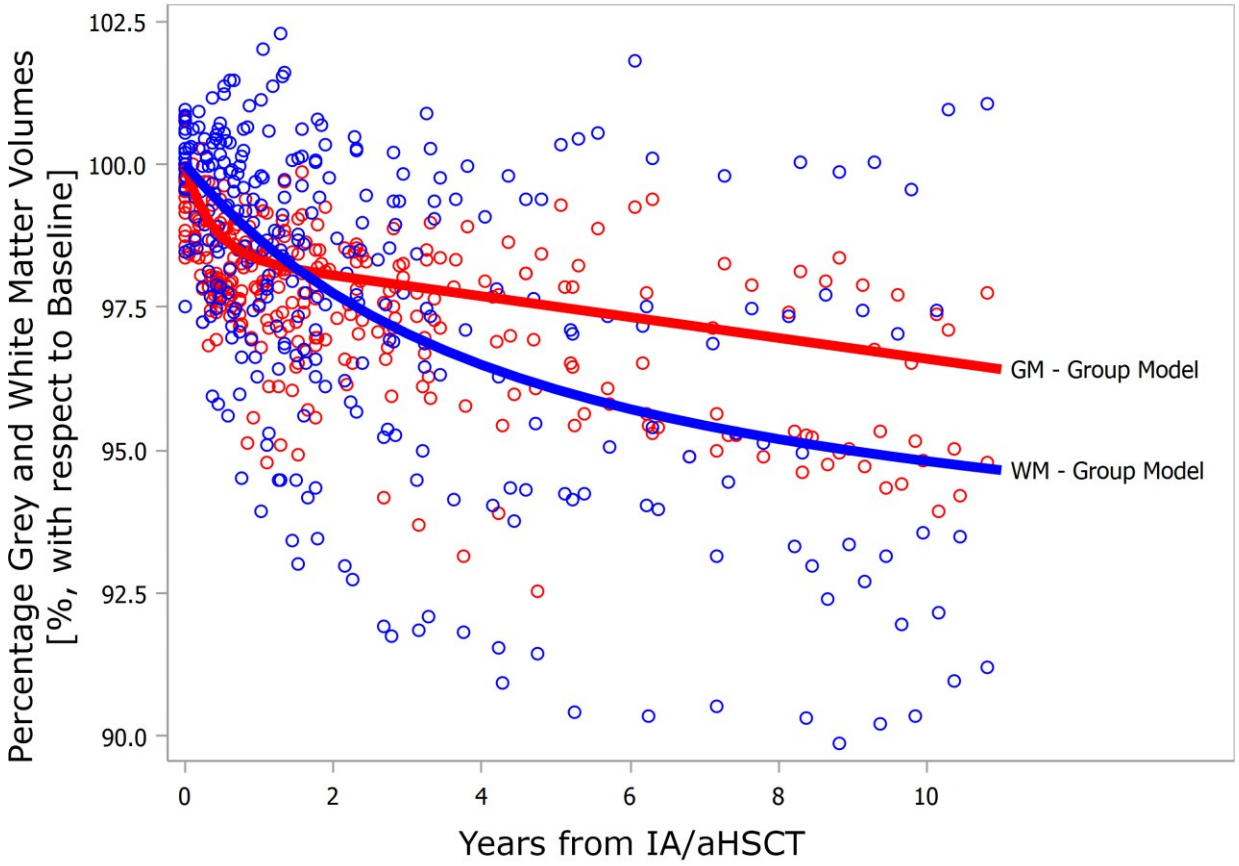

Figure 3-5: Comparison of grey and white matter atrophy after IA/aHSCT – Canadian MS-BMT

On average, the grey matter compartment was characterized by an accelerated period of atrophy that lasted about an year following IA/aHSCT. The accelerated period of atrophy was longer in the white matter compartment, resulting in a relatively greater loss of volume over time.

Legend:

GM – Group Model (solid red line) = Average group model of GM atrophy for all subjects, approximate half-life: 0.19y

WM – Group Model (solid blue line) = Average group model of WM atrophy for all subjects, approximate half-life: 1.18y

Red Dots = Actual GM percent volume change measurements from the baseline for the individual subjects

Blue Dots = Actual WM percent volume change measurements from the baseline for the individual subjects

The model term "LINRATE" indicated a consistent, small-scale GM atrophy throughout the follow-up with a small variance in the rates. WM atrophy rates were relatively more varied, but were close to zero on average. These patterns were similar to those found in normal-aging. Models of age-related changes have reported a linear decline in GM volume throughout adulthood, e.g. a 14% loss of cerebral cortex between the ages of 30-90 (-0.23%/year)[130] and a 12% loss between 30-80 years (-0.24%/year).[134] In WM, a polynomial-like growth-and-decline pattern that plateaued around the fourth-decade was found.[134] Considering the differences in the study methods, we cannot directly compare these values to our findings. Yet, we found a significant reduction in the GM atrophy rates after IA/aHSCT compared to the baseline rates. This was in the context of the slowing of average WB atrophy rate to that expected for normal-aging.[206] Neuroaxonal injury and degeneration in MS are closely associated with inflammatory damage in both the acute/relapsing and the progressive disease types.[190] Although IA/aHSCT arrested all measurable inflammatory activity in our patients,[195] it is unknown whether the treatment halted inflammation undetectable by clinical means or conventional imaging. However, one may deduce from the eventual reduction of GM atrophy that IA/aHSCT effectively halted inflammation that precedes neurodegeneration in MS.

The effect of the Siemens TIM upgrade on volume measurement was significant only in the WM compartment and not in the GM. Comparison of the means and the distributions revealed that the upgrade effect was indeed significantly larger in the WM than in the GM. As the TIM upgrade was a major upgrade that affected the gradient system, RF coil, and software, it is difficult with the current study to determine the exact cause of the systematic shift in the WM volumes. Previous findings have suggested that gradient nonlinearity,[216] intensity non-uniformity,[216] and number of head coil channels,[217] can cause significant bias on whole-brain volume measurements. A conceivable explanation for our finding is that the CSF/GM/WM contrasts was altered with the TIM upgrade. This would have affected GM and WM segmentation and led to systematic shifts in the compartmental volumes. A previous study has shown that images with higher WM/GM contrast lead to better delineation of the CSF/GM/WM boundaries.[182] Different MRI scanner hardware can provide a range of GM/WM volume ratio, and shifting of the CSF-GM-WM boundaries can occur when the same subject is scanned on different scanners; the specific behavior of such shifts differs across unique sets of scanner change combinations.[182] Therefore, our finding highlights the necessity to estimate and account for scanner changes,

especially in the study of brain compartment volumes, as the magnitude of the bias introduced by a hardware change can exceed those due to physiological or pathological processes. Previous studies have recommended mixed-effects models as a powerful method to analyze both cross-sectional and longitudinal data collected from multiple scanners.[185,186] We have used a similar approach to correct for the potential systematic bias on volumes, but how a scanner upgrade exactly affects atrophy measurements is an important topic that warrants further investigation.

We found no significant differences in the atrophy rates between the EDSS-progressed versus stable/improved patient groups. This was likely due to the paucity (N=4) of the progressed patients. A larger-scale controlled study is necessary to confirm not only whether reduction in the atrophy rates is paralleled by long-term stability or improvements in disability, but also whether IA/aHSCT possibly offers neuroprotection and stops GM and WM atrophy in MS.

Our results argue for the significant contribution of chemotherapy-related toxicity and secondary degeneration to the accelerated GM and WM atrophy after IA/aHSCT. The potential return of long-term atrophy rates to normal levels suggests that IA/aHSCT halted inflammatory activity, including that undetectable with MRI, and subsequent long-term chronic neurodegeneration.

3.2.6 Acknowledgements

Funding acknowledgements

This work was supported by a Multiple Sclerosis Scientific Research Foundation grant from the Multiple Sclerosis Society of Canada [grant number: OHRI REB # 2000374-O1H]. Hyunwoo Lee was supported by the Doctoral Training Award from Fonds de recherche du Québec.

Chapter 4 Brain atrophy in relapsing remitting multiple sclerosis following high-dose immunosuppressive therapy and autologous hematopoietic cell transplantation in the HALT-MS trial

4.1 Preface

As reviewed in chapter 1, different aHSCT trials for MS have reported variable outcomes in terms of controlling inflammatory MS activity, such as relapses and new white matter lesion formation. Various factors could have contributed to the differences, including inclusion criteria (e.g. age, MS subtype, baseline EDSS, DMT status) and the aHSCT procedure (e.g. intensity of the conditioning regimen).[218] One of the distinct characteristics of the Canadian MS-BMT study was the use of a high-dose, high-intensity busulfan/cyclophosphamide/antithymocyte globulin regimen. The results from chapters 2 and 3 suggest that chemotherapy-related CNS toxicity is likely responsible for the early acceleration of brain atrophy in these patients, and that stoppage of measurable inflammatory activity may be paralleled by long-term reduction in atrophy rates to rates expected with normal aging.

This chapter investigates the time courses of brain atrophy in patients enrolled in the HALT-MS trial. In doing so, it addresses another set of questions aimed at further understanding the effect of aHSCT on brain atrophy in MS. First, to what extent do the rates of brain atrophy accelerate following aHSCT in a cohort that received the intermediate-intensity BEAM chemotherapy regimen? Second, do the atrophy rates subsequently also slow down to rates expected for normal aging? It should be noted that a subset of patients in the HALT-MS trial continued to show signs of inflammatory activity after aHSCT.

In the following manuscript entitled “Brain atrophy in relapsing remitting multiple sclerosis patients following high-dose immunosuppressive therapy and autologous hematopoietic cell transplantation in the HALT-MS trial”, whole-brain, grey matter, and white matter atrophy are measured in the patients enrolled in the HALT-MS study. A non-linear mixed-effects model is

applied to model the time courses of atrophy. Two predictors are used in the model, including the total dose of BEAM chemotherapy regimen, and the baseline volume of T1-weighted white matter lesions. Long-term rates of atrophy are also estimated from the model.

4.2 Manuscript: Brain atrophy in relapsing remitting multiple sclerosis following high-dose immunosuppressive therapy and autologous hematopoietic cell transplantation in the HALT-MS trial

Hyunwoo Lee, B.Sc. [1], Kunio Nakamura, Ph.D. [3], Sridar Narayanan, Ph.D. [1], Robert A. Brown, Ph.D. [1], Richard A. Nash, M.D. [4], Douglas L. Arnold, M.D. [1,2]

[1] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[2] NeuroRx Research, Montreal, Quebec, Canada

[3] Department of Biomedical Engineering, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA

[4] Colorado Blood Cancer Institute, Denver, Colorado, USA

Manuscript in preparation

4.2.1 Abstract

Objective: To model the time-courses of whole-brain (WB), grey-matter (GM), and white-matter (WM) volume loss in relapsing-remitting multiple sclerosis (RRMS) patients treated with high-dose immunosuppressive therapy and autologous hematopoietic cell transplantation (HDIT/HCT) in order to better understand the mechanisms underlying brain atrophy associated with this treatment and MS.

Methods: We measured WB, GM, and WM volume changes over up to 5 years in twenty-four RRMS patients who underwent BEAM-based HDIT and HCT. A non-linear mixed-effects model was applied to the measurement data, with the total dose of BEAM chemotherapy (BEAMdose) and the baseline volume of T1-weighted WM lesion (T1LV) as the predictors of treatment-

related volume loss. We also estimated the rates of long-term volume loss not due to the treatment-related effects.

Results: Accelerated short-term brain volume loss of 0.98% (SE: 0.60), 1.2% (SE: 0.77), and 1.6% (SE: 0.84) occurred in WB, GM, and WM after HDIT/HCT, respectively. BEAMdose was a significant predictor of the accelerated volume loss in all compartments. T1LV was a significant predictor in the WB and the WM. The average rates of volume loss after the initial acceleration were -0.22%/y (SE: 0.075), -0.12%/y (SE: 0.081), and -0.15%/y (SE: 0.11) in the WB, GM, and WM, respectively.

Notably, patients with gadolinium-enhancing lesions at baseline had significantly higher short-term rates of GM volume loss (-1.83% vs. -0.40%, p=0.003) and WB volume loss (-1.90% vs. -0.76%, p=0.002) at 1-year follow-up, and also a significantly higher long-term rate of GM volume loss (-0.26%/y vs. -0.0051%/y, p=0.025) and a trend toward a higher long term rate of WB (-0.32%/y vs. -0.13%/y, p=0.081) volume loss rates compared to those without.

Conclusions: Brain volume loss may accelerate for months after HDIT/HCT due to the effects of chemotherapy-related toxicity and progression of WM lesion-related degeneration. However, over the long-term, adequate immunosuppression/immunoablation and HCT can reduce brain volume loss rates to those consistent with normal aging.

4.2.2 Introduction

Patients with multiple sclerosis (MS) suffer higher rates of brain atrophy (represented by brain volume loss (BVL), measured by magnetic resonance imaging (MRI)) compared to normal control subjects.[96] BVL begins early in the course of MS, and affects both the grey- (GM) and white-matter (WM).[99] BVL is associated with measures of disability such as the Expanded Disability Status Score (EDSS), such that a higher rate of volume loss is associated with a higher rate of EDSS progression.[96] Accordingly, BVL is frequently used as a marker of treatment response in the evaluation of disease-modifying treatments (DMTs).[219]

Some patients are refractory to the conventional DMTs and continue to have clinical relapses, MRI lesions, or EDSS progression despite being on therapy. In response, a small number of poor-prognosis patients have been treated with high-dose immunosuppressive therapy (HDIT)

and autologous hematopoietic stem-cell transplantation (HCT) based on the hypothesis that one-time ablation of a patient's immune system can eliminate autoreactive pathogenic clones, and subsequent reinfusion of purified hematopoietic stem-cells can repopulate an antigen-naïve system free of the autoimmune memory cells.[90] In the "High-Dose Immunosuppression and Autologous Transplantation for MS" (HALT-MS) trial, the 5-year progression-free, relapse-free, and MRI activity-free survivals were 91.3%, 86.9%, and 86.3%, respectively.[220] Seven of the 24 patients did not maintain event-free survival (EFS, defined as EDSS increase >0.5, or development of new relapses or new MRI lesions).[220]

Despite significant suppression of relapses or MRI lesions, multiple independent cohorts of MS patients treated with HDIT/HCT, including HALT-MS patients, have consistently shown accelerated whole-brain volume loss (WBVL) over the first two years after treatment.[174–177,206,220] However, the observation that the acceleration of WBVL starts immediately after treatment but lasts only a few years implies that it may be a temporary phenomenon associated with the treatment. A previous finding suggested that the cytotoxic effect of a chemotherapeutic regimen used for HDIT may be the main culprit for the accelerated WBVL, while pathological processes such as Wallerian degeneration may contribute as well.[206]

We analyzed the MRI follow-up dataset from the HALT-MS trial, with the objective of measuring WBVL, GM volume loss (GMVL), and WM volume loss (WMVL) and modeling their time-courses. The information obtained from statistical models can provide important insights into the dynamics of BVL after HDIT/HCT. First, we hypothesized that the combined effect of neurotoxicity from the chemotherapy used for HDIT and secondary degeneration resulting from axonal injury/transection in focal WM inflammatory lesions plays a major role in the temporary acceleration of BVL immediately after HDIT/HCT. Using the models, we estimated the amounts of volume losses attributable to these two predictors and the timeframes of the treatment-related accelerated BVL. The timeframe of the accelerated BVL remains an important piece of information to be gleaned from a longitudinal follow-up study because any slowing of BVL in response to treatment will be confounded by the early, treatment-related acceleration. Therefore, estimation of the rates of BVL in the patients under the influence of the reconstituted immune system will be more reliable and informative if this time-course can be defined and factored out of the atrophy trajectories. Second, we hypothesized that sustained

control of new inflammatory activity after HDIT/HCT would be associated with reduced rates of BVL, comparable to a level observed in normal-aging over the long term (after the early acceleration). To evaluate this, we used statistical models to estimate the rates of WBVL, GMVL, and WMVL while factoring out the treatment-related acceleration.

4.2.3 Methods

Subjects and imaging protocol

The HALT-MS study was approved by the institutional review boards at participating sites. All participants provided written informed consent. Detailed eligibility criteria and treatment procedures were previously described.[220] 25 relapsing-remitting MS (RRMS) patients were enrolled, but 24 underwent HDIT using high-dose BEAM (BiCNU, etoposide, ara-c, melphalan) and rabbit anti-thymocyte globulin, and then HCT. Maximum follow-up duration was five years with MRI twice at baseline and at months 2, 6, 12, 24, 36, 48, and 60 after HDIT/HCT. Table 4-1 summarizes the eligibility criteria and baseline subject characteristics. MRI scanning were done at three sites using 1.5T scanners from Philips, GE, and Siemens. Detailed MRI protocol and hardware upgrade information are described in the chapter appendix.

Table 4-1: Baseline subject characteristics – HALT-MS

Characteristics, N=25 unless otherwise stated	
Eligibility criteria: RRMS; EDSS 3.0-5.5 at baseline; Presence of brain MRI lesions; Less than 15 years of disease duration; failure of DMT[220]	
Mean age at mobilization (SD) [range], yr	37.3 (7.7) [27-53]
Sex, Female:Male	17:8
Mean disease duration (SD) [range], diagnosis to HDIT/HCT, yr	5.7 (3.7) [0.6-12.0]
Mean Expanded Disability Status Scale score (SD) [range]	4.4 (0.6) [3.0-5.5]
% of subjects with gadolinium-enhancing lesions during baseline evaluation prior to HDIT/HCT ("screening" or "baseline" timepoints), N=24 with MRI follow-up	46 (11 out of N=24)
Mean T1-weighted lesion volume (SD), ml, N=24 with MRI follow-up	1.2 (2.6)
Mean T2-weighted lesion volume (SD), ml, N=24 with MRI follow-up	10.9 (13.4)
Mean total BEAM dose (SD), mg, N=24 with MRI follow-up	3661 (345)

White-matter lesion volume measurement

A locally-developed Bayesian classifier was used to define areas of WM hyper-intensity on T2-weighted images, which were then manually corrected. T1 hypointense lesions were segmented on pre-contrast T1-weighted images that had an intensity of 85% or lower than surrounding normal-appearing WM (NAWM) and were also classified as T2 lesions.

Brain volume change measurement

Longitudinal pairwise Jacobian integration (PJI) was used to calculate WB, total GM, and total WM volume changes with respect to the last available baseline scan before HDIT/HCT ("baseline" for N=23, and "screening" for N=1 because the "baseline" scan was unreliable). Detailed steps of the PJI have been previously described.[117] Briefly, the input files were pre-contrast T1-weighted images and T2-weighted lesion masks. All T1-weighted images were pre-processed using a cross-sectional pipeline comprising intensity non-uniformity correction, lesion-filling, and standard-space registration. Then, each baseline-follow-up pair of pre-processed images was nonlinearly registered to determine the percent volume changes within the WB as well as the GM and WM masks segmented using Statistical Parametric Mapping (SPM).[221] Accordingly, we obtained three percent volume change values with respect to the baseline, for every follow-up scan from each subject.

Statistical analysis

We modeled the time-courses of WBVL, GMVL, and WMVL using a non-linear mixed-effects model approach, which we have previously applied to another trial.[206] Several independent MS cohorts treated with HDIT/HCT have consistently shown a distinct pattern of WBVL, that is, temporary acceleration of atrophy followed by eventual attenuation.[177,206] Our model captured this characteristic with an exponential term and a linear term, which are described as follows. The analysis was conducted using NLMIXED in SASv9.4:

$$Volume_{WB\ or\ GM\ or\ WM}(time\ from\ HCT) =$$

$$(a_1 * BEAMDose) * exp^{-time*(b_1*BEAMDose)} \quad \textit{[Component\_1]}$$

$$+ (a_2 * T1LV) * exp^{-time*(b_2*T1LV)} + AsymptoteVol \quad \textit{[Component\_2]}$$

$$+ (LINRATE + RANDOM_{SLOPE}) * time + RANDOM_{INTERCEPT} \quad \textit{[Component\_3]}$$

$$+ \left(EFFECT_{ScannerChg}\right) * CategoricalVar_{Scanner} \quad \textit{[Component\_4]}$$

$$+ error$$

Components 1 and 2: The exponential decay terms modeled the temporary acceleration of volume loss immediately after HDIT/HCT. Accordingly, the exponential coefficient estimated the amount of volume loss explained by the function, and the decay constant dictated how quickly the curve reached the half-point. We provided two exponential functions, related to each of the total dose of BEAM chemotherapy ("BEAMDose") and baseline T1-weighted WM lesion volume ("T1LV"). BEAMDose was used an index of the treatment-related neurotoxicity. It was calculated for each subject by combining the total doses (in mg) of BiCNU, etoposide, ara-C, and melphalan administered over the duration of immunosuppression. T1-weighted lesions represent significant tissue destruction and axonal loss,[52] and transected axons within these lesions may cause secondary Wallerian degeneration of the surrounding tissues.[30] T1LV was used to represent the amount of tissues that were irreversibly injured before treatment and may undergo degeneration regardless of the immunosuppressive therapy. T1LV for each subject was

calculated as the volume within the T1-weighted lesion mask at baseline. BEAMDose and T1LV were uncorrelated. Exponential functions are destined to level-off and reach the asymptote. The difference between the asymptote volume and the baseline volume would correspond to the amount of volume loss explained by the exponential models. We estimated this using the AsymptoteVol parameter.

Component 3: Rates of BVL in the patients continue to be non-zero due to normal-aging plus any other causes not accounted for by the exponential functions, such as MS-related atrophy that could persist even after treatment. We used a linear function with the slope term LINRATE to capture this. The rates estimated here were assumed to be constant throughout the follow-up period. Between-subject variability in the rates was expected to be present and modeled with the subject-specific random slope parameter $RANDOM_{SLOPE}$. The random intercept parameter $RANDOM_{INTERCEPT}$ accounted for inter-subject variability of the intercepts and intra-subject correlation of repeated measurements.

Component 4: For the patients who were affected by MRI scanner hardware changes/upgrades during the follow-up (details described in the chapter appendix), an additive correction was applied to address potential technical sources of volume fluctuation across the timepoint of upgrade. A categorical variable was assigned for the pair of scanners.

Subgroup comparisons

We compared the short-term accelerated WBVL, GMVL, and WMVL (represented by the raw percent WB, GM, WM volume changes at 1y follow-up; available for N=21 patients) and the long-term rates of WBVL, GMVL, and WMVL (represented by the LINRATE term, where the rates for each individual was estimated by adding the corresponding random slope to the mean; available for all patients) in two subgroup comparisons. First, we compared the rates between the patients with (Gd+, N=11) and without (Gd-, N=13) gadolinium-enhancing lesions during baseline evaluation prior to HDIT/HCT ("screening" or "baseline" timepoints). Second, we compared the rates in the patients with event-free survival maintained (EFS+, N=17) versus non-maintained (EFS-, N=7) by the end of follow-up. The types and timing of the subsequent events were previously reported,[220] and are summarized in the chapter appendix.

4.2.4 Results

Whole-brain (Figure 4-1, Table 4-2): The exponential models estimated an average of 0.98% (SE: 0.60) accelerated WBVL after HDIT/HCT. Both the BEAMDose and baseline T1LV were significant predictors of the accelerated WB volume loss, as represented by the exponential coefficients. However, neither were significant predictors of the decay rates, as represented by the decay constants. We calculated the average contribution of each predictor by inputting their mean values to the exponential model estimates: BEAMDose predicted about 0.89% volume loss with an approximate half-life of 0.1y, whereas baseline T1LV predicted about 0.16% with a half-life of 0.9y. The mean rate of the linear-change volume loss was -0.22%/y (SE: 0.075). Tables 4-2, 4-3 and 4-4 include the results of the additive corrections for the scanner upgrades for WB, GM and WM, respectively.

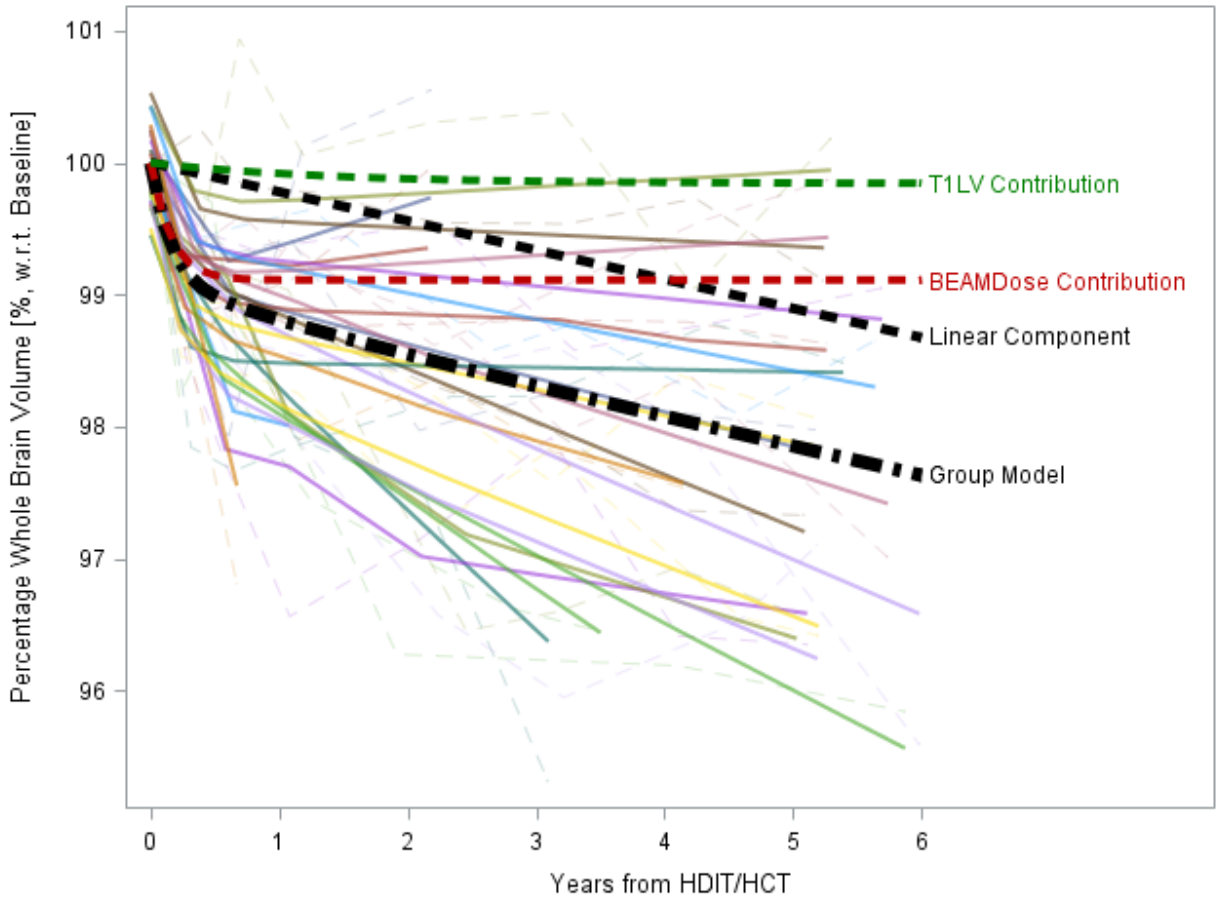

Figure 4-1: Evolution of whole-brain volume loss following HDIT/HCT – HALT-MS

Total BEAM dose (a marker of the chemotherapy-related toxicity) was a significant predictor, contributing about 0.89% whole-brain volume loss with an approximate half-life of 0.1y. Baseline T1-weighted lesion volume (an index of the amount of tissues lethally injured before treatment) was also a significant predictor, contributing about 0.16% whole-brain volume loss with an approximate half-life of 0.9y. The model of linear-change volume loss estimated about -0.22% loss per year.

Legend:

BEAMDose Contribution (dashed thick red line) = Exponential decay model associated with total BEAM dose

T1LV Contribution (dashed thick green line) = Exponential decay model associated with baseline T1-weighted WM lesion volume

Linear Component (dashed thick black line) = Model of the linear-change whole-brain volume loss, potentially due to normal-aging

Group Model (solid thick black line) = Net model for the cohort, i.e. the sum of T1LV and BEAMDose curves and the Linear component

Colored dashed transparent lines = Whole-brain volume change measurement data for the individual subjects

Colored solid thin lines = Fitted model for the individual subjects

Table 4-2: Parameter estimates for the model of whole brain volume loss after HDIT/HCT

Compartment	Parameter, df=22	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Whole Brain	**a_1: contribution of *BEAMDose* on exponential coefficient**	**0.00024**	**0.000049**	**4.99**	**<.0001**	**0.00014**	**0.00034**
	b_1: contribution of *BEAMDose* on decay constant	**0.0017**	**0.0013**	**1.3**	**0.21**	**-0.0010**	**0.0045**
	a_2: contribution of *T1LV* on exponential coefficient	**0.13**	**0.058**	**2.26**	**0.034**	**0.011**	**0.25**
	b_2: contribution of *T1LV* on decay constant	**0.61**	**0.57**	**1.07**	**0.30**	**-0.57**	**1.80**
	LINRATE: % change/year	**-0.22**	**0.075**	**-2.95**	**0.0074**	**-0.38**	**-0.065**
	Scanner Change Correction: Excite to Excite	-0.017	0.67	-0.03	0.98	-1.40	1.36
	Scanner Change Correction: Excite to HDx	-0.88	0.67	-1.32	0.20	-2.27	0.50
	Scanner Change Correction: Excite to HDxt	-1.38	0.76	-1.82	0.082	-2.94	0.19
	Scanner Change Correction: HDx to HDx	-0.18	0.63	-0.29	0.78	-1.49	1.13
	Scanner Change Correction: HDx to HDxt	-0.26	0.63	-0.41	0.69	-1.56	1.05
	Scanner Change Correction: HDxt to HDxt	0.38	0.63	0.6	0.55	-0.93	1.70
	Scanner Change Correction: Intera to Intera	-0.080	0.57	-0.14	0.89	-1.26	1.10
	Scanner Change Correction: Avanto to Avanto	-0.87	0.64	-1.37	0.18	-2.19	0.45
	AsymptoteVol	**99.02**	**0.60**	**164.17**	**<.0001**	**97.77**	**100.27**
	Model Variance	0.26	0.037	7.18	<.0001	0.19	0.34
	$RANDOM_{SLOPE}$	0.083	0.038	2.16	0.042	0.0033	0.16
	Random $Effect_{Covariance}$	-0.068	0.054	-1.26	0.22	-0.18	0.044
	$RANDOM_{INTERCEPT}$	0.17	0.10	1.63	0.12	-0.046	0.39

Grey matter (Figure 4-2, Table 4-3): The exponential models predicted about 1.2% (SE: 0.77) accelerated GMVL on average. Of the two predictors of the amount of volume loss, only BEAMDose was significant. Neither were significant predictors of the decay rates. Inputting the mean value, the exponential model for BEAMDose predicted about 0.98% loss with a half-life of 0.01y. Baseline T1LV was again not a significant predictor: the associated model predicted about 0.09% loss with a half-life of 1.4y. The mean rate of the linear-change volume loss was -0.12%/y (SE: 0.081).

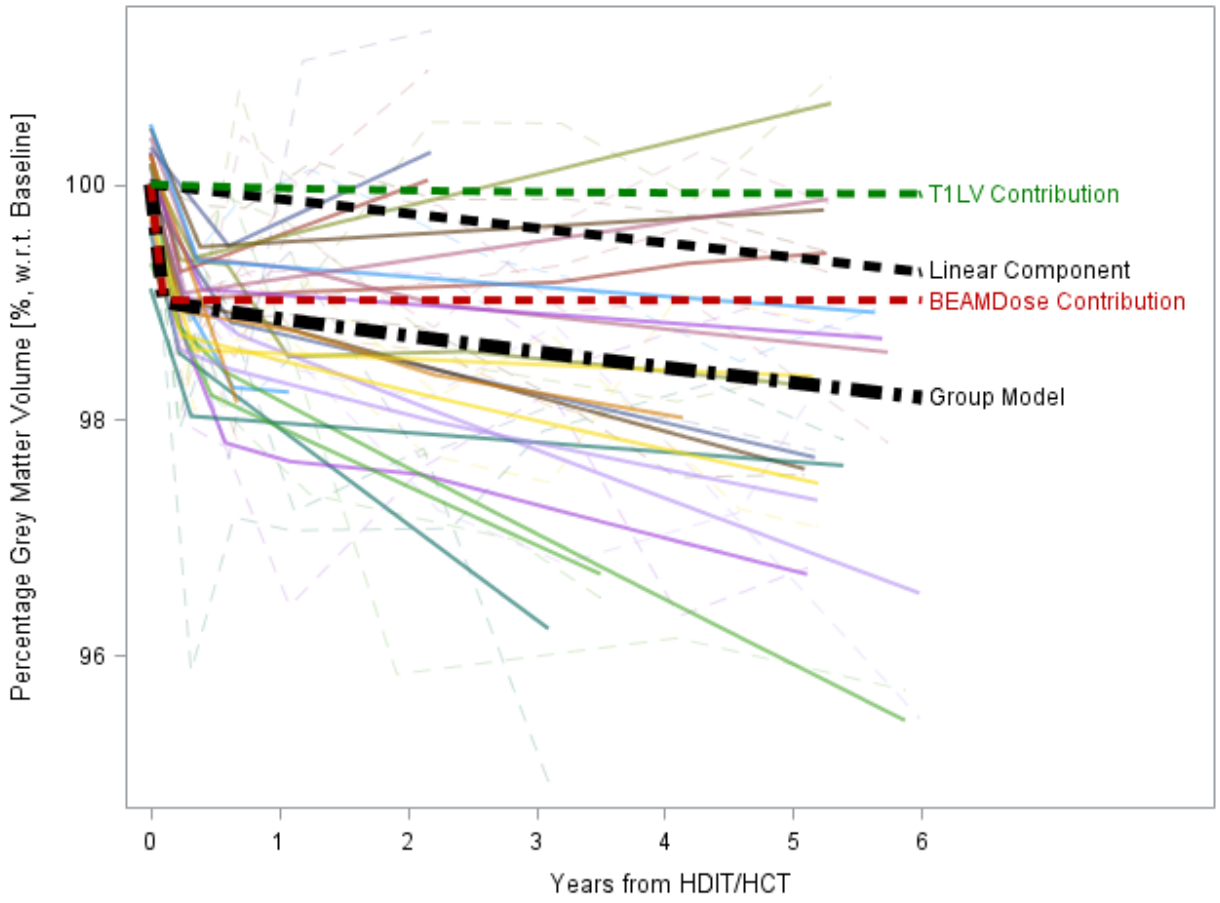

Figure 4-2: Evolution of grey matter volume loss following HDIT/HCT – HALT-MS

Total BEAM dose was a significant predictor, contributing about 0.98% volume loss with the approximate half-life of 0.01y. T1LV was not significant. The model of linear-change volume loss estimated about -0.12% loss per year.

Legend:

BEAMDose Contribution (dashed thick red line) = Exponential decay model associated with total BEAM dose

T1LV Contribution (dashed thick green line) = Exponential decay model associated with baseline T1-weighted WM lesion volume

Linear Component (dashed thick black line) = Model of the linear-change grey matter volume loss, potentially due to normal-aging

Group Model (solid thick black line) = Net model for the cohort, i.e. the sum of T1LV and BEAMDose curves and the Linear component

Colored dashed transparent lines = Grey matter volume change measurement data for the individual subjects

Colored solid thin lines = Fitted model for the individual subjects

Table 4-3: Parameter estimates for the model of grey matter volume loss after HDIT/HCT

Compartment	Parameter, df=22	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Grey Matter	**a_1: contribution of *BEAMDose* on exponential coefficient**	**0.00027**	**0.000048**	**5.58**	**<.0001**	**0.00017**	**0.00037**
	b_1: contribution of *BEAMDose* on decay constant	**0.015**	**0.23**	**0.06**	**0.95**	**-0.47**	**0.50**
	a_2: contribution of *T1LV* on exponential coefficient	**0.071**	**0.066**	**1.08**	**0.29**	**-0.066**	**0.21**
	b_2: contribution of *T1LV* on decay constant	**0.40**	**0.54**	**0.74**	**0.47**	**-0.72**	**1.51**
	LINRATE: % change/year	**-0.12**	**0.081**	**-1.53**	**0.14**	**-0.29**	**0.044**
	Scanner Change Correction: Excite to Excite	0.19	0.84	0.23	0.82	-1.54	1.93
	Scanner Change Correction: Excite to HDx	-0.54	0.86	-0.62	0.54	-2.33	1.25
	Scanner Change Correction: Excite to HDxt	-0.34	1.01	-0.34	0.74	-2.43	1.75
	Scanner Change Correction: HDx to HDx	0.16	0.81	0.20	0.84	-1.51	1.83
	Scanner Change Correction: HDx to HDxt	-0.0048	0.82	-0.010	1.00	-1.71	1.70
	Scanner Change Correction: HDxt to HDxt	0.24	0.81	0.30	0.77	-1.44	1.93
	Scanner Change Correction: Intera to Intera	0.078	0.73	0.11	0.92	-1.43	1.59
	Scanner Change Correction: Avanto to Avanto	-0.40	0.82	-0.49	0.63	-2.11	1.31
	AsymptoteVol	**98.83**	**0.77**	**128.85**	**<.0001**	**97.24**	**100.43**
	Model Variance	0.43	0.060	7.18	<.0001	0.31	0.56
	$RANDOM_{SLOPE}$	0.10	0.054	1.91	0.070	-0.0091	0.22
	Random $Effect_{Covariance}$	-0.0015	0.055	-0.03	0.98	-0.12	0.11
	$RANDOM_{INTERCEPT}$	0.15	0.089	1.73	0.098	-0.031	0.34

White matter (Figure 4-3, Table 4-4): On average, the exponential models predicted about 1.6% (SE: 0.84) accelerated WMVL. Both BEAMDose and baseline T1LV were significant predictors of the amount of WM loss. Further, BEAMDose was a significant predictor of the decay rates. Inputting the mean BEAMDose value, the exponential model for BEAMDose predicted about 1.8% loss with a half-life of 0.6y. Inputting the mean T1LV value, the model for T1LV predicted about 0.32% loss with a half-life of 0.9y. The mean rate of the linear-change volume loss was -0.15%/y (SE: 0.11).

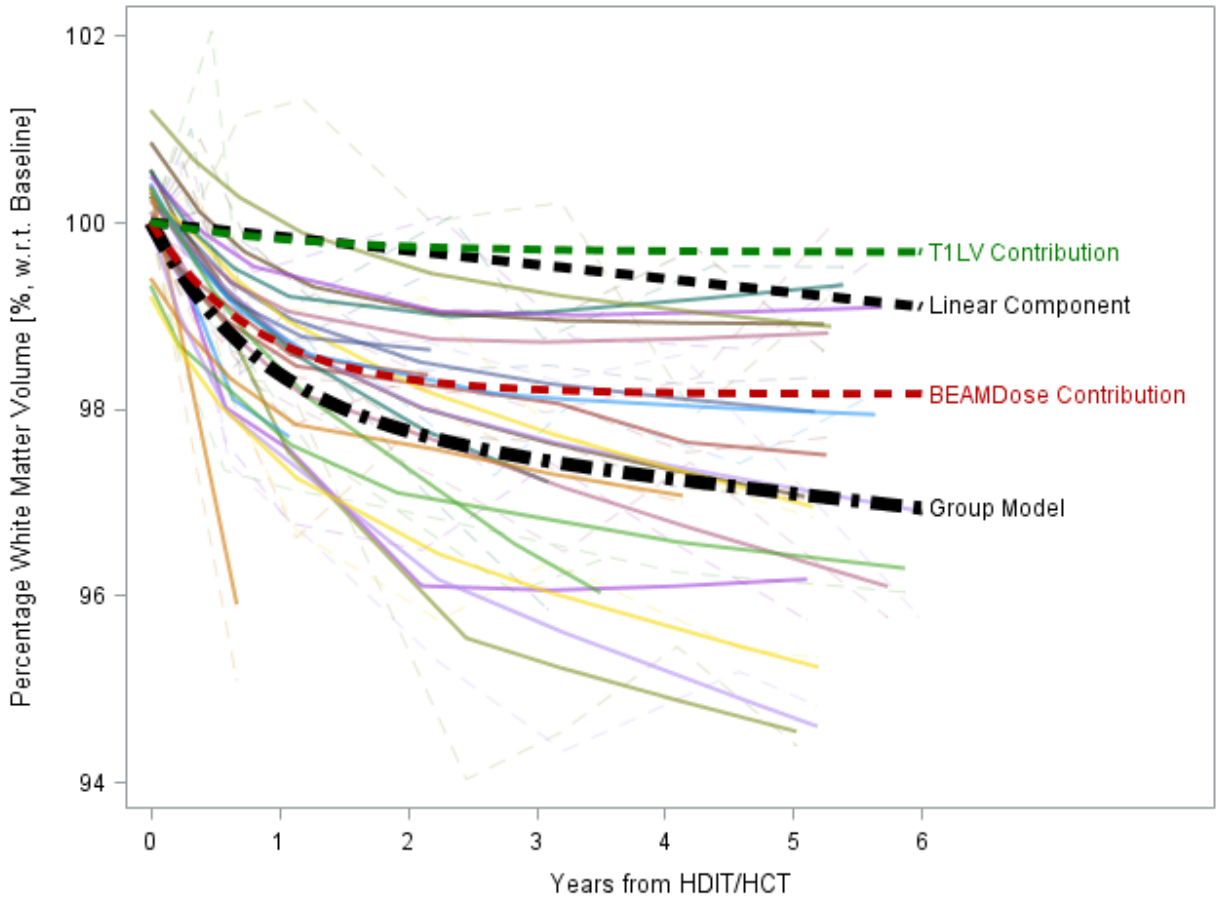

Figure 4-3: Evolution of white matter volume loss following HDIT/HCT – HALT-MS

Total BEAM dose was a significant predictor, predicting about 1.8% volume loss with the approximate half-life of 0.6y. The effect of T1LV was also significant, contributing about 0.32% WM volume loss with the approximate half-life of 0.9y. The model of linear-change volume loss estimated about -0.15% loss per year.

Legend:

BEAMDose Contribution (dashed thick red line) = Exponential decay model associated with total BEAM dose

T1LV Contribution (dashed thick green line) = Exponential decay model associated with baseline T1-weighted WM lesion volume

Linear Component (dashed thick black line) = Model of the linear-change white matter volume loss, potentially due to normal-aging

Group Model (solid thick black line) = Net model for the cohort, i.e. the sum of T1LV and BEAMDose curves and the Linear component

Colored dashed transparent lines = White matter volume change measurement data for the individual subjects

Colored solid thin lines = Fitted model for the individual subjects

Table 4-4: Parameter estimates for the model of white matter volume loss after HDIT/HCT

Compartment	Parameter, df=22	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
White Matter	**a_1: contribution of *BEAMDose* on exponential coefficient**	**0.00050**	**0.00011**	**4.52**	**0.00020**	**0.00027**	**0.00073**
	b_1: contribution of *BEAMDose* on decay constant	**0.00033**	**0.000097**	**3.41**	**0.0025**	**0.00013**	**0.00053**
	a_2: contribution of *T1LV* on exponential coefficient	**0.27**	**0.080**	**3.36**	**0.0028**	**0.10**	**0.43**
	b_2: contribution of *T1LV* on decay constant	**0.68**	**0.52**	**1.30**	**0.21**	**-0.40**	**1.76**
	LINRATE: % change/year	**-0.15**	**0.11**	**-1.44**	**0.16**	**-0.37**	**0.067**
	Scanner Change Correction: Excite to Excite	-0.39	0.93	-0.43	0.67	-2.31	1.52
	Scanner Change Correction: Excite to HDx	-1.23	0.90	-1.37	0.19	-3.11	0.64
	Scanner Change Correction: Excite to HDxt	-2.34	0.87	-2.69	0.013	-4.15	-0.54
	Scanner Change Correction: HDx to HDx	-0.93	0.86	-1.09	0.29	-2.71	0.85
	Scanner Change Correction: HDx to HDxt	-0.64	0.81	-0.79	0.44	-2.32	1.036
	Scanner Change Correction: HDxt to HDxt	0.68	0.81	0.84	0.41	-0.99	2.35
	Scanner Change Correction: Intera to Intera	-0.27	0.76	-0.36	0.72	-1.84	1.30
	Scanner Change Correction: Avanto to Avanto	-2.00	0.85	-2.35	0.028	-3.76	-0.24
	AsymptoteVol	**98.42**	**0.84**	**117.45**	**<.0001**	**96.68**	**100.16**
	Model Variance	0.48	0.064	7.51	<.0001	0.35	0.62
	$RANDOM_{SLOPE}$	0.074	0.035	2.09	0.048	0.00061	0.15
	Random $Effect_{Covariance}$	-0.20	0.096	-2.04	0.053	-0.40	0.0029
	$RANDOM_{INTERCEPT}$	0.71	0.33	2.13	0.045	0.017	1.40

BVL rate comparison between baseline Gd+ versus Gd- (mean, SD):

Short-term at 1y follow-up: WBVL and GMVL were significantly higher in the Gd+ patients (**WB**: Gd+: -1.90%, 0.76; Gd-: -0.76%, 0.69; p=0.002; **GM**: Gd+: -1.83%, 1.06; Gd-: -0.40%, 0.77; p=0.003). WMVL was not significantly different (Gd+: -1.97%, 0.94; Gd-: -1.34%, 1.20; p=0.20). All N=10 versus 11, two-tailed t-test.

Long-term: There was a trend of higher WBVL rates in the Gd+ patients (Gd+: -0.32%/y, 0.27; Gd-: -0.13%/y, 0.23; p=0.08). GMVL rates were significantly higher in the Gd+ patients (Gd+: -0.26%/y, 0.26; Gd-: -0.0051%/y, 0.27; p=0.03). WMVL rates were not significantly different (Gd+: -0.19%/y, 0.28; Gd-: -0.12%/y, 0.21; p=0.5). All N=11 versus 13, two-tailed t-test.

BVL rate comparison between EFS+ versus EFS- (mean, SD):

Short-term at 1y follow-up: There were no significant differences in the WB (EFS+: -1.11%, 0.98; EFS-: -1.68%, 0.66), GM (EFS+: -0.91%, 1.20; EFS-: -1.43%, 1.07), and WM volume losses (EFS+: -1.45%, 1.19; EFS-: -2.01%, 0.88). All N=14 versus 7, two-tailed t-test.

Long-term: There were no significant differences in the WB (EFS+: -0.22%/y, 0.27; EFS-: -0.22%/y, 0.25), GM (EFS+: -0.14%/y, 0.31; EFS-: -0.08%/y, 0.26), and WM volume loss rates (EFS+: -0.12%/y, 0.22; EFS-: -0.22%/y, 0.29). All N=17 versus 7, two-tailed t-test.

4.2.5 Discussion

We measured and modeled the time-courses of WB, GM, and WM volumes in a cohort of RRMS patients who received a chemotherapy-based immunosuppressive regimen and HCT. Accelerated volume loss was noted in all compartments at 1y follow-up. Of the two tested predictors, BEAMdose was the dominant predictor in all compartments; baseline T1LV was also a significant predictor in the WB and the WM. Factoring out the time-courses of the accelerated volume loss during the initial months, the average long-term rates of WBVL, GMVL, and WMVL after HDIT/HCT were -0.22%/y, -0.12%/y, and -0.15%/y, respectively. The patients with gadolinium-enhancing lesions at baseline had significantly higher short-term WB and GM volume losses as well as long-term GMVL rates. There were no significant differences in the rates of volume loss between the patients who maintained EFS versus those who did not.

Early acceleration of atrophy

Several observational cohort studies of MS patients following HDIT/HCT have reported accelerated loss of WB volume by the second year of follow-up, well beyond those expected in MS.[159,174–177,206,220] These were paralleled by significant reduction of relapses and MRI activity. Strikingly, studies utilizing frequent early follow-up scans revealed that most of the losses occurred during the first six months of follow-up or even earlier.[80,175,206] In fact, our previous study suggested that the cytotoxic effect of the busulfan/cyclophosphamide (Bu/Cy) regimen used for immunoablation was significantly associated with the accelerated WBVL in the patients enrolled in the Canadian MS-BMT trial.[206] The difference in the accelerated BVL in the Canadian BMT and HALT-MS trials suggests that the type and strength of the chemotherapeutic agent may be a relevant issue in the context of assessing brain atrophy after HDIT/HCT trials for MS.

Myeloablative regimens used in various HDIT/HCT trials for MS have included BEAM,[82,84,87,92,220] total-body irradiation/cyclophosphamide (TBI/Cy),[78,79,159,176] and Bu/Cy.[195] Each regimen has a different central nervous system (CNS) toxicity profile.[169] For example, carmustine and cytarabine (two of the agents constituting BEAM) are toxic to normal CNS cell types, including neuronal precursor cells, oligodendrocyte precursor cells, and non-dividing oligodendrocytes; their cell divisions are also compromised.[91] A busulfan-based regimen is generally considered more intense compared to a BEAM-based regimen[77] and can penetrate the CNS much better than the latter.[222] In keeping with this, MS patients who received BEAM-based HDIT/HCT had relatively less WBVL during the early follow-up compared to those treated with TBI/Cy or Bu/Cy. For example, WBVL during first year of follow-up in secondary-progressive MS (SPMS) patients treated with BEAM have included -1.87% (mean, SD: 2.19, N=10)[174] and -1.10% (mean, SD: 1.71, N=9)[177]. WBVL during first year of follow-up in SPMS patients treated with TBI/Cy have included -2.3% (median, SD: n/a, N=14)[176] and -2.1% (mean, SD: n/a, N=14)[159]. Notably, WBVL during first 2.4 months of follow-up in a group of SPMS patients treated with Bu/Cy was -3.2% (median, SD: n/a, N=5).[175]

Our results confirm that the effect of chemotherapy plays a significant role in the accelerated WB, GM, and WM volume loss during the initial follow-up months. The time-course was relatively rapid in GM, stabilizing within one year, and slower in WM, taking about two years to

stabilize. On average, the HALT-MS patients lost about 1.3% (SD: 0.91, mean raw percent volume change at 1y, N=21) of WB volume, during the first year of follow-up. This is comparable to WBVL reported in other MS cohorts treated with BEAM. Together, these results indicate that both the dose and type of immunosuppressive regimen affect the initial acceleration of volume loss after HDIT/HCT; a less aggressive regimen like BEAM may be associated with a shorter and milder course of treatment-related atrophy.

Loss of tissues committed to degenerate prior to treatment may progress for some time after HDIT/HCT and also may contribute to early volume loss. In keeping with this, the baseline volume of T1-weighted hypointense lesion was a significant predictor of the initially accelerated volume loss in the WM and the WB. The exponential model associated with T1LV may in part reflect Wallerian degeneration of axons in NAWM that are transected within focal WM lesions that formed before HDIT/HCT.[100] The temporal profile of WM volume loss was similar to that found in CNS Wallerian degeneration which, in humans, takes months to years.[201]

The question arises as to whether the early accelerated BVL could be secondary to resolution of inflammation, or so-called pseudoatrophy. However, when looked for, no evidence supporting pseudoatrophy has been found. For example, Chen and colleagues did not find evidence of a change in brain water content,[175] and Rocca and colleagues found evidence for acceleration of WB volume loss even after ignoring the first month of follow-up.[176] In addition, the significant difference in short-term volume loss was driven by the GM, whereas pseudoatrophy is suspected to occur mainly in the WM.[156]

Long-term rates of atrophy

We estimated the linear rates of WB, GM, and WM volume loss not due to short-term treatment-related effects and compared them to previously reported values obtained from cohorts in the similar age range.

The group rate of WBVL in HALT-MS was -0.22%/y (SD: 0.37, 95%CI: -0.38, -0.065), which is comparable to that found in nonelderly normal-aging, e.g. -0.27%/y (SD: 0.15, N=35 healthy controls, mean age: 37, measured with FSL-SIENA).[96] We used paired Jacobian integration, which has a smaller bias than FSL-SIENA, but the results are sufficiently robust that this does not affect the conclusion. The long-term rates of WBVL in this study were similar to those found

in the Canadian MS-BMT trial, e.g. -0.23%/y (SD: 0.74, N=19 RR/SPMS, FSL-SIENA). In comparison, the average rate of WBVL in MS is roughly two-fold higher than age-matched healthy controls.[96]

The group rate of GMVL in HALT-MS was -0.12%/y. This was comparable to that found in the Canadian MS-BMT trial, e.g. -0.18%/y.[223] Longitudinal analyses of GMVL in normal subjects have not been reported. Some estimates have been made from cross-sectional measurements over subjects of different ages, but these results are not directly comparable to ours.[130,134]

The group rate of WMVL in HALT-MS was -0.15%/y. This was about two-fold higher than that found in the Canadian MS-BMT trial, e.g. -0.07%/y (SD: 0.61).[223] This may be reflective of the less complete suppression of focal inflammation in the HALT-MS patients.

The long-term rates of GMVL were significantly higher in the patients who were Gd+ at baseline. One reason for this might be that the compromised blood-brain-barrier in the Gd+ patients allowed for a greater penetration of the BEAM agents, which normally have limited CNS penetration.[222] Another reason might be greater rates of MS-related atrophy. The presence of gadolinium-enhancing lesions at baseline is a significant predictor of GMVL in RRMS patients treated with DMTs.[99] These findings suggest that ongoing inflammation may influence the rate of GM atrophy.

We found no significant differences in the linear rates of WB, GM, and WM volume loss between the EFS+ and the EFS- subgroups of patients, possibly due to the small number of EFS- patients (N=7).

Study limitations

First, this was a single-arm study so most of the comparisons were made against historical controls. Comparison of the rates of BVL across different measurement techniques (e.g. FSL-SIENA and paired Jacobian integration) must be done carefully because of the different biases associated with the different techniques. An independent analysis found that the WBVL measurement output from PJI is highly correlated with that from FSL-SIENA, but the PJI rates are ~20% less than those of FSL-SIENA (K. Nakamura, personal communication). Second, the MRI scans for this study were not obtained at ultra-high field (7T), and so were insensitive to focal GM lesions, which could have been a predictor of GMVL. Third, we used a simple

addition of the doses of BiCNU, etoposide, ara-c, and melphalan to calculate the total dose of BEAM. This may not be an ideal way to calculate the dose effect.

Summary

Our results suggest several important points. First, chemotherapy-related cytotoxicity plays a significant role in the acceleration of brain volume loss commonly seen during the initial months of follow-up in MS patients treated with HDIT/HCT. This emphasizes the importance of selecting an immune-ablative regimen that can maximize the control of inflammation while minimizing neurotoxicity. Second, tissue loss, especially in the WM, may still progress for many months after HDIT/HCT due to existing, pre-treatment inflammatory injury that commits tissue to neurodegeneration. For these reasons, the long-term effect of HDIT/HCT on brain atrophy in MS should be assessed over intervals of 3 years and longer. Over the long-term, adequate immunosuppression/immunoablation and autologous hematopoietic stem cell transplantation can reduce BVL to rates consistent with normal aging.

4.2.6 Acknowledgments

Funding acknowledgements

The HALT-MS study was ponsored by the Division of Allergy, Immunology, and Transplantation, National Institute of Allergy and Infectious Diseases (DAIT-NIAID), NIH. Hyunwoo Lee was supported by the Doctoral Training Award from Fonds de recherché du Quebec.

4.2.7 Appendix

Table 4-5: MRI protocol for the images used in this analysis

Site #	125-BCM-2		203-FHC-2		210-OSU-1	
# of patients	4		16		4	
MRI scanner make/model	Philips Intera / Philips Achieva		GE Signa EXCITE / GE Signa HDx / GE Signa HDxt		Siemens Avanto	
Field strength	1.5T					
Contrast	T1-weighted Pre-Gadolinium	Axial T2-weighted / Proton density-weighted	T1-weighted Pre-Gadolinium	Axial T2-weighted / Proton density-weighted	T1-weighted Pre-Gadolinium	Axial T2-weighted / Proton density-weighted
Sequence	3D T1-FFE	2D TSE	3D SPGR	2D FSE-XL	3D FLASH	2D TSE
Repetition time (ms)	30	4000 (T2w) / 2400 (PDw)	30	5117 (T2w) / 2000 (PDw)	30	5750 (T2w) / 2980 (PDw)
Echo time (ms)	8	80 (T2w) / 15 (PDw)	9	83.5 (T2w) / 12 (PDw)	10	80 (T2w) / 15 (PDw)
Field of view (mm)	250	250	250	250	250	250
Number of Slices	60	60	60	60	60	60
Slice Thickness (mm)	3	3	3	3	3	3
Matrix	256x256	256x256	256x256	256x256	256x256	256x256
Flip angle	30	90	30	90	30	90
Echo train length (ETL)	n/a	8 (T2w) / 3(PDw)	n/a	8 (T2w) / 3(PDw)	n/a	7 (T2w) / 3(PDw)

MRI scanner hardware upgrade information

MRI scanning were done at three sites, using 1.5T scanners from Philips, GE, and Siemens. At one site, four patients were scanned on a Siemens Avanto without major hardware upgrade. At another site, four patients were scanned with a Philips Intera, except that one patient had a single timepoint scanned with a Philips Achieva and then returned to Intera. Finally, at another site, eight patients were scanned on a GE Signa HDxt; however the remaining eight patients were scanned on multiple GE scanners over the study (Excite to HDx to HDxt, N=2; Excite to HDx, N=1; HDx to HDxt, N=5). "Intera to Achieva" was the reference category used for the additive correction.

Types and timing of the subsequent MS-related events in the HALT-MS study

EDSS increase >0.5: N=2; 18.9m and 15.2m after HDIT/HCT

Development of new clinical relapse: N=3; 22.2m, 5.1m, and 32.6m after HDIT/HCT

Development of new MRI lesions: N=2; 45.6m and 48.4m after HDIT/HCT

Chapter 5 Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain atrophy measurements

5.1 Preface

The previous chapters examined longitudinal time courses of brain atrophy in two MS cohorts treated with aHSCT. Both studies were affected by mid-study MRI scanner upgrades, and the statistical models used included a covariate term to account for the potential technical bias that may have confounded the volume change measurements. In fact, mid-study MRI scanner upgrades or even scanner changes occur fairly often in multi-year follow-up studies. Also, unforeseen factors like patient re-location or scanner breakdown may force the remaining follow-up scans to be done on a different scanner. Surprisingly, it is still unclear how much impact these upgrades or changes can have on longitudinal measurements of brain volume change. They are likely not negligible, and a correction at the data analysis level may be necessary to more reliably estimate the courses of brain atrophy.

To provide further understanding on this issue, this chapter examines the effect of scanner changes in a large cohort study for which imaging data is publicly available – the Alzheimer's disease neuroimaging initiative (ADNI) 1.5T study. This longitudinal, multicenter dataset was deemed particularly suitable for this purpose because a standardized acquisition protocol was used on all the scanners, and different combinations of intra-vendor scanner upgrades and inter-vendor scanner changes occurred affecting sufficient numbers of subjects to allow statistical modeling. Also, a subset of these subjects was affected by a change in the standardized T1-weighted sequence at selected sites using GE scanners. In the following manuscript entitled "Estimating and accounting for the effect of MRI scanner hardware changes on longitudinal whole-brain atrophy measurements", a linear mixed-effects model is applied to estimate the effects of scanner and T1-sequence changes on whole-brain volume change measurements. I also assessed whether the inclusion of corrective terms can lead to a better model goodness-of-fit.

5.2 Manuscript: Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain atrophy measurements

Hyunwoo Lee, B.Sc. [1], Kunio Nakamura, Ph.D. [2], Sridar Narayanan, Ph.D. [1], Robert A. Brown, Ph.D. [1], Douglas L. Arnold, M.D. [1]

[1] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[2] Department of Biomedical Engineering, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA

Manuscript in preparation

5.2.1 Abstract

Objective: Longitudinal MRI studies are often subjected to mid-study scanner changes, which may alter image characteristics such as contrast, signal-to-noise ratio, contrast-to-noise ratio, and intensity non-uniformity. Measuring brain atrophy under these conditions can render the results potentially unreliable across the timepoint of the change. Estimating and accounting for this effect can provide more reliable estimates of brain atrophy rates.

Methods: We analyzed 237 subjects who were scanned at 1.5T for the Alzheimer's Disease Neuroimaging Initiative (ADNI) study, and were subject to intra-vendor or inter-vendor scanner changes during follow-up (up to 8 years). 63 subjects scanned on GE Signa HDx and HDxt platforms were also subject to a T1-weighted sequence change from MP-RAGE to IR-FSPGR, as part of the transition from ADNI-1 to ADNI-2/GO. Two-timepoint percentage brain volume changes (PBVCs) between the baseline "screening" and the follow-up scans were calculated using SIENA. A linear mixed-effects model with subject-specific random slopes and intercepts was applied to estimate the fixed effects of scanner hardware changes on the PBVC measures.

The same model also included a term to estimate the fixed effects of the T1-weighted sequence change.

Results: Different hardware upgrade or change combinations led to different offsets in the PBVC (SE; p):

GE Genesis Signa to Philips Intera, 0.98% (0.47, p=0.047);

GE Signa Excite to Signa HDx, 0.33% (0.095, p=0.0005);

GE Signa Excite to Signa HDxt, -0.049% (0.23, p=0.83);

GE Signa Excite to Signa HDx to Signa HDxt, 0.24% (0.095, p=0.011) and 0.26% (0.16, p=0.11), respectively;

GE Signa HDx to Signa HDxt, -0.25% (0.25, p=0.31);

Siemens Symphony to Symphony TIM -0.39% (0.17; p=0.021);

Philips Intera to Siemens Avanto -1.79% (0.29; p<0.001).

The sequence change from MP-RAGE to IR-SPGR was associated with an average -1.6% (0.1; p<0.0001) change.

Conclusion: Inter-vendor scanner changes generally led to greater effect sizes compared to intra-vendor scanner upgrades. The effect of T1-weighted sequence change was comparable to that from inter-vendor scanner changes. Inclusion of the corrective fixed-effects terms for the scanner hardware and T1-weighted sequence changes yielded better model goodness-of-fits, and thus, provided more reliable estimates of whole-brain atrophy rates.

5.2.2 Introduction

Measuring brain atrophy using magnetic resonance imaging (MRI) is a topic of significant interest in the study of neurological disorders such as Alzheimer's disease (AD). These diseases result in a range of pathological processes that lead to progressive neuronal, axonal, and dendritic degeneration, and ultimately, central nervous system (CNS) tissue atrophy.[224] For example, a longitudinal MRI study has reported that the annualized rates of atrophy in the whole-brain (WB) and hippocampus were respectively more than two-fold and four-fold higher on average, in AD

patients compared to age-matched healthy controls.[225] Indeed, higher rates of brain atrophy in patients with AD or mild cognitive impairment (MCI) are associated with higher rates of decline in cognitive measures.[226] For these reasons, MRI measures of brain atrophy are widely recognized as markers of progression of neurodegeneration.[227]

Various technical factors can influence MRI-based atrophy measurements, especially when calculating longitudinal changes using serial images. For example, head motion,[178] inconsistent image contrast,[178] different levels of noise,[178] gradient non-linearity,[216] intensity non-uniformity,[216] inconsistent subject positioning,[228] number of head coil channels,[217] and choice of image analysis methods[207,229] all can affect atrophy outcomes. Several single-site studies have investigated the effects of varying acquisition protocols on MRI outcomes. For example, Preboske and colleagues showed in a scan-rescan study that 1) implementing different flip angles and 2) switching from conventional to fast spoiled gradient echo (SPGR) sequence resulted in significant brain volume differences.[178] Han and colleagues showed that the average cortical thickness variability did not change significantly across an intra-vendor scanner upgrade from Siemens Sonata 1.5T to Avanto 1.5T.[179] However, different pulse sequences and image processing pipelines led to poorer scan-rescan reproducibility.[179] Jovicich and colleagues showed in a scan-rescan study that subcortical, ventricular, and intracranial volume reproducibility did not significantly differ across both intra-vendor scanner upgrade (i.e. Siemens Sonata to Avanto) and inter-vendor scanner change (i.e. Siemens Sonata to GE Signa) conditions; however, there were bias in certain regional volumes after the inter-vendor scanner change.[180]

The issue may become more significant when multiple MRI scanners are used, such as in a multi-center trial or when a significant mid-study change in the scanner occurs. For example, Kruggel and colleagues demonstrated that different 1.5T and 3.0T scanner platforms provide different levels of image quality, as measured by signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), and mutual information of the joint histogram, and that these affected brain volume measurements.[182] Therefore, analysis of data from multiple scanning platforms should be done carefully since that the effect of the platform change may constitute a proportion of the brain volume change, which could either over- or under-estimate the true rate of brain atrophy.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) study is a longitudinal, multi-center study that acquired MRI data using a variety of 1.5T and 3.0T scanners from General Electric

(GE), Siemens, and Philips.[230] In particular, the study focused on designing and implementing standardized acquisition methods, as well as performing a centralized image post-processing and quality control.[230] These qualities make the ADNI a suitable dataset to test the hypothesis that a change in the scanning platform can bias WB volume change measurements, and the magnitude of this effect between unique pairs of MRI scanners. To do this, we identified all subjects from the ADNI-1 1.5T study who had any MRI scanner change or upgrade during the follow-up. Also, we identified a subset of subjects who had a T1-weighted sequence change during the follow-up. Then, we used a linear mixed-effects (LME) model to estimate the rates of WB atrophy, as well as the effects of different MRI scanner change or upgrade combinations and T1-weighted sequence change on percentage brain volume change (PBVC) measurements.

5.2.3 Methods

Data acquisition

The Alzheimer's Disease Neuroimaging Initiative (ADNI) was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). MRI data used in the preparation of this article were downloaded from the ADNI database (adni.loni.usc.edu) on 2015-01-21. Detailed MRI protocols are reported on the ADNI protocol website: (http://adni.loni.usc.edu/methods/documents/mri-protocols).

Subject selection

We started with N=819 subjects (Screening diagnosis, Normal=229, MCI=401, AD=189; Baseline diagnosis, Normal=229, MCI=398, AD=192) officially enrolled in ADNI-1, who had baseline and follow-up visits conducted on 1.5T scanners using the ADNI-specified T1-weighted magnetization prepared rapid gradient echo (MP-RAGE) sequence. Subject demographics are shown in Table 5-1. A subset of these subjects continued with 1.5T MRI during the ADNI-Grand Opportunity (GO) and ADNI-2 phases. N=818 of these subjects coincided with the 1.5T ADNI-1 standard "screening-visits" dataset reported by Wyman and colleagues.[231] We excluded N=46

subjects with only a single timepoint. For the remaining N=773 subjects, the MRI scanner vendor (GE, Siemens, Philips) and the scanner model were identified for each timepoint. Scanner change was noted if the scanner models used during any of the follow-up timepoints did not match those of the baseline scan. Accordingly, N=271 (Normal=80, MCI=141, AD=50) subjects had scanner upgrades or changes (referred to as "Chg+" subgroup) versus N=502 (Normal=138, MCI=241, AD=123) who did not (referred to as "Chg-" subgroup) (Tables 5-2 and 5-3). Each timepoint had two back-to-back repeat 3-dimensional (3D) T1-weighted scans, and we analyzed the first scan whenever possible.

Table 5-1: Basic subject demographics – ADNI 1.5 T

Characteristics	**All subjects (N=773)**	**Normal control subjects (N=218)**	**Mild cognitive impairment subjects (N=382)**	**Alzheimer's disease subjects (N=173)**
Mean age at baseline (SD) [range], yr	75.3 (6.8) [55.2 – 91.0]	76.0 (5.1) [60.0 – 89.7]	74.9 (7.3) [55.2 – 89.4]	75.2 (7.6) [55.2 – 91.0]
Sex, Female:Male	322:451	104:114	137:245	81:92
	Subjects with no scanner change or upgrade (N=502)	**Normal control subjects (N=138)**	**Mild cognitive impairment subjects (N=241)**	**Alzheimer's disease subjects (N=123)**
Mean age at baseline (SD) [range], yr	75.3 (7.0) [55.2 – 91.0]	76.0 (5.6) [60.0 – 89.7]	74.9 (7.3) [56.2 – 89.4]	75.2 (7.7) [55.2 – 91.0]
Sex, Female:Male	211:291	64:74	89:152	58:65
	Subjects with scanner change or upgrade (N=271)	**Normal control subjects (N=80)**	**Mild cognitive impairment subjects (N=141)**	**Alzheimer's disease subjects (N=50)**
Mean age at baseline (SD) [range], yr	75.3 (6.5) [55.2 – 87.8]	76.1 (4.0) [70.0 – 87.7]	75.0 (7.3) [55.2 – 87.8]	75.1 (7.5) [56.7 – 85.6]
Sex, Female:Male	111:160	40:40	48:93	23:27

Table 5-2: MRI scanner information for subjects without MRI scanner change or upgrade

Subjects with no MRI scanner change or upgrade				
Baseline 1.5T MRI scanner models	**Total number of subjects (Total N=502)**	**Number of normal control subjects (Total N=138)**	**Number of MCI subjects (Total N=241)**	**Number of Alzheimer's disease subjects (Total N=123)**
GE Genesis Signa	38	10	21	7
GE Signa Excite	117	33	47	37
GE Signa HDx	10		5	5
Siemens Avanto	56	18	25	13
Siemens Sonata	84	22	45	17
Siemens SonataVision	6	1	4	1
Siemens Symphony	108	31	55	22
Philips Achieva	19	5	10	4
Philips Intera	64	18	29	17

Table 5-3: MRI scanner information for subjects with MRI scanner change or upgrade

Subjects with MRI scanner change or upgrade. The combinations included in the analysis are in boldface					
1.5T MRI scanner model combination; "Original scanner" To "Changed scanner"	**Total number of subjects (Total N=271)**	**Number of normal control subjects (Total N=105)**	**Number of MCI subjects (Total N=176)**	**Number of Alzheimer's disease subjects (Total N=50)**	**Type of scanner change**
GE Genesis Signa To Siemens Avanto	26	7	12	7	Excluded due to nonconvergence
GE Genesis Signa To Philips Intera	**7**	**1**	**5**	**1**	**Inter-vendor change; Included**
GE Genesis Signa To GE Signa HDx	1		1		Excluded due to nonconvergence
GE Genesis Signa To Siemens Symphony	1		1		Excluded due to nonconvergence
GE Signa Excite To GE Signa HDx	**85**	**15**	**39**	**31**	**Intra-vendor upgrade; Included**
GE Signa Excite To GE Signa HDxt	**24**	**12**	**12**		**Intra-vendor upgrade; Included**
GE Signa Excite To GE Signa HDx To GE Signa HDxt	**60**	**25**	**35**		**Intra-vendor upgrade; two upgrades; Included**
GE Signa HDx To GE Signa HDxt	**12**	**2**	**9**	**1**	**Intra-vendor upgrade; Included**
Siemens Avanto To GE Signa HDxt	1	1			Excluded due to nonconvergence
Siemens Avanto To Siemens SonataVision	1	1			Excluded due to nonconvergence
Siemens Sonata To Siemens Espree	4	1	2	1	Excluded due to nonconvergence
Siemens Symphony To Siemens Symphony TIM	**34**	**9**	**18**	**7**	**Intra-vendor upgrade; Included**
Philips Intera To Siemens Avanto	**15**	**6**	**7**	**2**	**Inter-vendor change; Included**

MRI scanner information

Baseline scans were distributed between scanners as follows: GE (N=381, 49.3%), Siemens (N=294, 38.0%), and Philips (N=98, 12.7%). A scanner change affected subjects who started scanning on one of seven scanner models distributed as follows: GE (N=216, 79.7%), Siemens (N=40, 14.8%), and Philips (N=15, 5.5%). 13 combinations of inter- or intra-vendor scanner upgrade or change occurred. The majority of the cases involved GE scanners (Table 5-3). Figure 5-1 shows an example pair of images acquired from a single subject using two different scanners from two different vendors (i.e. an inter-vendor scanner change). Similarly, Figure 5-2 provides an example from two scanning platforms from a single vendor (i.e. an intra-vendor upgrade).

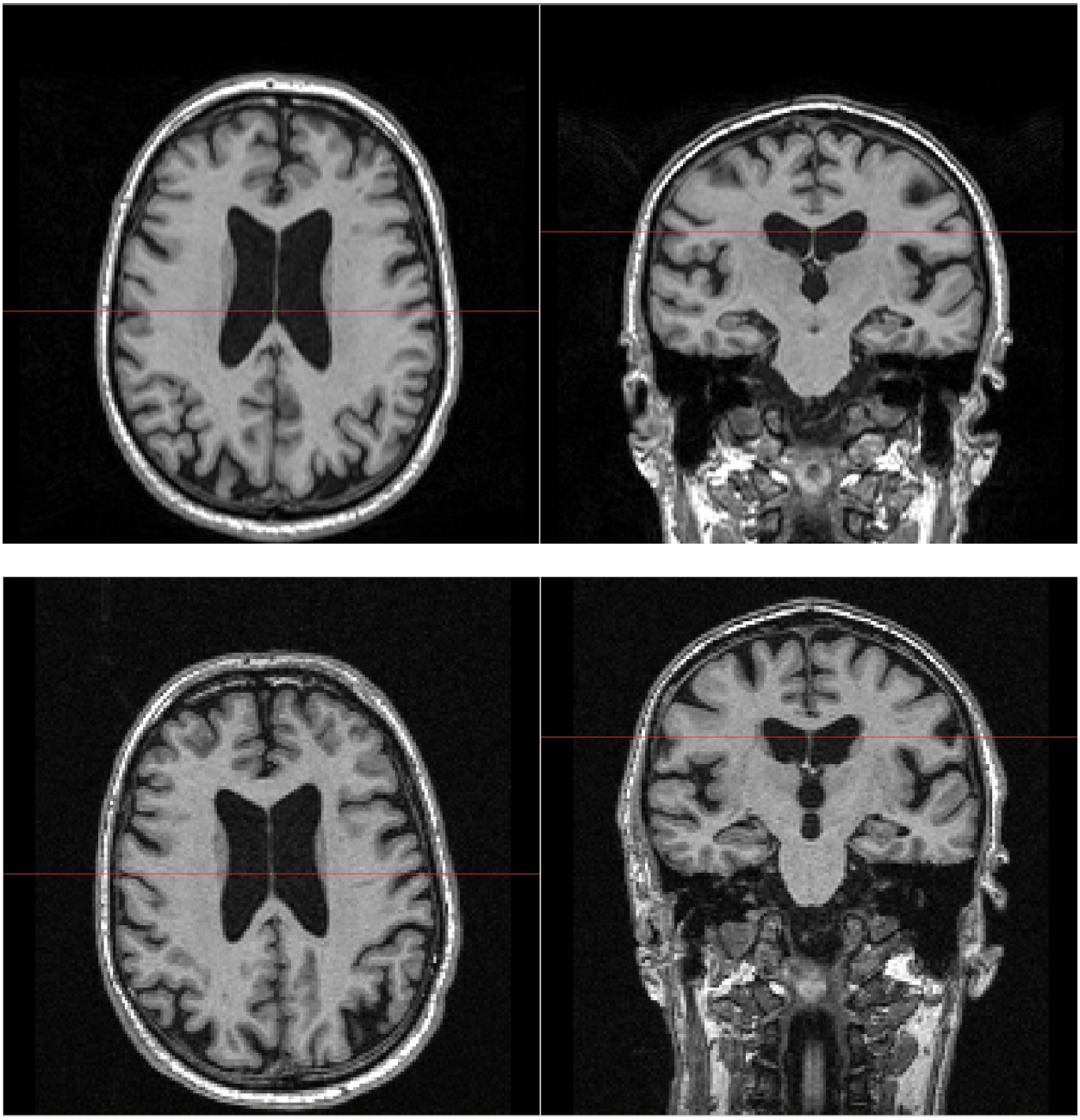

Figure 5-1: Example pair of images from a single subject (inter-vendor scanner change)

First row: 24 months follow-up, Philips Intera, MP-RAGE.

Second row: 36 months follow-up, Siemens Avanto, MP-RAGE.

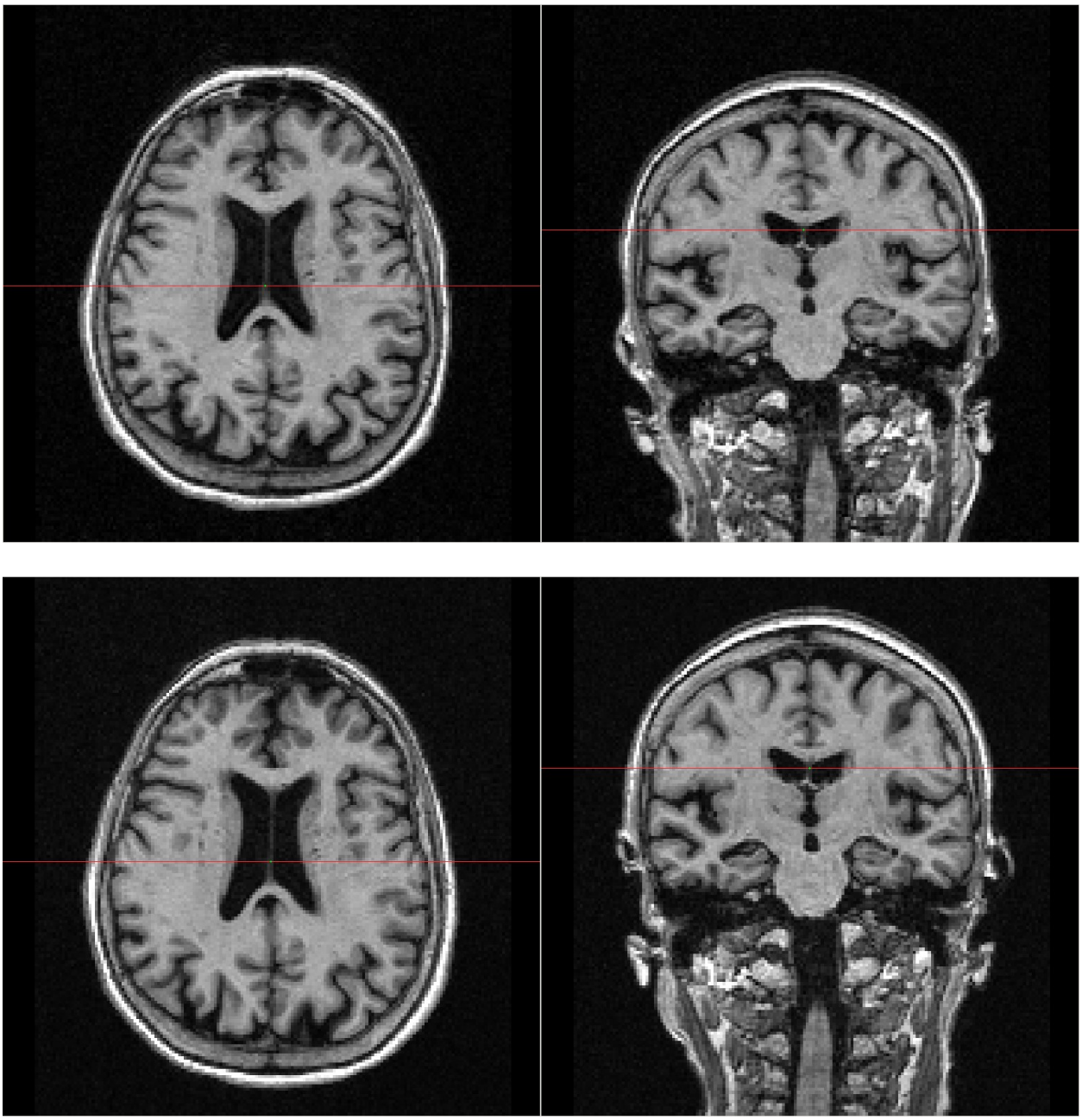

Figure 5-2: Example pair images from a single subject (intra-vendor scanner upgrade)

Top row: 12 months follow-up, Siemens Symphony, MP-RAGE.

Bottom row: 18 months follow-up, Siemens Symphony TIM, MP-RAGE.

3D T1-weighted sequence information

All Siemens and Philips scanners used the MP-RAGE sequence. It should be noted that GE scanners used a "works-in-progress" version of MP-RAGE during the ADNI-1 phase, and then switched to a Fast Spoiled Gradient Echo with IR Preparation (IR-FSPGR) sequence for the ADNI-GO and ADNI-2 phases.[232] This affected N=63 Chg+ subjects who had extended 1.5T follow-up (e.g. ADNI-GO month 48 and beyond, or ADNI-2) on the GE Signa HDx or HDxt scanners. Figure 5-3 shows an example pair of images acquired from a single subject using the two different sequences.

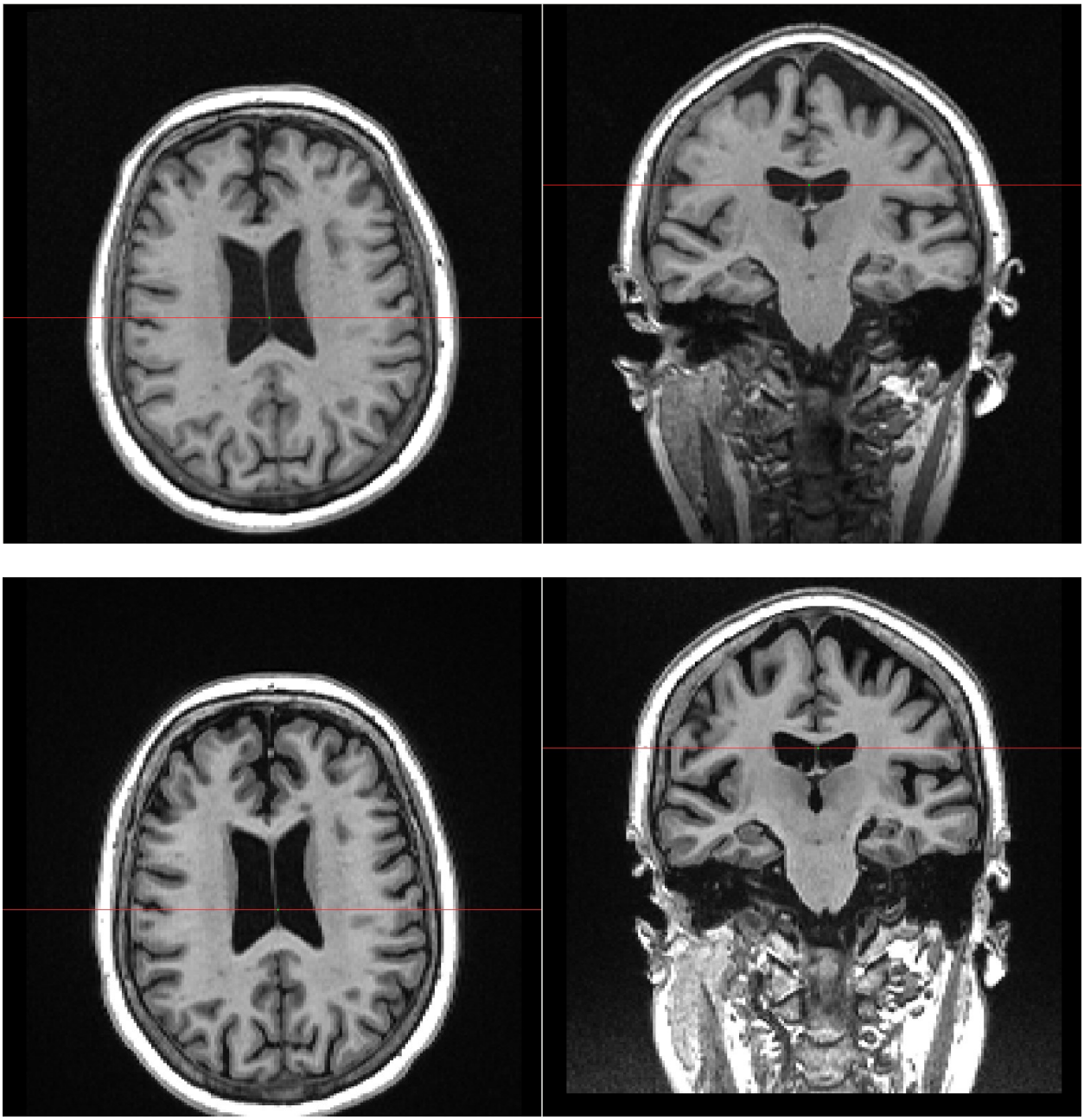

Figure 5-3: Example pair images from a single subject (T1-weighted sequence change)

Top row: 48 months follow-up, GE Signa HDxt, MP-RAGE.

Bottom row: 60 months follow-up, GE Signa HDxt, IR-FSPGR.

Image processing

We started with unpreprocessed 3D T1-weighted images (i.e. "Original") from the ADNI database. The images were preprocessed using the following steps: 1) nonparametric intensity non-uniformity normalization using N3,[233] 2) standard-space registration using the ICBM 2009c nonlinear symmetric template,[234] 3) brain extraction using BEaST.[235] Two-timepoint PBVCs were measured with SIENA,[120] part of FSL.[236] The baseline "screening" scans were designated as the reference (i.e. 100%), and all subsequent PBVCs were estimated with respect to the baseline. This produced a WB volume time course for each subject.

Statistical analysis

LME models with fixed (population-average) and random (subject-specific) effects have been frequently applied to model longitudinal brain atrophy outcomes.[185,186,225,237] We considered several previous findings in building our LME model. Notably, in the case of analyzing data from multiple T1-weighted protocols, it was shown to be advantageous to include a categorical fixed-effect term for the different protocols.[185] Another study tested various forms of LME models in terms of the Akaike Information Criterion (AIC) goodness-of-fit measure, and demonstrated the advantage of including subject-specific random intercepts and slopes with protocol-specific residual variance.[186] Also, having protocol as a fixed effect led to a better model fit as opposed to having protocol-by-study time interactions.[186] Finally, a study of N=713 ADNI subjects found no evidence of acceleration in WB atrophy rates over three years of follow-up.[225] Therefore, although many of the subjects in our study had more than three years of follow-up, we assumed linear time courses of atrophy.

The estimated PBVCs were modeled with an LME model that included subject-specific random slopes and intercepts, as well as fixed effects for the time from baseline, interaction between time and diagnosis group, MRI scanner models, and T1-weighted sequences. Scanner-specific residual variance was used. The model was as follows:

$$\Delta \boldsymbol{PBV} = \boldsymbol{\beta_0} + \boldsymbol{b_{oi}} + \boldsymbol{b_{1i}} * (\boldsymbol{time}) + \boldsymbol{\beta_{GroupRate}}(\boldsymbol{DiagnosisGroup} * \boldsymbol{time}) + \boldsymbol{\beta_{ScannerChg}}(\boldsymbol{ScannerModel}) + \boldsymbol{\beta_{SequenceChg}}(\boldsymbol{T1Sequence}) + \boldsymbol{\varepsilon_{ij}}$$

where

"ΔPBV" was a continuous variable for the percentage WB volume change from the baseline "screening" reference point;

"time" was a continuous variable for years from the baseline scan date;

"DiagnosisGroup" was a categorical variable for the diagnosis group, i.e. normal control, MCI, or AD;

"ScannerModel" was a categorical variable for the MRI scanner model changes shown in Table 5-3;

"T1Sequence" was a categorical variable for the 3D T1-weighted sequence, which was MP-RAGE for all Siemens and Philips scanners and either MP-RAGE or IR-FSPGR for GE scanners;

"β_0", "$\beta_{GroupRate}$", "$\beta_{ScannerChg}$", "$\beta_{SequenceChg}$" were the fixed-effects coefficients for the group intercept, atrophy rate associated with each diagnosis group, additive effect of MRI scanner change or upgrade on PBVCs, and additive effect of T1 sequence change on PBVCs, respectively;

"b_{oi}" and "b_{1i}" were the subject-specific random intercept and slope, respectively;

"ε_{ij}" was the error term.

This model was fitted using the MIXED procedure, SAS v9.4.

To assess whether including the corrective terms for the MRI scanner or the T1-weighted sequence changes leads to a better goodness-of-fit, an equivalent model without the corrective terms was also fitted. AIC was used to make the comparison.

5.2.4 Results

The N=773 subjects were divided into two subgroups: subjects who did not have a scanner upgrade or change during the follow-up ("Chg-", N=502) versus those who did ("Chg+", N=271). Overall, the modeled rates of WB atrophy between these two groups were not significantly different, Chg+, -1.15%/y (Standard Error: 0.05) vs. Chg-, -1.16%/y (0.04), p=0.84, F-test. The

rate for the Chg+ group estimated without regard to scanner model or sequence was -1.24%/y (0.05).

Chg- Subjects

Model-estimated WB atrophy rates by diagnosis group for the Chg- subjects are shown in Table 5-4 (Appendix). There were no effects of scanner model change or T1 sequence change in this subgroup. The average rates were AD: -1.69%/y (SE: 0.073), MCI: -1.25%/y (0.048), and normal controls: -0.67%/y (0.060). Figure 5-4 plots the average rates of atrophy for each diagnosis group overlaid on top of the actual PBVC measurement values for each subject.

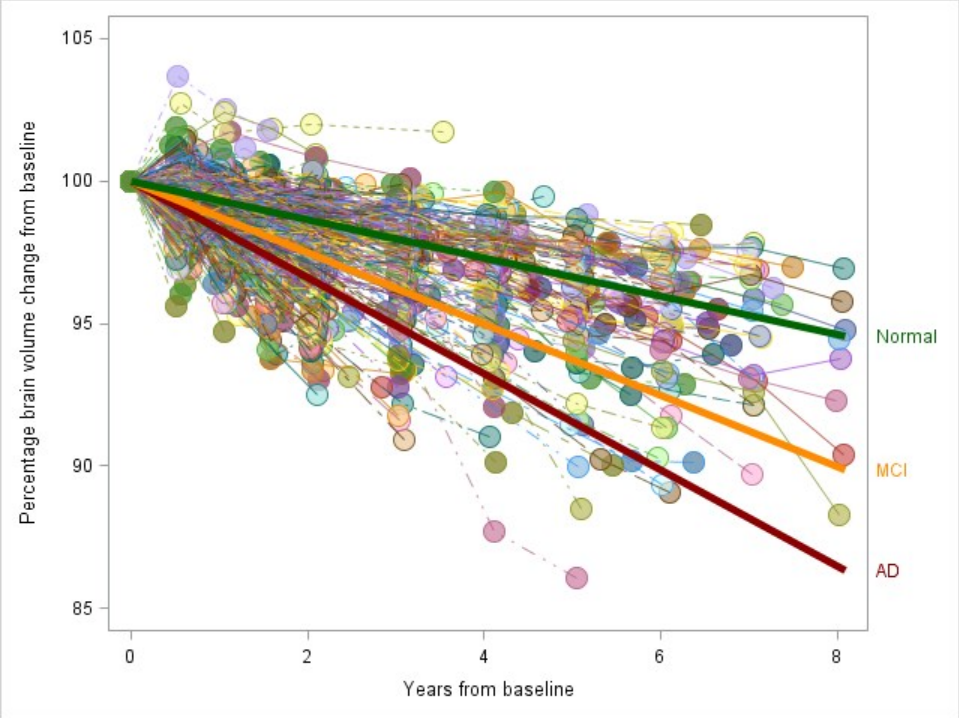

Figure 5-4: Group-average whole-brain atrophy trajectories, by diagnosis group (Chg- subjects)

Legend:

Colored dots and connected thin lines = actual PBVC measurement values with respect to baseline for each subject

Colored thick fitted lines = average rates of atrophy for each diagnosis group

Chg+ Subjects

Effects of the MRI scanner or T1-weighted sequence changes on PBVC measurements were estimated for seven combinations (subject N=237); the remaining six combinations had insufficient numbers of subjects or data points for model convergence. In all cases, the model with the corrective terms attained a lower AIC compared to that without the corrective terms. Figures 5-5 to 5-11 illustrate the additive effects of scanner upgrade or change on group-average time courses of atrophy (legend for all figures is placed below Figure 5-11). Detailed model outcomes can be found in Tables 5-5 to 5-11 (Appendix).

GE Genesis Signa to Philips Intera (Figure 5-5, Table 5-5): This inter-vendor scanner change led to an average increase of 0.98% (SE: 0.47) in PBVC, p=0.047.

Ge Signa Excite to GE Signa HDx (Figure 5-6, Table 5-6): This intra-vendor upgrade led to an average increase of 0.33% (0.095) in PBVC, p=0.0005. Also, the T1-weighted sequence change from MP-RAGE to IR-FSPGR led to an average decrease of -1.58% (0.35) in PBVC, p<0.0001.

Ge Signa Excite to GE Signa HDxt (Figure 5-7, Table 5-7): This intra-vendor upgrade led to an insignificant decrease of -0.049% (0.23) in PBVC, p=0.83. However, the T1-weighted sequence change from MP-RAGE to IR-FSPGR led to an average decrease of -2.50% (0.33) in PBVC, p<0.0001.

GE Signa Excite to GE Signa HDx to GE Signa HDxt (Figure 5-8, Table 5-8): The intra-vendor upgrade from Signa Excite to Signa HDx led to an average increase of 0.24% (0.095) in PBVC, p=0.0107. There was a trend of PBVC increase (0.26%, (0.16), p=0.108) when calculating changes from Signa Excite to Signa HDxt. The T1-weighted sequence change from MP-RAGE to IR-FSPGR led to an average decrease of -1.41% (0.15) in PBVC, p<0.0001.

GE Signa HDx to GE Signa HDxt (Figure 5-9, Table 5-9): This intra-vendor upgrade led to an insignificant decrease of -0.25% (0.25) in PBVC, p=0.31. However, the T1-weighted sequence change from MP-RAGE to IR-FSPGR led to an average decrease of -2.11% (0.32) in PBVC, p<0.0001.

Siemens Symphony to Siemens Symphony Total Imaging Matrix (TIM) (Figure 5-10, Table 5-10): This major intra-vendor upgrade led to an average decrease of -0.39% (0.17) in PBVC, p=0.021.

Philips Intera to Siemens Avanto (Figure 5-11, Table 5-11): This inter-vendor scanner change led to an average decrease of -1.79% (0.29) in PBVC, p<0.0001.

The 7 combinations analyzed above comprised N=237 Chg+ subjects. Model-based group-average atrophy rates by diagnosis group for these subjects, adjusted for scanner and sequence changes, were AD: -1.97%/y (0.11), MCI: -1.19%/y (0.058), and normal controls: -0.69%/y (0.073). Detailed model outcomes are shown in Table 5-12 (Appendix). The average effect T1-weighted sequence change from MP-RAGE to IR-FSPGR was -1.63% (0.12) in PBVC, p<0.0001 (Figure 5-12). The model with the corrective terms for scanner and sequence changes provided a lower AIC (3936.7) compared to the model without the corrective terms (AIC: 4274.8). Also, the average rates of atrophy in AD, MCI, and normal controls differed between these two models by 3%, 7%, and 16%, respectively.

Figure 5-5: GE Genesis Signa to Philips Intera

Figure 5-6: GE Signa Excite to GE Signa HDx

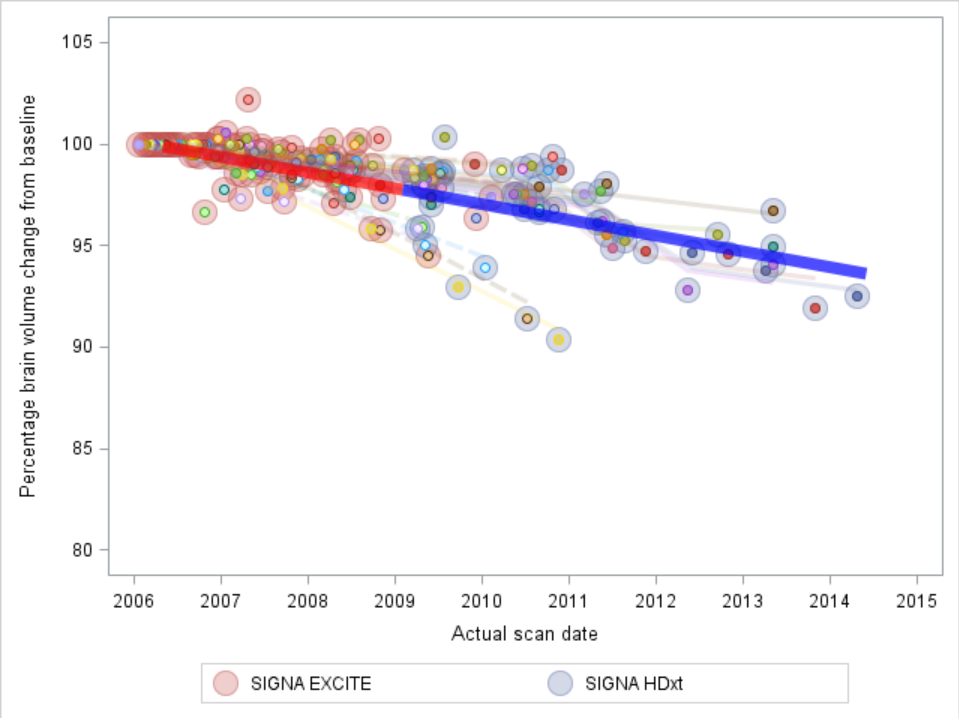

Figure 5-7: GE Signa Excite to GE Signa HDxt

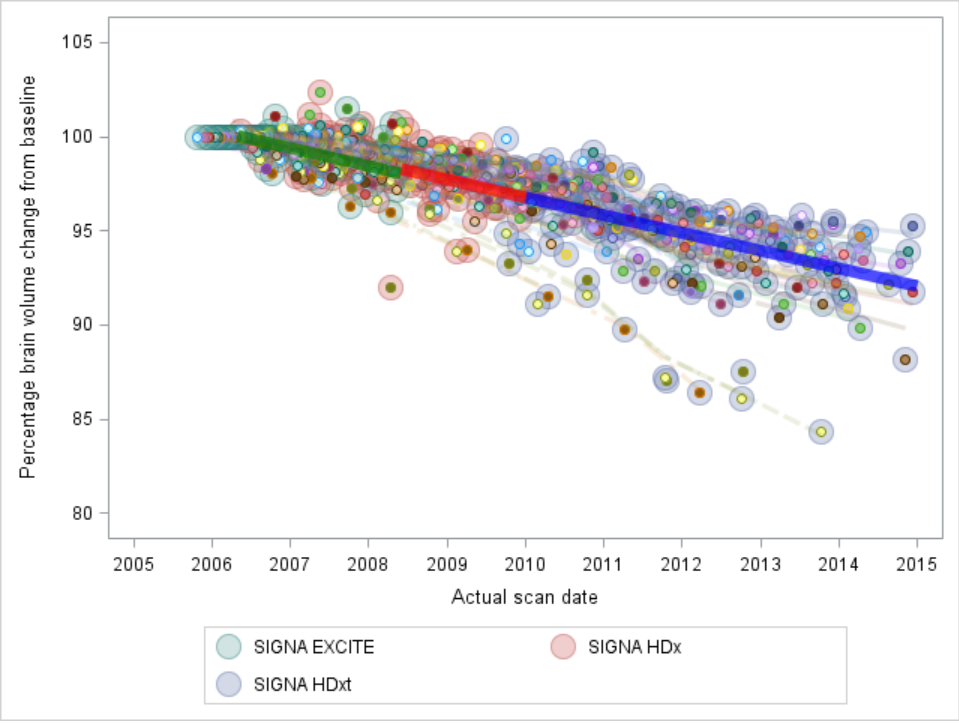

Figure 5-8: GE Signa Excite to GE Signa HDx to GE Signa HDxt

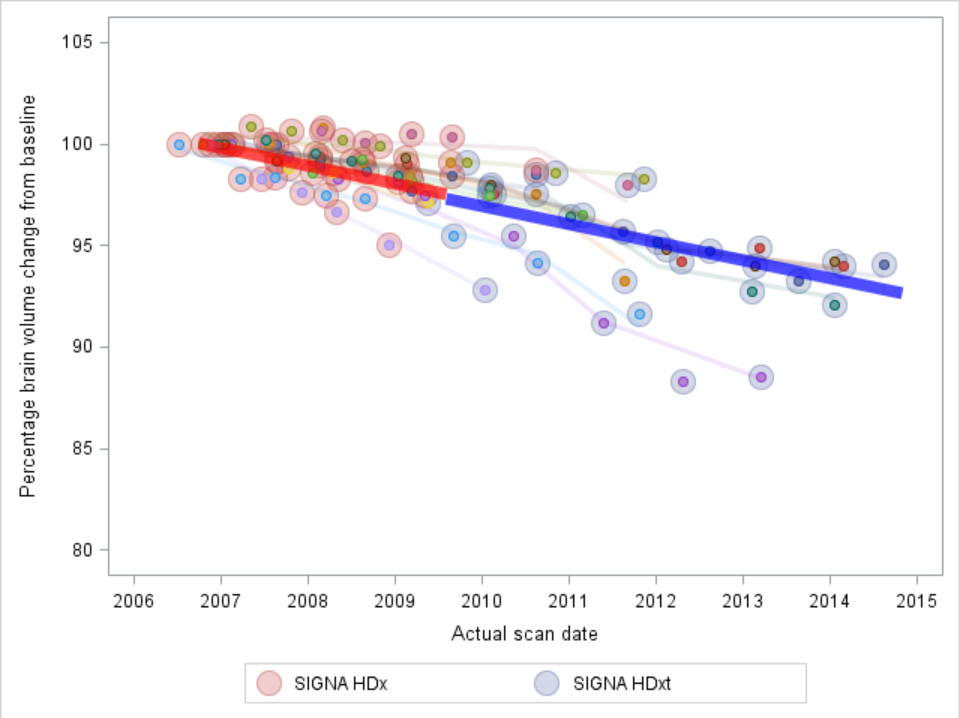

Figure 5-9: GE Signa HDx to GE Signa HDxt

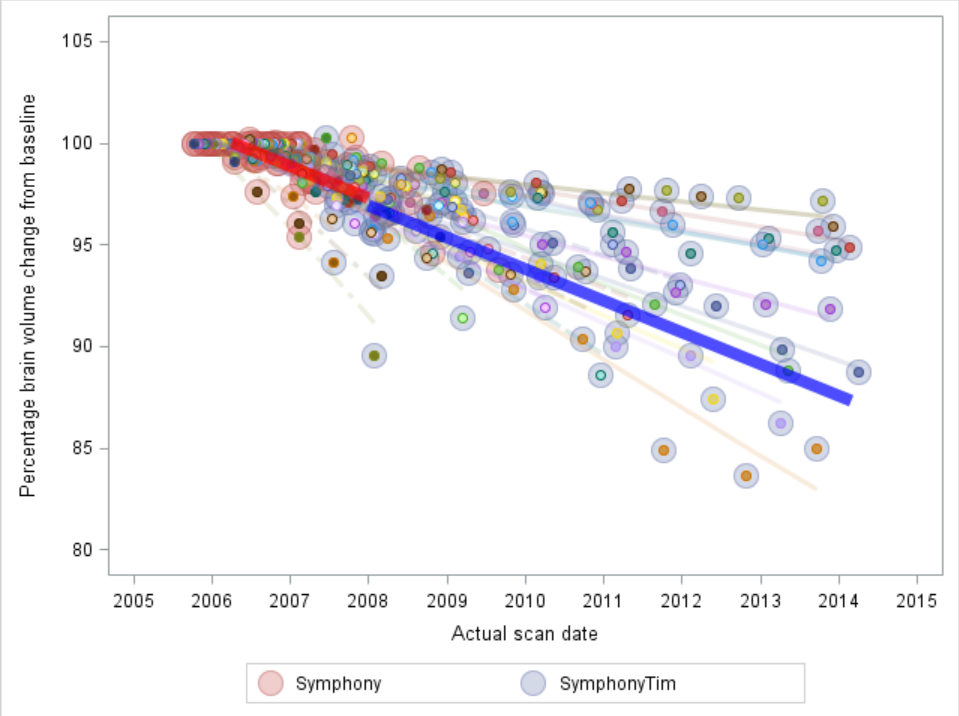

Figure 5-10: Siemens Symphony to Siemens Symphony TIM

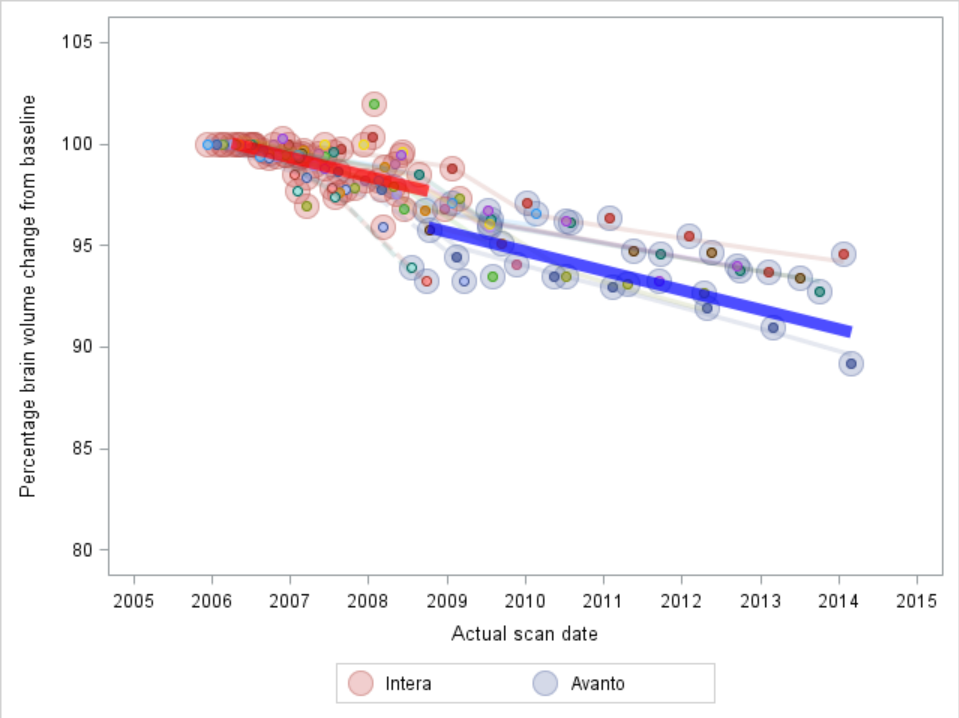

Figure 5-11: Philips Intera to Siemens Avanto

Legend for Figures 5-5 to 5-11:

Model-fitted lines, plotted against actual scan dates for better visualization of the scanner change effects.

Colored large circles = Each color denotes the specific MRI scanner used at that timepoint.

Colored thick lines = Average model for all subjects, grouped by MRI scanner used during follow-up. Each color represents each MRI scanner. The line discontinuity represents the scanner upgrade/change effect.

Colored small dots and thin lines = Actual PBVC measurement values with respect to baseline for each subject, with the thin lines representing the fitted model for each subject. Each color represents each subject. The thin-line discontinuities ('jump downs') at later scan dates on figures 5-6, 5-7, 5-8 and 5-9 (involving Signa HDx or HDxt) represent the effects of T1-weighted sequence change, in addition to the scanner changes.

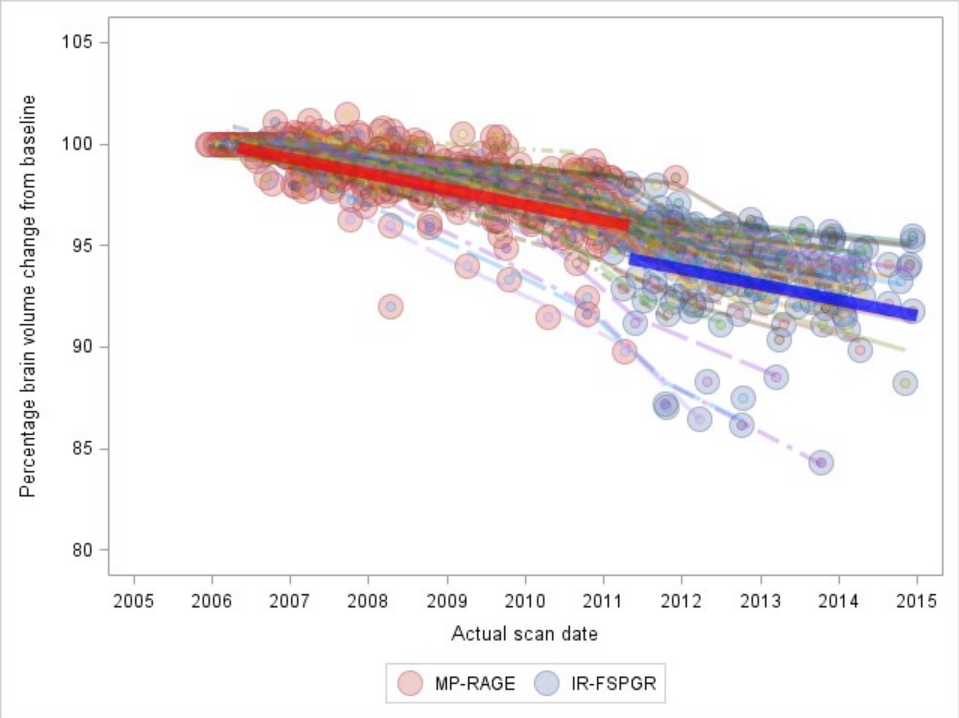

Figure 5-12: Illustration of the effect of T1-weighted sequence change from MP-RAGE to IR-FSPGR

Legend for Figure 5-12:

Model-fitted lines, plotted against actual scan dates for better visualization of the sequence change effects.

Colored large circles = Each color denotes the specific T1-weighted sequence used at that timepoint.

Colored thick lines = Average model for all subjects, grouped by T1-weighted sequence. Each color represents each sequence. The line discontinuity represents the sequence change effect.

Colored small dots and thin lines = Actual PBVC measurement values with respect to baseline for each subject, with the thin lines representing the fitted model for each subject. Each color represents each subject.

5.2.5 Discussion

We surveyed 819 normal control, MCI, and AD subjects enrolled in the ADNI-1 1.5T study and identified those who had an MRI scanner upgrade or change during follow-up. Longitudinal PBVCs, with respect to the baseline, were measured from serial MRIs that reached up to 8 years of follow-up. An LME model was applied to model the time courses of WB atrophy while estimating the effects of inter-vendor scanner change, intra-vendor scanner upgrade, and T1-weighted sequence change from MP-RAGE to IR-FSPGR (subset of GE scanners only). The change of sequence from MP-RAGE to IR-FSPGR was associated with an average of -1.63% change in PBVC. Artifactual changes in PBVC were found across different scanner hardware upgrade or change combinations. Inclusion of the corrective terms for scanner and sequence changes always led to a better model fit (i.e. lower AIC).

We first modeled the time courses of WB atrophy in the Chg- subgroup to explore rates unaffected by changes in the scanning hardware or T1-weighted sequence. The average rate of atrophy in AD patients was 1.35x higher than those of MCI patients and 2.52x higher than those of normal controls (Table 5-4). This is in line with a previous study by Leung and colleagues, in which they reported LME model-estimate WB atrophy rates over 3-years follow-up in the ADNI subjects.[225] Their model did not include terms for scanner or sequence, but these factors did not affect our Chg- group. Using a version of boundary-shift integral (KN-BSI), they found that the average rate in AD was 1.40x higher than those of MCI patients and 2.24x higher than those of normal controls. The actual rates differed due to the fact that SIENA tends to systematically give about 20% larger values compared to BSI,[122].

The Chg+ subgroup comprised N=237 subjects in which each was followed-up with one of the seven 1.5T MRI scanner model combinations. These seven combinations could be broadly classified into three categories: inter-vendor scanner change (GE to Philips; Philips to Siemens), intra-vendor scanner upgrade (GE to GE; Siemens to Siemens), and T1-weighted sequence change (GE to GE). Presuming greater degrees of changes in hardware configuration lead to a larger effect on image characteristics, we hypothesized that inter-vendor scanner changes would induce greater bias on PBVC measurements compared to intra-vendor scanner upgrades. Furthermore, we hypothesized that the change in T1-weighted sequence from MP-RAGE to IR-FSPGR would have had a direct impact on image contrast and thus the PBVC measurements,

even though the change occurred while on consistent MRI hardware platforms (GE Signa HDx or HDxt). These indeed were the cases in our study.

We analyzed two cases of inter-vendor scanner change combinations: GE Genesis Signa to Philips Intera, and Philips Intera to Siemens Avanto. Both cases represented a complete change in the hardware, including the main magnet and the coil. The average scanner change effects of +0.98% (Signa to Intera, $p=0.047$) and -1.79% (Intera to Avanto, $p<0.0001$) were significant and were roughly equivalent to a year's worth of atrophy in MCI and AD, respectively. As demonstrated in Figure 5-1, contrast differences between images of the same subjects scanned on these different scanners from different vendors were subtle but present. When inter-vendor scanner changes occur, the bias due to the scanner change may exceed the magnitude of the main effect of interest depending on the study.

There were five cases of intra-vendor scanner upgrade combinations: GE Signa Excite to HDx, GE Signa Excite to HDxt, GE Signa HDx to HDxt, GE Signa Excite to HDx to HDxt, and Siemens Symphony to Symphony TIM.

The upgrade from Excite to HDx included both hardware (e.g. receive chain architecture) and software components, whereas that from HDx to HDxt was mainly software-related. Upgrading from Excite to HDx exerted a significant effect on PBVC (+0.33%, $p=0.0005$), whereas going from Excite to HDxt did not (-0.049%, $p=0.83$). The average effect of a minor upgrade from HDx to HDxt was not significant (-0.25%, $p=0.31$). A similar pattern was observed in the group of subjects who had two upgrades from Excite to HDx to HDxt; upgrading from Excite to HDx led to a significant effect (+0.24%, $p=0.0107$) whereas going from Excite to HDxt did not (+0.26%, $p=0.108$). The Siemens TIM upgrade was a major hardware change that affected the gradient system, radiofrequency coil, and software. This upgrade led to a significant effect (-0.39%, $p=0.021$) on PBVC, comparable to half a year's worth of normal aging in this group, and about 20% of the annual change in this population of AD subjects. Direct hardware changes led to effects with both positive and negative directions. Overall, we found no evidence that the software-related upgrade from HDx to HDxt led to a significant systematic bias on PBVC.

Intra-vendor upgrades are the most common scenarios that occur in longitudinal studies, and our results suggest that, although the effects may differ from one upgrade to another, they may be ignorable if the rates of brain volume loss of interest are not very small, such as in AD.

Understanding the exact circumstances in which these effects are best ignored is not trivial, as it depends not only on the magnitude of the effect of interest, but also on study design factors such as the number of subjects affected by the change in comparison to the total number of subjects in the study.

The transition from ADNI-1 to ADNI-2/GO protocols included a change in the 3D T1-weighted sequence from MP-RAGE to IR-FSPGR in N=63 subjects scanned on select 1.5T GE platforms (N=4 HDx and N=59 HDxt). The earlier "works-in-progress" version of MP-RAGE focused on maximizing inter-vendor protocol standardization, but at the expense of replicability of the exact ADNI methods on other GE scanners; for the ADNI-2/GO phases, a complete switch was made to the manufacturer-available IR-FSPGR sequence.[232] This would have resulted in an alteration of the SNR and CNR.[238] Figure 5-3 shows example images from a single subject who had scans available from both sequences. Indeed, the T1-sequence change effect was significant in all cases with the average effect of -1.63% on PBVC, $p<0.0001$, estimated from the model that included all Chg+ subjects (Figure 5-12). This type of sequence change was specific to the ADNI study design and is unlikely to occur in a typical prospective longitudinal study. However, such changes may affect retrospective studies or studies in which standardized acquisitions are not performed. Our results demonstrate that the effect of sequence changes, even within the same hardware platform, can even exceed that observed in inter-vendor scanner changes and produce significant errors in brain volume measurements.

It is likely that the step changes in PBVC resulted from a change in the brain-CSF boundary delineation after the scanner or sequence change, since the measurement of brain volume changes over time generally depends on the detection of edge motion between registered scans. For example, SIENA uses the derivative of the gradient across the brain-CSF boundary to estimate the brain/non-brain edge motion between two timepoints and converts the mean edge displacement into the PBVC value.[120] An example scenario would be an improved boundary delineation (e.g. due to increased contrast or a reduction in the brain-CSF partial volume effects) which may result in an apparently reduced WB volume; in this case, the pre-upgrade volume would have been an overestimation due to the partial CSF volume being included in the brain volume.

Our observation has important implications for longitudinal studies of brain volume change. In all cases models with corrective terms for scanner and sequence changes yielded a lower AIC compared to those without, despite the two additional parameters being included in the model. This suggests that the atrophy rates estimated using the model with the corrective terms better represent the data, and that the rates from models without the corrective terms may be over- or under-estimated, depending on the direction of the effect of the change. For example, the average model-estimated rates in the subjects who had the intra-vendor upgrade from Symphony to Symphony TIM were 8%, 5%, and 16% different for AD, MCI and normal respectively, compared to the rates estimated ignoring the changes. In the subjects who switched from Philips Intera to Siemens Avanto, these differences were 29%, 28%, and 42% for AD, MCI and normal respectively. These results suggest that image pre- and post-processing steps alone may not be sufficient to remove the image variability originating from changes in scanning platforms. Additional steps, such as incorporating corrective terms into the statistical analysis model, may be necessary to attain a more reliable outcome. However, the modeling requires that there be sufficient numbers of subjects affected by any modeled change to be able to estimate the effect. This can be a challenge in some multi-center drug trials with many sites and not many subjects per site.

Study limitations and future directions: There were several limitations to our study. First, the ADNI study provided a valuable large-scale, multi-site dataset acquired using a standardized protocol and quality control. We surveyed the data 'as-is', and could analyze seven different scanner upgrade/change combinations. This approach to subject selection inherently resulted in an unbalanced design. This issue was partially alleviated by the use of the LME model, which can accommodate unbalanced data.[188] Second, the average age at baseline was around 75 for our subjects, and this is when the rate of normal aging-related WB atrophy begins to accelerate.[137] We kept the model as parsimonious as possible and did not take the potentially non-linear pattern of WB atrophy into account. Although this effect may not be apparent over the short-term,[225] it could have affected subjects who had 8 years of follow-up. Third, it is unknown whether the specific estimates of the effects for different scanner changes obtained from this study are generalizable, as various designs exist with regards to image acquisition, processing, and analysis pipeline. Yet, our study pipeline can be fairly easily replicated. The ADNI acquisition protocol is widely available and is increasingly being used in clinical trials, and our pre-

processing steps and the SIENA method also have been commonly used. Whether our specific estimates apply to different image processing pipelines or atrophy measurement techniques (e.g. BSI, Jacobian Integration) needs to be further explored. Our study analyzed only 1.5T scans, but 3.0T systems are rapidly being adopted. In fact, newly enrolled subjects in the ADNI-2/GO have been entirely scanned at 3.0T.[232] Potential effects of 3.0T scanner change/upgrade need to be investigated. Finally, there is a growing interest in measuring grey- and white-matter volumes separately, as grey-matter atrophy may be better correlated with disability progression and cognitive impairment than WB atrophy.[99] The measurement of grey-matter atrophy itself is technically challenging, and any scanner upgrade or change can add further complexity to the analysis. Kruggel and colleagues demonstrated significant within-subject variability of grey- and white-matter compartmental volumes (and thus, WB volume also) on different 1.5T and 3.0T scanners used in the ADNI.[182] Moreover, Nakamura and colleagues revealed that the presence of white-matter lesions can significantly bias grey- and white-matter segmentation.[128] Further research on this important topic is warranted.

In conclusion, we demonstrated that different scanner hardware upgrades can exert different bias effects on PBVC. Inter-vendor scanner changes generally led to greater effects compared to intra-vendor scanner upgrades. Change in the 3D T1-weighted sequence from MP-RAGE to IR-FSPGR, within the same scanning platform, also led to a significant effect, comparable to that from inter-vendor scanner changes. Modeling brain volume loss with an LME model that includes corrective terms for scanner and sequence changes yields better model fits and more reliable estimates of WB atrophy rates.

5.2.6 Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli

Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Hyunwoo Lee was supported by the Doctoral Training Award from Fonds de recherché du Quebec.

5.2.7 Appendix

Table 5-4: Model-estimated group-average whole-brain atrophy rates for subjects who did not have MRI scanner upgrade or change during follow-up (Chg-)

Effect	Diagnosis group	Estimate	Standard error
Estimated whole-brain atrophy rates by diagnosis group	**Normal**	**-0.67%/y**	**0.060**
	MCI	**-1.25%/y**	**0.048**
	AD	**-1.69%/y**	**0.073**

Table 5-5: Effect of inter-vendor scanner change from GE Genesis Signa to Philips Intera (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				99.80%	0.43	<.0001	Model AIC: 167.3
Reference T1-weighted sequence		MP-RAGE only		.	.	.	**Model without the corrective terms for scanner and sequence changes**: AIC: 169.9 Normal: -0.13%/y MCI: -0.81%/y AD: -1.09%/y
MRI scanner change	**GE Genesis Signa to Philips Intera**			**0.98%**	**0.47**	**0.047**	
Reference MRI scanner	Genesis Signa			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.29%/y**	**0.44**	**0.56**	
			MCI	**-1.03%/y**	**0.23**	**0.0055**	
			AD	**-1.57%/y**	**0.75**	**0.050**	

Table 5-6: Effect of intra-vendor scanner upgrade from GE Signa Excite to GE Signa HDx (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				99.97%	0.047	<.0001	Model AIC: 983.0
T1-weighted sequence change		**IR-FSPGR**		**-1.58%**	**0.35**	**<.0001**	**Model without the corrective terms for scanner and sequence changes**: AIC: 1021.0 Normal: -0.72%/y MCI: -1.18%/y AD: -1.80%/y
Reference T1-weighted sequence		MP-RAGE		.	.	.	
MRI scanner upgrade	**GE Signa Excite to GE Signa HDx**			**0.33%**	**0.095**	**0.0005**	
Reference MRI scanner	GE Signa Excite			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.77%/y**	**0.22**	**0.0006**	
			MCI	**-1.30%/y**	**0.14**	**<.0001**	
			AD	**-1.94%/y**	**0.16**	**<.0001**	

Table 5-7: Effect of intra-vendor scanner upgrade from GE Signa Excite to GE Signa HDxt (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				99.89%	0.097	<.0001	Model AIC: 431.0
T1-weighted sequence change		**IR-FSPGR**		**-2.50%**	**0.33**	**<.0001**	**Model without the corrective terms for scanner and sequence changes**: AIC: 482.1 Normal: -0.70%/y MCI: -1.13%/y
Reference T1-weighted sequence		MP-RAGE		.	.	.	
MRI scanner upgrade	**GE Signa Excite to GE Signa HDxt**			**-0.049%**	**0.23**	**0.83**	
Reference MRI scanner	GE Signa Excite			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.57%/y**	**0.18**	**0.0049**	
			MCI	**-0.97%/y**	**0.19**	**<.0001**	

Table 5-8: Effects of two intra-vendor scanner upgrades from GE Signa Excite to GE Signa HDx to GE Signa HDxt (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				100.02%	0.066	<.0001	Model AIC: 1227.5 **Model without the corrective terms for scanner and sequence changes**: AIC: 1353.2 Normal: -0.91%/y MCI: -1.19%/y
T1-weighted sequence change		**IR-FSPGR**		**-1.41%**	**0.15**	**<.0001**	
Reference T1-weighted sequence		MP-RAGE		.	.	.	
Second MRI scanner upgrade	**GE Signa Excite to GE Signa HDxt**			**0.26%**	**0.16**	**0.108**	
First MRI scanner upgrade	**GE Signa Excite to GE Signa HDx**			**0.24%**	**0.095**	**0.0107**	
Reference MRI scanner	GE Signa Excite			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.78%/y**	**0.089**	**<.0001**	
			MCI	**-1.07%/y**	**0.078**	**<.0001**	

Table 5-9: Effect of intra-vendor scanner upgrade from GE Signa HDx to GE Signa HDxt (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				100.04%	0.15	<.0001	Model AIC: 208.1
T1-weighted sequence change		**IR-FSPGR**		**-2.11%**	**0.32**	**<.0001**	**Model without the corrective terms for scanner and sequence changes**: AIC: 250.3 Normal: -0.92%/y MCI: -1.27%/y AD: -1.42%/y
Reference T1-weighted sequence		MP-RAGE		.	.	.	
MRI scanner upgrade	**GE Signa HDx to GE Signa HDxt**			**-0.25%**	**0.25**	**0.31**	
Reference MRI scanner	GE Signa HDx			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.46%/y**	**0.45**	**0.33**	
			MCI	**-0.96%/y**	**0.21**	**0.0011**	
			AD	**-1.17%/y**	**0.72**	**0.13**	

Table 5-10: Effect of intra-vendor scanner upgrade from Siemens Symphony to Siemens Symphony TIM (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				100.05%	0.073	<.0001	Model AIC: 610.6
Reference T1-weighted sequence		MP-RAGE only		.	.	.	**Model without the corrective terms for scanner and sequence changes**: AIC: 628.5 Normal: -0.68%/y MCI: -1.81%/y AD: -2.87%/y
MRI scanner upgrade	**Siemens Symphony to Siemens Symphony TIM**			**-0.39%**	**0.17**	**0.021**	
Reference MRI scanner	Siemens Symphony			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.58%/y**	**0.20**	**0.012**	
			MCI	**-1.73%/y**	**0.16**	**<.0001**	
			AD	**-2.65%/y**	**0.31**	**<.0001**	

Table 5-11: Effect of inter-vendor scanner change from Philips Intera to Siemens Avanto (Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				100.05%	0.14	<.0001	Model AIC: 252.1
Reference T1-weighted sequence		MP-RAGE only		.	.	.	**Model without the corrective terms for scanner and sequence changes**: AIC: 281.3 Normal: -1.01%/y MCI: -1.40%/y AD: -2.97%/y
MRI scanner change	**Philips Intera to Siemens Avanto**			**-1.79%**	**0.29**	**<.0001**	
Reference MRI scanner	Philips Intera			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.66%/y**	**0.100**	**0.004**	
			MCI	**-1.06%/y**	**0.13**	**0.0015**	
			AD	**-2.21%/y**	**0.29**	**<.0001**	

Table 5-12: Model-estimated group-average whole-brain atrophy rates for subjects who had MRI scanner upgrade or change during follow-up, adjusted for the effects of T1-weighted sequence change and scanner change or upgrade (N=237 Chg+)

Effect	MRI scanner model	T1-weighted sequence	Diagnosis group	Estimate	Standard error	Pr > |t|	Note
Intercept				99.82%	0.25	<.0001	Model AIC: 3936.7
T1-weighted sequence change		IR-FSPGR		-1.63%	0.12	<.0001	**Model without the corrective terms for scanner and sequence changes: AIC: 4274.8 Normal: -0.81%/y MCI: -1.28%/y AD: -2.03%/y**
Reference T1-weighted sequence		MP-RAGE		.	.	.	
MRI scanner	Siemens Symphony to Siemens Symphony TIM			-0.41%	0.30	0.17	
MRI scanner	Siemens Symphony			0.19%	0.26	0.47	
MRI scanner	GE Signa HDx to GE Signa HDxt			0.097%	0.39	0.80	
MRI scanner	GE Signa HDx			0.30%	0.28	0.28	
MRI scanner	GE Signa Excite to GE Signa HDxt			0.41%	0.28	0.14	
MRI scanner	GE Signa Excite to GE Signa HDx			0.40%	0.26	0.12	
MRI scanner	GE Signa Excite			0.17%	0.25	0.51	
MRI scanner	Philips Intera to Siemens Avanto			-1.58%	0.39	<.0001	
MRI scanner	Philips Intera			0.23%	0.28	0.42	
MRI scanner	GE Genesis Signa to Philips Intera			1.15%	0.48	0.02	
Reference MRI scanner	GE Genesis Signa			.	.	.	
Estimated whole-brain atrophy rates by diagnosis group			**Normal**	**-0.69%/y**	**0.073**	**<.0001**	
			MCI	**-1.19%/y**	**0.058**	**<.0001**	
			AD	**-1.97%/y**	**0.11**	**<.0001**	

Chapter 6 General Discussion

The overarching aim of this thesis was to further understand the impact of aHSCT on brain atrophy in MS patients who have high disease activity despite being on DMTs. Using multi-year MRI follow-up data from two independent MS cohorts treated with aHSCT, time courses of whole-brain, grey matter, and white matter atrophy were measured and modeled in these patients. Both studies suffered from mid-study MRI scanner upgrades, which were accounted for at the statistical analysis level. To further investigate the impact of scanner hardware changes on whole-brain atrophy measurements, an independent dataset from the ADNI study was analyzed. This chapter begins by discussing the main findings and their implications. Then, study limitations and potential future directions for further research are discussed. The chapter ends with the final conclusion.

6.1 Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis

IA/aHSCT can induce significant long-term reduction[220], or even stoppage[195] of relapses and new focal white matter lesion formation in MS patients with aggressive inflammatory disease activity. This, however, is paradoxically paralleled by an early acceleration in the rate of brain volume loss following treatment. The goal of chapter 2 was to explain this phenomenon at the whole-brain level, by 1) examining the factors potentially associated with the early, accelerated volume loss, 2) estimating the timeframes for each of these processes, and 3) estimating the long-term rates of volume loss to assess the impact of aHSCT on whole-brain atrophy in MS.

To achieve this, I modeled the time courses of whole-brain volume change in the Canadian MS-BMT patients. The early acceleration was modeled in terms of two hypothesized factors: (1) the dose of busulfan chemotherapy regimen (an index of chemotherapy-related neurotoxicity) and (2) baseline volume of T1-weighted white matter lesions (a marker of the amount of focally injured tissue that may be committed to degeneration prior to aHSCT).

Modelling showed both were significant factors, but the busulfan dose was associated with a greater proportion of the accelerated whole-brain volume loss. The additional contribution from

the white matter lesion-related process suggested that ongoing degeneration of lethally injured tissues continued to progress after aHSCT. The assessment of changes in the brain tissue water content did not provide evidence that the acceleration was due to resolution of edema. The accelerated atrophy slowed continuously over approximately 2.5 years, after which the average rate of whole-brain atrophy was consistent with the rate observed in normal aging.

These findings suggest that the early, accelerated brain volume loss in these patients reflect actual tissue loss, largely because of the toxicity due to the chemotherapy drugs. Cytotoxicity to normal CNS cells is likely an unavoidable consequence of a treatment program that involves high-dose chemotherapy. Therefore, further investigation is warranted to determine a strategy that can achieve an appropriate level of immunoablation (for more complete control of inflammatory activity) while minimizing side-effects such as brain atrophy.

Another important observation associated with the complete ablation of focal inflammatory activity was that the rate of brain atrophy progressively decreased to the rate expected with normal aging. This finding may serve as evidence that stopping the immune dysregulation not only halts ongoing focal inflammatory processes, but also can slow or even stop degenerative processes in MS. However, it may take up to three years before the cytotoxic effects subside and the benefit of aHSCT on whole-brain atrophy becomes apparent.

6.2 Impact of immunoablation and autologous hematopoietic stem cell transplantation on grey and white matter atrophy in multiple sclerosis

Processes underlying the time course of whole-brain atrophy can be further elucidated by examining compartmental volume changes separately. Grey matter and white matter consist of different types of cells (e.g. primarily neuronal cell bodies and glial cells in the grey matter, myelinated axons in the white matter), and the treatment-related atrophy of each tissue compartment may proceed differently, as different underlying processes can contribute to atrophy. Also, it is unknown whether the slowing of whole-brain atrophy after aHSCT is reflected in both compartments. The goal of chapter 3 was to expand upon the findings of chapter 2, by assessing the impact of aHSCT on grey and white matter separately.

To achieve this, I modeled the time courses of grey and white matter atrophy in the Canadian MS-BMT patients. As in the assessment of whole-brain atrophy in this cohort, I used two hypothesized factors: the dose of the busulfan chemotherapy regimen and the baseline volume of T1-weighted white matter lesions. Both were significant predictors of the early, accelerated white matter atrophy following IA/aHSCT, whereas only the busulfan dose was a significant predictor of the grey matter atrophy. In keeping with what I have reported for whole-brain atrophy, the rates of both grey and white matter atrophy subsequently slowed to levels seen in normal-aging, although in different timeframes: atrophy in the grey matter slowed over the first 1 to 2 years, whereas that in the white matter slowed over approximately the first 2 to 3 years.

This finding suggests that a common mechanism related to acute chemotoxicity operates across the grey and white matter. This is in line with the finding that clinically relevant doses of chemotherapy agents can still kill normal neuronal and glial progenitor cells, as well as non-dividing oligodendrocytes.[91] Grey matter atrophy as a side-effect of chemotherapy is something to be further investigated because it may be associated with the so-called "chemo-fog" phenomenon, and confound the clinical evaluation of aHSCT on MS in some patients for an unknown amount of time.[210]

The latter part of the accelerated brain atrophy appears to be driven primarily by the white matter, possibly due to compartment-specific processes such as Wallerian degeneration. This could be a consequence of significant white matter damage that occurred before treatment, and it may take several years before it fully resolves.

Notably, there was a significant reduction in the long-term rates of grey matter atrophy compared to the baseline rates. Also, the slowing in white and grey matter atrophy occurred in conjunction with the halting of relapses and new white matter lesion formation in these patients; these findings further support that addressing immune dysregulation can not only halt focal inflammatory activity, but also decrease volume loss in both white matter and grey matter; in particular, grey matter volume loss is a potentially important marker of disease progression in MS.

6.3 Brain atrophy in relapsing remitting multiple sclerosis following high-dose immunosuppressive therapy and autologous hematopoietic cell transplantation in the HALT-MS trial

The goal of chapter 4 was to confirm the observations made in the Canadian MS-BMT trial in the HALT-MS trial, which also involved aHSCT but with an intermediate-intensity BEAM chemotherapy regimen.

To achieve this, I modeled the time courses of whole-brain, grey matter, and white matter atrophy in the HALT-MS cohort. As in the assessment of the Canadian MS-BMT cohort, I used two hypothesized factors to describe the early, accelerated atrophy in all compartments: the total dose of BEAM chemotherapy regimen and the baseline volume of T1-weighted white matter lesions. Both were significant predictors of the early, accelerated whole-brain and white matter atrophy. For the grey matter, only the BEAM dose was a significant predictor of atrophy. Overall, the BEAM dose was the dominant predictor in all compartments.

The HALT-MS cohort experienced shorter and milder courses of the early, accelerated brain volume loss compared to the Canadian MS-BMT cohort; on average, the whole-brain, grey matter, and white matter volume losses due to the early acceleration in HALT-MS were, respectively, about 3-fold, 2-fold, and 2-fold lower compared to that in the Canadian MS-BMT. The accelerated whole-brain atrophy slowed continuously over approximately the first year post-treatment; atrophy in the grey matter also slowed over the first year, and that in the white matter slowed over a longer period, approximately the first 1 to 2 years.

Over the long term, the average rates of whole-brain and grey matter atrophy in the HALT-MS cohort were comparable to those found in the Canadian MS-BMT cohort. However, this was not true for white matter, for which the average rate was about two-fold higher than that in the Canadian MS-BMT cohort, potentially due to the less complete suppression of focal inflammation in the HALT-MS patients. Still, these results suggest that immunoablation or adequate immunosuppression and autologous hematopoietic stem cell transplantation can reduce brain atrophy rates to those consistent with normal aging.

One reason for the different time courses of atrophy might be the usage of an intermediate-intensity BEAM regimen with lower toxicity profile, as opposed to a high-intensity busulfan-

based regimen used in the Canadian MS-BMT trial. This suggests that informed decisions need to be made with regards to the strategy that can provide the best outcome for the patients. On one hand, high intensity drugs, such as busulfan/cyclophosphamide, may better eliminate autoreactive processes, providing better control of focal inflammatory activity and degeneration in MS; but they are inherently more toxic and cause more severe side effects including treatment-related brain atrophy. On the other hand, lower intensity drugs can be better tolerated, but may not be as effective in controlling focal inflammation or long-term disease related degeneration. A randomized comparison of chemotherapy regimens would be a feasible way to objectively assess the benefits and toxicity profiles, and to devise an optimized strategy.

6.4 Estimating and accounting for the effect of MRI scanner hardware changes on longitudinal whole-brain atrophy measurements

Although the focus of this thesis was to examine the courses of brain atrophy in MS patients treated with immunoablation/immunosuppression and aHSCT, another important contribution has been the assessment of the effect of MRI scanner changes on longitudinal atrophy measurements. As discussed in chapters 2 to 4, both the Canadian MS-BMT and the HALT-MS studies were affected by mid-study MRI scanner upgrades, which were accounted for in the work presented here by including a corrective term in the statistical model. In fact, mid-study MRI scanner upgrades or changes frequently occur in longitudinal studies like the abovementioned trials. However, at least in brain atrophy studies, this issue is not always dealt with in a consistent manner, or is sometimes even ignored. One reason might be that the effects of scanning platform changes on longitudinal measurement of atrophy are not clear; further information on the magnitude of these effects will be instrumental in interpreting atrophy measurements from multiple scanning platforms. The goal of chapter 5 was to estimate the size of effects of different MRI scanner upgrade or change combinations, and of T1-weighted sequence changes, on whole-brain volume change measurements, using a large cohort from the ADNI study.

To achieve this, I analyzed subjects who were subject to scanner upgrades, separating these into *intra*-vendor upgrades or *inter*-vendor changes during follow-up. A subset of these subjects were also affected by a change in the standardized T1-weighted sequence. A linear mixed-effects

model was applied to estimate the effects of the scanner and the sequence changes on the whole-brain volume change measurements. AIC was used to assess whether including the corrective terms in the model led to improvements in goodness-of-fit measures.

Generally, the effects due to inter-vendor scanner changes were relatively large and roughly equivalent to a year's worth of atrophy in AD and MCI. Calculating rates of brain atrophy without considering this effect would result in various levels of over- or under-estimations and confound the true effect of interest. It should not be presumed that image-processing steps would adequately account for this technical bias. In principle, it may be corrected using living phantoms to estimate the scanner change effect and calibrate the measurements. However, this is generally not feasible. At a minimum, the effect of inter-vendor scanner changes should be considered at the statistical analysis level.

The effects due to intra-vendor scanner upgrades were smaller: the magnitude of the largest effect was roughly comparable to half a year's worth of normal aging in this group, and about 20% of the annual change in the AD subjects. Depending on the changes in atrophy rates being looked for, it may be practical to ignore these effects. However, if the rates being examined are small, similar arguments as the above paragraph apply.

The effect of the examined T1-weighted sequence change was large, and was comparable to that from inter-vendor scanner changes. Ideally, images from two different sequences should not be directly compared, as their contrasts may differ significantly. Yet, in practice, this sometimes needs to be done in situations such as retrospective studies. The sequence change effect should be carefully considered during the analysis and needs to be reported along with the results.

Inclusion of the corrective fixed-effects terms led to better model goodness-of-fits, and thus, provided more reliable estimates of whole-brain atrophy rates. This finding demonstrates the overall advantage of considering scanner changes or upgrades in the analysis of brain atrophy.

Overall, this thesis focused on the understanding of brain atrophy processes in MS, using unique datasets from MS cohorts treated with immunoablation/immunosuppression and aHSCT. I characterized the courses of whole-brain, grey matter, and white matter atrophy and demonstrated that the early acceleration of brain atrophy in these patients is likely due to

chemotherapy-related CNS toxicity as well as loss of white matter tissues already in the process of degenerating due to MS-related injury. The results obtained from the studies may also serve as evidence that stopping focal white matter inflammation can eventually lead to substantial slowing of brain atrophy in MS. In addition, I have confirmed that the effects of MRI scanner upgrade or change on brain atrophy measurement can be significant depending on the situation, and it would be beneficial to account for these effects during statistical analysis.

6.5 Study Limitations and Future Directions

This thesis raises several key issues that could serve as a basis for future research. Building upon these issues, this section proposes several potential studies that may aid in further understanding of the mechanisms of brain atrophy in MS.

6.5.1 A randomised comparison of different immunoablative/immunosuppressive regimens

In chapters 2 to 4, I showed that the early acceleration in the grey matter and white matter (and thus, whole-brain) atrophy are likely due to the neurotoxic effects of chemotherapeutic regimens. Based on the differences in the accelerated volume losses between the Canadian MS-BMT and HALT-MS cohorts, I suggested that a higher-intensity chemotherapy agent with higher toxicity profile may be associated with a higher degree of treatment-related brain atrophy, and *vice versa*. However, this claim can only be adequately confirmed with a randomised comparison of different chemotherapy regimens.

The current opinion is that early treatment on young patients in relapsing phases of MS can provide the best outcome.[218] An example study design would be to randomly assign this population of subjects into aHSCT therapies using different chemotherapy regimens, ranging from low to high intensity. Doing so, between-group comparisons of treatment-related brain atrophy and clinical outcomes can be made. If differences are evident, their potential associations with long-term MRI and clinical outcomes can also be assessed. This would help clinicians make informed decisions as to which treatment strategy can best suit the needs of the patients.

6.5.2 A controlled comparison of chemotherapy-related brain atrophy in MS subjects versus non-MS subjects

One of the goals of this thesis was to determine whether MS patients, who have already suffered damage to brain tissue, have a predisposition to chemotherapy-related brain atrophy that is greater than that in subjects without pre-existing brain pathology. In chapter 3, I investigated this question to a certain extent by testing whether the atrophy associated with busulfan dose could be more severe in patients with indications of more injured brain (i.e. higher baseline T1-weighted white matter lesion volume). However, this question can be best answered by making a direct comparison of chemotherapy-related brain atrophy between MS subjects and non-MS subjects. An ideal non-MS control group would be those who will be receiving the same types of chemotherapy, for example, Hodgkin or non-Hodgkin lymphoma patients who are treated with BEAM or busulfan-based therapies. Between-group comparisons of treatment-related brain atrophy would reveal the potential differences, which would serve as evidence whether or not MS patients are further vulnerable to chemotherapy-related toxicity. This will help to determine whether MS patients require additional pre-emptive measures to minimize side-effects of the immunoablative therapy.

6.5.3 A controlled comparison of long-term brain atrophy rates in MS subjects treated with aHSCT versus normal controls

Both the Canadian MS-BMT and the HALT-MS trials were single-arm studies without randomised control groups. As a result, the long-term rates of whole-brain atrophy in these patients were compared to values in the literature obtained using a similar imaging protocol and the same atrophy measurement technique. However, the best approach to determine whether the long-term atrophy rates in these patients slowed to normal aging level, would have been to compare them to the rates observed in a concurrent healthy control arm.

Nonetheless, healthy controls are infrequently included in longitudinal clinical trials, at least in the MS field. Treatment effects on brain atrophy usually are assessed based on comparison to placebo arms or patients on other therapy.[68,239] But based on the findings in chapters 2 to 4, it would be informative to include a set of healthy control subjects in the MRI protocol of future trials where a very strong treatment effect is expected, e.g. immunoablation and aHSCT.[240]

Doing so would make it possible to determine whether the treatment can truly return brain atrophy rates in MS to normal.

6.5.4 Higher field MRI and markers of focal grey matter pathology

Grey matter lesions are present in all stages of MS,[45] affect a significant portion of the cortex,[45] and are associated with cognitive deficits.[241] Although these lesions are most prominent in the progressive phases of MS,[34] they may have affected the poor-prognosis subjects enrolled in the Canadian MS-BMT and the HALT-MS trials. Unfortunately, both trials used 1.5 T scanners that were insensitive to focal grey matter lesions, and therefore it was not possible to confirm whether aHSCT had any effect on stopping grey matter lesions. Although I used T1-weighted white matter lesions as a potential predictor of atrophy in all compartments, grey matter lesions could also be a relevant predictor of grey matter atrophy.[242]

Detection of grey matter lesions, however, is a challenge by itself. Techniques like double-inversion recovery can only detect a small portion of cortical lesions, especially the leukocortical types.[112] 7 T scanners can detect intracortical and subpial lesions with better sensitivity compared to 3 T scanners, but overall still miss more than 40% of histopathologically-confirmed cortical lesions.[114] Therefore, it may be many years before grey matter lesions can be reliably detected using MRI, and be routinely assessed in clinical trials.

The above suggestions would be of high interest assuming there is still an important role for immunoablation and aHSCT, despite the recent development of highly effective DMTs such as natalizumab, alemtuzumab and ocrelizumab. An NIH-funded multicenter clinical trial comparing aHSCT with best available therapy is being planned,[240] and will answer this question.

6.5.5 Accounting for MRI scanner upgrades in longitudinal studies

As discussed in chapter 5, different scanner hardware upgrades or changes can exert different bias effects on whole-brain volume change measurements. I showed that intra- and inter-vendor scanner changes, as well as changes in T1-weighted sequence have effects on brain volume

change measurements, and that these can be quite large. This analysis could be extended in various ways; for example, when is it better to exclude data across a scanner change as opposed to not excluding it. This question depends on many variables, such as the total number of subjects, the number of scanners, the magnitude of the effect of interest (e.g. atrophy rates), and the effect of the scanner upgrade, and the alternative to keeping the data (e.g. excluding affected patients from the analysis or imputing the missing data) to name a few. This question is worth investigating as it has important implications for multicenter clinical trials.

6.6 Conclusion

In conclusion, this work contributes to the understanding of the potential mechanisms underlying brain atrophy in MS patients treated with aHSCT. First, the current generation of immunoablative chemotherapy drugs cause an unwanted side effect of brain atrophy in all tissue compartments. The next generation treatment protocols should try to optimize the balance between maximizing efficacy and minimizing toxicity. Second, loss of injured white matter tissues, which are already in the process of degenerating, may continue to progress for years after aHSCT. This suggests that, in future clinical trials, at least three years of follow-up monitoring is necessary to detect the full beneficial effect of aHSCT on brain atrophy. Third, suppressing all measurable focal white matter inflammatory disease activity in MS patients results in a subsequent reduction of brain atrophy rates to levels comparable to normal aging. This supports the hypothesis that targeting inflammation can provide substantial benefits to MS patients, especially for those with high inflammatory disease activity. Finally, a mid-study scanner change can significantly confound the interpretation of brain atrophy measurements. This issue is not always dealt with appropriately, but is crucial in reliably assessing brain atrophy not only in MS, but in all applicable disciplines.

References

1. Browne P, Chandraratna D, Angood C, et al. Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology* 2014; 83: 1022–4.

2. Beck C a, Metz LM, Svenson LW, et al. Regional variation of multiple sclerosis prevalence in Canada. *Mult Scler* 2005; 11: 516–519.

3. Dunn SE, Gunde E, Lee H. Sex-Based Differences in Multiple Sclerosis (MS): Part II: Rising Incidence of Multiple Sclerosis in Women and the Vulnerability of Men to Progression of this Disease. *Curr Top Behav Neurosci* 2015; 26: 57–86.

4. Weinshenker BG, Bass B, Rice GP, et al. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. *Brain* 1989; 112: 133–146.

5. Dutta R, Trapp BD. Mechanisms of neuronal dysfunction and degeneration in multiple sclerosis. *Prog Neurobiol* 2011; 93: 1–12.

6. Gourraud PA, Harbo HF, Hauser SL, et al. The genetics of multiple sclerosis: An up-to-date review. *Immunol Rev* 2012; 248: 87–103.

7. Compston A, Coles A. Multiple sclerosis. *Lancet* 2008; 372: 1502–1517.

8. Ramagopalan S V., Dobson R, Meier UC, et al. Multiple sclerosis: risk factors, prodromes, and potential causal pathways. *Lancet Neurol* 2010; 9: 727–739.

9. Belbasis L, Bellou V, Evangelou E, et al. Environmental risk factors and multiple sclerosis: An umbrella review of systematic reviews and meta-analyses. *Lancet Neurol* 2015; 14: 263–273.

10. Noseworthy JH, Lucchinetti C, Rodriguez M, et al. Multiple sclerosis. *N Engl J Med* 2000; 343: 938–52.

11. Lublin FD, Reingold SC, Tiqwa P. Defining the clinical course of multiple sclerosis : *Neurology* 1996; 46: 907–911.

12. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 2014; 83: 278–86.

13. Miller DH, Chard DT, Ciccarelli O. Clinically isolated syndromes. *Lancet Neurol* 2012; 11: 157–69.

14. Miller D, Barkhof F, Montalban X, et al. Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. *Lancet Neurol* 2005; 4: 281–8.

15. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011; 69: 292–302.

16. Ransohoff RM, Hafler DA, Lucchinetti CF. Multiple sclerosis—a quiet revolution. *Nat Rev Neurol* 2015; 11: 134–142.

17. Tremlett H, Zhao Y, Rieckmann P, et al. New perspectives in the natural history of multiple sclerosis. *Neurology* 2010; 74: 2004–2015.

18. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983; 33: 1444–1452.

19. Rudick R a, Polman CH, Cohen J a, et al. Assessing disability progression with the Multiple Sclerosis Functional Composite. *Mult Scler* 2009; 15: 984–97.

20. Polman CH, Rudick RA. The multiple sclerosis functional composite: a clinically meaningful measure of disability. *Neurology* 2010; 74 Suppl 3: S8-15.

21. Matthews PM, Roncaroli F, Waldman A, et al. A practical review of the neuropathology and neuroimaging of multiple sclerosis. *Pract Neurol* 2016; practneurol-2016-001381.

22. Filippi M, Rocca MA, Barkhof F, et al. Association between pathological and MRI findings in multiple sclerosis. *Lancet Neurol* 2012; 11: 349–360.

23. Adams CW, Poston RN, Buk SJ, et al. Inflammatory vasculitis in multiple sclerosis. *J Neurol Sci* 1985; 69: 269–83.

24. Lucchinetti C. Pathological heterogeneity of idiopathic central nervous system inflammatory demyelinating disorders. *Curr Top Microbiol Immunol* 2008; 318: 19–43.

25. Brück W, Porada P, Poser S, et al. Monocyte/macrophage differentiation in early multiple sclerosis lesions. *Ann Neurol* 1995; 38: 788–96.

26. Patrikios P, Stadelmann C, Kutzelnigg A, et al. Remyelination is extensive in a subset of multiple sclerosis patients. *Brain* 2006; 129: 3165–72.

27. Bö L, Mörk S, Kong PA, et al. Detection of MHC class II-antigens on macrophages and microglia, but not on astrocytes and endothelia in active multiple sclerosis lesions. *J Neuroimmunol* 1994; 51: 135–46.

28. Bitsch A, Schuchardt J, Bunkowski S, et al. Acute axonal injury in multiple sclerosis. Correlation with demyelination and inflammation. *Brain* 2000; 123 (Pt 6: 1174–83.

29. Ferguson B, Matyszak MK, Esiri MM, et al. Axonal damage in acute multiple sclerosis lesions. *Brain* 1997; 120 (Pt 3: 393–9.

30. Trapp BD, Peterson J, Ransohoff RM, et al. Axonal transection in the lesions of multiple sclerosis. *N Engl J Med* 1998; 338: 278–285.

31. Trapp BD, Nave K-A. Multiple Sclerosis: An Immune or Neurodegenerative Disorder? *Annu Rev Neurosci* 2008; 31: 247–269.

32. Kornek B, Storch MK, Weissert R, et al. Multiple sclerosis and chronic autoimmune encephalomyelitis: a comparative quantitative study of axonal injury in active, inactive, and remyelinated lesions. *Am J Pathol* 2000; 157: 267–76.

33. Moll NM, Rietsch AM, Thomas S, et al. Multiple sclerosis normal-appearing white matter: pathology-imaging correlations. *Ann Neurol* 2011; 70: 764–73.

34. Kutzelnigg A, Lucchinetti CF, Stadelmann C, et al. Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain* 2005; 128: 2705–2712.

35. Evangelou N, Esiri MM, Smith S, et al. Quantitative pathological evidence for axonal loss in normal appearing white matter in multiple sclerosis. *Ann Neurol* 2000; 47: 391–395.

36. Evangelou N, Konz D, Esiri MM, et al. Regional axonal loss in the corpus callosum correlates with cerebral white matter lesion volume and distribution in multiple sclerosis. *Brain* 2000; 123 (Pt 9: 1845–1849.

37. BROWNELL B, HUGHES JT. The distribution of plaques in the cerebrum in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 1962; 25: 315–20.

38. Bø L, Vedeler CA, Nyland HI, et al. Subpial demyelination in the cerebral cortex of multiple sclerosis patients. *J Neuropathol Exp Neurol* 2003; 62: 723–32.

39. Huitinga I, De Groot CJ, Van der Valk P, et al. Hypothalamic lesions in multiple sclerosis. *J Neuropathol Exp Neurol* 2001; 60: 1208–18.

40. Vercellino M, Plano F, Votta B, et al. Grey Matter Pathology in Multiple Sclerosis. *J Neuropathol Exp Neurol* 2005; 64: 1101–1107.

41. Geurts JJG, Bö L, Roosendaal SD, et al. Extensive hippocampal demyelination in multiple sclerosis. *J Neuropathol Exp Neurol* 2007; 66: 819–827.

42. Gilmore CP, Donaldson I, Bö L, et al. Regional variations in the extent and pattern of grey matter demyelination in multiple sclerosis: a comparison between the cerebral cortex, cerebellar cortex, deep grey matter nuclei and the spinal cord. *J Neurol Neurosurg Psychiatry* 2009; 80: 182–7.

43. Peterson JW, Bö L, Mörk S, et al. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Ann Neurol* 2001; 50: 389–400.

44. Wegner C, Esiri MM, Chance SA, et al. Neocortical neuronal, synaptic, and glial loss in multiple sclerosis. *Neurology* 2006; 67: 960–7.

45. Kutzelnigg A, Lassmann H. Cortical demyelination in multiple sclerosis: a substrate for cognitive deficits? *J Neurol Sci* 2006; 245: 123–6.

46. Magliozzi R, Howell O, Vora A, et al. Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. *Brain* 2007; 130: 1089–1104.

47. Howell OW, Reeves CA, Nicholas R, et al. Meningeal inflammation is widespread and linked to cortical pathology in multiple sclerosis. *Brain* 2011; 134: 2755–2771.

48. Choi SR, Howell OW, Carassiti D, et al. Meningeal inflammation plays a role in the pathology of primary progressive multiple sclerosis. *Brain* 2012; 135: 2925–2937.

49. Lucchinetti CF, Popescu BFG, Bunyan RF, et al. Inflammatory cortical demyelination in early multiple sclerosis. *N Engl J Med* 2011; 365: 2188–97.

50. Absinta M, Vuolo L, Rao A, et al. Gadolinium-based MRI characterization of leptomeningeal inflammation in multiple sclerosis. *Neurology* 2015; 85: 18–28.

51. Nitz WR, Reimer P. Contrast mechanisms in MR imaging. *Eur Radiol* 1999; 9: 1032–46.

52. van Walderveen MA, Kamphorst W, Scheltens P, et al. Histopathologic correlate of hypointense lesions on T1-weighted spin-echo MRI in multiple sclerosis. *Neurology* 1998; 50: 1282–8.

53. van Walderveen MA, Barkhof F, Hommes OR, et al. Correlating MRI and clinical disease activity in multiple sclerosis: relevance of hypointense lesions on short-TR/short-TE (T1-weighted) spin-echo images. *Neurology* 1995; 45: 1684–90.

54. Truyen L, van Waesberghe JH, van Walderveen MA, et al. Accumulation of hypointense lesions ('black holes') on T1 spin-echo MRI correlates with disease progression in multiple sclerosis. *Neurology* 1996; 47: 1469–76.

55. van Walderveen MA, Lycklama A Nijeholt GJ, Adèr HJ, et al. Hypointense lesions on T1-weighted spin-echo magnetic resonance imaging: relation to clinical characteristics in subgroups of patients with multiple sclerosis. *Arch Neurol* 2001; 58: 76–81.

56. van Waesberghe JH, van Walderveen MA, Castelijns JA, et al. Patterns of lesion development in multiple sclerosis: longitudinal observations with T1-weighted spin-echo and magnetization transfer MR. *AJNR Am J Neuroradiol* 1998; 19: 675–83.

57. Brück W, Bitsch A, Kolenda H, et al. Inflammatory central nervous system demyelination: correlation of magnetic resonance imaging findings with lesion pathology. *Ann Neurol* 1997; 42: 783–93.

58. Katz D, Taubenberger JK, Cannella B, et al. Correlation between magnetic resonance imaging findings and lesion development in chronic, active multiple sclerosis. *Ann Neurol* 1993; 34: 661–669.

59. He J, Grossman RI, Ge Y, et al. Enhancing patterns in multiple sclerosis: evolution and persistence. *AJNR Am J Neuroradiol* 2001; 22: 664–9.

60. Bekiesińska-Figatowska M. T2-hyperintense foci on brain MR imaging. *Med Sci Monit*

2004; 10 Suppl 3: 80–7.

61. van Waesberghe JH, Kamphorst W, De Groot CJ, et al. Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability. *Ann Neurol* 1999; 46: 747–54.

62. Fisher E, Chang A, Fox RJ, et al. Imaging correlates of axonal swelling in chronic multiple sclerosis brains. *Ann Neurol* 2007; 62: 219–228.

63. Li DKB, Held U, Petkau J, et al. MRI T2 lesion burden in multiple sclerosis: a plateauing relationship with clinical disability. *Neurology* 2006; 66: 1384–9.

64. Bakshi R, Ariyaratana S, Benedict RH, et al. Fluid-attenuated inversion recovery magnetic resonance imaging detects cortical and juxtacortical multiple sclerosis lesions. *Arch Neurol* 2001; 58: 742–8.

65. Brusaferri F, Candelise L. Steroids for multiple sclerosis and optic neuritis: a meta-analysis of randomized controlled clinical trials. *J Neurol* 2000; 247: 435–42.

66. Comi G, Radaelli M. Oral corticosteroids for multiple sclerosis relapse. *Lancet* 2015; 386: 937–938.

67. Galea I, Ward-Abel N, Heesen C. Relapse in multiple sclerosis. *Bmj* 2015; 350: h1765–h1765.

68. Montalban X, Hauser SL, Kappos L, et al. Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N Engl J Med* 2017; 376: 209–220.

69. Dörr J, Paul F. The transition from first-line to second-line therapy in multiple sclerosis. *Curr Treat Options Neurol* 2015; 17: 354.

70. Hartung DM, Bourdette DN, Ahmed SM, et al. The cost of multiple sclerosis drugs in the US and the pharmaceutical industry. *Neurology* 2015; 84: 2185–2192.

71. Wingerchuk DM, Carter JL. Multiple sclerosis: Current and emerging disease-modifying therapies and treatment strategies. *Mayo Clin Proc* 2014; 89: 225–240.

72. Cross AH, Naismith RT. Established and novel disease-modifying treatments in multiple sclerosis. *J Intern Med* 2014; 275: 350–363.

73. Fassas A, Anagnostopoulos A, Kazis A, et al. Peripheral blood stem cell transplantation in the treatment of progressive multiple sclerosis: first results of a pilot study. *Bone Marrow Transplant* 1997; 20: 631–8.

74. Currò D, Mancardi G. Autologous hematopoietic stem cell transplantation in multiple sclerosis: 20 years of experience. *Neurol Sci* 2016; 37: 857–65.

75. Atkins H. Hematopoietic SCT for the treatment of multiple sclerosis. *Bone Marrow Transplant* 2010; 45: 1671–1681.

76. Mancardi G, Saccardi R. Autologous haematopoietic stem-cell transplantation in multiple sclerosis. *Lancet Neurol* 2008; 7: 626–636.

77. Saccardi R, Kozak T, Bocelli-Tyndall C, et al. Autologous stem cell transplantation for progressive multiple sclerosis: update of the European Group for Blood and Marrow Transplantation autoimmune diseases working party database. *Mult Scler* 2006; 12: 814–23.

78. Samijn JPA. Intense T cell depletion followed by autologous bone marrow transplantation for severe multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2006; 77: 46–50.

79. Nash RA, Bowen JD, McSweeney PA, et al. High-dose immunosuppressive therapy and autologous peripheral blood stem cell transplantation for severe multiple sclerosis. *Blood* 2003; 102: 2364–72.

80. Nash RA, Hutton GJ, Racke MK, et al. High-Dose Immunosuppressive Therapy and Autologous Hematopoietic Cell Transplantation for Relapsing-Remitting Multiple Sclerosis (HALT-MS). *JAMA Neurol* 2015; 72: 159.

81. Mancardi G, Sormani M, Di Gioia M, et al. Autologous haematopoietic stem cell transplantation with an intermediate intensity conditioning regimen in multiple sclerosis: the Italian multi-centre experience. *Mult Scler J* 2012; 18: 835–842.

82. Burman J, Iacobaeus E, Svenningsson A, et al. Autologous haematopoietic stem cell transplantation for aggressive multiple sclerosis: the Swedish experience. *J Neurol Neurosurg Psychiatry* 2014; 85: 1116–21.

83. Shevchenko JL, Kuznetsov AN, Ionova TI, et al. Autologous hematopoietic stem cell transplantation with reduced-intensity conditioning in multiple sclerosis. *Exp Hematol* 2012; 40: 892–8.

84. Krasulová E, Trneny M, Kozák T, et al. High-dose immunoablation with autologous haematopoietic stem cell transplantation in aggressive multiple sclerosis: a single centre 10-year experience. *Mult Scler* 2010; 16: 685–693.

85. Freedman MS, Atkins HL, Bowman M. Long-term outcome of the Canadian multiple sclerosis BMT study: efficacy and safety of treating aggressive multiple sclerosis with immunoablation and autologous stem cell transplantation. In: *29th Congress of the European Committee for Treatment and Research in Multiple Sclerosis*. Copenhagen, Denmark, 2013.

86. Burt RK, Balabanov R, Han X, et al. Association of nonmyeloablative hematopoietic stem cell transplantation with neurological disability in patients with relapsing-remitting multiple sclerosis. *JAMA* 2015; 313: 275–84.

87. Shevchenko JL, Kuznetsov AN, Ionova TI, et al. Long-term outcomes of autologous hematopoietic stem cell transplantation with reduced-intensity conditioning in multiple sclerosis: physician's and patient's perspectives. *Ann Hematol* 2015; 94: 1149–57.

88. Fassas a, Kimiskidis VK, Sakellari I, et al. Long-term results of stem cell transplantation for MS: a single-center experience. *Neurology* 2011; 76: 1066–70.

89. Mancardi GL, Saccardi R, Filippi M, et al. Autologous hematopoietic stem cell transplantation suppresses Gd-enhanced MRI activity in MS. *Neurology* 2001; 57: 62–68.

90. Atkins HL, Freedman MS. Hematopoietic Stem Cell Therapy for Multiple Sclerosis: Top 10 Lessons Learned. *Neurotherapeutics* 2012; 68–76.

91. Dietrich J, Han R, Yang Y, et al. CNS progenitor cells and oligodendrocytes are targets of chemotherapeutic agents in vitro and in vivo. *J Biol* 2006; 5: 22.

92. Mancardi GL, Sormani MP, Gualandi F, et al. Autologous hematopoietic stem cell transplantation in multiple sclerosis: a phase II trial. *Neurology* 2015; 84: 981–8.

93. Sormani MP, Muraro P. Updated views on autologous hematopoietic stem cell transplantation for treatment of multiple sclerosis. *Expert Rev Neurother* 2016; 7175: 1–2.

94. Saccardi R, Freedman MS, Sormani MP, et al. A prospective, randomized, controlled trial of autologous haematopoietic stem cell transplantation for aggressive multiple sclerosis: a position paper. *Mult Scler* 2012; 18: 825–34.

95. Atkins H, Freedman M. Immune Ablation Followed by Autologous Hematopoietic Stem Cell Transplantation for the Treatment of Poor Prognosis Multiple Sclerosis. In: Gordon D, Scolding NJ (eds) *Neural Cell Transplantation*. Humana Press, pp. 231–246.

96. De Stefano N, Stromillo ML, Giorgio A, et al. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2016; 87: 93–9.

97. Fisher E, Rudick RA, Cutter G, et al. Relationship between brain atrophy and disability: an 8-year follow-up study of multiple sclerosis patients. *Mult Scler* 2000; 6: 373–377.

98. De Stefano N, Giorgio A, Battaglini M, et al. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology* 2010; 74: 1868–76.

99. Fisher E, Lee J-C, Nakamura K, et al. Gray matter atrophy in multiple sclerosis: A longitudinal study. *Ann Neurol* 2008; 64: 255–265.

100. Dziedzic T, Metz I, Dallenga T, et al. Wallerian degeneration: a major component of early axonal pathology in multiple sclerosis. *Brain Pathol* 2010; 20: 976–85.

101. Kezele IB, Chen JT, Arnold DL, et al. The relation of focal white matter signal abnormality and focal volume loss in multiple sclerosis. *Mult Scler* 2007; 13: 809–813.

102. Sepulcre J, Goñi J, Masdeu JC, et al. Contribution of white matter lesions to gray matter atrophy in multiple sclerosis: evidence from voxel-based analysis of T1 lesions in the visual pathway. *Arch Neurol* 2009; 66: 173–9.

103. Mühlau M, Buck D, Förschler A, et al. White-matter lesions drive deep gray-matter atrophy in early multiple sclerosis: support from structural MRI. *Mult Scler* 2013; 19: 1485–92.

104. Bendfeldt K, Kuster P, Traud S, et al. Association of regional gray matter volume loss and

progression of white matter lesions in multiple sclerosis - A longitudinal voxel-based morphometry study. *Neuroimage* 2009; 45: 60–67.

105. Bendfeldt K, Blumhagen JO, Egger H, et al. Spatiotemporal distribution pattern of white matter lesion volumes and their association with regional grey matter volume reductions in relapsing-remitting multiple sclerosis. *Hum Brain Mapp* 2010; 31: 1542–55.

106. Fisniku LK, Chard DT, Jackson JS, et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann Neurol* 2008; 64: 247–54.

107. Chard DT, Griffin CM, Rashid W, et al. Progressive grey matter atrophy in clinically early relapsing-remitting multiple sclerosis. *Mult Scler* 2004; 10: 387–391.

108. Steenwijk MD, Geurts JJG, Daams M, et al. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain* 2015; 139: 115–126.

109. Geurts JJG, Calabrese M, Fisher E, et al. Measurement and clinical effect of grey matter pathology in multiple sclerosis. *Lancet Neurol* 2012; 11: 1082–1092.

110. Calabrese M, Rocca M a, Atzori M, et al. Cortical lesions in primary progressive multiple sclerosis: a 2-year longitudinal MR study. *Neurology* 2009; 72: 1330–6.

111. van de Pavert SHP, Muhlert N, Sethi V, et al. DIR-visible grey matter lesions and atrophy in multiple sclerosis: partners in crime? *J Neurol Neurosurg Psychiatry* 2015; 1–7.

112. Seewann A, Kooi E-J, Roosendaal SD, et al. Postmortem verification of MS cortical lesion detection with 3D DIR. *Neurology* 2012; 78: 302–8.

113. Simon B, Schmidt S, Lukas C, et al. Improved in vivo detection of cortical lesions in multiple sclerosis using double inversion recovery MR imaging at 3 Tesla. *Eur Radiol* 2010; 20: 1675–83.

114. Kilsdonk ID, Jonkman LE, Klaver R, et al. Increased cortical grey matter lesion detection in multiple sclerosis with 7 T MRI: a post-mortem verification study. *Brain* 2016; 139: 1472–81.

115. Vidal-Jordana A, Sastre-Garriga J, Pérez-Miralles F, et al. Early brain pseudoatrophy while on natalizumab therapy is due to white matter volume changes. *Mult Scler* 2013; 19:

1175–1181.

116. Vidal-Jordana A, Sastre-Garriga J, Pérez-Miralles F, et al. Brain Volume Loss During the First Year of Interferon-Beta Treatment in Multiple Sclerosis: Baseline Inflammation and Regional Brain Volume Dynamics. *J Neuroimaging* 2016; 26: 532–8.

117. Nakamura K, Guizard N, Fonov VS, et al. Jacobian integration method increases the statistical power to measure gray matter atrophy in multiple sclerosis. *NeuroImage Clin* 2014; 4: 10–7.

118. Losseff NA, Wang L, Lai HM, et al. Progressive cerebral atrophy in multiple sclerosis. A serial MRI study. *Brain* 1996; 119: 2009–2019.

119. Rudick RA, Fisher E, Lee JC, et al. Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS. Multiple Sclerosis Collaborative Research Group. *Neurology* 1999; 53: 1698–704.

120. Smith SM, Zhang Y, Jenkinson M, et al. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. *Neuroimage* 2002; 17: 479–489.

121. Freeborough P a, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Trans Med Imaging* 1997; 16: 623–629.

122. Smith SM, Rao A, De Stefano N, et al. Longitudinal and cross-sectional analysis of atrophy in Alzheimer's disease: cross-validation of BSI, SIENA and SIENAX. *Neuroimage* 2007; 36: 1200–6.

123. Ashburner J, Friston KJ. Voxel-Based Morphometry—The Methods. *Neuroimage* 2000; 11: 805–821.

124. Ashburner J, Friston KJ. Morphometry. *Hum brain Funct* 2003; 1–21.

125. Smith SM, De Stefano N, Jenkinson M, et al. Normalized accurate measurement of longitudinal brain change. *J Comput Assist Tomogr* 2001; 25: 466–475.

126. Freeborough PA, Woods RP, Fox NC. Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. *J Comput Assist*

Tomogr; 20: 1012–22.

127. Reuter M, Fischl B. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* 2011; 57: 19–21.

128. Nakamura K, Fisher E. Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *Neuroimage* 2009; 44: 769–76.

129. 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: 2062–2071.

130. Jernigan TL, Archibald SL, Fennema-Notestine C, et al. Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol Aging* 2001; 22: 581–94.

131. Ge Y, Grossman RI, Babb JS, et al. Age-related total gray matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis. *AJNR Am J Neuroradiol* 2002; 23: 1327–1333.

132. Sowell ER, Peterson BS, Thompson PM, et al. Mapping cortical change across the human life span. *Nat Neurosci* 2003; 6: 309–15.

133. Liu RSN, Lemieux L, Bell GS, et al. A longitudinal study of brain morphometrics using quantitative magnetic resonance imaging and difference image analysis. *Neuroimage* 2003; 20: 22–33.

134. Allen JS, Bruss J, Brown CK, et al. Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. *Neurobiol Aging* 2005; 26: 1245-60-82.

135. van Haren NEM, Hulshoff Pol HE, Schnack HG, et al. Progressive brain volume loss in schizophrenia over the course of the illness: evidence of maturational abnormalities in early adulthood. *Biol Psychiatry* 2008; 63: 106–13.

136. Fox NC, Jenkins R, Leary SM, et al. Progressive cerebral atrophy in MS: a serial study using registered, volumetric MRI. *Neurology* 2000; 54: 807–812.

137. Scahill RI, Frost C, Jenkins R, et al. A longitudinal study of brain volume changes in

normal aging using serial registered magnetic resonance imaging. *Arch Neurol* 2003; 60: 989–94.

138. Henley SMD, Frost C, MacManus DG, et al. Increased rate of whole-brain atrophy over 6 months in early Huntington disease. *Neurology* 2006; 67: 694–6.

139. Ge Y, Grossman RI, Udupa JK, et al. Brain atrophy in relapsing-remitting multiple sclerosis and secondary progressive multiple sclerosis: longitudinal quantitative analysis. *Radiology* 2000; 214: 665–670.

140. Lin X, Blumhardt LD, Constantinescu CS. The relationship of brain and cervical cord volume to disability in clinical subtypes of multiple sclerosis: a three-dimensional MRI study. *Acta Neurol Scand* 2003; 108: 401–6.

141. Tedeschi G, Lavorgna L, Russo P, et al. Brain atrophy and lesion load in a large population of patients with multiple sclerosis. *Neurology* 2005; 65: 280–5.

142. Pagani E, Rocca M a., Gallo A, et al. Regional brain atrophy evolves differently in patients with multiple sclerosis according to clinical phenotype. *Am J Neuroradiol* 2005; 26: 341–346.

143. Horakova D, Cox JL, Havrdova E, et al. Evolution of different MRI measures in patients with active relapsing-remitting multiple sclerosis over 2 and 5 years: a case-control study. *J Neurol Neurosurg Psychiatry* 2008; 79: 407–14.

144. Chu R, Tauhid S, Glanz BI, et al. Whole Brain Volume Measured from 1.5T versus 3T MRI in Healthy Subjects and Patients with Multiple Sclerosis. *J Neuroimaging* 2015; 26: 62–7.

145. Sormani MP, Rovaris M, Valsasina P, et al. Measurement error of two different techniques for brain atrophy assessment in multiple sclerosis. *Neurology* 2004; 62: 1432–4.

146. Jasperse B, Minneboo A, de Groot V, et al. Determinants of cerebral atrophy rate at the time of diagnosis of multiple sclerosis. *Arch Neurol* 2007; 64: 190–4.

147. Anderson VM, Bartlett JW, Fox NC, et al. Detecting treatment effects on brain atrophy in

relapsing remitting multiple sclerosis: sample size estimates. *J Neurol* 2007; 254: 1588–94.

148. Kappos L, Radue E-W, O'Connor P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010; 362: 387–401.

149. Di Filippo M, Anderson VM, Altmann DR, et al. Brain atrophy and lesion load measures over 1 year relate to clinical status after 6 years in patients with clinically isolated syndromes. *J Neurol Neurosurg Psychiatry* 2010; 81: 204–8.

150. Barkhof F, Hulst HE, Drulovic J, et al. Ibudilast in relapsing-remitting multiple sclerosis: a neuroprotectant? *Neurology* 2010; 74: 1033–40.

151. Popescu V, Agosta F, Hulst HE, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2013; 84: 1082–1091.

152. Miller DH, Soon D, Fernando KT, et al. MRI outcomes in a placebo-controlled trial of natalizumab in relapsing MS. *Neurology* 2007; 68: 1390–1401.

153. Fisher E, Nakamura K, Lee J-C, et al. Effect of intramuscular interferon beta-1a on gray matter atrophy in relapsing-remitting multiple sclerosis: A retrospective analysis. *Mult Scler* 2015; 1–9.

154. Vercellino M, Masera S, Lorenzatti M, et al. Demyelination, inflammation, and neurodegeneration in multiple sclerosis deep gray matter. *J Neuropathol Exp Neurol* 2009; 68: 489–502.

155. Duning T, Kloska S, Steinsträter O, et al. Dehydration confounds the assessment of brain atrophy. *Neurology* 2005; 64: 548–550.

156. Nakamura K, Brown R a, Araujo D, et al. Correlation between brain volume change and T2 relaxation time induced by dehydration and rehydration: Implications for monitoring atrophy in clinical studies. *NeuroImage Clin* 2014; 6: 166–70.

157. Lassmann H, van Horssen J, Mahad D. Progressive multiple sclerosis: pathology and pathogenesis. *Nat Rev Neurol* 2012; 8: 647–656.

158. Dwyer MG, Zivadinov R, Tao Y, et al. Immunological and short-term brain volume

changes in relapsing forms of multiple sclerosis treated with interferon beta-1a subcutaneously three times weekly: an open-label two-arm trial. *BMC Neurol* 2015; 15: 232.

159. Petzold A, Mondria T, Kuhle J, et al. Evidence for acute neurotoxicity after chemotherapy. *Ann Neurol* 2010; 68: 806–815.

160. Molyneux PD, Kappos L, Polman C, et al. The effect of interferon beta-1b treatment on MRI measures of cerebral atrophy in secondary progressive multiple sclerosis. *Brain* 2000; 123: 2256–2263.

161. Rovaris M, Comi G, Rocca M a, et al. Short-term brain volume change in relapsing-remitting multiple sclerosis: Effect of glatiramer acetate and implications. *Brain* 2001; 124: 1803–1812.

162. Frank JA, Richert N, Bash C, et al. Interferon-??-1b slows progression of atrophy in RRMS: Three-year follow-up in NAb- and NAb+ patients. *Neurology* 2004; 62: 719–725.

163. Hommes OR, Sørensen PS, Fazekas F, et al. Intravenous immunoglobulin in secondary progressive multiple sclerosis: randomised placebo-controlled trial. *Lancet (London, England)* 2004; 364: 1149–56.

164. Hardmeier M, Wagenpfeil S, Freitag P, et al. Rate of brain atrophy in relapsing MS decreases during treatment with IFNbeta-1a. *Neurology* 2005; 64: 236–240.

165. Mikol DD, Barkhof F, Chang P, et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. *Lancet Neurol* 2008; 7: 903–914.

166. O'Connor P, Filippi M, Arnason B, et al. 250 ??G or 500 ??G Interferon Beta-1B Versus 20 Mg Glatiramer Acetate in Relapsing-Remitting Multiple Sclerosis: a Prospective, Randomised, Multicentre Study. *Lancet Neurol* 2009; 8: 889–897.

167. Montalban X, Sastre-Garriga J, Tintoré M, et al. A single-center, randomized, double-blind, placebo-controlled study of interferon beta-1b on primary progressive and transitional multiple sclerosis. *Mult Scler* 2009; 15: 1195–1205.

168. Radue EW, Stuart WH, Calabresi PA, et al. Natalizumab plus interferon beta-1a reduces lesion formation in relapsing multiple sclerosis. *J Neurol Sci* 2010; 292: 28–35.

169. Rinne ML, Lee EQ, Wen PY. Central Nervous System Complications of Cancer Therapy. *J Support Oncol* 2012; 10: 133–141.

170. McDonald BC, Conroy SK, Ahles TA, et al. Gray matter reduction associated with systemic chemotherapy for breast cancer: a prospective MRI study. *Breast Cancer Res Treat* 2010; 123: 819–828.

171. Inagaki M, Yoshikawa E, Matsuoka Y, et al. Smaller regional volumes of brain gray and white matter demonstrated in breast cancer survivors exposed to adjuvant chemotherapy. *Cancer* 2007; 109: 146–156.

172. Koppelmans V, de Ruiter MB, van der Lijn F, et al. Global and focal brain volume in long-term breast cancer survivors exposed to adjuvant chemotherapy. *Breast Cancer Res Treat* 2012; 132: 1099–106.

173. De Ruiter MB, Reneman L, Boogerd W, et al. Late effects of high-dose adjuvant chemotherapy on white and gray matter in breast cancer survivors: Converging results from multimodal magnetic resonance imaging. *Hum Brain Mapp* 2012; 33: 2971–2983.

174. Inglese M, Mancardi GL, Pagani E, et al. Brain tissue loss occurs after suppression of enhancement in patients with multiple sclerosis treated with autologous haematopoietic stem cell transplantation. *J Neurol Neurosurg Psychiatry* 2004; 75: 643–4.

175. Chen JT, Collins DL, Atkins HL, et al. Brain atrophy after immunoablation and stem cell transplantation in multiple sclerosis. *Neurology* 2006; 66: 1935–1937.

176. Rocca MA, Mondria T, Valsasina P, et al. A Three-Year Study of Brain Atrophy after Autologous Hematopoietic Stem Cell Transplantation in Rapidly Evolving Secondary Progressive Multiple Sclerosis. *Am J Neuroradiol* 2007; 28: 1659–1661.

177. Roccatagliata L, Rocca M, Valsasina P, et al. The long-term effect of AHSCT on MRI measures of MS evolution: a five-year follow-up study. *Mult Scler* 2007; 13: 1068–70.

178. Preboske GM, Gunter JL, Ward CP, et al. Common MRI acquisition non-idealities

significantly impact the output of the boundary shift integral method of measuring brain atrophy on serial MRI. *Neuroimage* 2006; 30: 1196–1202.

179. Han X, Jovicich J, Salat D, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *Neuroimage* 2006; 32: 180–194.

180. Jovicich J, Czanner S, Han X, et al. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage* 2009; 46: 177–192.

181. Stonnington CM, Tan G, Kl??ppel S, et al. Interpreting scan data acquired from multiple scanners: A study with Alzheimer's disease. *Neuroimage* 2008; 39: 1180–1185.

182. Kruggel F, Turner J, Muftuler LT, et al. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. *Neuroimage* 2010; 49: 2123–2133.

183. Fennema-Notestine C, Gamst AC, Quinn BT, et al. Feasibility of multi-site clinical structural neuroimaging studies of aging using legacy data. *Neuroinformatics* 2007; 5: 235–245.

184. Bernal-Rusiel JL, Greve DN, Reuter M, et al. Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. *Neuroimage* 2013; 66: 249–260.

185. Jones BC, Nair G, Shea CD, et al. Quantification of multiple-sclerosis-related brain atrophy in two heterogeneous MRI datasets using mixed-effects modeling. *NeuroImage Clin* 2013; 3: 171–179.

186. Chua AS, Egorova S, Anderson MC, et al. Handling changes in MRI acquisition parameters in modeling whole brain lesion volume and atrophy data in multiple sclerosis subjects: Comparison of linear mixed-effect models. *NeuroImage Clin* 2015; 8: 606–610.

187. Zeger SL, Liang KY. An Overview of Methods for the Analysis of Longitudinal Data. *Stat Med* 1992; 11: 1825–1839.

188. Fitzmaurice GM, Ravichandran C. A primer in longitudinal data analysis. *Circulation* 2008; 118: 2005–10.

189. Pinheiro JC, Bates DM. *Mixed-effects models in S and S-PLUS*. New York: Springer New York, 2000.

190. Frischer JM, Bramow S, Dal-Bianco A, et al. The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain* 2009; 132: 1175–1189.

191. Comi G, Radaelli M, Soelberg Sørensen P. Evolving concepts in the treatment of relapsing multiple sclerosis. *Lancet (London, England)* 2017; 389: 1347–1356.

192. Wingerchuk DM, Weinshenker BG. Disease modifying therapies for relapsing multiple sclerosis. *BMJ* 2016; 354: i3518.

193. Yong HY, Yong VW. Stop inflammation and you stop neurodegeneration in MS - YES. *Mult Scler* 2017; 8: 1352458517707266.

194. Kalincik T. Stop inflammation and you stop neurodegeneration in MS - NO. *Mult Scler* 2017; 8: 1352458517707267.

195. Atkins HL, Bowman M, Allan D, et al. Immunoablation and autologous haemopoietic stem-cell transplantation for aggressive multiple sclerosis: a multicentre single-group phase 2 trial. *Lancet (London, England)* 2016; 388: 576–85.

196. Wiebe VJ, Smith BR, DeGregorio MW, et al. Pharmacology of agents used in bone marrow transplant conditioning regimens. *Crit Rev Oncol Hematol* 1992; 13: 241–270.

197. Dietrich J, Monje M, Wefel J, et al. Clinical patterns and biological correlates of cognitive dysfunction associated with cancer therapy. *Oncologist* 2008; 13: 1285–1295.

198. McDonald BC, Conroy SK, Smith DJ, et al. Frontal gray matter reduction after breast cancer chemotherapy and association with executive symptoms: A replication and extension study. *Brain Behav Immun* 2013; 30: S117–S125.

199. Han R, Yang YM, Dietrich J, et al. Systemic 5-fluorouracil treatment causes a syndrome of delayed myelin destruction in the central nervous system. *J Biol* 2008; 7: 12.

200. Fu L, Matthews PM, De Stefano N, et al. Imaging axonal damage of normal-appearing

white matter in multiple sclerosis. *Brain* 1998; 121: 103–13.

201. Vargas ME, Barres BA. Why Is Wallerian Degeneration in the CNS So Slow? *Annu Rev Neurosci* 2007; 30: 153–179.

202. De Stefano N, Airas L, Grigoriadis N, et al. Clinical relevance of brain volume measures in multiple sclerosis. *CNS Drugs* 2014; 28: 147–56.

203. Roosendaal SD, Bendfeldt K, Vrenken H, et al. Grey matter volume in a large cohort of MS patients: relation to MRI parameters and disability. *Mult Scler* 2011; 17: 1098–106.

204. Dutta R, McDonough J, Yin X, et al. Mitochondrial dysfunction as a cause of axonal degeneration in multiple sclerosis patients. *Ann Neurol* 2006; 59: 478–89.

205. Klaver R, Popescu V, Voorn P, et al. Neuronal and axonal loss in normal-appearing gray matter and subpial lesions in multiple sclerosis. *J Neuropathol Exp Neurol* 2015; 74: 453–8.

206. Lee H, Narayanan S, Brown RA, et al. Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis. *Mult Scler* 2017; 23: 420–431.

207. Popescu V, Schoonheim MM, Versteeg A, et al. Grey Matter Atrophy in Multiple Sclerosis: Clinical Interpretation Depends on Choice of Analysis Method. *PLoS One* 2016; 11: e0143942.

208. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp* 2002; 17: 143–155.

209. Dietrich J, Prust M, Kaiser J. Chemotherapy, cognitive impairment and hippocampal toxicity. *Neuroscience* 2015; 309: 224–32.

210. Scherling CS, Smith A. Opening up the window into and 'chemobrain': A neuroimaging review. *Sensors (Switzerland)* 2013; 13: 3169–3203.

211. Jindahra P, Petrie A, Plant GT. The time course of retrograde trans-synaptic degeneration following occipital lobe damage in humans. *Brain* 2012; 135: 534–541.

212. Steenwijk MD, Daams M, Pouwels PJW, et al. What Explains Gray Matter Atrophy in Long-standing Multiple Sclerosis? *Radiology* 2014; 272: 132708.

213. Calabrese M, Magliozzi R, Ciccarelli O, et al. Exploring the origins of grey matter damage in multiple sclerosis. *Nat Rev Neurosci* 2015; 16: 147–158.

214. Bermel R a, Innus MD, Tjoa CW, et al. Selective caudate atrophy in multiple sclerosis: a 3D MRI parcellation study. *Neuroreport* 2003; 14: 335–9.

215. Houtchens MK, Benedict RHB, Killiany R, et al. Thalamic atrophy and cognition in multiple sclerosis. *Neurology* 2007; 69: 1213–23.

216. Takao H, Abe O, Hayashi N, et al. Effects of gradient non-linearity correction and intensity non-uniformity correction in longitudinal studies using structural image evaluation using normalization of atrophy (SIENA). *J Magn Reson Imaging* 2010; 32: 489–92.

217. Krueger G, Granziera C, Jack CR, et al. Effects of MRI scan acceleration on brain volume measurement consistency. *J Magn Reson Imaging* 2012; 36: 1234–40.

218. Muraro PA, Pasquini M, Atkins HL, et al. Long-term Outcomes After Autologous Hematopoietic Stem Cell Transplantation for Multiple Sclerosis. *JAMA Neurol* 2017; 74: 459–469.

219. Sormani MP, Arnold DL, De Stefano N. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. *Ann Neurol* 2014; 75: 43–9.

220. Nash RA, Hutton GJ, Racke MK, et al. High-dose immunosuppressive therapy and autologous HCT for relapsing-remitting MS. *Neurology* 2017; 0: 1–12.

221. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005; 26: 839–51.

222. Cheng T, Forsyth P, Chaudhry A, et al. High-dose thiotepa, busulfan, cyclophosphamide and ASCT without whole-brain radiotherapy for poor prognosis primary CNS lymphoma. *Bone Marrow Transplant* 2003; 31: 679–85.

223. Lee H, Nakamura K, Narayanan S, et al. Impact of immunoablation and autologous hematopoietic stem cell transplantation on grey and white matter atrophy in multiple sclerosis. *Mult Scler J*; in press.

224. Jack CR, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in

Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol* 2013; 12: 207–16.

225. Leung KK, Bartlett JW, Barnes J, et al. Cerebral atrophy in mild cognitive impairment and Alzheimer disease: Rates and acceleration. *Neurology* 2013; 80: 648–654.

226. Evans MC, Barnes J, Nielsen C, et al. Volume changes in Alzheimer's disease and mild cognitive impairment: cognitive associations. *Eur Radiol* 2010; 20: 674–82.

227. Frisoni GB, Fox NC, Jack CR, et al. The clinical use of structural MRI in Alzheimer disease. *Nat Rev Neurol* 2010; 6: 67–77.

228. Caramanos Z, Fonov VS, Francis SJ, et al. Gradient distortions in MRI: characterizing and correcting for their effects on SIENA-generated measures of brain volume change. *Neuroimage* 2010; 49: 1601–11.

229. Nakamura K, Guizard N, Fonov VS, et al. MRI-based Simulation of Central Brain Atrophy for Evaluation of Brain Atrophy Measurement Methods. In: *Proceedings of the International Society for Magnetic Resonance in Medicine*. 2014, p. 4314.

230. Jack CR, Barnes J, Bernstein MA, et al. Magnetic resonance imaging in Alzheimer's Disease Neuroimaging Initiative 2. *Alzheimers Dement* 2015; 11: 740–56.

231. Wyman BT, Harvey DJ, Crawford K, et al. Standardization of analysis sets for reporting results from ADNI MRI data. *Alzheimers Dement* 2013; 9: 332–7.

232. Jack CR, Bernstein MA, Borowski BJ, et al. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. *Alzheimers Dement* 2010; 6: 212–20.

233. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 1998; 17: 87–97.

234. Fonov V, Evans A, McKinstry R, et al. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage* 2009; 47: S102.

235. Eskildsen SF, Coupé P, Fonov V, et al. BEaST: brain extraction based on nonlocal segmentation technique. *Neuroimage* 2012; 59: 2362–73.

236. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 2004; 23 Suppl 1: S208-19.

237. Nakamura K, Brown RA, Narayanan S, et al. Diurnal fluctuations in brain volume: Statistical analyses of MRI from large populations. *Neuroimage* 2015; 118: 126–32.

238. Lin C, Watson RE, Ward HA, et al. MP-RAGE compared to 3D IR SPGR for Optimal T1 Contrast and Image Quality in the Brain at 3T. In: *Proc. Intl. Sco. Mag. Reson. Med.* 2006, p. 981.

239. Coles AJ, Twyman CL, Arnold DL, et al. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. *Lancet (London, England)* 2012; 380: 1829–39.

240. Racke MK, Imitola J. Selection of Patients With Multiple Sclerosis to Undergo Autologous Hematopoietic Stem Cell Transplantation. *JAMA Neurol* 2017; 74: 392–394.

241. Roosendaal S, Moraal B, Pouwels P, et al. Accumulation of cortical lesions in MS: relation with cognitive impairment. *Mult Scler* 2009; 15: 708–714.

242. Calabrese M, Poretto V, Favaretto A, et al. Cortical lesion load associates with progression of disability in multiple sclerosis. *Brain* 2012; 135: 2952–61.