

Diplomarbeit

**Änderungen des zerebralen Volumens bei Multipler
Sklerose**

**Vergleich visuelles Rating und computerunterstützte quantitative
Analyse**

eingereicht von

Johanna Leitgeb

zur Erlangung des akademischen Grades

Doktorin der gesamten Heilkunde

(Dr. med. univ.)

an der

Medizinischen Universität Graz

ausgeführt an der

Universitätsklinik für Neurologie

unter der Anleitung von

Assoz. Prof. Priv.-Doz. Dr. med. univ. MBA Christian Enzinger

und

Priv.- Doz. Dr. med. univ. et scient. med. Alexander Pichler

Graz, 23.09.2019

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am 23.09.2019

Johanna Leitgeb eh

Danksagungen

An dieser Stelle möchte ich allen danken, die mich bei meiner Diplomarbeit unterstützt haben.

Ein besonderer Dank gilt meinen beiden Betreuern Assoz. Prof. Priv.-Doz. Dr. med. univ. MBA Christian Enzinger und Priv.- Doz. Dr. med. univ. et scient. med. Alexander Pichler, für die Möglichkeit an der Universitätsklinik für Neurologie meine Diplomarbeit verfassen zu können, für ihre nette Betreuung und die Korrektur der Arbeit.

Herzlichen Dank auch meinen Eltern, ohne deren Hilfe und Unterstützung ich mein Studium nicht hätte absolvieren können. Auch meinen Freunden möchte ich danken, die mir während meines Studiums stets beigestanden sind. Danke für die schöne Zeit!

Zusammenfassung

Hintergrund

Die Abnahme von Hirnvolumen bei PatientInnen mit Multipler Sklerose (MS) wird als potenzieller prognostischer Faktor für Verlauf und Therapieerfolg diskutiert. Visuelle Methoden zur Quantifizierung von Hirnvolumenänderungen wurden bisher jedoch kaum in Betracht gezogen.

Zielstellung

In dieser Diplomarbeit wurde anhand einer semiquantitativen Rating Skala die Sensitivität der visuellen Quantifizierung von Hirnvolumenänderung bei MS PatientInnen geprüft und mit Software-basierten Messmethoden verglichen. Zudem wurde der Frage nach einem „Cut-off“ zur Unterscheidung zwischen physiologischer und pathologischer Abnahme von Hirnvolumen nachgegangen.

Methoden

Die Daten der 120 PatientInnen stammen aus einer Studie an der Abteilung für Allgemeine Neurologie der Medizinischen Universität Graz (Pichler *et al.*, 2016). Das verwendete MRT Gerät war ein Siemens 3-Tesla Scanner. Das Hirnvolumen wurde mit Hilfe der Software-Programme SIENA und SIENAX der FMRIB Software Library gemessen. Das visuelle Rating wurde von Seiten eines Neurologen und einer Medizinstudentin durchgeführt.

Resultate

Der Untersuchungszeitpunkt konnte in 75% richtig, in 11% falsch und in 14% nicht zugeordnet werden. Das Interrater-Agreement zeigte eine schwache ($k = 0.36$), das Intrarater-Agreement eine sehr gute Übereinstimmung ($k = 1$). Die durchschnittliche Änderung des Hirnvolumens pro Jahr (pBVC) der richtigen Zuordnungen lag bei -0.45% ($-0.74, -0.2$), das pBVC der nicht zuordenbaren Scans bei -0.33% ($-0.46, -0.19$). Es erfolgte eine Gruppierung der MRT-Untersuchungen je nach pBVC pro Jahr in vier Quartile: die Scans mit einer jährlichen pBVC zwischen -2.10% und -0.65% wurden in 96.7% visuell richtig zugeordnet, bei den Scans in der Gruppe mit einer Hirnvolumenabnahme zwischen -0.64% und -0.42% war die Zuordnung in 76.7% der Fälle richtig. Außer der Dauer des Scanintervalls konnte für keine der klinischen Variablen ein statistisch signifikanter Unterschied zwischen richtig, falsch und nicht zuordenbaren Scans gefunden werden.

Interpretation

Die Ergebnisse dieser Arbeit weisen darauf hin, dass das visuelle Rating anhand von semiquantitativen Rating Skalen neben den computerbasierten Methoden eine einfache, kostengünstige und schnelle Möglichkeit bietet, MRT Untersuchungen von MS PatientInnen auszuwerten und dabei zumindest pathologische Hirnatrophie zu erkennen und zu quantifizieren.

Abstract

Background

Brain volume loss in patients with multiple sclerosis (MS) has been proposed as an important prognostic factor for disease progression and therapy effects. However, there are no actual studies about the use of visual rating scales for quantification of brain volume loss in MS.

Objective

This study aimed to determine the sensitivity of a semiquantitative rating scale for quantification of brain volume change in patients with MS. Therefore, we compared the results of the visual analysis with those obtained by software-based methods. In addition, we tried to establish a possible cut-off value to distinguish visually pathological from physiological brain volume loss.

Methods

We reanalyzed the data used in a previous study by the neurological department of the Medical University of Graz (Pichler *et al.*, 2016). We evaluated MRI scans of 120 patients acquired using a Siemens 3-Tesla Scanner. The brain volume measures were calculated using the software programs SIENA and SIENAX from the FMRIB Software Library. The visual rating was performed by a neurologist and a medical student.

Results

We were able to classify baseline and follow-up for 75% of the scan-pairs correctly, 14% were incorrectly and in 11% of the cases it was not possible to see any difference between the two scans. The strength of interrater agreement for the assignments was fair ($k = 0.36$), for the intrarater agreement the strength was very good ($k = 1$). The mean percentage of annual brain volume change (pBVC) for the correct assignments was -0.45% ($-0.74, -0.2$). The annual pBVC for the scans without visible difference was -0.33% ($-0.46, -0.19$). To identify a possible cut-off for correct assignments we used quartiles of annual pBVC: in the quartile with the highest range of annual pBVC (from -2.10% to -0.65%) 96.7% of the scan pairs were assigned correctly. In the quartile with an annual pBVC between -0.64% and -0.42% the assignment of 76.7% of the scan pairs was correct. Only the interval between baseline and follow-up MRI showed a significant difference between correctly, wrong and not assigned MRI scans.

Conclusion

Visual rating using semiquantitative rating scales could offer an easy, cheap and quick alternative to computer-based methods in analyzing MRI scans of patients with MS. It is likely that at least pathological brain atrophy can be identified and quantified.

Inhaltsverzeichnis

Danksagungen	III
Zusammenfassung.....	IV
Abstract.....	VI
Inhaltsverzeichnis.....	VIII
Abkürzungen und Glossar	X
Abbildungsverzeichnis	XI
Tabellenverzeichnis.....	XII
1 Einleitung	1
1.1 Epidemiologie	2
1.2 Ätiologie.....	3
1.2.1 Genetik	3
1.2.2 Umweltfaktoren.....	5
1.3 Pathogenese.....	6
1.4 Verlaufsformen.....	7
1.5 Symptome.....	8
1.6 Diagnose.....	11
1.7 Therapie.....	13
1.7.1 Therapiewahl	15
1.7.2 Monitoring.....	16
1.8 Prognose	16
1.9 Hirnatrophie bei Multipler Sklerose.....	18
1.9.1 Definition.....	18
1.9.2 Ursachen	19
1.9.3 Lokalisation	20
1.9.4 Klinische Relevanz.....	22
1.9.5 Hirnatrophie in der MRT	23
1.9.6 Computerbasierte Messung von Hirnvolumen und Hirnatrophie	24
1.9.7 Messfehler und Störfaktoren.....	26
1.9.8 Visuelles Rating von Hirnatrophie	27
1.10 Fragestellung und Zielsetzung.....	28
2 Material und Methoden	29
2.1 Patientinnen und Patienten	29
2.2 MRT-Protokoll.....	29
2.3 Visuelles Rating	30
2.4 SIENA-Auswertung	33
2.5 Statistik.....	33
3 Ergebnisse	34
3.1 Demographische und klinische Daten der Population.....	34
3.2 Ergebnisse des visuellen Atrophie-Ratings	37
3.2.1 Inter-Rater Variability	40

3.2.2	Intra-Rater Variability	42
3.3	Vergleich visuelles Rating und Ergebnisse aus SIENA/SIENAX.....	42
3.4	Klinische Daten und visuelles Rating.....	50
4	Diskussion.....	53
4.1	Das visuelle Rating.....	53
4.2	Visuelles Rating und SIENA/SIENAX	55
4.3	Visuelles Rating im klinischen Alltag	58
4.4	Limitationen	60
4.5	Zusammenfassung	60
5	Literaturverzeichnis	61
6	Anhang	68
6.1	Rating Sheet	68
6.2	SIFAP-Rating Skala	69

Abkürzungen und Glossar

annual pBVC	prozentuelle Abnahme von Gehirnvolumen pro Jahr
BI	Baseline Scan (erster MRT-Scan)
CD4⁺	Lymphozyten: Cluster of Differentiation
CIS	klinisch isoliertes Syndrom („clinically isolated syndrome“)
Converter	PatientInnen die zwischen Baseline und Follow-up Scan eine definitive MS entwickelt haben
DIS	Dissemination im Raum
DIT	Dissemination in der Zeit
DMT	krankheitsmodifizierende Therapie (disease modifying therapy)
EBV	Epstein-Barr-Virus
EDSS	Extended disability status scale
FAST	FMRIB's Automated Segmentation Tool
FLAIR	fluid-attenuated inversion recovery
FLIRT	FMRIB's Linear Image Registration Tool
FMRIB	Functional MRI of the Brain
FS	Funktionssysteme
FSL	FMRIB Software Library
Fu	Follow-up Scan (zweiter MRT-Scan)
LADIS	Leukoaraiosis and Disability Study Group
MHC	Major Histability Complex
MPRAGE	magnetization-prepared rapid gradient echo
MRT	Magnetresonanztomographie
MS	Multiple Sklerose
MSFC	Multiple Sclerosis Functional Composite
MSSS	Multiple Sclerosis Severity Score
NBV	normalisiertes Hirnvolumen (normalized brain volume)
NEDA	keine klinische oder in der MRT sichtbare Krankheitsprogression (no evidence of disease activity)
pBVC	prozentuelle Abnahme von Gehirnvolumen (percentage of brain volume change)
PPMS	primär progrediente Multiple Sklerose (primary progressive multiple sclerosis)
RRMS	schubförmig-remittierende Multiple Sklerose („relapsing-remitting multiple sclerosis“)
SIENA	Structural Image Evaluation, Using Normalization, of Atrophy
SIENAX	Structural Image Evaluation, Using Normalization, of Atrophy Cross-Sectional
SIFAP	Stroke in Young Fabry Patients
SPMS	sekundär progrediente Multiple Sklerose (secondary progressive multiple sclerosis)
TE	Echozeit (echo time)
TR	Repetitionszeit (repetition time)
ZNS	zentrales Nervensystem

Abbildungsverzeichnis

Abbildung 1: Stufentherapie der Multiple Sklerose	14
Abbildung 2: T1-gewichteter MRT-Scan eines ausgewählten Patienten bzw. einer ausgewählten Patientin in axialer Schnittebene. Die Pfeile markieren die Höhe der Ausmessung des 2. Ventrikels an der dünnsten Stelle des Septum pellucidum (Doppelpfeil) und des gesamten Gehirns auf derselben Höhe (rechter und linker Pfeil)	31
Abbildung 3: T1-gewichteter MRT-Scan eines ausgewählten Patienten bzw. einer ausgewählten Patientin in axialer Schnittebene. Die Pfeile markieren die Höhe der Ausmessung des 3. Ventrikels am anteroposterioren Mittelpunkt.	32
Abbildung 4: Die Grafik vergleicht die Anzahl der richtig, falsch und nicht zuordenbaren Scan-Paare aus den Auswertungen von Rater 1 und Rater 2	41
Abbildung 5: Die Boxplots zeigen die prozentuelle Änderung des Hirnvolumens über den gesamten Untersuchungszeitraum der richtig, falsch und nicht-zuordenbaren MRT-Untersuchungen	43
Abbildung 6: Die Boxplots zeigen das NBV der richtig zugeordneten Baseline Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die ventrikuläre Weite.....	48
Abbildung 7: Die Boxplots zeigen das NBV der richtig zugeordneten Baseline Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die sulcale Weite	48
Abbildung 8: Die Boxplots zeigen das NBV der richtig zugeordneten Follow-up Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die ventrikuläre Weite.....	49
Abbildung 9: Die Boxplots zeigen das NBV der richtig zugeordneten Follow-up Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die sulcale Weite	49

Tabellenverzeichnis

Tabelle 1: Häufige Symptome der MS	9
Tabelle 2: Expanded Disability Status Scale (EDSS)	10
Tabelle 3: Diagnosekriterien DIS und DIT	12
Tabelle 4: McDonald Kriterien für die Diagnose einer MS bei PatientInnen mit einem Schub bei Erkrankungsbeginn	12
Tabelle 5: Diagnosekriterien für die primär progrediente MS (PPMS)	13
Tabelle 6: Demographische und klinische Daten	35
Tabelle 7: Verteilung Volumenmaße SIENA/SIENAX.....	36
Tabelle 8: Ergebnis Zuordnung zum Untersuchungszeitpunkt	37
Tabelle 9: Ergebnisse des visuellen Ratings	38
Tabelle 10: Interrater Agreement	41
Tabelle 11: Intrarater Agreement	42
Tabelle 12: pBVC über den gesamten Untersuchungszeitraum und pBVC pro Jahr bei den richtig, falsch und nicht zuordenbaren MRT-Scans	43
Tabelle 13: Zuordnung zum Untersuchungszeitpunkt innerhalb der Quartile der pBVC pro Jahr	44
Tabelle 14: Verteilung des „Sicherheitsgrads der Zuordnung“ der richtig zugeordneten Scans innerhalb der Quartile der pBVC pro Jahr	44
Tabelle 15: Zuordnung zum Untersuchungszeitpunkt zwischen den Quartilen Q1+Q2 und Q3+Q4 der pBVC pro Jahr.....	45
Tabelle 16: NBV (normalized brain volume, normalisiertes Hirnvolumen) innerhalb der visuellen Rating Kategorien für die ventrikuläre und sulcale Weite der richtig zugeordneten Scans (n = 90)	46
Tabelle 17: Zuordnung zum Untersuchungszeitpunkt in Bezug auf klinische und demographische Variablen	51
Tabelle 18: Sicherheitsgrad der Zuordnungen der richtig zugeordneten Scans (n = 90) in Bezug auf klinische und demographische Variablen	52

1 Einleitung

Multiple Sklerose (MS) ist eine chronisch entzündliche, neurodegenerative und demyelinisierende Erkrankung des zentralen Nervensystems und zählt zu den häufigsten Ursachen für nicht traumatisch bedingte, neurologische Beeinträchtigung im jungen Erwachsenenalter (De Stefano *et al.*, 2014). Weltweit sind ca. 2.5 Millionen Menschen von MS betroffen, die Mehrzahl davon sind Frauen (Dendrou, Fugger and Friese, 2015).

Die Krankheit verläuft entweder in schubförmig-remittierender Form oder primär bis sekundär chronisch-progredient. Zusätzlich zu der progredienten körperlichen Behinderung leiden die PatientInnen unter psychischen Symptomen wie Depression, Angstzuständen und Fatigue (Galeazzi *et al.*, 2005; Beiske *et al.*, 2008; Compston and Coles, 2008).

Neben der klinischen Symptomatik ist seit den letzten Jahren auch die MRT ein wichtiger diagnostischer Faktor der MS. Fokale Läsionen, aber auch globale Hirnatrophie kennzeichnen das radiologische Bild der MS ab dem Beginn der Erkrankung (De Stefano *et al.*, 2014; Rocca *et al.*, 2017).

Aus diesem Grund beschäftigt sich die Forschung bereits seit einiger Zeit mit dem Zusammenhang zwischen der Abnahme von Hirnvolumen und der Krankheitsprogredienz bzw. der Zunahme der klinischen Symptomatik. Ziel ist es, anhand der Erfassung der zerebralen Atrophie über die Änderung des Hirnvolumens Prognose und Therapieerfolg besser abschätzen zu können (Sormani, Arnold and De Stefano, 2014; Vidal-Jordana *et al.*, 2015).

Bisher wurden vor allem computerbasierte Software Programme in teilweise sehr aufwändigen Verfahren verwendet, um die Änderung des Hirnvolumens zu messen, und es gibt kaum Studie zur Verwendung von visuellen Rating Skalen bei der MS.

Diese Diplomarbeit versucht, anhand einer solchen Rating Skala Änderungen des Hirnvolumens bei MS-PatientInnen visuell zu beurteilen und zu quantifizieren, und anschließend die Ergebnisse mit den Messdaten von Software-basierten Programmen (SIENA und SIENAX) zu vergleichen.

Im ersten Teil dieser Arbeit wird allgemein auf die Erkrankung selbst und spezifisch auf die bei MS auftretende Hirnatrophie eingegangen. Anschließend werden die Methoden und die Ergebnisse der durchgeführten Studie beschrieben und diskutiert.

1.1 Epidemiologie

Laut einer im Jahr 2013 durchgeführten Untersuchung der internationalen Föderation für Multiple Sklerose leiden weltweit circa 2.3 Millionen Menschen an MS (Thompson and Baneke, 2013), andere Studien sprechen bereits von 2.5 Millionen (Dendrou, Fugger and Friese, 2015). Im Jahr 2008 wurden noch 2.1 Millionen Fälle gezählt; seitdem stieg die weltweite Prävalenz von 30 auf 33/100,000 Einwohnern an. Es ist jedoch nicht klar, ob diese Zunahme auf verbesserte Diagnose- und Dokumentationstechniken oder auf andere Ursachen zurückzuführen ist.

Am häufigsten tritt MS in Europa auf; an zweiter Stelle steht Amerika (140 und 108 pro 100,000). Die niedrigste Prävalenz findet sich im Sub-Sahara-Afrika Raum und Ostasien (2.1 und 2.2 pro 100,000) (Thompson and Baneke, 2013).

Obwohl die Datenlage in einigen der untersuchten Länder noch sehr ungenau ist, konnte in einem weltweiten „MS-Atlas“ eine klare demographische Abhängigkeit der MS Prävalenz aufgezeigt werden.

Es wurde festgestellt, dass in Abhängigkeit mit der Entfernung vom Äquator - sowohl in südliche als auch in nördliche Richtung - die Zahl der an MS erkrankten Menschen ansteigt. Möglicherweise ist dies auf die geringere Sonneneinstrahlung in diesen Ländern und auf den damit verbundene verminderten Vitamin D Haushalt der EinwohnerInnen zurückzuführen (Alonso and Hernán, 2008; Simpson Jr. *et al.*, 2011).

Zudem wurden Unterschiede zwischen ethnischen Gruppen festgestellt. So erkrankten beispielweise in Kanada im Vergleich zur restlichen Bevölkerung weit weniger Ureinwohner an MS. Dasselbe gilt für die Anzahl an Erkrankungsfällen in der hispanischen und nicht-hispanisch-schwarzen Bevölkerung in Texas im Vergleich zu den nicht-hispanisch Weißen. Diese Tatsache stützt die These eines genetischen Einflusses auf die Entstehung von MS (Evans *et al.*, 2013; Kingwell *et al.*, 2013).

Andererseits haben Migrationsstudien gezeigt, dass Personen, die aus einem Land mit einer niedrigen Prävalenz für MS in ein Land mit hoher Prävalenz ziehen, das Erkrankungsrisiko ihres Herkunftslandes beibehalten. Dieses überträgt sich jedoch nicht auf ihre im „Zielland“ geborenen Kinder. Umgekehrt sinkt die Wahrscheinlichkeit an MS zu erkranken für MigrantInnen aus einem Hochrisiko- in ein Niedrigrisikoland, besonders wenn diese schon im frühen Kindesalter auswandern. Die genetische Komponente scheint demnach nicht als einzige das Erkrankungsrisiko zu beeinflussen (Young, 2011).

Frauen sind laut verschiedener Studien in Europa und den USA zwei- bis dreimal häufiger betroffen als Männer (Kingwell *et al.*, 2013). Am häufigsten wird MS zwischen dem 20. und dem 40. Lebensjahr diagnostiziert, es können jedoch auch Kinder und ältere Personen daran erkranken (Banwell *et al.*, 2007; Dargahi *et al.*, 2017).

In Österreich wurden 2011 in einer landesweiten Erhebung 12,500 MS Fälle gezählt, davon waren über 70% Frauen. Auch hierzulande stieg die Zahl der Erkrankungsfälle an. Betrug die Prävalenz 1999 noch 99/100.000 EinwohnerInnen, so waren es 2011 bereits 149 und 2017 158.9 pro 100,000 Einwohner. Das durchschnittliche Alter bei klinischer Erstmanifestation liegt in Österreich bei 30 Jahren und deckt sich somit mit den weltweiten Daten (Baumhackl, 2014; Salhofer-Polanyi *et al.*, 2017).

1.2 Ätiologie

Die genaue Ursache für Multiple Sklerose ist bis heute ungeklärt. Die Forschung ist sich jedoch weitgehend einig, dass die Erkrankung von verschiedenen Faktoren abhängig ist. Besonders das Zusammenspiel zwischen genetischer Disposition und verschiedenen Umweltfaktoren wird diskutiert (Compston and Coles, 2008; Dendrou, Fugger and Friese, 2015).

1.2.1 Genetik

Bereits Anfang der 1980er Jahre haben eine Vielzahl von epidemiologischen Studien auf eine familiäre Häufung von Multiple Sklerose Fällen hingewiesen. Seitdem wurde ein vermehrtes Auftreten von MS in Familien von bereits Betroffenen im Vergleich zur Durchschnittsbevölkerung beobachtet. Unter diesem Gesichtspunkt wurde auch untersucht, ob diese Häufung zudem mit dem Grad der Verwandtschaft zwischen Erkrankten und Familienmitgliedern variiert (z.B. Zwillingen, Geschwistern, Halb- und Stiefgeschwistern, adoptierten Kindern, Ehepaaren und Nachkommen). All diesen Untersuchungen bestätigten ein vermehrtes Auftreten der Erkrankung innerhalb von bestimmten Familien, wodurch eine gewisse genetische Komponente als Ursache für die Entstehung von MS immer weiter in den Mittelpunkt rückt (O’Gorman *et al.*, 2013; Sawcer, Franklin and Ban, 2014).

Eine Metaanalyse aus dem Jahr 2013 fasste die Daten von über 500 Risikostudien zusammen, mit dem Ziel, das Erkrankungsrisiko für Familienangehörige von MS

PatientInnen zu berechnen. Die Beteiligung von genetischen Faktoren an der Entstehung der Erkrankung liegt demnach bei 54% (O’Gorman *et al.*, 2013).

Das Erkrankungsrisiko für Familienangehörige von MS PatientInnen sinkt mit zunehmendem Verwandtschaftsgrad; so haben Eltern und Kinder von Betroffenen ein geringeres Risiko als deren Geschwister. Die Erkrankungswahrscheinlichkeit für monozygote Zwillinge ist mit 18.2% am höchsten, jene für dizygoten Zwillinge liegt bei ca. 5% und ist größer als für Geschwister. Dies könnte eventuell auf temporäre Umwelteinflüsse oder auf intrauterine Effekte zurückzuführen sein (O’Gorman *et al.*, 2013).

Bisher konnten noch kein spezifisches Gen für MS ermittelt werden. Die Forschung geht vielmehr von einem polygenetischen Vererbungsmuster aus, bei dem das Erkrankungsrisiko von einem Hochrisiko- und von mehreren Niedrigrisikoallelen abhängig ist (Sawcer, Franklin and Ban, 2014).

Der Haupthistokompatibilitätskomplex (major histocompatibility complex, MHC) auf dem kurzen Arm von Chromosom 6 (Genlocus 6p21.3) wurde bereits vor über 30 Jahren als Risiko-Genlocus für MS identifiziert (Lill, 2014). Diese Gene kodieren für die sogenannten „human leucocyte antigens“ (HLA) oder Leukozytenantigene, antigen-präsentierende Moleküle, die auf fast allen Zellen vorkommen und eine entscheidende Rolle für die Immunabwehr, vor allem bei der Unterscheidung zwischen körpereigenem und körperfremdem Gewebe spielen.

Unter den zahlreichen Genen und Gen-Subtypen, die in dieser Region vorkommen, konnte das HLA-DRB1*15:01 Gen bereits in mehreren Studien mit einem erhöhten Risiko für MS in Verbindung gebracht werden. Seit 2007 verwendet man Genomweite Assoziationsstudien (GWAS) für die Identifikation von weiteren Risikoallelen. GWAS erlauben es, tausende von „Single Nucleotide Polymorphismen“, d.h. verschiedene Gen-Varianten, bei Erkrankten und Gesunden zu untersuchen und miteinander zu vergleichen (Lill, 2014; Sawcer, Franklin and Ban, 2014).

Neben den HLA-Allelen konnten dadurch weitere Risiko-Genloci außerhalb des MHC identifiziert werden. Darunter auch Genloci, die für den Interleukin-2-Rezeptor und den Interleukin-7-Rezeptor-alpha kodieren. Letzterer scheint bei der Entstehung von autoreaktiven T-Zellen bei MS-PatientInnen eine Rolle zu spielen. (International Multiple Sclerosis Genetics Consortium *et al.*, 2007).

Die Beteiligung an der genetischen Heritabilität aller bekannten MS assoziierten Genloci wird auf ein Viertel geschätzt, 11% davon fallen allein auf das HLA-DRB1*15:01 Allel. Damit bleibt über die Hälfte der errechneten Erblichkeit von MS unerklärt; dies wird in der

Wissenschaft als „missing heritability“ bezeichnet. Eine mögliche Erklärung dafür wären epigenetische Veränderungen des Genoms, wie zum Beispiel DNA Methylierung oder Modifikationen von Histonen. Die dadurch veränderte Expression von Genen könnte die Reaktion auf Umwelteinflüsse und somit auch die Suszeptibilität für MS beeinflussen (Lill, 2014; Zhou *et al.*, 2014).

1.2.2 Umweltfaktoren

Die global gesehen sehr heterogene Verteilung von MS deutet darauf hin, dass auch verschiedene Umweltfaktoren das Risiko einer MS-Erkrankung mitbeeinflussen. Neben diversen Krankheitserregern werden auch nicht-infektiöse Agentia diskutiert. Auch eine verminderte Sonnenlichtexposition und der damit verbundene geringere Vitamin D Haushalt wurden in Zusammenhang mit einem erhöhten MS Risiko gebracht (Ascherio and Munger, 2007a, 2007b; Correale and Gaitán, 2015).

Unter den infektiöser Risikofaktoren für MS scheint vor allem das EBV Virus eine pathogenetischen Rolle zu spielen (Ascherio and Munger, 2007a). Laut einer Studie haben Personen ohne EBV-Antikörper, die also seronegativ sind, ein extrem geringes Erkrankungsrisiko. Im Vergleich dazu ist die Wahrscheinlichkeit für Personen, die in ihrer frühen Kindheit an EBV erkrankt sind, 10mal und für Personen, die erst spät an Mononukleose erkrankt sind, sogar 20 mal höher (Thacker, Mirzaei and Ascherio, 2006).

Es gibt viele Theorien über die genaue Pathophysiologie des Zusammenspiels zwischen EBV und MS. Die momentan führende Hypothese vermutet, dass es bei der Immunantwort auf EBV zu einer Kreuzreaktion mit Myelinantigenen kommt. An einer derart fehlgeleiteten Immunantwort könnten sowohl T-Lymphozyten als auch Antikörper beteiligt sein: Forscher konnten im Liquor von MS PatientInnen eine erhöhte Anzahl an EBV spezifischen CD4⁺ T-Zellen und eine gegen zwei EBV Proteine gerichtete Immunantwort nachweisen (Cepok *et al.*, 2005; Lünemann *et al.*, 2006).

Der Anstieg der MS Prävalenz mit zunehmendem Breitengrad deutet darauf hin, dass sich eine verminderte UV Strahlung und ein damit verbundener, niedriger Vitamin D Spiegel negativ auf das Risiko für MS auswirken (Ascherio and Munger, 2007b; Simpson Jr. *et al.*, 2011).

Ein Mangel an Vitamin D scheint auch die Progression der Erkrankung zu begünstigen und in Zusammenhang mit einer erhöhten Krankheitsaktivität zu stehen (van der Mei *et al.*, 2007; Ascherio *et al.*, 2014).

In einer Metaanalyse aus dem Jahr 2016, in der insgesamt 26 Studien aus den letzten 30 Jahren zusammengefasst worden sind, wurde der Zusammenhang zwischen Rauchen und MS untersucht. Das Ergebnis zeigte ein erhöhtes MS-Risiko, vor allem für männliche Raucher. Auch für Passivraucher ergab sich eine größere Erkrankungswahrscheinlichkeit als für nicht Nikotin-exponierte Personen (Zhang *et al.*, 2016).

1.3 Pathogenese

Obwohl ein eindeutiger Beweis bis heute noch nicht erbracht werden konnte, geht die Forschung davon aus, dass es sich bei MS um eine Autoimmunerkrankung handelt, bei der autoreaktive Lymphozyten ZNS-spezifische Autoantigene angreifen (Hohlfeld and Wekerle, 2004).

Warum es zur Aktivierung dieser Immunzellen kommt ist unklar. Aktuell werden Mechanismen wie Molekulares Mimikry, Antigenpräsentation von fremden, aber den körpereigenen Peptiden sehr ähnlichen, Antigenen und die sogenannte „Bystander“-Aktivierung diskutiert. Ob diese Prozesse der primäre Auslöser (CNS extrinsic Modell) oder ob sie erst sekundär nach ZNS internen Ereignissen (CNS intrinsic Modell) auftreten, ist ebenfalls noch Gegenstand der Forschung (Dendrou, Fugger and Friese, 2015).

Bei Letzterem könnte es sich entweder um eine Entzündungsreaktion auf ein bisher unbekanntes ZNS-Virus oder um die sogenannte primäre Neurodegeneration handeln. Dies ist der progressive Untergang von Axonen und Nervenzellen, ausgelöst durch ZNS eigene Zellen, und wurde schon bei Alzheimer und Parkinson Erkrankungen beobachtet (Dendrou, Fugger and Friese, 2015).

Entzündungsreaktionen und der damit verbundene Permeabilitätsverlust der Blut-Hirn-Schranke charakterisieren das akute Stadium von MS. In der Peripherie aktivierte Entzündungszellen, allen voran Makrophagen, aber auch CD8⁺ und CD4⁺ T-Zellen, sowie B-Zellen und Plasmazellen, invadieren das ZNS (Dendrou, Fugger and Friese, 2015). Diese Zellen produzieren inflammatorische und neurotoxische Mediatoren, welche die Myelinscheiden der Nervenfasern sowie die Oligodendrozyten angreifen. Es bilden sich MS-typische, konfluierende, demyelinisierende Areale, sogenannte Läsionen oder Plaques. Anfangs beschränken sich diese weitgehend auf die weiße Substanz. Die Anteile außerhalb der fokalen Plaques nennt man „normal-appearing white matter“, weiße Substanz ohne sichtbare Entzündungszeichen (Dendrou, Fugger and Friese, 2015; Mahad, Trapp and Lassmann, 2015).

Diese akuten Läsionen können durch das enorme Regenerationspotenzial des ZNS nach einigen Wochen wieder verschwinden und mit ihnen auch die klinischen Symptome der PatientInnen. Höchst wahrscheinlich sind es Vorläuferzellen der Oligodendrozyten, die durch Remyelinisierung die Integrität der geschädigten Axone wiederherstellen (Chandran *et al.*, 2008; Mahad, Trapp and Lassmann, 2015).

Nach Erschöpfung des Regenerationspotenzials kommt es aufgrund des chronischen Verlusts von weißer und grauer Substanz zu einer Gliose, einer chronischen Aktivierung von Mikrogliazellen. Diese produzieren freie Sauerstoffradikale, welche Mutationen in der mitochondrialen DNA der Neuronen induzieren und damit deren Energiehaushalt schädigen. Ein progressiver neuroaxonaler Verlust ist die Folge, mit zunehmender Behinderung der PatientInnen. Zusätzlich bestimmen Hirnatrophie und Ventrikelerweiterung das Krankheitsbild (Compston and Coles, 2008; Mahad, Trapp and Lassmann, 2015).

1.4 Verlaufsformen

Die Unterscheidung zwischen verschiedenen Verlaufsformen einer Erkrankung ist sowohl im klinischen Alltag (Dokumentation, Prognose, Therapieentscheidungen) als auch für die Forschung wichtig. Ein internationales Komitee bestehend aus Experten der US-amerikanischen National MS Society (NMSS) und des European Committee for Treatment and Research in MS (ECTRIMS) veröffentlichte 2013 eine Überarbeitung der MS-Klassifikation von Lublin und Reingold aus dem Jahr 1996 (Lublin *et al.*, 2014).

Diese neue Einteilung unterscheidet anhand des zeitlichen Verlaufs der Krankheitsaktivität, charakterisiert durch klinische Progredienz z.B. als Schub oder in Form von neuen MS-typische MRT-Veränderungen, zwischen schubförmigen und progredienten Phänotypen. Der Schub ist definiert als das Neuauftreten oder die Verschlechterung von MS-typischen Symptomen, die über mindestens 24 Stunden bestehen und in keinem Zusammenhang mit Fieber oder einer Infektion stehen. Zwischen zwei Schüben ist ein Zeitintervall von mindestens 30 Tagen notwendig (Deutsche Gesellschaft für Neurologie, 2015; Thompson *et al.*, 2018).

Die schubförmige Verlaufsform RRMS (relapsing remitting MS) ist charakterisiert durch eine Phase der neurologischen Stabilität ohne Symptomenprogredienz zwischen den Schüben (Lublin *et al.*, 2014).

Bei der progredient verlaufenden MS nimmt die neurologische Symptomatik stetig und unabhängig von Schüben zu. Fluktuationen, Phasen der Stabilität oder überlagerte Schübe

sind möglich. Dieser Phänotyp kann sich aus einer RRMS entwickeln, dann spricht man von einer sekundär progressiven MS (SPMS, Secondary progressive multiple sclerosis) oder bereits seit Krankheitsbeginn bestehen, primär progrediente MS (PPMS, primary progressive multiple sclerosis, primary) (Thompson *et al.*, 2018).

CIS (Klinisch isoliertes Syndrom)

Das klinisch isolierte Syndrom gilt als erste klinische Manifestation von MS-typischen Symptomen bei PatientInnen ohne bekannte MS, welche auf einen fokalen oder multifokalen demyelinisierenden Vorgang im ZNS hindeuten (Lublin *et al.*, 2014). Es handelt sich dabei um eine monophasische Episode mit akutem oder subakutem Beginn, die mindestens 24 Stunden andauert, mit möglicher Remission, und nicht mit Fieber oder einem Infekt assoziiert ist (Thompson *et al.*, 2018).

Das Risiko für PatientInnen mit CIS eine definitive MS zu entwickeln ist abhängig von der Anzahl der in der MRT sichtbaren Läsionen sowie von deren Lokalisation (Miller, Chard and Ciccarelli, 2012). Beim Übergang in eine definitive MS entspricht das CIS dem ersten Schub des Patienten bzw. der Patientin (Lublin *et al.*, 2014).

RIS (Radiologisch isoliertes Syndrom)

Das radiologisch isolierte Syndrom (Radiologically isolated syndrome, RIS) ist kein eigenständiger Phänotyp von MS. Vielmehr handelt es sich dabei um zufällig entdeckte ZNS Veränderungen in der MRT, die auf einen demyelinisierenden Prozess hindeuten, ohne dass ein klinisches Korrelat vorhanden ist. Studien haben gezeigt, dass PatientInnen mit RIS ein erhöhtes Risiko für MS bzw. radiologische Progression aufweisen. RIS könnte daher eine präklinische Form der MS sein (Okuda *et al.*, 2009).

1.5 Symptome

Bereits Mitte des 19. Jahrhunderts fasste der französische Neurologe Jean-Martin Charcot die Kombination aus Tremor, Dysarthrie und Nystagmus in der nach ihm benannten Charcot-Trias zusammen. Heute ist die Palette der beschriebenen klinischen Manifestationen von MS weitaus größer, jedoch gibt es kein spezifisches Symptom für die Erkrankung. Je nach Lokalisation der entzündlichen Herde im ZNS treten die Symptome in den unterschiedlichsten Ausprägungen auf und verändern sich auch im Verlauf der

Erkrankung (Poser and Brinar, 2001; Schmidt, 2015, pp. 55–56). In Tabelle 1 sind die häufigsten Symptome der MS aufgelistet.

Trotz der großen Variabilität werden bestimmte Symptome besonders häufig beobachtet. So leiden 40–84% der PatientInnen an progressiv zunehmenden spastischen Paresen, die sich in einem erhöhtem Muskeltonus bis hin zu Rigor und verminderter Mobilität äußern (Flachenecker, Henze and Zettl, 2014). Auch ataktische Symptome wie Intentions- und Haltetremor werden oft beobachtet (Koch *et al.*, 2007). Des weiteren leiden die PatientInnen gehäuft an Sensibilitätsstörungen wie Parästhesie und Schmerzen, die besonders die Extremitäten sowie Hände und Füße betreffen, (Beiske *et al.*, 2004) und an Störungen des autonomen Nervensystems, vor allem an Blasenfunktionsstörungen (Kisic Tepavcevic *et al.*, 2017).

Unter den neuropsychischen Symptomen fallen Fatigue und Depressionen besonders schwer ins Gewicht, da diese eine enorme Belastung für das alltägliche Leben der PatientInnen darstellen (Galeazzi *et al.*, 2005; Siegert and Abernethy, 2005; van Kessel and Moss-Morris, 2006).

Tabelle 1: Häufige Symptome der MS

Symptome	Beschreibung
Motorik- und Koordinationsstörungen	Paresen Spastik und Pyramidenbahnzeichen Ataxie Tremor
Hirnnervenstörungen	Optikusneuritis Visusminderung Nystagmus internukleäre Ophthalmoplegie Diplopie Sprach- und Schluckstörungen
Sensibilitätsstörungen, Schmerzen	Parästhesien und Dystästhesien an Händen und Füßen Störungen der Oberflächen- und Tiefensensibilität Allodynie Trigeminusneuralgie Lhermitte-Zeichen
Vegetative Funktionsstörungen	Störungen der Blasen- und Darmentleerung Störung der Sexualfunktion
Neuropsychologische Störungen	kognitive Defizite

	Fatigue Depressive Störungen
Modifiziert nach (Schmidt, 2015, p. 56)	

Die objektive Beurteilung des neurologischen Defizits erfolgt mittels verschiedener Einteilungen. Seit 1983 gilt die EDSS-Skala (Expanded Disability Status Scale) von John Kurtzke als Goldstandard für den klinischen Alltag.

Mittels einer standardisierten neurologischen Untersuchung werden auf einer 6-gradigen Skala acht verschiedene Funktionssysteme (FS) bewertet, wobei 0 für keinen und 6 für einen schweren Funktionsverlust steht. Anhand dieser Punkte können PatientInnen einem der 20 Grade auf der EDSS Skala zugeordnet werden. EDSS-Werte unter 4.0 stehen für PatientInnen mit nahezu vollständig erhaltener Mobilität, PatientInnen mit höheren Werten werden anhand der von ihnen noch bewältigbaren Gehstrecke evaluiert. In Tabelle 2 ist die EDSS nach Kurtzke aufgelistet (Schmidt, 2015, pp. 65–55).

Tabelle 2: Expanded Disability Status Scale (EDSS)

Expanded Disability Status Scale (EDSS, nach Kurtzke 1983)	
FS = Funktionssysteme: Motorik, Kleinhirn, Hirnstamm, Sensibilität, Blasen- und Mastdarmfunktionen, Sehfunktion, zerebrale Funktionen, andere Funktionen	
0.0	Normale neurologische Untersuchung
1.0	Keine Behinderung, minimale Symptomatik in 1 FS
1.5	Keine Behinderung, minimale Symptomatik in 1 als einem FS
2.0	Minimale Behinderung in 1 FS
2.5	Minimale Behinderung in 2 FS
3.0	Mäßige Behinderung in 1 FS oder leichte Behinderung in 3 oder 4 FS, keine Einschränkung der Gehfähigkeit
3.5	Keine Einschränkung der Gehfähigkeit, in 1 FS mäßige und in 1 oder 2 FS leichte Behinderung oder in 2 FS mäßige oder in 5 FS leichte Behinderung
4.0	Gehfähigkeit ohne Hilfe min. 500m, selbstständig, 12h/d aktiv trotz relativ schwerer Behinderung
4.5	Gehfähigkeit ohne Hilfe min. 300m, Arbeitsfähigkeit ganztägig, meiste Zeit des Tages aktiv mit geringen Einschränkungen oder Hilfe, relativ schwerer Behinderung
5.0	Gehfähigkeit ohne Hilfe min. 200m, schwere Behinderung die tägliche Aktivität beeinträchtigt
5.5	Gehfähigkeit ohne Hilfe min. 100m, schwere Behinderung, normale tägliche Aktivität nicht möglich
6.0	Intermittierende oder konstante Gehilfe für 100m ohne Rast
6.5	Konstante Gehilfe beidseits für 20m ohne Rast
7.0	Ohne Hilfe max. 5m, meist Rollstuhl notwendig (ca. 12h/d), Mobilität und Transfer ohne Hilfe

7.5	Nicht mehr als einige Schritte möglich, Rollstuhl notwendig, Hilfe beim Transfer, alleine mobil, nicht den ganzen Tag im Rollstuhl
8.0	An Bett gebunden, sitzt tagsüber auch außerhalb, alleinige Pflege möglich, guter Gebrauch der Arme
8.5	Auch tagsüber meist ans Bett gebunden, eigener, nützlicher Gebrauch der Arme erhalten, Selbstpflege möglich
9.0	Bettlägerig, hilflos, essen und kommunizieren möglich
9.5	Bettlägerig, vollständig auf Hilfe angewiesen, essen, schlucken und kommunizieren nicht möglich
10	Tod infolge von MS
Modifiziert nach (Schmidt, 2015, p. 66)	

Eine weitere Bewertungsmöglichkeit bietet die Multiple Sclerosis Functional Composite (MSFC). Diese Skala beinhaltet zusätzlich noch die Funktion der Extremitäten sowie eine Beurteilung der Kognition und gilt Studien zufolge als sensitiverer und besserer Marker für den therapeutischen Effekt als die EDSS Skala (Ozakbas, Ormeci and Idiman, 2005).

1.6 Diagnose

Seit 2001 werden die sogenannten „McDonald-Kriterien“ für die Diagnose der MS verwendet. Diese wurden vom Internationalen Gremium für Multiple Sklerose (International Panel on MS) erstellt und sind bisher dreimal, 2005, 2010 und 2017, durch dasselbe Gremium überarbeitet worden (Thompson *et al.*, 2018).

Durch diese Kriterien hat man versucht die Diagnosestellung von MS sowohl für den klinischen Gebrauch als auch für klinische Studien zu vereinheitlichen. Als bildgebendes Verfahren wurde die MRT aufgrund sehr hoher Sensitivität und Spezifität in die Untersuchung integriert und ist seitdem neben der klinischen Beurteilung der PatientInnen für eine korrekte Diagnose unentbehrlich (Filippi *et al.*, 2016).

Nach der letzten Revision im Jahr 2017 basiert die Diagnose der MS in erster Linie auf dem Nachweis einer örtlichen und zeitlichen Dissemination von fokal neurologischen Defiziten und dem Ausschluss von möglichen Differentialdiagnosen. Im frühen Stadium (z.B. bei PatientInnen mit unklarem CIS) wird zudem der Nachweis von oligoklonalen Banden im Liquor zur Diagnosesicherung empfohlen (Aktas *et al.*, 2018).

Die örtliche Dissemination gilt als Hinweis auf einen multifokalen Prozess im ZNS und ist definiert als die Entwicklung von Läsionen an bestimmten anatomischen Regionen des ZNS.

Unter der zeitlichen Dissemination versteht man das Auftreten solcher ZNS-Läsionen im zeitlichen Verlauf (siehe Tabelle 3) (Thompson *et al.*, 2018).

Tabelle 3: Diagnosekriterien DIS und DIT

Örtliche Dissemination, DIS (dissemination in space)	Zeitliche Dissemination, DIT (dissemination in time)
mindestens eine T2-hyperintense Läsion in mindestens 2 MS-typischen Arealen: <ul style="list-style-type: none"> – periventrikulär – kortikal oder juxakortikal – infratentoriell – spinal (Rückenmark) bei Personen über 50 Jahren oder bei Personen mit erhöhtem vaskulärem Risiko ist eine einzelne Läsion möglicherweise nicht aussagekräftig	Gleichzeitiger und zeitpunktunabhängiger Nachweis von kontrastmittelaufnehmenden und nicht-kontrastmittelaufnehmenden Läsionen oder Nachweis einer neuen T2-hyperintensiven Läsion oder kontrastmittelaufnehmenden Läsion in einer Verlaufs-MRT unabhängig vom Zeitpunkt der Referenz-MRT
Modifiziert nach (Thompson <i>et al.</i> , 2018)	

Die zahlreichen Differenzialdiagnosen der MS, wie beispielsweise die Neuromyelitis optica oder die akute disseminierte Enzephalomyelitis, sollten vor Diagnosestellung ausgeschlossen werden. Erst wenn es keine „andere Erklärung“ der klinischen Symptomatik gibt und ein eindeutiger Befund eines typischen CIS oder einer PPMS mit allen notwendigen MRT-Kriterien vorliegt, kann die Diagnose einer MS gestellt werden (Aktas *et al.*, 2018; Thompson *et al.*, 2018).

In Tabelle 4 sind die Diagnosekriterien der MS und in Tabelle 5 jene der PPMS aufgelistet.

Tabelle 4: McDonald Kriterien für die Diagnose einer MS bei PatientInnen mit einem Schub bei Erkrankungsbeginn

	Anzahl der klinisch relevanten Läsionen	Notwendige, zusätzliche Diagnosekriterien
≥ 2 Schübe	≥ 2	Keine
≥ 2 Schübe	1 (bei objektivem Nachweis eines früheren Schubs, der ein anderes ZNS Areal miteinschließt)	Keine
≥ 2 Schübe	1	DIS (siehe Tabelle 3) in der MRT oder Nachweis eines zusätzlichen Schubs, der ein anderes ZNS Areal miteinschließt
1 Schub	≥ 2	DIT (siehe Tabelle 3) in der MRT oder Nachweis eines zusätzlichen Schubs, der ein anderes ZNS Areal miteinschließt oder der

		Nachweis von oligoklonalen Banden im Liquor
1 Schub	1	DIS (siehe Tabelle 3) in der MRT oder Nachweis eines zusätzlichen Schubs, der ein anderes ZNS Areal miteinschließt und DIT (siehe Tabelle 3) in der MRT oder Nachweis eines zusätzlichen Schubs, der ein anderes ZNS Areal miteinschließt oder den Nachweis von oligoklonalen Banden im Liquor*
*Der Nachweis von oligoklonalen Banden im Liquor ist kein Beweis für DIT, kann aber statt dem Nachweis einer DIT (siehe Tabelle 3) zur Diagnosestellung verwendet werden.		
Modifiziert nach (Thompson <i>et al.</i> , 2018)		

Tabelle 5: Diagnosekriterien für die primär progrediente MS (PPMS)

PPMS (Primär progrediente MS)
<ol style="list-style-type: none"> 1. mindestens 1 Jahr progredienter Erkrankungsverlauf (retrospektive oder prospektive Diagnose) unabhängig von klinischen Schüben <p>zusätzlich 2 der 3 folgenden Kriterien:</p> <ol style="list-style-type: none"> 2. mehr als eine T2-hyperintense Läsion in mehr als einem MS-typischen ZNS-Areal (periventrikulär, kortikal oder juxakortikal oder infratentoriell) 3. mehr als 2 T2-hyperintense Läsionen im Rückenmark 4. Nachweis von isolierten oligoklonalen Banden im Liquor
Modifiziert nach (Thompson <i>et al.</i> , 2018)

1.7 Therapie

Die Behandlung von MS PatientInnen besteht in den meisten Fällen aus einer verlaufsmodifizierender Langzeittherapie und ggf. der Therapie des akuten Schubs, die je nach Erkrankungsstadium, Verlaufsform und Krankheitsdynamik mit einer symptomatischen Therapie kombiniert wird. Faktoren wie Lebensstil, Kinderwunsch, Begleiterkrankungen, Alter und Geschlecht spielen bei der Therapieentscheidung ebenfalls eine wichtige Rolle (Schmidt, 2015, p. 216).

Abbildung 1: Stufentherapie der Multiple Sklerose

(Deutsche Gesellschaft für Neurologie, 2015)

Indikation	CIS ¹		RRMS ¹			SPMS ¹	
Verlaufsmodifizierende Therapie	(Hoch-) aktive Verlaufsform		1. Wahl - Alemtuzumab - Fingolimod - Natalizumab	2. Wahl - Mitoxantron (- Cyclophosphamid) ²	3. Wahl - Experimentelle Verfahren	mit aufgesetzten Schüben	ohne aufgesetzte Schübe
	Milde/moderate Verlaufsform	- Glatirameracetat - Interferon-β 1a im - Interferon-β 1a sc - Interferon-β 1b sc	- Dimethylfumarat - Glatirameracetat - Interferon-β 1a im - Interferon-β 1a sc - Interferon-β 1b sc - PEG-IFN-β 1a sc - Teriflunomid (- Azathioprin) ² (- IFN) ²	- Interferon-β 1a sc - Interferon-β 1b sc - Mitoxantron (- Cyclophosphamid) ²			- Mitoxantron (- Cyclophosphamid) ²
Schubtherapie	2. Wahl - Plasmaseparation						
	1. Wahl - Methylprednisolonpuls						

Bei Versagen einer verlaufsmodifizierenden Therapie bei milder/moderater Verlaufsform einer MS werden diese Patienten wie eine aktive MS behandelt.

¹ Substanzen in alphabetischer Reihenfolge; die hier gewählte Darstellung impliziert KEINE Überlegenheit einer Substanz gegenüber einer anderen innerhalb einer Indikationsgruppe (dargestellt innerhalb eines Kastens)

² zugelassen wenn Interferon-β nicht möglich oder unter Azathioprin-Therapie stabiler Verlauf erreicht

³ Einsatz nur *postpartal* im Einzelfall gerechtfertigt, insbesondere vor dem Hintergrund fehlender Behandlungsalternativen

⁴ zugelassen für bedrohlich verlaufende Autoimmunkrankheiten, somit lediglich nur für fulminante Fälle als Ausweichtherapie vorzusehen, idealerweise nur an ausgewiesenen MS-Zentren

Die aktuellen Leitlinien für die verlaufsmodifizierende Therapie von MS wurden 2014 von der deutschen Gesellschaft für Neurologie überarbeitet. Im neuen Stufenschema (siehe Abbildung 1) der MS-Therapie wird zwischen milden/moderaten und (hoch)aktiven Verlaufsformen der Erkrankung unterschieden.

Bei ersterer werden als Basistherapie rekombinante Interferon-beta-Präparate oder Glatirameracetat verwendet. Diese Medikamente sind für alle schubförmigen Subtypen der MS einschließlich CIS-PatientInnen mit hohem Konversionsrisiko zu einer klinisch gesicherten MS zugelassen. Glatirameracetat ist für den SPMS-Phänotyp nicht zugelassen. PatientInnen mit RRMS können auch mit den oralen Basistherapeutika Dimethylfumarat und Teriflunomid behandelt werden. Im Falle von Therapieversagen oder Unverträglichkeiten der Immunmodulatoren können diese durch Mitoxantron, einem sogenannten „unselektiven Immunsuppressivum“, ersetzt werden. Dieses Präparat darf bei PatientInnen mit SPMS oder aktiver RRMS und einem EDSS-Score zwischen 3 und 6 eingesetzt werden. Auch Azathioprin kann als Ersatz für eine immunmodulatorische Therapie verwendet werden.

Bei (hoch-)aktiven Verlaufsformen von RRMS werden vorwiegend ein immunmodulatorisches Medikament namens Fingolimod und die monoklonalen Antikörper Natalizumab und Alemtuzumab empfohlen. Diese können entweder nach Misserfolg der

zuvor beschriebenen Basistherapie oder im Falle von besonders gravierender Verschlechterung der Symptomatik angewendet werden. Als 2. Wahl wird das bereits erwähnte Präparat Mitoxantron genannt, danach muss auf experimentelle Verfahren umgestiegen werden.

Cyclophosphamid, Methotrexat und intravenöse Gabe von Immunglobulinen sind nicht für die MS-Therapie zugelassen und sollten nur in Ausnahmefällen und nach individueller Anpassung verwendet werden.

Die Therapie der ersten Wahl für den akuten MS-Schub sind Glukokortikosteroide (GKS), Methylprednisolon oder Dexamethason als Kurzinfusion, die durch eine immunsuppressive Wirkung rasch zu einer Rückbildung der Symptome führen. Ein langfristiger Effekt konnte jedoch in Studien noch nicht belegt werden. Findet nach zweiwöchiger Behandlung keine Besserung statt, wird eine Plasmapherese bzw. Immunadsorption angewandt (Deutsche Gesellschaft für Neurologie, 2015).

Die symptomatische Therapie besteht aus medikamentösen und nicht-medikamentösen Maßnahmen. Ziel ist es, funktionelle Störungen möglichst zu beheben sowie deren Verschlechterung zu verlangsamen und Komplikationen zu vermeiden. Im Vordergrund der MS-Therapie steht eine Verbesserung der Lebensqualität der PatientInnen (Deutsche Gesellschaft für Neurologie, 2015).

1.7.1 Therapiewahl

Die Wahl der richtigen Therapie erfolgt abhängig von Krankheitsverlauf und den persönlichen Risiken individuell für jeden Erkrankten.

Nach dem Prinzip der Steigerungstherapie (escalation therapy) werden PatientInnen zuerst mit dem Medikament mit dem besten Nutzen/Risiko Verhältnis therapiert. Medikamente mit höherer Toxizität bei meist gleichbleibendem Wirkungsgrad aber mit anderen Wirkmechanismen kommen bei Therapieversagen sukzessive zum Einsatz.

Bei früh fulminantem Krankheitsverlauf kann nach der weit aggressiveren Induktions- und Erhaltungstherapie (induction and maintenance therapy) behandelt werden. Mit dem Ziel, das Immunsystem komplett wiederherzustellen, wird für möglichst kurze Zeit mit hochpotenten Immunsuppressiva therapiert. Sobald eine Remission der Krankheitsaktivität eingetreten ist, wird auf eine sicherere und bessere tolerierbare immunmodulatorische Therapie umgestellt (Rieckmann, 2009; Wingerchuk and Carter, 2014).

1.7.2 Monitoring

Die Bewertung der Krankheitsprogression sowie der Therapiewirkung anhand von Krankheitsaktivität und Verschlechterung der Symptomatik erfolgt durch klinische und MRT-basierte Parameter. Diese Parameter werden im sogenannten „no evidence of disease activity“ (NEDA) Konzept zusammengefasst, welches bisher vor allem in klinischen Studien verwendet wurde (Kappos *et al.*, 2016; Storelli *et al.*, 2018). Die Abwesenheit von Schüben, die Progression der Beeinträchtigung der PatientInnen durch die Krankheit und der Nachweis von kontrastmittelaufnehmenden T2-Läsionen zählen zu dem NEDA-3 Prinzip (Giovannoni *et al.*, 2015).

Da die bereits früh im Krankheitsverlauf auftretende progressive Abnahme des Hirnvolumens ebenfalls ein Marker für die Behinderungsprogression und kognitive Verschlechterung ist, wurde diese in Studien zu dem neuen NEDA-4 Prinzip hinzugefügt (Kappos *et al.*, 2016). Laut einer Studie von 2016 kann von einer pathologischen Hirnatrophie ab einer jährlichen Rate von über -0.4% ausgegangen werden. Da PatientInnen mit einer höheren Volumenabnahme auch einen weitaus schlechteren EDSS-Status aufwiesen, kann dieser „Cut-off“ als klinisch relevant angesehen werden (De Stefano *et al.*, 2016).

1.8 Prognose

Eine klare Prognose über den individuellen Verlauf von MS ist für die Behandlung und Therapie von PatientInnen von großer Bedeutung. Neben potenziellen Prädiktoren wie Alter bei Erkrankungsbeginn, Geschlecht oder Nachweis von oligoklonalen Banden im Liquor, sind vor allem die in der MRT sichtbaren und für MS typischen Veränderungen der Hirnstrukturen als prognostisch relevante Marker in den Mittelpunkt aktueller Forschungen gerückt (Filippi *et al.*, 2013; Tintore *et al.*, 2015).

Im Rahmen der natürlichen Krankheitsprogression kommt es neben den klassischen MS Läsionen zu einem erhöhten Verlust von Hirnvolumen bei nahezu allen MS Phänotypen, einschließlich CIS-PatientInnen (Vollmer *et al.*, 2015). Diese degenerativen Prozesse treten im gesamten ZNS auf und könnten je nach Lokalisation als eventuelle prognostische Parameter zu Rate gezogen werden.

Besonders bestimmte Strukturen in der grauen Substanz scheinen eine Schlüsselrolle für die progressive motorische und kognitive Verschlechterung von MS PatientInnen zu spielen (Filippi *et al.*, 2013).

Bei PatientInnen mit CIS, RRMS und SPMS korreliert laut Studien vor allem die Atrophie des Thalamus mit der zunehmenden Behinderung der PatientInnen (Rocca *et al.*, 2010; Hasan *et al.*, 2011). Atrophische Veränderungen in Cortex und Putamen standen ebenfalls in Relation mit einer Progression der Beeinträchtigung nach 5, nicht aber nach 10 Jahren (Jacobsen *et al.*, 2014).

Der Einfluss auf das Fortschreiten der physischen Einschränkungen sowie das Auftreten in den frühen Erkrankungsstadien von MS macht die Rate an Hirnatrophie in diesen Strukturen der grauen Substanz zu einem potenziellen Prädiktor für den späteren Krankheitsverlauf (Zivadinov *et al.*, 2013).

Ähnliche Ergebnisse zeigten sich für die Atrophie der weißen Substanz. So konnte in einer Studie die frühe Schädigung des Corpus Callosum als relevante Determinante für die motorische Beeinträchtigung und die zukünftigen kognitiven Störungen bei einer PatientInnengruppe mit PPMS identifiziert werden (Bodini *et al.*, 2013).

Einen möglichen Übergang von einer RRMS in eine SPMS im Vorfeld prognostizieren zu können, wäre vor allem für den klinischen Bereich von großer Bedeutung und könnte die richtige Therapiewahl deutlich erleichtern. Besonders MRT-basierte Parameter scheinen sowohl für Kurzzeit-, als auch für Langzeitprognosen geeignet zu sein. Auch hier scheint die regionale Volumenänderung in Thalamus und Corpus Callosum eine wichtige Rolle zu spielen (Uher *et al.*, 2016).

Laut der sogenannten „threshold hypothesis“ („Schwellen-Hypothese“) handelt es sich bei RRMS um eine vorwiegend in der weißen Substanz stattfindenden Erkrankung. Obwohl bereits pathologische Prozesse in der grauen Substanz stattfinden, beeinflussen diese aufgrund der enormen Regenerationsfähigkeit der kortikalen Neurone kaum das klinische Krankheitsbild. Erst nach Überschreitung einer bestimmten Schwelle, nimmt die Plastizität des Kortex ab und es kommt aufgrund der permanenten Schädigung der grauen Substanz zu einem Übergang in eine SPMS. Zusätzlich scheinen auch die Anzahl an kortikalen Läsionen und das kortikale Volumen des Kleinhirns von prognostischer Relevanz zu sein (Calabrese *et al.*, 2013).

1.9 Hirnatrophie bei Multipler Sklerose

1.9.1 Definition

Grundsätzlich unterscheidet man zwischen einfacher und numerischer Atrophie. Unter einfacher Atrophie versteht man eine reversible Zellverkleinerung aufgrund der Abnahme der Zellmasse der einzelnen Zellen, numerische Atrophie hingegen ist die Verkleinerung eines Organs oder Gewebes durch Zellzahlverminderung. Die zugrundeliegenden Ursachen können physiologischen oder pathologischen Ursprungs sein (Böcker, 2012, p. 24).

Im Gehirn unterscheidet man je nach Verteilungsmuster zwischen diffuser und generalisierter Hirnatrophie. In beiden Fällen kommt es in Folge des Untergangs von Neuronen zu einer Abnahme von Hirnvolumen. Dieser Vorgang ist einerseits Teil des physiologischen Alterungsprozesses und ohne Krankheitswert. Andererseits kommt es bei neurodegenerativen Erkrankungen zur Zellschädigung mit anschließender Atrophie in den betroffenen Hirnarealen. Ist ein bestimmtes funktionelles System im Gehirn betroffen spricht man von Systematrophie (z.B. bei Morbus Parkinson), sind mehrere funktionelle Neuronensysteme gleichzeitig befallen von Multisystematrophie (Böcker, 2012, p. 245).

Obwohl die Multiple Sklerose primär zu den entzündlichen Erkrankungen des ZNS gezählt wird, kann zudem von einer neurodegenerativen Komponente mit daraus resultierender Hirnatrophie ausgegangen werden (Siffrin *et al.*, 2010).

In der Fachliteratur gibt es keine klare Unterscheidung zwischen Hirnatrophie und Abnahme von Hirnvolumen bei MS PatientInnen. Prinzipiell basieren Veränderungen des Hirnvolumens auf zwei verschiedene Mechanismen: entweder sind diese durch Schwankungen der Menge an Hirnödemen oder durch Gewebsveränderung, in Form von Degeneration (Myelin, Axone, Zellen) bzw. Regeneration (Remyelinisierung) bedingt (Zivadinov *et al.*, 2008).

Wasseransammlungen im Hirngewebe entstehen aufgrund einer für MS charakteristischen Entzündungsreaktion; diese Volumenschwankungen werden als Pseudoatrophie bezeichnet und gelten als Störfaktor bei Atrophie-Messungen (De Stefano *et al.*, 2014).

Die Abnahme des Hirnvolumens aufgrund von Gewebeschwund ist eine Folge von neurodegenerativen Prozessen, welche bereits ab Beginn der Erkrankung auftreten und den gesamten Krankheitsverlauf beeinflussen (Zivadinov *et al.*, 2008). Da es sich dabei um eine Form der Hirnatrophie handelt, wird in dieser Diplomarbeit der Verlust von Hirnvolumen synonym zum Begriff der Hirnatrophie verwendet.

1.9.2 Ursachen

Die Pathogenese der MS kann in zwei Hauptphasen unterteilt werden. Im initialen Stadium der Erkrankung kommt es durch eine erhöhte Permeabilität der Blut-Hirn-Schranke zu einer Einwanderung von CD4⁺ Lymphozyten und Makrophagen in das ZNS. Dort bilden sich die für die Erkrankung charakteristischen, multifokalen, demyelinisierenden Läsionen. Im progressiven Stadium klingt die Entzündung weitgehend ab und neurodegenerative Prozesse dominieren das Krankheitsbild (Minagar *et al.*, 2004).

Demyelinisierung sowie neuro-axonaler Verlust führen letztendlich zu einer in der MRT messbaren Hirnatrophie. Vermutlich beginnt der Verlust von Hirnvolumen bereits in den frühesten Erkrankungsstadien und erstreckt sich über den gesamten Krankheitsverlauf. Während der ersten Erkrankungsphasen könnten jedoch Reparations- und Regenerationsmechanismen (z.B. Remyelinisierung durch humorale und zelluläre Immunantwort) diesen Verlust zumindest teilweise ausgleichen (Trapp *et al.*, 1998; Minagar *et al.*, 2004; Popescu *et al.*, 2015).

N-acetyl Aspartat (NAA), ein Stoffwechselprodukt des oxidativen Metabolismus der Mitochondrien, kann in der MRT-Spektroskopie als Biomarker für axonale Integrität herangezogen werden. MRT Untersuchungen von MS PatientInnen wiesen im Vergleich zu Kontrollgruppen sowohl in den akuten MS Läsionen, wie auch in der in der MRT normal erscheinenden weißen und grauen Substanz verminderte NAA Levels auf (Chard *et al.*, 2002; Sastre-Garriga *et al.*, 2005).

Besonders interessant ist die Atrophie der grauen Substanz, da diese in direktem Zusammenhang mit der klinischen Beeinträchtigung und kognitiven Funktion der PatientInnen steht (Filippi *et al.*, 2013; Popescu *et al.*, 2013). Bisher ist noch unklar, ob es sich dabei um eine primär unabhängige Neurodegeneration oder um eine Begleiterscheinung der pathologischen Prozesse der weißen Substanz handelt (Mühlau *et al.*, 2013; Steenwijk *et al.*, 2015).

Durch Studien konnte bisher bestätigt werden, dass es zumindest einen Zusammenhang zwischen MS Läsionen und der Atrophie der grauen Substanz gibt. Besonders hohe Atrophie-Raten wiesen kortikale Regionen mit einer Vielzahl von Verbindungen zu anderen Regionen auf. Dies spricht für die These, dass eine Unterbrechung der Hirnbahnen durch MS Plaques zumindest teilweise den Verlust von Hirngewebe fördert. Dennoch kann eine intrinsisch bedingte erhöhte Vulnerabilität aufgrund von lokalen Krankheitsprozessen nicht ausgeschlossen werden (Charil *et al.*, 2007).

Möglicherweise wird Atrophie der grauen Substanz nur während der frühen MS-Stadien durch entzündliche Vorgänge in der umliegenden weißen Substanz induziert. Auch hier gilt axonale Durchtrennung infolge einer fokalen Schädigung durch inflammatorische Noxen als potenzieller Atrophie-Trigger. Dieser führt zum Absterben des distalen Endes der Nervenfasern (Waller Degeneration). Durch transneuronale retro- oder anterograde Degeneration kommt es anschließend zum Verlust der Axone außerhalb der Entzündungsherde sowie zum Untergang der Neuronen selbst (Minagar *et al.*, 2004).

Im Gegensatz zur weißen Substanz steigt die Atrophie-Rate der grauen Substanz mit dem Fortschreiten der Erkrankung an (Fisher *et al.*, 2008). Besonders in den kortikalen Regionen wiesen PatientInnen mit SPMS im Vergleich zu jenen mit RRMS eine ausgeprägtere Atrophie auf, während zwischen PPMS und RRMS PatientInnen kein signifikanter Unterschied nachgewiesen werden konnte (Steenwijk *et al.*, 2015). Zudem scheint die Hirnatrophie in der grauen Substanz im späteren, progressiven MS Stadien einen von der weißen Substanz weitgehend unabhängiger Prozess zu sein. Welche anderen Faktoren in diesem Stadium entscheidend sind, ist aktuell noch nicht geklärt (Fisher *et al.*, 2008; Steenwijk *et al.*, 2015).

Auch innerhalb der grauen Substanz kann es zur Bildung von Läsionen kommen, welche ebenfalls den Atrophie-Prozess beeinflussen könnten (Calabrese and Gallo, 2009). Besonders für das CIS und die frühe RRMS wurde eine Korrelation zwischen dem Auftreten von neuen kortikalen Läsionen und kortikaler Volumenabnahme nachgewiesen, die im Verlauf der Erkrankung jedoch nachlässt. In späteren Stadien scheinen auch hier andere Faktoren das Fortschreiten der Hirnatrophie zu beeinflussen (Calabrese *et al.*, 2015).

1.9.3 Lokalisation

Atrophie der grauen Substanz spielt neben den für MS charakteristischen fokalen Prozessen in der weißen Substanz eine Schlüsselrolle für die klinische Ausprägung der Erkrankung. Insbesondere während des späteren progredienten Stadiums (SPMS) dominieren neurodegenerative Vorgänge in der grauen Substanz das Krankheitsgeschehen. Atrophie der grauen Substanz gilt als unabhängiger Marker für körperliche und klinische Beeinträchtigung der PatientInnen und ist als solcher weitaus robuster als die für MS typischen Läsionen der weißen Substanz (Calabrese *et al.*, 2013; Lansley *et al.*, 2013; Eshaghi *et al.*, 2014)

Die genaue Lokalisation dieser pathologischen Vorgänge in den kortikalen und subkortikalen Hirnstrukturen scheint nicht zufällig zu sein. Im Gegenteil, höchst wahrscheinlich handelt es sich um einen Regionen-spezifischen, dynamischen Prozess, der sich je nach Erkrankungsstadium und Krankheitsaktivität verschieden präsentiert (Amiri *et al.*, 2018).

Je nach Hirnregion schreitet der Atrophie-Prozess zu einem anderen Erkrankungszeitpunkt und mit einer anderen Geschwindigkeit voran. Derartige Unterschiede bestätigten Studien auch für die PPMS. In einer longitudinalen Studie wurde über einen Zeitraum von fünf Jahren die schnellste Volumenabnahme im Gyrus cinguli und die langsamste im Gyrus precentralis gemessen (Eshaghi *et al.*, 2014).

Bestimmte Regionen des Cortex scheinen abhängig vom Erkrankungsstadium anfälliger für fokale (Läsionen) und diffuse (kortikale Verdünnung) Schädigungen zu sein. Bei CIS und frühen RRMS PatientInnen sind eher fronto-temporale Regionen, insbesondere der Hippocampus, der Gyrus parahippocampalis, die Insula und der Gyrus cinguli betroffen. Im späteren Krankheitsverlauf (späte RRMS und SPMS PatientInnen) befällt der neurodegenerative Prozess auch den Gyrus pre- und postcentralis sowie das Kleinhirn (Calabrese *et al.*, 2015).

Auch innerhalb einer Struktur variiert die Abnahme des Volumens im Verlauf der Erkrankung. In einer Studie über Atrophie im Hippocampus wurde bei RRMS in der CA1 Region ein erhöhter Volumenverlust im Vergleich zur Kontrollgruppe registriert, bei SPMS PatientInnen war zusätzlich der Bereiche CA2/3 betroffen (Sicotte *et al.*, 2008).

Bisher ist noch unklar, ob die Volumenabnahme von grauer Substanz eher ein diffuser bzw. globaler Prozess ist oder ob es sich dabei um eine fokale Degeneration in prädominanten Hirnregionen handelt. Die ausschlaggebenden Faktoren für eine erhöhte Vulnerabilität bestimmter Hirnregionen sind ebenfalls noch nicht gefunden. Im Fall der Thalamus-Atrophie werden beispielweise die extrem hohe Anzahl an Verbindungen zu kortikalen und subkortikalen Strukturen als potentielle Ursache für eine erhöhte Sensitivität für pathologische Prozesse diskutiert (Lansley *et al.*, 2013; Minagar *et al.*, 2013).

Obwohl die Beteiligung der kritischen Hirnregionen der grauen Substanz anfangs eine hohe klinische Relevanz aufweist, scheint im Spätstadium der MS die globale Schädigung der gesamten grauen Substanz eine übergeordnete Rolle zu spielen (Filippi *et al.*, 2013).

1.9.4 Klinische Relevanz

Veränderungen des Hirnvolumens von MS PatientInnen prägen ab der Frühphase der Erkrankung den gesamten Verlauf und gelten als potentielle Marker für die Schädigung von Hirngewebe. Mittels MRT-basierter Methoden ist es möglich, den Verlust von Hirnvolumen im Verlauf zu messen und Rückschlüsse auf die globale sowie regionale Hirnatrophie zu ziehen. Dies kann zur Quantifizierung der zugrundeliegenden destruktiven, pathologischen Prozesse, wie irreversible Neurodegeneration und Demyelinisierung, verwendet werden und könnte daher für diagnostische und therapeutische Entscheidungen von Interesse sein (De Stefano *et al.*, 2014).

Die klinische Relevanz von Hirnatrophie ergibt sich aus dem dringenden Bedarf PatientInnen mit schlechtem klinischen Langezeit-Outcome frühzeitig zu identifizieren, um so deren Therapie und Management zu verbessern (Filippi *et al.*, 2013). Die Relevanz des Hirnvolumenschwunds als früher prognostischer Parameter für klinische Beeinträchtigung wurde bereits mehrfach durch Studien bestätigt. So konnte von der Verlustrate in einem Jahr auf die Progression der klinischen Beeinträchtigung über 10 Jahre später rückgeschlossen werden (Popescu *et al.*, 2013). Ebenso zeigten CIS-PatientInnen, die eine definitive MS entwickelten, sowie PatientInnen, die an progressiver klinischer Beeinträchtigung litten, einen höheren Verlust von globalem Hirnvolumen und vor allem von grauer Substanz als nicht konvertierte CIS- bzw. stabile MS PatientInnen (Dalton *et al.*, 2004; Pérez-Miralles *et al.*, 2013; Lavorgna *et al.*, 2014).

Wie bereits erwähnt, scheint die Atrophie-Messung in der grauen Substanz verglichen mit jener der weißen Substanz eher mit der klinischen und auch kognitiven Beeinträchtigung der PatientInnen zu korrelieren. Bei RRMS-PatientInnen war die Schädigung der grauen Substanz sogar der einzige Prädiktor für kognitive Verschlechterung nach 13 Jahren (Filippi *et al.*, 2013).

Globale und regionale Atrophie der grauen Substanz nimmt im Verlauf der Erkrankung stetig zu und könnte daher als Unterscheidungsmerkmal zwischen verschiedenen MS Phänotypen verwendet werden. Zudem können anhand der Atrophie einzelner Regionen in der grauen Substanz, spezifische klinische Defizite und die Verschlechterung kognitiver Funktionen erklärt werden (Rocca *et al.*, 2017).

In Studien wurde die Hirnatrophie-Rate bereits als Outcome-Marker für den Behandlungseffekt von sogenannten krankheitsmodifizierenden Medikamenten (DMT, disease modifying therapy) verwendet. Die gesamte Hirnatrophie gilt hierbei als ein äußerst

verlässlicher Parameter für die Abschätzung von Neuroprotektion, eine wichtige Anforderung an neue MS Medikation (Barkhof *et al.*, 2009; Battaglini *et al.*, 2009; Healy *et al.*, 2009). In einer Metaanalyse konnte durch Auswirkung von DMT auf die Hirnatrophie und auf fokale Läsionen in der MRT eine verlangsamte Zunahme der klinischen Beeinträchtigung bei RRMS PatientInnen über einen Zeitraum von 2 Jahren erklärt werden (Sormani, Arnold and De Stefano, 2014).

Neben Krankheitsprogression, Schubrate und Läsionen in der MRT wurde deshalb die jährliche Hirnatrophie-Rate als neues Therapieziel von DMT definiert. Diese vier Parameter werden im sogenannten NEDA-4-Prinzip (no evidence of disease activity) zusammengefasst, das in Studien bereits Verwendung findet (Kappos *et al.*, 2016).

Obwohl man aus heutiger Sicht anhand der Messung von Hirnatrophie viele für die Therapie und Prognose von MS relevante Informationen erhält, sind für eine sichere Anwendung im klinischen Alltag noch weitere Studien notwendig (Rocca *et al.*, 2017).

1.9.5 Hirnatrophie in der MRT

2001 wurde die MRT zum Nachweis von Dissemination in Zeit und Raum erstmals formel in die Diagnosestellung der MS integriert (Filippi *et al.*, 2016). Neben den charakteristischen Läsionen der weißen und grauen Substanz kann in MRT-Aufnahmen von MS PatientInnen mittlerweile auch globale und regionale Hirnatrophie bestimmt werden, was für die Diagnose der Erkrankung jedoch derzeit keine Rolle spielt. Als Marker für Neurodegeneration ist die Hirnatrophie in Studien ein wichtiger Parameter zur Überwachung des Krankheitsverlaufs von MS; insbesondere die Abnahme der grauen Substanz steht in direktem Zusammenhang mit progressiven klinischen und kognitiven Beeinträchtigungen der PatientInnen (Popescu, 2015, p. 10).

Die Mehrzahl der volumetrischen Messungen erfolgt auf T1 gewichteten MRT Scans, die einen optimalen Kontrast zwischen dem hell erscheinenden Parenchym und dem dunklen Liquor sowie zwischen grauer und weißer Substanz aufweisen (Rocca *et al.*, 2017). Für die Quantifizierung von Hirnatrophie wurden verschiedene Softwareprogramme entwickelt, die auf komplexen, automatisierten Gewebe- und Struktursegmentationsmethoden basieren (Rocca *et al.*, 2017).

Je nach anatomischer Zielstruktur gibt es Programme für die Berechnung von globalem oder lokalem Hirnvolumen, d.h. eines bestimmten Bereichs wie z.B. der subkortikalen Nuclei oder nur der weißen Substanz. Zusätzlich existieren auch Software-Programme zur Messung

der kortikalen Dicke. Globale Messungen (z.B. SIENA oder FreeSurfer) weisen meist eine größere Präzision und statistische „Power“ auf, während anatomische Veränderungen besser auf lokaler Ebene erkennbar sind (z.B. Registration-based cortical thickness oder CIVET) (Rocca *et al.*, 2017).

Die Beschreibung aller existierenden Programme würde den Rahmen dieser Arbeit allerdings sprengen. Daher wird im folgenden Kapitel nur auf die für diese Diplomarbeit relevanten Softwareprogramme eingegangen.

1.9.6 Computerbasierte Messung von Hirnvolumen und Hirnatrophie

Für die Quantifizierung des Hirnvolumens stehen sowohl querschnittsförmige als auch longitudinale Messmethoden zur Verfügung. In diesem Kapitel werden die in dieser Diplomarbeit verwendete Software Programme SIENA und SIENAX beschrieben.

Querschnittsförmige Methoden messen das Volumen von Hirngewebe anhand einer einzigen MRT-Aufnahme, indem der gemessene Wert mit einem Standardwert verglichen wird. SIENAX (Structural Image Evaluation, Using Normalization, of Atrophy Cross-Sectional) ist ein Beispiel für eine solche querschnittsförmige Messmethoden. Als Ergebnis erhält man das sogenannte normalisierte Hirnvolumen (Smith *et al.*, 2002).

Bei den longitudinalen Methoden werden zwei zeitlich voneinander getrennt akquirierte MRT Serien miteinander verglichen, um so Änderung des Volumens zu registrieren. Die Software SIENA (Structural Image Evaluation, Using Normalization, of Atrophy) verwendet diese longitudinale Methode, das Ergebnis beschreibt die Veränderung des Hirnvolumens in Prozent (Smith *et al.*, 2002).

SIENA und SIENAX repräsentieren vollautomatisierte Softwareprogramme und gehören zu der für den akademischen Gebrauch frei zugänglichen FMRIB Software Library FSL (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki>). In beiden Fällen werden normalerweise T1-gewichtete MRT-Scans ohne Kontrastmittel verwendet (Neacsu *et al.*, 2008).

Im Programm SIENA durchlaufen beide MRT-Scans folgende Segmentierungsschritte: Zuerst erfolgt die Trennung zwischen Hirn- und Nicht-Hirngewebe durch das sogenannten BET (Brain Extraction Tool). Dabei wird ein mosaikartiges Netz exakt an die Hirnoberfläche angepasst und über ein auf Intensitätsunterschieden basierendes Histogramm eine binäre Maske (Objekt und Hintergrund) errechnet. Durch diesen Schritt wird das gesamte Hirngewebe extrahiert und für die Messung nicht relevantes Gewebe wie Augenbulbi, Haut,

Fett, Muskeln usw. entfernt. Zusätzlich kann mit diesem Programm die Außenoberfläche des Schädelknochens ermittelt werden (Smith *et al.*, 2004).

Anschließend wird die Geometrie in beiden Bildern durch das FMRIB's Linear Image Registration Tool (FLIRT) aneinander angeglichen (Smith *et al.*, 2002).

Die zweite Segmentation erfolgt mittels des FMRIB's Automated Segmentation Tool (FAST) und dient zur Differenzierung zwischen Liquor, Hirngewebe (auch zwischen grauer und weißer Substanz), potenziellen Läsionen und dem Hintergrund. Dafür werden mindestens drei bis vier Klassen von Voxeln unterschiedlicher Intensitäten je nach Gewebetyp definiert. Schlussendlich werden beide MRT-Aufnahmen miteinander verglichen und über die Berechnung der durchschnittlichen Randverschiebung die prozentuelle Veränderung des Hirnvolumens mit einer Genauigkeit bis zu 0.2% ermittelt (Smith *et al.*, 2004).

Mit der Software SIENAX kann das Hirnvolumen anhand einer einzelnen MRT-Aufnahme berechnet werden. Der Algorithmus ist ähnlich dem von SIENA, anstatt des zweiten Scans wird ein Standardbild für die Berechnung des totalen Hirnvolumens verwendet, welches multipliziert mit dem sogenannten „volumetric scaling factor“ das normalisierte Hirnvolumen (normalized brain volume) ergibt (Smith *et al.*, 2002; Durand-Dubief *et al.*, 2012; Popescu, 2015, p. 12; Rocca *et al.*, 2017).

Programme wie SIENA eignen sich besser für longitudinale Studien, da die Volumenänderung im Verlauf ermittelt wird. Eine Software wie SIENAX hingegen wird vor allem in querschnittsförmigen Studien verwendet (De Stefano *et al.*, 2014).

Manche Läsionen sind auf T1-gewichteten MRT-Aufnahmen ähnlich hypointens und auf T2-gewichteten Bildern ähnlich hyperintens wie das Gewebe der grauen Substanz. Aus diesem Grund durchlaufen die Scans, unabhängig von der für die Atrophiemessung verwendeten Software, vorher meist noch den Schritt des sogenannten „Lesion-filling“ (Auffüllen der Läsionen). Dabei werden die Voxel der Läsionen markiert und durch weiße Substanz ersetzt, damit anschließend eine genauere Quantifizierung der grauen Substanz möglich ist. Für diese Bildbearbeitung wird eine Läsions-Maske auf dem T1-gewichteten Bild benötigt, für T2-gewichtete Bilder steht die sogenannte Lesion Segmentation Toolbox Software zu Verfügung (Popescu, 2015, pp. 11–21; Pareto *et al.*, 2016; Rocca *et al.*, 2017).

1.9.7 Messfehler und Störfaktoren

Die MRT-basierte Messung von Hirnvolumen gilt im Allgemeinen als sensitive und reproduzierbare Methode zur Erfassung von Hirnatrophie im Verlauf der MS. Dennoch sollten bei der Interpretation der Daten potenzielle Störfaktoren miteinbezogen werden, da diese die Ergebnisse beeinflussen und zu einer Über- oder Unterschätzung des Hirnvolumens führen können (De Stefano *et al.*, 2014).

Diese Störgrößen können physiologischer Natur, durch die Krankheit selbst bedingt oder vom verwendeten MRT-Gerät bzw. dem Software-Programmen abhängig sein (Amiri *et al.*, 2018).

Zu den physiologischen Faktoren zählen unter anderem typische Einflussgrößen wie Alter, Geschlecht und Schädelgröße (gesamtes Hirnvolumen), die in den meisten klinischen Studien bereits berücksichtigt werden (Barnes *et al.*, 2010).

Da computerbasierte Messungen in der Lage sind, sehr kleine Volumenänderungen von 0.2-0.13% des Hirnvolumens zu detektieren, können auch der Hydratationsstatus der PatientInnen und die kontinuierliche Abnahme des Hirnvolumens über den Tag das Ergebnis verändern (Duning *et al.*, 2005; Nakamura *et al.*, 2015).

Zusätzlich haben Lifestyle-Faktoren (Rauchen und Alkohol), genetische Disposition (Expression von Apolipoprotein E) oder Komorbiditäten wie Diabetes oder kardiovaskuläre Erkrankungen eine potenzielle Auswirkung auf die Messung (De Stefano *et al.*, 2014).

Pathologische Veränderungen außerhalb der Zielgröße, die durch die Krankheit selbst bedingt sind, wirken sich ebenfalls auf die computerbasierte Atrophie-Messung aus. Neben den bereits erwähnten Läsionen der weißen Substanz zählen auch Veränderungen der grauen Substanz und paradoxerweise auch die Hirnatrophie selbst dazu, da diese Zwischenschritte des Software-Algorithmus beeinflussen kann (Amiri *et al.*, 2018).

Einige krankheitsmodifizierende Medikamente (DMT), wie Natalizumab, Fingolimod und hohe Dosen von Interferon-beta, können in den ersten Monaten der Therapie eine vermeintliche Reduktion von Hirnvolumen bewirken. Dieser als Pseudoatrophie bekannte Effekt betrifft vor allem die weiße Substanz und ist wahrscheinlich eine Folge der anti-entzündlichen Wirkung der DMTs (Zivadinov *et al.*, 2008; Vidal-Jordana *et al.*, 2015).

Nicht zuletzt hängt die Atrophie-Messung auch von den gegebenen technischen Voraussetzungen ab, sowohl auf der Ebene der Bildentstehung wie auf jener der Bildbearbeitung.

Studien berichten von Schwankungen der Werte für das Hirnvolumen bei Verwendung von verschiedenen MRT-Scannern, aber auch bei Verwendung desselben Scanners (Biberacher *et al.*, 2016; Zivadinov *et al.*, 2018). Vergleiche zwischen den verschiedenen Softwareprogrammen zur Datenbearbeitung ergaben auch hier unterschiedliche Werte (Storelli *et al.*, 2018). Eine größere Genauigkeit erzielt man demzufolge mit Programmen wie SIENA, die Segmentierungsverfahren verwenden, als mit Bildregistrierungsverfahren wie z.B. SIENAX (Durand-Dubief *et al.*, 2012).

Einflussfaktoren, die den Scan-Vorgang direkt betreffen, wie beispielsweise die Positionierung der PatientInnen (Caramanos *et al.*, 2010) im Gerät oder Kopfbewegungen (Reuter *et al.*, 2015) während des Aufnahmevorgangs, werden ebenfalls zu den potentiellen Störfaktoren gezählt.

1.9.8 Visuelles Rating von Hirnatrophie

In den letzten zwei Jahrzehnten hat sich die klinische Relevanz von Hirnatrophie bei MS-PatientInnen in einer Vielzahl von Studien bestätigt. Stetig werden neue Methoden für eine genauere und verlässlichere Beurteilung von Hirnvolumen entwickelt. Dennoch ist man von der Anwendung im klinischen Alltag noch weit entfernt (Rocca *et al.*, 2017).

Dafür wären vor allem schnelle, einfache, verlässliche und kostengünstige Methoden notwendig. Diese Voraussetzungen könnte ein visuelles Rating von MRT-Aufnahmen durch den Neurologen bzw. die Neurologin erfüllen. Studien mit AlzheimerpatientInnen haben für visuelle Rating-Skalen im Vergleich zu computerbasierten Techniken bereits sehr gute und genaue Ergebnisse erzielt (Wahlund, 2000; Bresciani *et al.*, 2005; Harper *et al.*, 2016).

In der Literatur wurde die Verwendung von solchen semiquantitative Rating-Skalen zur visuellen Erfassung von Läsionen der weißen Substanz bereits diskutiert (Fazekas *et al.*, 2002; Prins *et al.*, 2004) und deren Ergebnisse mit computerbasierten Methoden verglichen (Olsson *et al.*, 2013).

Derzeit gibt es keine eigene Rating-Skala für die Beurteilung von Hirnatrophie in MS PatientInnen. Das im Rahmen dieser Diplomarbeit entworfene Rating Sheet, beinhaltet neben anderen Parametern die quantitative Skala des SIFAP-Ratings zur Beurteilung von Ventrikel- und Sulcusweite. Eine genauere Beschreibung findet sich im Methodenteil.

1.10 Fragestellung und Zielsetzung

Neben Demyelinisierungsprozessen und Entzündungsreaktionen prägen auch Neurodegeneration und axonaler Verlust das Krankheitsbild von MS (De Stefano *et al.*, 2014). Die daraus resultierende Hirnatrophie tritt bereits in den frühen Stadien der Erkrankung auf und setzt sich über den gesamten Verlauf der MS fort (De Stefano *et al.*, 2014; Popescu *et al.*, 2015; Filippi *et al.*, 2016; Kappos *et al.*, 2016).

In Studien wurde eine Korrelation zwischen dem Ausmaß an Hirnatrophie und dem klinischen Schweregrad sowie dem Auftreten von kognitiven Funktionsstörungen und neurologischer Beeinträchtigung gefunden (Rocca *et al.*, 2017). Dies macht die Messung der Hirnatrophie zu einem potenziellen Marker für Krankheitsprognose und Krankheitsverlauf (NEDA-4 Konzept). Zudem könnte dadurch auch die Wirksamkeit einer Therapie abgeschätzt werden (Kappos *et al.*, 2016).

Bisher gestaltet sich die systematische Erfassung des Hirnvolumens als zeit- und ressourcenaufwändiges Verfahren und scheint daher oftmals weitgehend ungeeignet für eine Eingliederung in den klinischen Alltag zu sein. In dieser Arbeit soll daher untersucht werden, ob und ab welchem Ausmaß der Hirnsubstanzänderung ein visuelles Rating mit modernen, computerbasierten Messungen (SIENA Structural Image Evaluation, Using Normalization, of Atrophy Cross- Sectional) korreliert.

Das Ziel der vorliegenden Arbeit ist es, die Sensitivität eines visuellen Ratings von Hirnatrophie zu prüfen und mit Software-basierten Messmethoden zu vergleichen. Zudem wird versucht, einen möglichen Schwellenwert („Cut-off“) zu definieren, ab dem Hirnatrophie visuell sicher erfasst werden kann. Dabei soll auch der Frage nach einer Anwendung eines solchen radiologischen „Eye-balling“ im klinischen Alltag zu nachgegangen werden.

Diese Fragestellungen wurden an einer Kohorte von 120 PatientInnen mit gesicherter MS Diagnose untersucht. Der Datensatz stammt aus der MS Ambulanz der Universitätsklinik für Neurologie Graz. Die visuellen Rating-Systeme basieren auf jenen, die in den sogenannten LADIS und der SIFAP Studien angewandt worden sind (Pantoni *et al.*, 2005; Fazekas *et al.*, 2013), für die computerbasierte Messung wurden die Programme SIENA und SIENAX verwendet (Jokinen *et al.*, 2012; Pichler *et al.*, 2016).

2 Material und Methoden

2.1 Patientinnen und Patienten

Die PatientInnen Daten für diese Diplomarbeit stammen von einer Studie an der Abteilung für Allgemeine Neurologie der Medizinischen Universität Graz (Pichler *et al.*, 2016). Dabei wurden PatientInnen der MS-Ambulanz, die sich zwischen 2006 und 2012 sowohl einem Baseline als auch einem Follow-up MRT-Scan unterzogen haben, ausgewählt. Alle Aufnahmen stammen von demselben 3-Tesla MRT-Scanner und folgen einem einheitlich festgelegten Protokoll. Weitere Einschlusskriterien waren:

1. Ein diagnostiziertes CIS (klinisch isoliertes Syndrom) mit Verdacht auf MS (Miller *et al.*, 2008) ($n = 63$) oder eine definitive MS-Diagnose nach den McDonald-Kriterien (RRMS = 53 und SPMS = 4) (Polman *et al.*, 2005);
2. Eine genaue neurologische Untersuchung, einschließlich der EDSS, beim Baseline und beim Follow-up MRT-Scan (Kurtzke, 1983);
3. Eine vollständig dokumentierte Krankengeschichte, inklusive der Therapie mit DMT;
4. Mindestens 18 Monate Zeitabstand zwischen Baseline und Follow-up Scan;
5. Mindestens 8 Wochen Abstand zwischen der Gabe von Steroiden und der MRT-Aufnahme, um dadurch bedingte Volumenschwankungen zu vermeiden.

Nach diesen Kriterien wurden 120 PatientInnen in die Studie eingeschlossen (Pichler *et al.*, 2016).

Diese Studie wurde von Seiten der Ethikkommission der Medizinischen Universität Graz genehmigt. Die zugehörige EK-NR lautet 17-046 ex 05/06.

2.2 MRT-Protokoll

Alle MRT-Aufnahmen stammen vom gleichen 3-Tesla-Scanner (Siemens Tim Trio, Siemens Healthcare, Erlangen, Germany). Es wurde eine phased-array Kopfspule mit 12 Empfängerelementen nach einem einheitlichen MRT Protokoll verwendet.

Die Aufnahmen beinhalteten eine FLAIR (fluid-attenuated inversion recovery) Sequenz (Repetitionszeit (TR = repetition time) zu Echozeit (TE = echo time) zu Inversionszeit (TI = inversion time) gleich 9000 ms/69 ms/2500 ms und einen inplane Resolution von $0.9 \times$

0.9 mm) mit einer Schichtdicke von 3mm und einer Aufnahmezeit von 4 Minuten und 22 Sekunden und eine T1-gewichtete 3-dimensionale (3D) MPRAGE (magnetization-prepared rapid gradient echo) Sequenz, mit einer Isotropen-Auflösung von 1mm (TR/TE/TI/Kippwinkel gleich 1900 ms/2.19 ms/900 ms/9°) und einer Aufnahmezeit von 6 Minuten und 1 Sekunde (Pichler *et al.*, 2016).

2.3 Visuelles Rating

Für das visuelle Rating der MRT Bilder wurde ein eigens dafür konzipiertes Protokoll verwendet (siehe Anhang: Rating Sheet). Die Auswertung wurde von Seiten eines Neurologen und einer Medizinstudentin durchgeführt. Die RaterInnen waren während des gesamten Ratings bezüglich Baseline und Follow-up Aufnahme verblindet und erhielten den Datensatz in randomisierter Form, ohne Kenntnis der Reihenfolge der Scans. Zum Aufrufen der MRT-Scans wurde das Programm Mango 4.0.1 (Research Imaging Institute, UTHSCSA) verwendet.

Im erste Rating-Schritt erfolgte die Zuordnung der beiden MRT-Aufnahme zum Untersuchungszeitpunkt (entweder Baseline oder Follow-up Scan) und die Einschätzung des Sicherheitsgrads dieser Zuordnung (möglicher – sicherer – kein Unterschied). Anschließend wurde die Bild-Qualität beider Scans beurteilt (gut – ausreichend – nicht beurteilbar).

Basierend auf der SIFAP-Ratings Skala (siehe SIFAP-Rating Skala) wurden in der axialen Schnittbildansicht die Weite des 2. Ventrikels und die der Sulci der Großhirnrinde auf beiden Aufnahmen bewertet. Je nach Ausprägung sieht diese Skala eine Einteilung der Ventrikel- bzw. sulcalen Weite von 1 bis 8 vor, wobei 1 für schmale und 8 für sehr breite Ventrikel bzw. Sulci steht.

Zusätzlich hatten die RaterInnen die globale Änderung von Ventrikel- und sulcaler Weite zwischen Baseline und Follow-up Scan in „keine, gering, mäßig oder ausgeprägt“ einzustufen.

Der letzte Punkt beinhaltete die Ausmessung der linearen Ventrikelweite. Für den III-Ventrikel wurde die axialen Schnittbildebene, in der dieser am besten und größten sichtbar ist, eingestellt und dessen Weite auf Höhe des anteroposterioren Mittelpunkts gemessen. Die Weite der Seitenventrikel wurden auf der Schnittebene mit der dünnsten Darstellung des Septum pellucidum bestimmt. Als Referenz diente jeweils das Hirnparenchym entlang derselben Linie der Messungen (d.h. der Abstand zwischen zwei Punkten auf der Kortexoberfläche auf Höhe der Messung der Ventrikelweite) (Simon *et al.*, 1999).

Zur Messung wurde ein im Bildprogramm integriertes Lineal [mm] verwendet. Siehe Abbildung 2 und Abbildung 3.

Abbildung 2: T1-gewichteter MRT-Scan eines ausgewählten Patienten bzw. einer ausgewählten Patientin in axialer Schnittebene. Die Pfeile markieren die Höhe der Ausmessung des 2. Ventrikels an der dünnsten Stelle des Septum pellucidum (Doppelpfeil) und des gesamten Gehirns auf derselben Höhe (rechter und linker Pfeil)

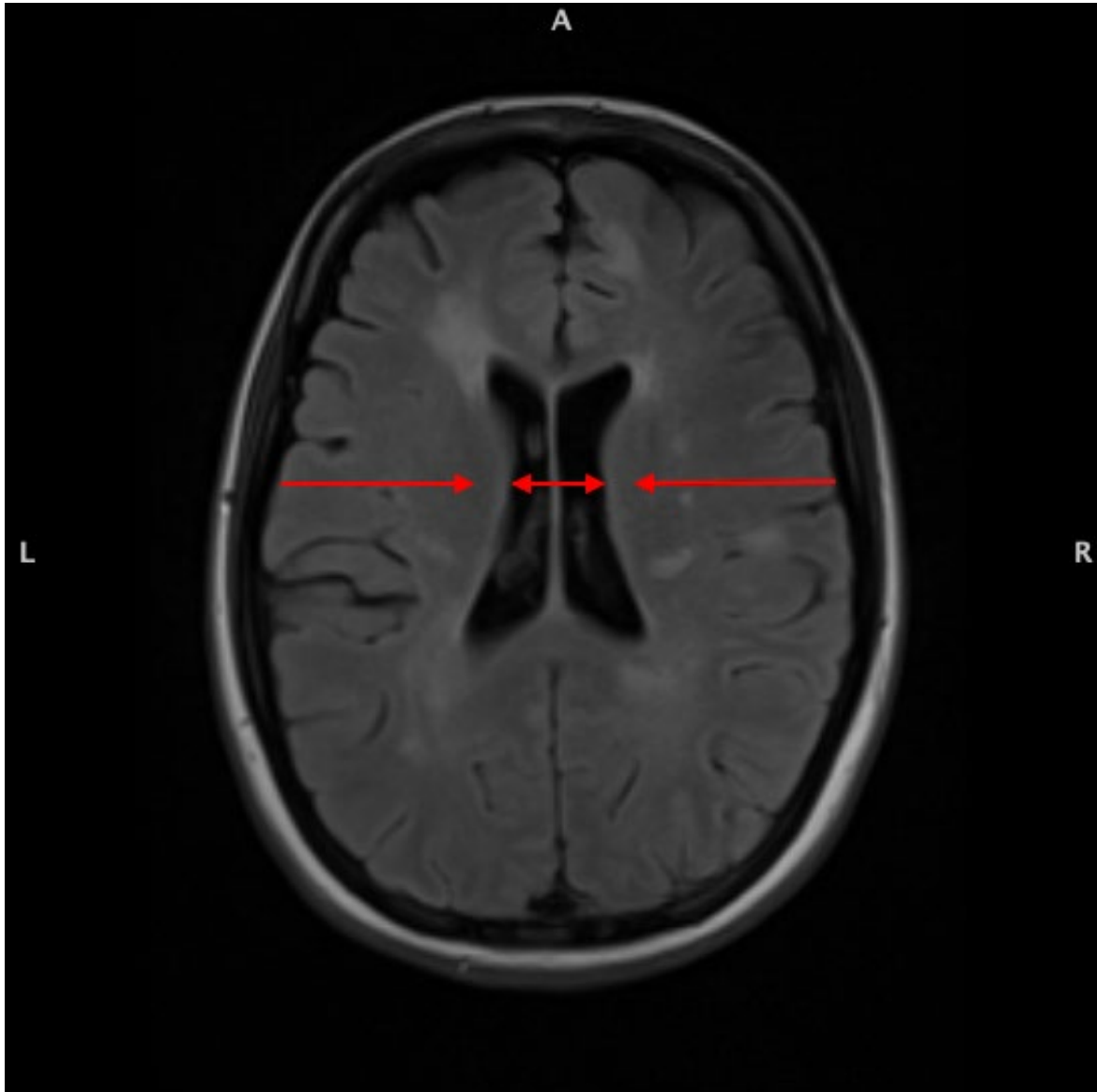
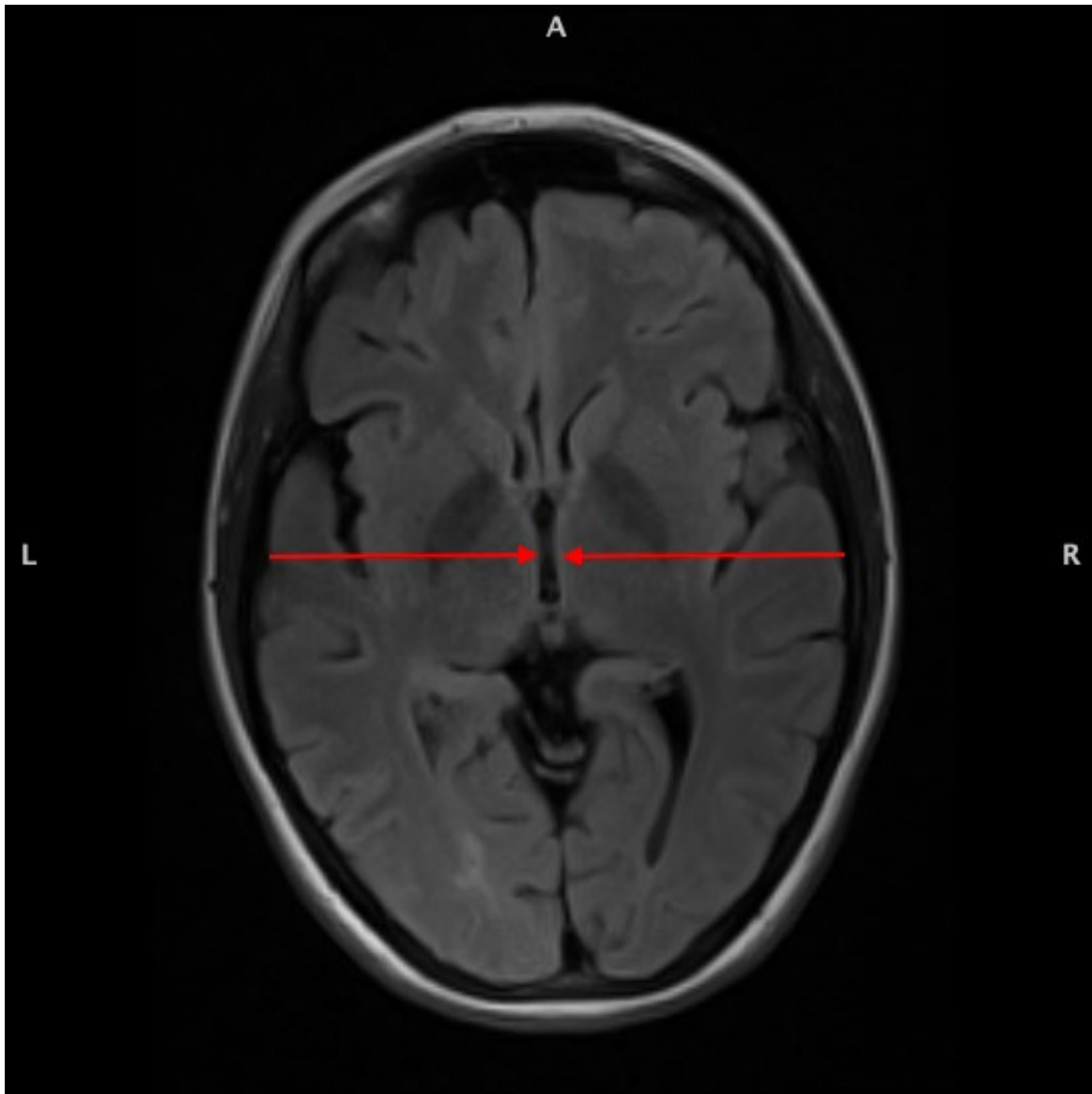


Abbildung 3: T1-gewichteter MRT-Scan eines ausgewählten Patienten bzw. einer ausgewählten Patientin in axialer Schnittebene. Die Pfeile markieren die Höhe der Ausmessung des 3. Ventrikels am anteroposterioren Mittelpunkt.



Nach dem oben beschriebenen Verfahren wurden alle Aufnahmen von einer Medizinstudentin mindestens einmal ausgewertet. Nach einem Zeitraum von über einem Monat bewertete die Medizinstudentin die Scans von 10 zufällig ausgewählten PatientInnen ein weiteres Mal. Als Vergleich wurden zusätzlich noch die MRT-Scans von insgesamt 21 wiederum zufällig ausgewählten PatientInnen von Seiten eines Neurologen ausgewertet.

2.4 SIENA-Auswertung

Für die computerbasierte Messung der prozentuellen Abnahme des Gehirnvolumens über die Zeit (pBVC), d.h. zwischen Baseline und Follow-up Scan, wurde das Programm SIENA verwendet. Das globale Hirnvolumen in einem einzelnen Scan (Normalized Brain Volume, NBV) konnte mit Hilfe von SIENAX berechnet werden. Diese Software-Programme wurde an der Oxford Universität in England entwickelt und sind Teil der für den akademischen Gebrauch frei zugänglichen FSL (FMRIB Software Library <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL>) (Smith *et al.*, 2004).

Die Daten dieser Messungen stammen aus einer Publikation der Abteilung für Allgemeine Neurologie der Medizinischen Universität Graz (Pichler *et al.*, 2016).

2.5 Statistik

Für die statistische Auswertung wurde IBM SPSS Statistics® Version 25 verwendet. Die demographische Beschreibung der Population erfolgte mittels deskriptiver Statistik.

Die Daten wurden anhand von Histogrammen und des Kolmogorov-Smirnov Test auf Normalverteilung geprüft. Normalverteilte Daten wurde mit dem Student's t-Test oder dem Welch-Test, im Falle von ungleichen Varianzen, und der Varianzanalyse ANOVA untersucht. Folgten die Daten keiner Normalverteilung, so wurde der Mann Whitney U Test und der Kruskal Wallis Test angewandt.

Unterschiede zwischen qualitativen Gruppen wurden in Kreuztabellen durch den Chi-Quadrat Test bzw. den Fisher Exakt Test analysiert.

Um die Stärke des Inter- und Intrarater Agreement zu ermitteln wurde der kappa-Koeffizient berechnet (Altman Douglas G., 1990, pp. 403–406).

3 Ergebnisse

3.1 Demographische und klinische Daten der Population

Die in dieser Diplomarbeit untersuchte Kohorte umfasste 120 PatientInnen, davon waren 65% (n = 78) weiblich und 35% (n = 42) männlich. Durchschnittlich lag die klinische Erstmanifestation der Erkrankten im Alter von 29 (\pm 9.1) Jahren, das Durchschnittsalter beim Baseline Scan lag bei 33.8 (\pm 9.4) Jahren.

Zum diesem Zeitpunkt wiesen 52.5% (n = 63) der PatientInnen ein CIS und 47.5% (n = 57) eine definitive MS auf; Innerhalb der MS-Gruppe verteilten sich die Subtypen folgendermaßen: 43.3% (n = 52) RRMS und 3.3% (n = 4) SPMS. Ein Fall (0.8%) litt an einer PPMS. Die mittlere Schubrate der RRMS-PatientInnen pro Jahr lag bei 0.29 (\pm 0.33). Während eines durchschnittlichen Scan-Intervalls von 43 (30, 53.8) Monaten entwickelten 27.5% (n = 33) der CIS-PatientInnen eine definitive MS. Der mittlere EDSS-Wert veränderte sich von 1.75 (1, 2.5) auf 1.5 (0, 2.5) und die Anzahl der PatientInnen mit einer DMT stieg von 38.3% (n = 46) auf 66.7% (n = 80) (siehe Tabelle 6).

In Tabelle 7 ist die Verteilung der Volumenmaße innerhalb der gesamten Kohorte und der verschiedenen Untergruppen aufgelistet. Das pBVC wurde mit Hilfe des Software Programms SIENA bestimmt, das NBV mit dem Programm SIENAX. Alle diese Daten wurden aus der Studie von Pichler et al. 2016 (Pichler *et al.*, 2016) übernommen.

Tabelle 6: Demographische und klinische Daten

Gesamt, n	120 (100%)
Geschlecht	
weiblich, n (%)	78 (65%)
männlich, n (%)	42 (35%)
Alter, Jahre	
Erkrankungsbeginn	29 ± 9,1
Baseline Scan	33.8 ± 9.4
Klinischer Phänotyp zum Baseline Scan	
CIS, n (%)	63 (52.5%)
MS, n (%)	57 (47.5%)
RRMS, n (%)	52 (43.3%)
SPMS, n (%)	4 (3.3%)
PPMS, n (%)	1 (0.8%)
Klinischer Phänotyp zum Follow-up Scan	
CIS (non- converter), n (%)	30 (25%)
Converter to MS, n (%)	33 (27.5%)
MS Phänotyp zum Follow-up Scan	
RRMS, n (%)	81 (67.5%)
SPMS, n (%)	8 (6.7%)
PPMS, n (%)	1 (0.8%)
EDSS	
Baseline Scan	1.75 (1, 2.5)
Follow-up Scan	1.5 (0, 2.5)
DMT	
Baseline Scan, n (%)	46 (38.3%)
Follow-up Scan, n (%)	80 (66.7%)
Annualisierte Schubrate bei RRMS	0.29 ± 0.33
Intervall zwischen Baseline und Follow-up Scan, 43 (30, 53.8) Monate	
<p>Normalverteilte Daten werden als Mittelwert (± Standardabweichung), nicht normalverteilte Daten als Median (Interquartilsabstand) angegeben. n = Anzahl, CIS = klinisch isoliertes Syndrom, MS = Multiple Sklerose, RRMS= Schubförmige Multiple Sklerose (Relapsing Remitting Multiple Sclerosis), SPMS = Sekundär chronisch progredient Multiple Sklerose (secondary progressive Multiple Sclerosis), PPMS = Primär progredient Multiple Sklerose (primary progressive multiple sclerosis), EDSS = Expanded Disability Status Scale, DMT = krankheitsmodifizierende Therapie (Disease modifying therapies).</p>	

Tabelle 7: Verteilung Volumenmaße SIENA/SIENAX

	pBVC gesamt	% annual pBVC%	NBV BI [cm ³]	NBV FU [cm ³]	delta NBV [cm ³]	delta NBV annual [cm ³]
alle (n = 120)	-1.25 (-2.44, -0.67)	-0.42 (-0.65, -0.19)	1555.51 (1496.96, 1611.40)	1545.39 (1480.55, 1590.83)	-18.74 (-39.17, -1.82)	-5.18 (-10.77, -0.54)
männlich (n = 42)	-1.09 (-2.44, -0.56)	-0.37 (-0.63, -0.18)	1549.17 (1471.19, 1590.12)	1536.28 (1450.37, 1579.06)	-21.94 (-39.68, -1.24)	-6.60 (-10.92, -0.40)
weiblich (n = 78)	-1.28 (-2.48, -0.71)	-0.49 (-0.65, -0.20)	1563.35 (1513.49, 1619.95)	1549.53 (1493.75, 1603.45)	-17.56 (-38.87, -2.29)	-4.94 (-10.73, -0.73)
Alter BI Scan ≤33a (n = 83)	-1.41 (-2.48, -0.70)	-0.42 (-0.67, -0.19)	1561.08 (1496.59, 1617.72)	1546.96 (1487.60, 1606.46)	-20.56 (-39.58, -1.65)	-5.19 (-11.31, -0.53)
Alter BI Scan >33a (n = 37)	-1.09 (-2.44, -0.57)	-0.37 (-0.61, -0.17)	1554.89 (1494.41, 1589.91)	1544.60 (1469.09, 1584.28)	-10.93 (-32.87, -2.25)	-5.06 (-10.22, -0.61)
CIS BI (n = 63)	-1.05 (-2.32, -0.54)	-0.36 (-0.58, -0.15)	1567.30 (1524.12, 1626.64)	1553.58 (1501.43, 1590.95)	-17.25 (-39.98, -2.04)	-4.83 (-10.63, -0.59)
MS BI (n = 57)	-1.64 (-2.64, -0.83)	-0.45 (-0.77, -0.22)	1532.79 (1470.08, 1598.33)	1515.34 (1446.74, 1591.95)	-22.33 (-39.02, -1.70)	-6.12 (-11.03, -0.53)
non- converter (n = 30)	-0.73 (-1.29, -0.09)	-0.22 (-0.43, -0.02)	1563.11 (1514.12, 1594.00)	1546.68 (1497.13, 1590.35)	-13.66 (-29.77, -2.36)	-4.51 (-9.79, -0.62)
converter (n = 33)	-1.36 (-2.66, -0.63)	-0.45 (-0.69, -0.19)	1574.03 (1524.70, 1649.20)	1560.25 (1521.87, 1604.45)	-19.63 (-48.25, 6.22)	-4.83 (-12.46, 2.24)
DMT BI (n = 46)	-1.47 (-2.55, -0.73)	-0.44 (-0.72, -0.19)	1516.98 (1473.80, 1588.90)	1496.65 (1450.98, 1554.04)	-20.18 (-33.14, -1.72)	-5.03 (-11.28, -0.53)
DMT FU (n = 80)	-1.54 (-2.61, 0.775)	-0.44 (-0.68, -0.20)	1545.09 (1489.59, 1597.95)	1536.28 (1459.94, 1578.22)	-22.83 (-39.51, -1.94)	-6.09 (-11.30, -0.56)

Daten nicht normalverteilt, als Median (Interquartilsabstand) angegeben.

n = Anzahl, pBVC = prozentuelle Änderung des Hirnvolumens (percentage of brain volume change), annual pBVC = pBVC pro Jahr, NBV = normalisiertes Hirnvolumen (normalized brain volume), BI = Baseline Scan, FU = Follow-up Scan, delta NBV = Änderung des NBV zwischen BI und FU Scan, CIS = klinisch isoliertes Syndrom, MS = Multiple Sklerose, DMT = krankheitsmodifizierende Therapie (Disease modifying therapies).

3.2 Ergebnisse des visuellen Atrophie-Ratings

Beim visuellen Atrophie-Rating der insgesamt 120 PatientInnen wurde der Untersuchungszeitpunkt in 90 Fällen (75%) richtig zugeordnet, 13 Fälle (11%) wurden falsch zugeordnet und bei 17 PatientInnen (14%) war es nicht möglich zwischen Baseline und Follow-up Scan zu unterscheiden ($p < 0.001$) (siehe Tabelle 8).

Tabelle 8: Ergebnis Zuordnung zum Untersuchungszeitpunkt

Zuordnung der Scans zu einem Untersuchungszeitpunkt (Baseline, Follow-up) n = 120 (100%)	
Zuordnung richtig	90 (75%)
Zuordnung falsch	13 (11%)
keine Zuordnung	17 (14%)
n = Anzahl	

Unter dem ersten Punkt des Auswertungsbogens sollte eine Aussage zum „Sicherheitsgrad der Zuordnung zum Untersuchungszeitpunkt“ anhand des visuellen Unterschieds zwischen den beiden Scans getroffen werden (siehe Rating Sheet).

Dieser wurde bei den richtig zugeordneten Scans in 10 (11%) Fällen als „sicher“ und in 80 (89%) Fällen als „möglich“ klassifiziert. Bei allen falsch zugeordneten MRT-Scans ($n = 13$) wurde der Unterschied als „möglich“ eingestuft und bei 17 MRT-Untersuchungen konnte keine Zuordnung vorgenommen werden.

Die Bildqualität der richtig zugeordneten Baseline Scans wurde bei 65 (72%) als gut und bei 25 (28%) als ausreichend beurteilt, bei den Follow-up Scans waren es jeweils 85 (94%) und 5 (6%) Untersuchungen.

In 92% ($n = 12$) der falsch zugeordneten Baseline MRT-Untersuchungen wurde die Bildqualität als gut und in einem Fall als ausreichend eingestuft. Bei den falsch zugeordneten Follow-up Scans wurde die Bildqualität in 10 Fällen (77%) als gut und in 3 Fällen (23%) als ausreichend bewertet. Ein Fall aus den als „nicht zuordenbar“ klassifizierten Scans konnte aufgrund der schlechten Bildqualität nicht beurteilt werden, bei den restlichen 16 Scans (94%) wurde die Bildqualität als ausreichend eingestuft.

Die Ventrikelweite und Sulcusweite wurde anhand der acht Kategorien der SIFAP Rating-Skala (siehe Rating Sheet) beurteilt. Für die Auswertung wurden diese acht Kategorien zusätzlich noch in zwei Gruppen (Kategorie 1-4 und Kategorie 5-8) zusammengefasst. Neben der Beurteilung der einzelnen Scans wurde als letzter Punkt noch eine globale Änderung der Ventrikel- und Sulcusweite anhand von 4 qualitativen Kategorien bewertet

(Tabelle 9). Sowohl die Ventrikelweite als auch die Sulcusweite und deren globale Änderung konnten nur in 119 Fällen bestimmt werden, da bei einem Scan-Paar die Qualität der Bilder als nicht beurteilbar eingestuft worden ist.

Während des visuellen Ratings wurden zusätzlich noch bei jedem Scan die Weite des III-Ventrikels und der Seitenventrikel gemessen. Der zusätzliche Datenumfang dieser „linearen Maße“ würde den Rahmen der Diplomarbeit jedoch sprengen, weshalb auf deren Auswertung verzichtet worden ist.

Tabelle 9: Ergebnisse des visuellen Ratings

	Zuordnung richtig n = 90 (100%)	Zuordnung falsch n = 13 (100%)	keine Zuordnung n = 17 (100%)	p-Wert
Sicherheitsgrad der Zuordnung (n = 120)				
möglicher Unterschied	80 (89)	13 (100)	0	< 0,001 ^a
sicherer Unterschied	10 (11)	0	0	
kein Unterschied	0	0	17 (100)	
Bildqualität Baseline Scan (n = 120)				
gut	65 (72)	12 (92)	16 (94)	NS ^a
ausreichend	25 (28)	1 (8)	0	
nicht beurteilbar	0	0	1 (6)	
Bildqualität Follow-up Scan (n = 120)				
gut	85 (94)	10 (77)	16 (94)	NS ^a
ausreichend	5 (6)	3 (23)	0	
nicht beurteilbar	0	0	1 (6)	
Rating Ventrikelweite (n = 119)				
Baseline Scan				
V1	25 (28)	6 (46)	4 (25)	NS ^a
V2	23 (26)	3 (23)	5 (31)	
V3	19 (21)	3 (23)	5 (31)	
V4	12 (13)	0	1 (6)	
V5	10 (11)	1 (8)	1 (6)	
V6	1 (1)	0	0	
V7	0	0	0	
V8	0	0	0	
V1	25 (28)	6 (46)	4 (25)	
Baseline Scan zusammengefasst				
V1-V4	79 (88)	12 (92)	15 (94)	NS ^a
V5-V8	11 (12)	1 (8)	1 (6)	
Follow-up Scan				
V1	3 (3)	1 (8)	4 (25)	NS ^a
V2	22 (24)	6 (46)	5 (31)	
V3	24 (27)	2 (15)	5 (31)	
V4	21 (23)	3 (23)	1 (6)	
V5	9 (10)	1 (8)	1 (6)	

V6	10 (11)	0	0	
V7	1 (1)	0	0	
V8	0	0	0	
Follow-up Scan zusammengefasst				
V1-V4	70 (78)	12 (92)	15 (94)	NS ^a
V5-V8	20 (22)	1 (8)	1 (6)	
Rating Sulcusweite (n = 119)				
Baseline Scan				
S1	15 (17)	N.A.	2 (13)	NS ^a
S2	37 (41)	5 (39)	5 (31)	
S3	19 (21)	6 (46)	6 (38)	
S4	11 (12)	2 (15)	2 (13)	
S5	2 (2)	0	1 (6)	
S6	2 (2)	0	0	
S7	2 (2)	0	0	
S8	2 (2)	0	0	
Baseline Scan zusammengefasst				
S1-S4	82 (91)	13 (100)	15 (94)	NS ^a
S5-S8	8 (9)	0	1 (6)	
Follow-up Scan				
S1	0	0	2 (12.5)	NS ^a
S2	16 (18)	2 (15)	5 (31)	
S3	34 (38)	4 (31)	6 (38)	
S4	20 (22)	6 (46)	2 (12.5)	
S5	10 (11)	1 (8)	1 (6)	
S6	3 (3)	0	0	
S7	3 (3)	0	0	
S8	4 (4)	0	0	
Follow-up Scan zusammengefasst				
S1-S4	70 (78)	12 (92)	15 (94)	NS ^a
S5-S8	20 (22)	1 (8)	1 (6)	
Globale Änderung Ventrikelweite (n = 119)				
keine	7 (8)	3 (23)	16 (100)	< 0,001 ^a
gering	80 (89)	10 (77)	0	
mäßig	2 (2)	0	0	
ausgeprägt	1 (1)	0	0	
Globale Änderung Sulcusweite (n = 119)				
keine	7 (8)	4 (31)	16 (100)	< 0,001 ^a
gering	72 (80)	9 (69)	0	
mäßig	10 (11)	0	0	
ausgeprägt	1 (1)	0	0	
Absoluter und prozentueller (spaltenweise) Wert angeben.				
^a Chi Quadrat Test				
n = Anzahl, NS = nicht signifikant				

3.2.1 Inter-Rater Variability

Von 21 zufällig ausgewählten PatientInnen wurden die MRT-Untersuchungen auch von einem zweiten Rater ausgewertet. Um die Übereinstimmung zwischen den beiden Ratern, das sogenannte Interrater Agreement, zu untersuchen, wurde einerseits der Chi Quadrat bzw. der Fisher Exakt Test durchgeführt und andererseits der kappa-Koeffizient berechnet.

Beim Punkt "Zuordnung zu Untersuchungszeitpunkt" stimmten beide Rater in 85.7% (n = 18) überein. Davon waren 88.9% (n = 16) richtig zugeordnet und 11.1% (n = 2) falsch zugeordnet (siehe Abbildung 4). Der Chi Quadrat Test ergab keinen statistisch signifikanten Unterschied zwischen den Ergebnissen der Rater. Der kappa-Koeffizient von 0.36 (p = 0.004) zeigte eine schwache Übereinstimmung an.

Ventrikelweite und Sulcusweite wurden für diese Analyse in 2 Gruppen unterteilt: V1-V4 bzw. S1-S4 wurde als geringe Ventrikel- bzw. Sulcusweite und V5-V8 bzw. S5-S8 als große Ventrikel- bzw. Sulcusweite klassifiziert.

Betrachtet man die Übereinstimmung zwischen den Ratern bezüglich dieser Kategorien, so erhält man für die Ventrikelweite der Baseline und Follow-up Scan mit dem Fisher Exakt Test einen statistisch signifikanten Unterschied. Der kappa-Wert beschreibt für die Ventrikelweite zum Baseline Scan (k = 0.89) und für die Ventrikelweite zum Follow-up Scan (k = 0.81) eine sehr gute Übereinstimmung.

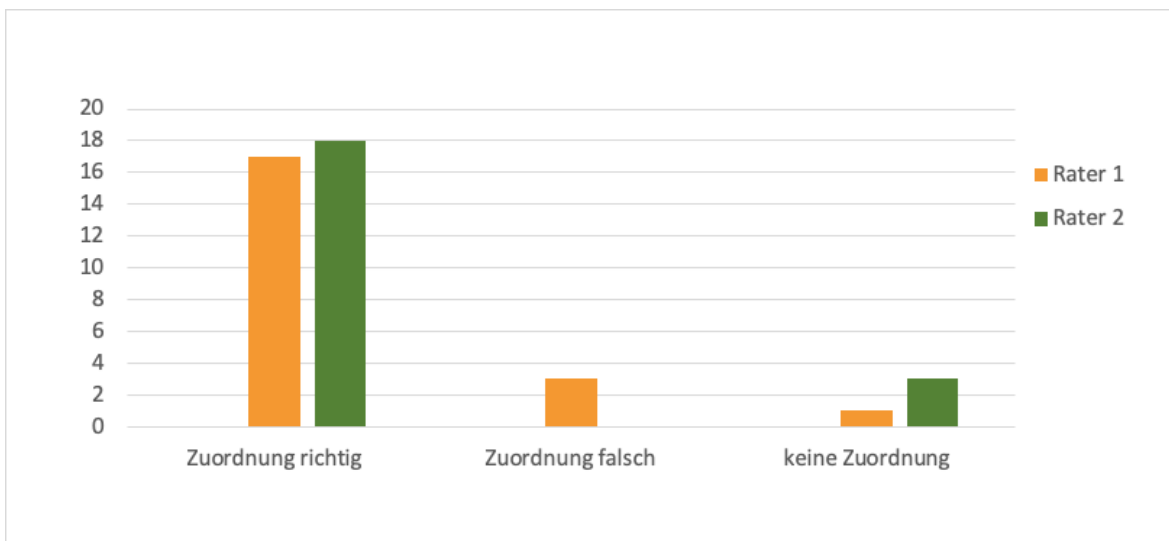
Bei der Bewertung der Sulcusweite ergab sich im Fisher Exakt Test nur für die Follow-up Scans ein statistisch signifikanter Unterschied (p = 0.019). Der k-Koeffizient dafür lag bei 0.46 (moderate Stärke der Übereinstimmung). Für die Bewertung des Baseline Scans ergab sich ein k-Koeffizient von 0.42, welcher ebenfalls eine moderate Stärke der Übereinstimmung anzeigt (siehe Tabelle 10) (Altman Douglas G., 1990, p. 404).

Tabelle 10: Interrater Agreement

	p-Wert	kappa	p-Wert (kappa)
Zuordnung zu einem Untersuchungszeitpunkt	NS ^a	0.36	0.004
Ventrikelweite Baseline Scan	<0.001 ^b	0.89	<0.001
Ventrikelweite Follow-up Scan	<0.001 ^b	0.81	<0.001
Sulcusweite Baseline Scan	NS ^b	0.42	0.019
Sulcusweite Follow-up Scan	0.019 ^b	0.46	0.012

^aChi Quadrat Test
^bFisher Exakt Test
 NS = nicht signifikant
 Stärke der Übereinstimmung je nach k-Wert: k < 0.20 sehr schwach, k = 0.21-0.40 schwach, k = 0.41-0.60 moderat, k = 0.61-0.80 gut, k = 0.81-1.00 sehr gut.

Abbildung 4: Die Grafik vergleicht die Anzahl der richtig, falsch und nicht zuordenbaren Scan-Paare aus den Auswertungen von Rater 1 und Rater 2



3.2.2 Intra-Rater Variability

Tabelle 11 zeigt das Intrarater Agreement von zehn zufällig ausgewählten Scans, die einem zweiten Rating durch die Medizinstudentin unterzogen wurden. Die „Zuordnung zu einem Untersuchungszeitpunkt“ und „Ventrikelweite des Baseline Scans“ zeigten eine sehr gute Übereinstimmung ($k = 1$). Die Stärke der Übereinstimmung bei der „Ventrikelweite des Follow-up Scans“, der „Sulcusweite des Baseline Scans“ und der „Sulcusweite des Follow-up Scans“ war moderat ($k = 0.6$, $k = 0.41$, $k = 0.58$).

Tabelle 11: Intrarater Agreement

	p-Wert	kappa	p-Wert (kappa)
Zuordnung zu einem Untersuchungszeitpunkt	0.002 ^a	1	0.002
Ventrikelweite Baseline Scan	0.008 ^b	1	0.002
Ventrikelweite Follow-up Scan	NS ^b	0.6	0.038
Sulcusweite Baseline Scan	NS ^b	0.41	NS
Sulcusweite Follow-up Scan	NS ^b	0.58	0.065
^a Chi Quadrat Test ^b Fisher Exakt Test NS = nicht signifikant Stärke der Übereinstimmung je nach k-Wert: $k < 0.20$ sehr schwach, $k = 0.21-0.40$ schwach, $k = 0.41-0.60$ moderat, $k = 0.61-0.80$ gut, $k = 0.81-1.00$ sehr gut;			

3.3 Vergleich visuelles Rating und Ergebnisse aus SIENA/SIENAX

Die absolute prozentuelle Veränderung des Gehirnvolumens (pBVC) zwischen Baseline und Follow-up Scan unterscheidet sich statistisch signifikant zwischen korrekt zugeordneten, falsch zugeordneten und MRT-Scans ohne mögliche Zuordnung ($p < 0.001$). In den paarweise durchgeführten Vergleichen wurde nur zwischen den Gruppen „richtig und falsch“ ($p < 0.001$) und „richtig und keine Zuordnung“ ($p = 0.021$) ein statistisch signifikanter Unterschied gefunden.

Auch für die prozentuelle Veränderung des Gehirnvolumens pro Jahr berechnet (annual pBVC) konnte mit dem Kruskal Wallis Test ein signifikanter Unterschied zwischen den drei Gruppen ermittelt werden ($p = 0.003$). Ein solcher fand sich zwischen den Paaren nur für

den Vergleich der zwei Kategorien „richtig zugeordnet“ und „falsch zugeordnet“ ($p = 0.006$).

Die richtig zugeordneten Scans wiesen die höchste pBVC über den gesamten Zeitraum auf (-1.65% (-2.78, -0.76)). Falsche Zuordnungen zeigten im Mittel eine geringere Abnahme von -0.5% (-0.99, 0.28), bei den nicht zuordenbaren Scans nahm das Gehirnvolumen im Durchschnitt um -0.84% (-1.43 -0.56) ab.

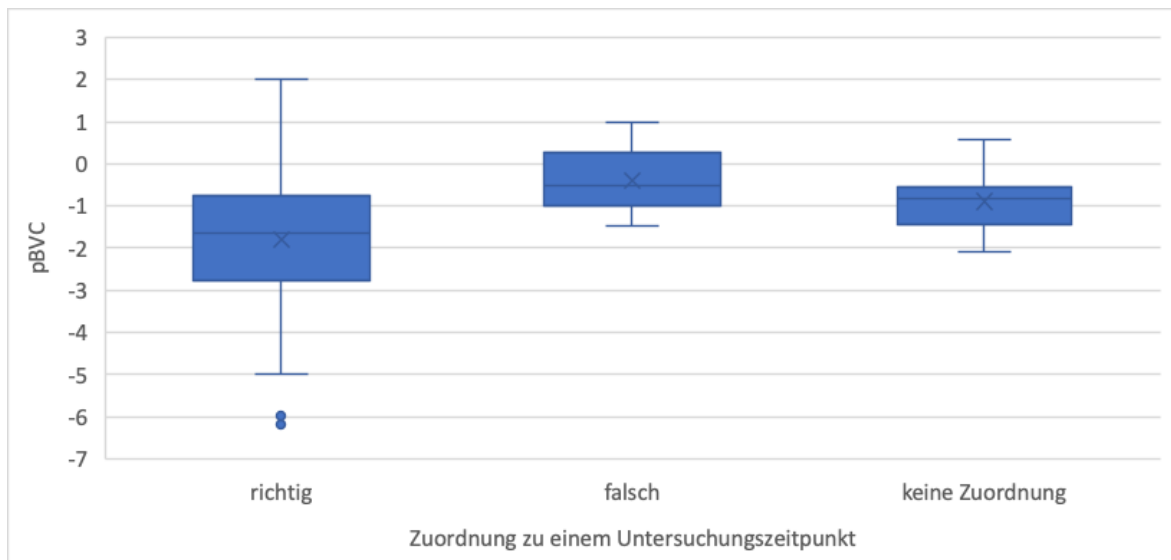
Auch die pBVC pro Jahr war am höchsten bei den richtigen Zuordnungen mit -0.45% (-0.74, -0.2), gefolgt von den nicht zuordenbaren Scans mit -0.33% (-0.46, -0.19). Die geringste Abnahme verzeichneten die falschen Zuordnungen (-0.17% (-0.4, 0.09)) (siehe Tabelle 12).

Tabelle 12: pBVC über den gesamten Untersuchungszeitraum und pBVC pro Jahr bei den richtig, falsch und nicht zuordenbaren MRT-Scans

	Zuordnung richtig n = 90	Zuordnung falsch n = 13	keine Zuordnung möglich n = 17	p-Wert
pBVC (%)	-1.65 (-2.78, -0.76)	-0.5 (-0.99, 0.28)	-0.84 (-1.43, -0.56)	< 0.001 ^e
annual pBVC (%)	-0.45 (-0.74, -0.2)	-0.17 (-0.4, 0.09)	-0.33 (-0.46, -0.19)	0.003 ^e

Daten nicht normalverteilt, als Median (Interquartilsabstand) angegeben.
^eKruskal Wallis Test
n = Anzahl, PBVC = prozentuelle Änderung des Hirnvolumens (percentage of brain volume change), annual pBVC = pBVC pro Jahr.

Abbildung 5: Die Boxplots zeigen die prozentuelle Änderung des Hirnvolumens über den gesamten Untersuchungszeitraum der richtig, falsch und nicht-zuordenbaren MRT-Untersuchungen



Um den Schwellenwert zwischen visuell erkennbarer und visuell nicht erkennbarer Veränderung von Hirnvolumen in diesem Rating zu ermitteln, wurde die gesamte Kohorte

nach der jährlichen pBVC in vier Quartile zu jeweils 30 PatientInnen unterteilt. Diese Einteilung wurde aus der Studie von Pichler et al. 2016 übernommen (Pichler *et al.*, 2016). Tabelle 13 zeigt die Verteilung der „Zuordnung zum Untersuchungszeitpunkt“ innerhalb dieser Quartile. Die Anzahl der richtig zugeordneten Scans nimmt vom Quartil mit der höchsten annual pBVC (Q1) zu dem mit der niedrigsten annual pBVC (Q4) leicht ab, die der falschen Zuordnungen bzw. die nicht zuordenbaren Scans hingegen steigt leicht an ($p = 0.01$).

Tabelle 13: Zuordnung zum Untersuchungszeitpunkt innerhalb der Quartile der pBVC pro Jahr

	Q 1 (-2.10 -0.65)	Q 2 (-0.64 -0.42)	Q 3 (-0.41 -0.19)	Q 4 (-0.18 0.78)	p-Wert
Zuordnung richtig (n = 90)	29 (96.7)	23 (76.7)	19 (63.3)	19 (63.3)	0.01 ^a
Zuordnung falsch (n = 13)	0	3 (10)	3 (10)	7 (23.3)	
keine Zuordnung (n = 17)	1 (3.3)	4 (13.3)	8 (26.7)	4 (13.3)	
Anzahl der PatientInnen pro Quartil= 30. Absoluter und prozentueller (spaltenweise) Wert angegeben. ^a Chi Quadrat Test n = Anzahl, NS = nicht signifikant					

Betrachtet man die Verteilung der richtig zugeordneten Scans (n = 90) innerhalb der vier Quartile nach dem Punkt „Sicherheitsgrad der Zuordnung“, so weist das Quartil mit der größten pBVC pro Jahr die höchste Anzahl an Zuordnungen mit einem visuell sicher beurteilbaren Unterschied zwischen den beiden Scans auf. Der Unterschied zwischen den Quartilen ist statistisch gesehen nicht signifikant (siehe Tabelle 14).

Tabelle 14: Verteilung des „Sicherheitsgrads der Zuordnung“ der richtig zugeordneten Scans innerhalb der Quartile der pBVC pro Jahr

	Q 1 (-2.10 -0.65)	Q 2 (-0.64 -0.42)	Q 3 (-0.41 -0.19)	Q 4 (-0.18 0.78)	p-Wert
Zuordnung richtig (n = 90)	n = 29	n = 23	n = 19	n = 19	
möglicher Unterschied (n = 80)	23 (79.3)	22 (95.7)	19 (100)	16 (84.2)	NS ^a
sicherer Unterschied (n = 10)	6 (20.7)	1 (4.3)	0	3 (15.8)	
Absoluter und prozentueller (spaltenweise) Wert angegeben. ^a Chi Quadrat Test n = Anzahl, NS = nicht signifikant					

Der Grenzwert zwischen den beiden Quartilen Q2 und Q3 entspricht gerundet dem im NEDA-4 Konzept vorgeschlagenem Schwellenwert für eine pathologische Hirnatrophie von -0.4% (De Stefano *et al.*, 2016). Um die Zuordnungen des visuellen Ratings in Bezug auf diesen Grenzwert zu untersuchen, wurden die Scans aus den Quartalen Q1 und Q2 sowie Q3 und Q4 jeweils zu einer Gruppe zusammengefasst. Jede dieser Gruppen enthält genau die Hälfte der PatientInnen.

Tabelle 15 zeigt die Ergebnisse dieser Analyse: In allen drei möglichen Gruppierungen für die Zuordnung zu einem Untersuchungszeitpunkt konnte ein statistisch signifikanter Unterschied gefunden werden. Die MRT-Untersuchungen jener 60 PatientInnen mit einem jährlichen pBVC über -0.42% wurden öfters dem richtigen Untersuchungszeitpunkt zugeordnet als die Scans in denen die Änderung des Hirnvolumens kleiner als -0.41% war (siehe Tabelle 15).

Tabelle 15: Zuordnung zum Untersuchungszeitpunkt zwischen den Quartilen Q1+Q2 und Q3+Q4 der pBVC pro Jahr

annual pBVC	< -0.42	> -0.41	p-Wert
richtig (n = 90)	52 (87)	38 (63)	0.011 ^a
falsch (n = 13)	3 (5)	10 (17)	
keine Zuordnung (n = 17)	5 (8)	12 (20)	
Anzahl der PatientInnen pro Gruppe: n = 60. Absoluter und prozentueller (spaltenweise) Wert angegeben. ^a Chi Quadrat Test n = Anzahl			

Als weiterer Parameter wurde das NBV der richtig zugeordneten Baseline und Follow-up Scans untersucht. Wie aus Tabelle 16 ersichtlich, wurden von den acht Stufen der SIFAP-Rating Skala (siehe Anhang: SIFAP-Rating Skala) für die visuelle Beurteilung der Ventrikelweite bzw. der Sulcusweite jeweils zwei zusammengefasst und das NBV zwischen den daraus resultierenden vier Gruppen verglichen. Damit soll eine zu geringe Fallzahl innerhalb einer Gruppe ausgeglichen werden.

Sowohl bei den Baseline, als auch bei den Follow-up Scans nimmt das NBV bei einer visuell beurteilten, größeren Weite von Ventrikel bzw. Sulci ab. Baseline Scans, die in die SIFAP Kategorien V1 und V2 eingeordnet worden sind, wiesen durchschnittlich ein NBV von 1582.19cm³ (1535.9, 1646.91) auf, wohingegen das NBV jener mit einer Ventrikelweite von V5 bis V6 im Durchschnitt 1433.76cm³ (1366.18, 1528.11) betrug. Auch bei den Follow-up Scans konnte eine Abnahme des NBV von 1555.43cm³ (1519.18, 1613.70) bei V1-V2

kategorisierten Bildern auf 1423.75cm³ bei dem einzigen MRT-Scan in der Kategorie V7-V8 festgestellt werden.

Ein ähnlicher Trend zeigte sich bei der Sulcusweite. Baseline Scans in der SIFAP-Skala S1-S2 gingen mit einem NBV von 1595.91cm³ (1556.71, 1646.91) einher, Bilder mit S7-S8 mit 1377.22cm³ (1104.00, 1421.02). Bei den Follow-up Scans nahm das NBV von 1593.73cm³ (1544.69, 1620.87) in den Kategorien S1-S2 auf 1404.15cm³ (1304.09, 1424.32) in Kategorie S7-S8 ab. Der Kruskal Wallis Test bestätigte in allen vier Fällen einen statistisch signifikanten Unterschied.

Nicht signifikant waren hingegen die Unterschiede der totalen und der jährlichen Änderung des NBV zwischen Baseline und Follow-up Scan, sei es bei der globalen Änderung der Ventrikelweite wie auch bei der Änderung der Sulcusweite (siehe Tabelle 16).

Tabelle 16: NBV (normalized brain volume, normalisiertes Hirnvolumen) innerhalb der visuellen Rating Kategorien für die ventrikuläre und sulcale Weite der richtig zugeordneten Scans (n = 90)

Ventrikelweite Bl	V1-V2 (n = 48)	V3-V4 (n = 31)	V5-V6 (n = 11)	V7-V8 (n = 0)	p-Wert
NBV Bl Scan [cm ³]	1582.19 (1535.89, 1646.91)	1549.83 (1482.65, 1586.59)	1433.76 (1366.18, 1528.11)	N.A.	< 0.001 ^e
Ventrikelweite FU	V1-V2 (n = 25)	V3-V4 (n = 45)	V5-V6 (n = 19)	V7-V8 (n = 1)	
NBV FU [cm ³]	1555.43 (1519.18, 1613.70)	1551.88 (1458.24, 1593.30)	1479.44 (1404.15, 1516.58)	1423.67	0.002 ^e
Sulcusweite Bl	S1-S2 (n = 52)	S3-S4 (n = 30)	S5-S6 (n = 4)	S7-S8 (n = 4)	
NBV Bl [cm ³]	1595.91 (1556.71, 1646.91)	1517.65 (1481.94, 1547.09)	1433.36 (1313.60, 1444.80)	1377.22 (1104.00, 1421.02)	< 0.001 ^e
Sulcusweite FU	S1-S2 (n = 16)	S3-S4 (n = 54)	S5-S6 (n = 13)	S7-S8 (n = 7)	
NBV FU [cm ³]	1593.73 (1544.69, 1620.87)	1552.49 (1500.30, 1593.30)	1457.74 (1414.80, 1490.98)	1404.15 (1304.09, 1424.32)	< 0.001 ^e
globale Änderung Ventrikelweite	keine (n = 7)	gering (n = 80)	mäßig (n = 2)	ausgeprägt (n = 1)	

NBV [cm ³]	-23.48 (-36.67, 7.28)	-26.42 (-47.12, -7.27)	-16.32 (-24.50)	-47.50	NS ^e
globale Änderung Sulcusweite	keine (n = 7)	gering (n = 72)	mäßig (n = 10)	ausgeprägt (n = 1)	
NBV [cm ³]	-20.56 (-27.78, 7.86)	-24.41 (-45.79, -5.69)	-36.00 (-72.30, -21.20)	-47.50	NS ^e
globale Änderung Ventrikelweite	keine (n = 7)	gering (n = 80)	mäßig (n = 2)	ausgeprägt (n = 1)	
delta NBV annual [cm ³]	-5.19 (-8.66, 1.21)	-7.52 (-12.85, -1.59)	-5.93 (-7.75)	-9.79	NS ^e
globale Änderung Sulcusweite	keine (n = 7)	gering (n = 72)	mäßig (n = 10)	ausgeprägt (n = 1)	
delta NBV annual [cm ³]	-4.26 (-14.54, 2.59)	-7.18 (-12.07, -1.19)	-8.27 (-15.61, -6.44)	-9.79	NS ^e
Normalverteilte Daten als Mittelwert (\pm Standardabweichung), nicht normalverteilte Daten als Median (Interquartilsabstand angegeben).					
°Kruskal Wallis Test					
n = Anzahl, NBV = normalisierte Hirnvolumen (normalized brain volume), BI = Baseline Scan, FU = Follow-up Scan, delta NBV = Änderung des NBV zwischen BI und FU, delta NBV annual = Änderung des NBV zwischen BI und FU pro Jahr					

Die Abbildungen 6-10 zeigen die mit Hilfe des Programms SIENAX gemessenen Hirnvolumina zum Baseline bzw. zum Follow-up Scan (NBV) innerhalb der visuell zugeordneten Kategorien für die ventrikuläre und sulcale Weite.

Abbildung 6: Die Boxplots zeigen das NBV der richtig zugeordneten Baseline Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die ventrikuläre Weite

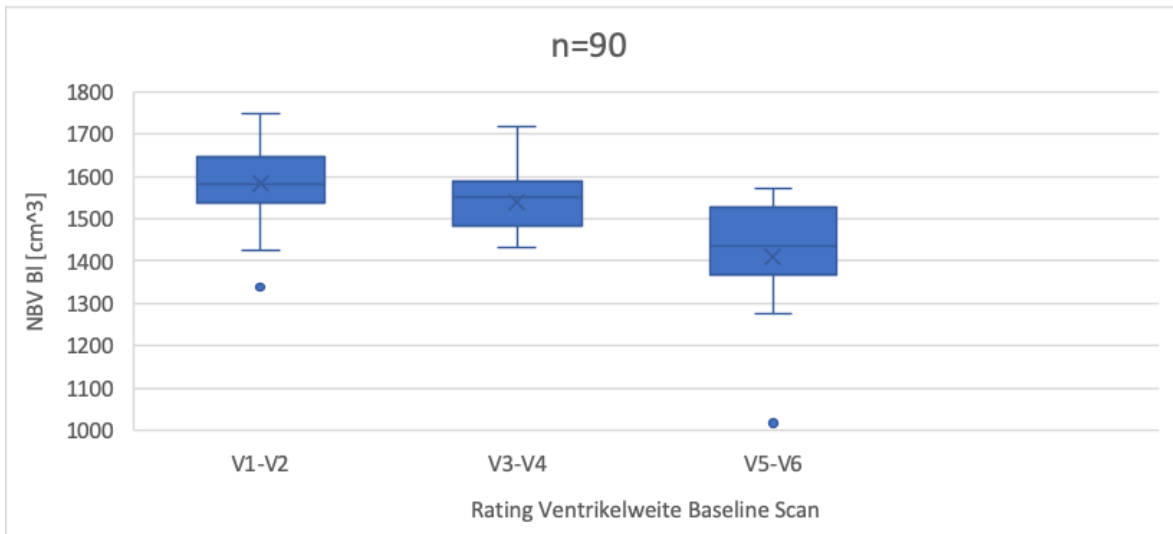


Abbildung 7: Die Boxplots zeigen das NBV der richtig zugeordneten Baseline Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die sulcale Weite

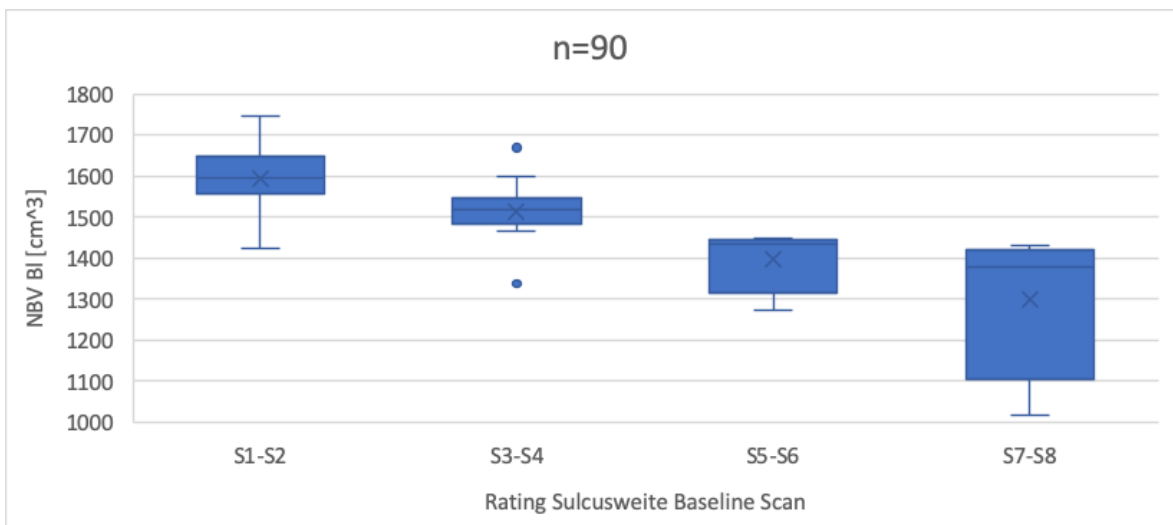


Abbildung 8: Die Boxplots zeigen das NBV der richtig zugeordneten Follow-up Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die ventrikuläre Weite

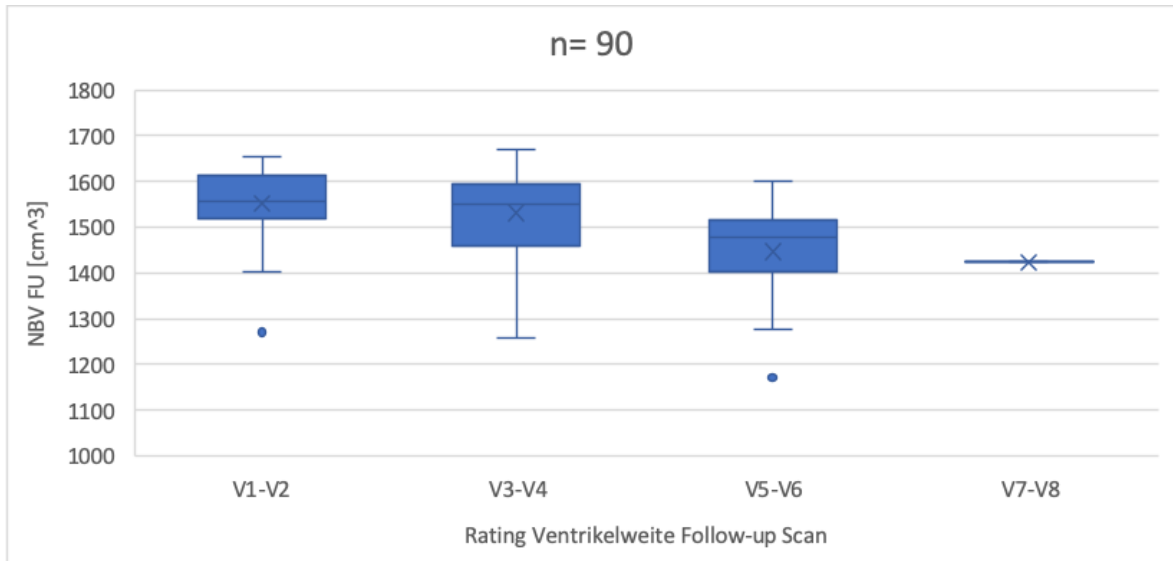
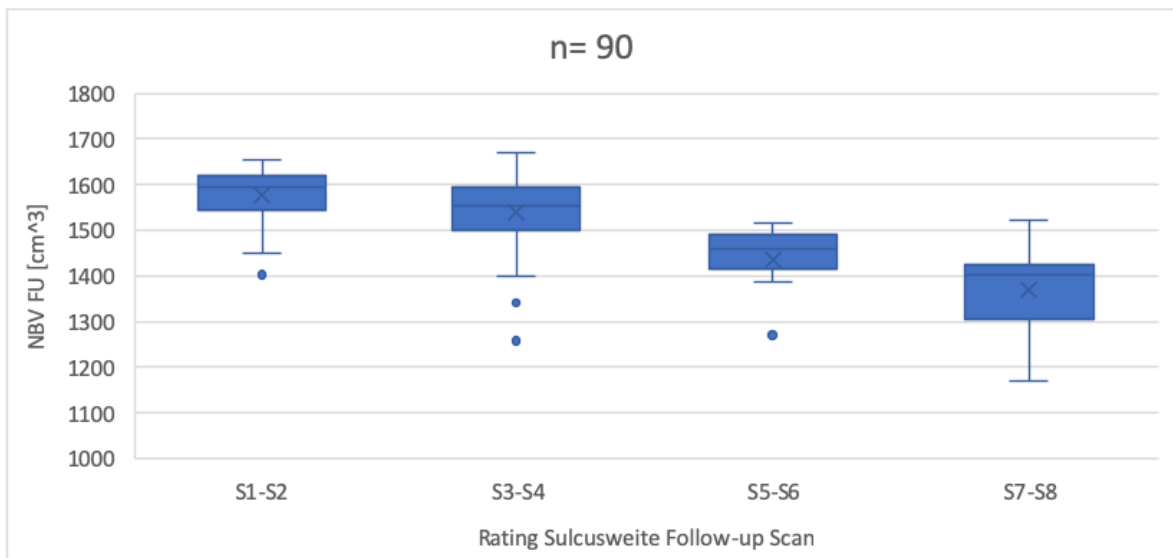


Abbildung 9: Die Boxplots zeigen das NBV der richtig zugeordneten Follow-up Scans innerhalb der zu 4 Gruppen zusammengefassten visuellen Kategorien für die sulcale Weite



3.4 Klinische Daten und visuelles Rating

Innerhalb der Gruppe der PatientInnen mit gesicherter MS zum Zeitpunkt der ersten MRT-Untersuchung wurden 84% der Scans dem richtigen Untersuchungszeitpunkt zugeordnet, bei den CIS PatientInnen war die Zuordnung in 67% der Fälle korrekt ($p = 0.035$).

Kein statistisch signifikanter Unterschied ergab sich zwischen PatientInnen mit einer Konversion der Erkrankung zwischen erstem und zweitem Scan von CIS zu definitiver MS, und den „non-converter“, d.h. PatientInnen mit unverändertem Stadium zum zweiten Scan.

Zwischen den Geschlechtern zeigte sich im Hinblick auf eine korrekte Zuordnung der Scans zum Untersuchungszeitpunkt kein statistisch signifikanter Unterschied.

Der durchschnittliche Zeitraum zwischen dem ersten und dem zweiten Scan betrug bei den richtig zugeordneten Scans 44 (33.75, 58) Monate, bei den falsch zugeordneten Scans waren es 38 (23.5, 47) Monate und bei den nicht zuordenbaren Scans 30 (24.5, 43.5) ($p = 0.004$).

Bei allen anderen klinischen Variablen konnte kein signifikanter Unterschied zwischen den richtigen und falschen Zuordnungen ermittelt werden (siehe Tabelle 17).

Die Analyse der richtig zugeordneten Scans unterteilt nach dem Sicherheitsgrad der Zuordnung (möglich oder sicher) ergab für keine der klinischen Variablen einen statistisch signifikanten Unterschied (siehe Tabelle 18).

Tabelle 17: Zuordnung zum Untersuchungszeitpunkt in Bezug auf klinische und demographische Variablen

Zuordnung der Scans zu einem Untersuchungszeitpunkt (Baseline, Follow-up)	richtig (n = 90)	falsch (n = 13)	nicht möglich (n = 17)	p-Wert
MS-Phänotyp Baseline Scan				
CIS (n = 63, %)	42 (67)	12 (19)	9 (14)	0.008 ^a
MS (n = 57, %)	48 (84)	1 (2)	8 (14)	
converter, MS zum Follow-up (n = 33, %)	24 (73)	5 (15)	4 (12)	NS ^a
non-converter, CIS zum Follow-up (n = 30, %)	18 (60)	7 (23)	5 (17)	
Geschlecht				
weiblich (n = 78, %)	57 (73)	9 (12)	12 (15)	NS ^a
männlich (n = 42, %)	33 (79)	4 (10)	5 (12)	
Alter Baseline Scan	34.1 ± 10	33.2 ± 6.1	32.6 ± 8.1	NS ^g
Alter Erkrankungsbeginn	28.3 ± 9.3	30.4 ± 7.5	28.6 ± 9.8	NS ^g
Krankheitsdauer Baseline Scan, Monate	20 (2, 97)	7 (0.5, 38.5)	6 (2.5, 100)	NS ^e
Intervall Baseline und Follow-up Scan, Monate	44 (33.75, 58)	38 (23.5, 47)	30 (24.5, 43.5)	0.004 ^e
EDSS Baseline Scan	1,8 (1, 3)	1 (0.5, 2.3)	2 (0, 1)	NS ^e
EDSS Follow-up Scan	2 (0.8, 2.6)	1 (0, 2.3)	1.5 (0, 2.3)	NS ^e
Annualisierte Schubrate	0.09 (0, 0.49)	0 (0, 0.13)	0 (0, 0.56)	NS ^e
Änderung EDSS von Baseline auf Follow up Scan				
Verbesserung (n = 48, %)	38 (79)	5 (10)	5 (10)	NS ^a
Gleich (n = 31, %)	22 (71)	4 (13)	5 (16)	
Verschlechterung (n = 41, %)	30 (73)	4 (10)	7 (17)	
DMT Baseline Scan (n = 46, %)	36 (78)	5 (11)	5 (11)	NS ^a
DMT Follow-up Scan (n = 80, %)	62 (79)	9 (11)	9 (11)	NS ^a
Absoluter und prozentueller (zeilenweise) Wert angegeben.				
Normalverteilte Daten als Mittelwert (± Standardabweichung), nicht normalverteilte Daten als Median (Interquartilsabstand) angegeben.				
*Chi Quadrat Test				
^e Kruskal Wallis Test				
^g ANOVA				
n = Anzahl, NS = nicht signifikant, CIS = klinisch isoliertes Syndrom, MS = Multiple Sklerose, EDSS = Expanded Disability Status Scale, DMT = krankheitsmodifizierende (Disease modifying therapies).				

Tabelle 18: Sicherheitsgrad der Zuordnungen der richtig zugeordneten Scans (n = 90) in Bezug auf klinische und demographische Variablen

Sicherheitsgrad der Zuordnung: Unterschied	möglich (n = 80)	sicher (n =10)	p-Wert
MS- Subtyp Baseline Scan			
MS (n = 48, %)	44 (92)	4 (8)	NS ^a
CIS (n = 42, %)	36 (86)	6 (14)	
converter (n = 18, %)	15 (83)	3 (17)	
non-converter (n = 24, %)	21 (88)	3 (13)	NS ^a
Geschlecht			
weiblich (n = 64, %)	58 (91)	6 (9)	
männlich (n = 26, %)	22 (85)	4 (15)	NS ^a
Alter Baseline Scan	34.1 ± 9,9	33.8 ± 11.1	NS ^d
Alter Erkrankungsbeginn	28.9 ± 8.8	29.4 ± 13	NS ^d
Krankheitsdauer Baseline Scan, Monate	22 (2, 94.75)	6 (0.75, 113.25)	NS ^c
Intervall Baseline und Follow up Scan	44.5 (34.25, 58.75)	36 (28.5, 54.35)	NS ^c
EDSS Baseline Scan	2 (1, 2.9)	1 (0.75, 3.5)	NS ^c
EDSS Follow-up Scan	2 (0.25, 3)	1.75 (0.75, 2)	NS ^c
Annualisierte Schubrate	0 (0, 0.48)	0.3 (0, 0.78)	NS ^c
Änderung EDSS von Baseline auf Follow up Scan			
Verbesserung (n =38, %)	34 (89.5)	4 (10.5)	NS ^a
Gleich (n = 22, %)	19 (86)	3 (14)	
Verschlechterung (n = 30, %)	27 (90)	3 (10)	
DMT Baseline Scan (n = 32, %)	31 (97)	1 (3)	NS ^b
DMT Follow-up Scan (n = 57, %)	50 (88)	7 (12)	NS ^b
Absoluter und prozentueller (zeilenweise) Wert angegeben.			
Normalverteilte Daten als Mittelwert (±Standardabweichung), nicht normalverteilte Daten als Median (Interquartilsabstand) angegeben.			
^a Chi Quadrat Test			
^c Mann Whitney U Test			
^d Welch Test			
n = Anzahl, NS = nicht signifikant, CIS = klinisch isoliertes Syndrom, MS = Multiple Sklerose, EDSS = Expanded Disability Status Scale, DMT = krankheitsmodifizierende Therapie (Disease modifying therapies).			

4 Diskussion

Die zentrale Fragestellung dieser Diplomarbeit befasst sich mit den Möglichkeiten und Grenzen einer visuellen Beurteilung von Hirnvolumenänderung bei MS PatientInnen anhand der kraniellen MRT. Im Vergleich mit Software-basierten Messmethoden wurde der Frage nach einem möglichen Schwellenwert für das visuelle Erkennen „pathologischer“ Hirnatrophie in MRT-Scans nachgegangen.

Zudem wurden die Relevanz einer semiquantitativen Rating-Skala im klinischen Alltag geprüft und Vor- und Nachteile erörtert.

Um diesen Forschungsfragen nachzugehen wurden jeweils zwei MRT Untersuchungen von insgesamt 120 PatientInnen anhand eines eigens dafür erstellten visuellen Rating-Sheets (siehe Anhang: Rating Sheet) ausgewertet und mit den computerbasiert ermittelten Werten der Software-Programme SIENA und SIENAX verglichen. Ein zweiter Rater beurteilte zusätzlich 21 zufällig ausgewählte Fälle.

4.1 Das visuelle Rating

Anhand des visuellen Ratings konnte in 75% (n = 90) der Fälle ein Unterschied zwischen den zwei MRT-Scans erkannt und Baseline und Follow-up Scan richtig zugeordnet werden. Bei 14% (n = 17) der Bildpaare konnte der Untersuchungszeitpunkt nicht bestimmt werden, da kein Unterschied zwischen den beiden Bildern ersichtlich war. Nur 11% (n = 13) der Zuordnungen waren nicht korrekt.

Der Unterschied zwischen falsch und nicht zugeordneten Fälle kann folgendermaßen interpretiert werden: „keine Zuordnung möglich“ bedeutet, dass die auswertende Person visuell keinen Unterschied zwischen den beiden MRT-Bildern feststellen konnte, und das Hirnvolumen zwischen Baseline und Follow-up Scan visuell nicht abnimmt. Im Fall einer nicht korrekten Zuordnung hingegen, wurden Baseline und Follow-up Scan vertauscht und die Volumina beider Scans falsch eingeschätzt. Die möglichen Fehlerursachen dafür werden in einem später folgendem Teil dieser Diskussion erörtert. Dementsprechend sind diese beiden Ergebnisse gesondert voneinander zu interpretieren.

Die Fehlerquote dieses Rating-Durchgangs lag somit bei 11% und kann bei einer daraus folgenden Trefferquote von 75% bzw. insgesamt 89% durchaus als gering eingestuft werden.

Dies wurde auch durch die subjektive Einschätzung bezüglich der Sicherheit der Zuordnung bestätigt: unter den richtig zugeordneten Scans wurde der Unterschied zwischen Baseline und Follow-up bei immerhin 11% der Bildpaare als sicher gewertet.

Ein Einflussfaktor für eine korrekte Zuordnung schien die Beurteilung der globalen Änderung der Ventrikelweite bzw. der Sulcusweite zu sein. Gab es große Unterschiede in der Weite der Ventrikel bzw. Sulci zwischen Baseline und Follow-up Scan, so wurden diese auch eher dem richtigen Untersuchungszeitpunkt zugeordnet.

Die Qualität der Scans hingegen spielte für ein visuelles Rating kaum eine Rolle. Nur in einem Fall konnten die MRT-Bilder aufgrund der schlechten Bildqualität nicht ausgewertet werden.

Um einen Vergleich zwischen zwei verschiedenen Ratern zu erhalten, wurde das sogenannte Interrater-Agreements berechnet. Dabei wurde die Auswertung von 21 zufällig ausgewählten Scan-Paaren durch zwei voneinander unabhängigen Ratern bzw. Raterinnen verglichen. Der kappa-Koeffizient für die Zuordnung zu einem Untersuchungszeitpunkt ergab dabei nur eine schwache Übereinstimmung.

Anders als beim Interrater-Agreement variierte das Ergebnis bei einer wiederholten Auswertung durch denselben Rater kaum. Der kappa-Koeffizient von $k = 1$ für das sogenannte Intrarater-Agreement für die Zuordnung zu einem Untersuchungszeitpunkt steht beispielsweise für eine sehr gute Übereinstimmung zwischen erstem und zweitem Rating-Durchgang.

Hier sei jedoch noch angemerkt, dass die Fallzahl sowohl für das Inter- als auch für das Intrarater-Agreement, 21 bei ersterem und 10 bei letzterem, wahrscheinlich zu gering war, um ein generalisierbares Ergebnis zu erhalten. Mit einer größeren Fallzahl könnte auch der sogenannte „weighted k-Koeffizient“ verwendet werden, bei dem zusätzlich noch das Ausmaß der Abweichungen zwischen den verschiedenen Rating-Durchgängen miteinberechnet wird.

Zusammenfassend konnte mit diesem ersten Versuch eines visuellen Ratings von Hirnatrophie gezeigt werden, dass es für das menschliche Auge prinzipiell möglich ist, mit relativ gutem Ergebnis, zwischen einem Baseline und einem Follow-up MRT-Scan zu unterscheiden. Die durchschnittlich benötigte Zeit für das visuelle Rating betrug zwischen 3 und 5 Minuten pro PatientIn.

Je weiter fortgeschritten die Veränderungen im Gehirn der PatientInnen, und je größer der sichtbare Unterschied zwischen den beiden Scans ist, desto eher wurden diese dem richtigen Untersuchungszeitpunkt zugeordnet. Die Bildqualität schien das Rating hingegen nicht sehr

stark zu beeinflussen. Dies könnte eventuell ein Vorteil gegenüber computerbasierten Methoden sein.

Dennoch darf nicht außer Acht gelassen werden, dass es sich hierbei um das Rating einer einzelnen Person handelt, das zwar zwischen erstem und zweitem Durchgang nur geringe Schwankungen aufwies, jedoch von den Ergebnissen eines zweiten Raters stark abwich und deshalb nicht als allgemein gültig angesehen werden kann.

4.2 Visuelles Rating und SIENA/SIENAX

Da das Outcome des visuellen Ratings allein kaum aussagekräftig ist, wurden die computerbasierten Messdaten der Programme SIENA und SIENAX zwischen richtig, falsch und nicht zugeordneten Scans miteinander verglichen.

Die durchschnittliche Änderung des Hirnvolumens über den gesamten Scanintervall (pBVC) der richtig zugeordneten Scans lag bei -1.65% (-2.78, -0.76), und jenes der nicht zuordenbaren Scans bei -0.84% (-1.43, -0.56). Vergleicht man nur die beiden Mediane miteinander, so unterscheidet sich die Abnahme des Hirnvolumens der Scans ohne erkennbaren Unterschied von dem pBVC der Scans mit visuell sichtbarem Unterschied um den Faktor 1.96, die Abnahme des Hirnvolumens war folglich beinahe halb so groß.

Bezogen auf die pBVC pro Jahr nahm das Hirnvolumen bei den richtigen Zuordnungen um -0.45% (-0.74, -0.2), und bei den nicht zuordenbaren Scans um -0.33% (-0.46, -0.19) ab; in diesem Fall war die Abnahme der korrekten Zuordnungen noch 1.4-mal so groß.

Demnach ist ein solches Ergebnis ein Hinweis darauf, dass es zumindest innerhalb dieses Ratingdurchgangs möglich gewesen war, Änderungen des Hirnvolumens ab einem bestimmten Wert visuell zu erfassen.

Falsch zugeordnete Scans wiesen sowohl bezüglich des gesamten pBVC als auch des pBVC pro Jahr im Durchschnitt die geringsten Veränderungen auf. Zudem fanden sich auch nur in dieser Gruppe innerhalb des Interquartilsabstand Scan-Paare mit einer positiven Volumenänderung (siehe Abbildung 5). Dies könnte möglicherweise der falschen Zuordnung zugrunde liegen. Andererseits sind diese positiven Volumenschwankungen bis maximal +0.28% sehr gering und somit kaum visuell erfassbar. Die Ursache der Falschzuordnungen liegt daher wahrscheinlich vor allem in der mangelnden Expertise der Auswertenden begründet.

Im Rahmen des visuellen Ratings wurde von jedem Scan die Weite der Ventrikel und der Sulci anhand einer Scala mit Vergleichsbildern in 8 Kategorien eingeteilt (siehe SIFAP-Rating Skala). Im Vergleich der durchschnittlichen Werte des NBV dieser acht Kategorien (für die Auswertung wurden die MRT-Scans von jeweils 2 Kategorien zusammengefasst), zeigte sich deutlich eine Abnahme des normalisierten Hirnvolumens mit zunehmender Ventrikel- bzw. Sulcusweite. Dieser Trend konnte bei allen richtig ausgewerteten Baseline und Follow-up Scans sowohl bezüglich Ventrikelweite als auch für die Sulcusweite beobachtet werden. Die globale Änderung der ventrikulären bzw. sulcalen Weite stimmte allerdings kaum mit der Änderung des NBV überein.

Demnach kann davon ausgegangen werden, dass mit Hilfe von Vergleichsbildern, wie der in dieser Diplomarbeit verwendeten SIFAP-Rating Skala (siehe Anhang: SIFAP-Rating Skala), das Hirnvolumen visuell exakter erfasst werden kann als durch eine globale Beurteilung ohne jegliche Referenz.

Diese Gegenüberstellung von visuellen Ergebnissen und computerbasierten Messdaten der SIENA/SIENAX Programme zeigt, dass es innerhalb der angeführten Spanne möglich ist, Änderungen des Hirnvolumens mit freiem Auge zu erkennen und mit Hilfe von Rating-Skalen und Vergleichsbildern in Kategorien einzuteilen.

Um genauer zu untersuchen, in welcher Bandbreite ein relevanter Schwellenwert für visuell erfassbare Hirnvolumenänderung liegen könnte, wurde die jährliche pBVC aller 120 Scan-Paare in 4 Quartile zu je 30 Fällen unterteilt. Diese Einteilung lehnt sich dem Paper von Pichler et al., 2016 (Pichler *et al.*, 2016) an. Analysiert man die Zuordnung innerhalb dieser Quartile, so fällt auf, dass im ersten Quartil (pBVC pro Jahr von -2.10% bis -0.65%) 96.9% der Scans richtig zugeordnet worden sind; ein Scan-Paar konnte nicht zugeordnet werden, es gab keine falschen Zuordnungen. Dies deutet darauf hin, dass eine Änderung des Hirnvolumens in der Bandbreite des Q1 Intervalls mit dem freien Auge gut erkennbar war. Im Quartil Q2 (pBVC pro Jahr von -0.64% bis -0.42%) wurden immerhin noch 76.6% der Scans richtig zugeordnet und 13.3% konnten nicht zugeordnet werden. 10% der Scans wurden falsch zugeordnet. Da die Anzahl der nicht zuordenbaren Scans im darauffolgenden Quartil Q3 (-0.41% bis -0.19%), bei gleichbleibenden falschen Zuordnungen, weiter ansteigt, und die der richtigen Zuordnungen abnimmt, sinkt die Wahrscheinlichkeit, dass bei einem pBVC pro Jahr unter -0.41% eine Volumenänderung der Gehirns noch verlässlich visuell erfasst werden können.

Auch die subjektive Einschätzung des Sicherheitsgrads der Zuordnung aller richtigen Auswertungen ergibt innerhalb des Quartils Q1 mit 20.7% die höchste Anzahl an MRT-Untersuchungen mit sicherem Unterschied zwischen Baseline und Follow-up Scan. Im 2. Quartil wurde nur mehr zwischen einem Scan-Paar ein sicherer Unterschied gesehen und im 3. Quartil wurde der Unterschied von allen richtig zugeordneten MRT-Scans nur mehr als möglich gewertet. Allerdings erreichte der Unterschied zwischen den Quartilen für den „Sicherheitsgrad der Zuordnung“ keine statistische Signifikanz.

Nach der oben beschriebenen Verteilung der Zuordnungen zwischen den vier Quartilen für das jährliche pBVC könnte man den Schwellenwert zwischen visuell erkennbarer und nicht erkennbarer Hirnvolumenänderung am ehesten zwischen dem zweiten und dritten Quartil annehmen. Innerhalb von Q1 und Q2 (pBVC pro Jahr > -0.42) konnten 87% ($n = 60$) der Scans richtig zugeordnet werden, nur bei 8% konnte kein Unterschied festgestellt werden. In Q3 und Q4 waren nur mehr 63% aller Scans richtig zugeordnet worden, bei 20% war keine Volumenänderung sichtbar.

Hier sei angemerkt, dass ein anhand der vorliegenden Daten festgelegter „Cut-off“ für das pBVC nur für dieses Rating gültig wäre, und kaum auf andere etwaige RaterInnen übertragen werden könnte. Wie bereits durch die Auswertung des Interrater-Agreements gezeigt wurde, wichen die Ergebnisse von zwei unterschiedlichen RaterInnen stark voneinander ab, anders als beim Intrarater Agreement, wo diese kaum variierten.

In Hinblick auf die klinische Relevanz könnte eine mögliche Grenze zwischen visuell sichtbarer und visuell nicht sichtbarer Änderung von -0.4% von Relevanz sein. Dies würde nämlich dem im NEDA-4 Konzept festgelegten Grenzwert für eine pathologische Abnahme von Hirnvolumen entsprechen (De Stefano *et al.*, 2016).

Das bisher nur in Studien verwendete NEDA (no evidence of disease) Prinzip wurde als ein Therapieziel der MS konzipiert und beinhaltet neben Krankheitsprogression, Schubrate und T2 Läsionen in der MRT auch die jährliche Abnahme von Hirnvolumen (Kappos *et al.*, 2016).

In der gesunden Kontrollgruppe konnte in einer Studie eine durchschnittliche pBVC pro Jahr von 0.1 bis 0.3% gemessen werden (Kappos *et al.*, 2016). Für MS PatientInnen rangierte die pBVC zwischen -0.5% und -1.35% pro Jahr (Vollmer *et al.*, 2015). Aus den Ergebnissen dieser Diplomarbeit kann man schließen, dass eine Volumenänderung dieser Größenordnung im MRT mit freiem Auge erkennbar wäre.

An der klinischen Anwendbarkeit der Hirnatrophie als Marker für die Beurteilung der Krankheitsprogression und des Therapieeffekt wird bereits seit mehreren Jahren intensiv geforscht, mit dem Ziel, die individuelle Betreuung von PatientInnen zu optimieren (Beadnall *et al.*, 2019). Umso mehr wäre ein visuelles Ratingverfahren zur Unterscheidung zwischen pathologischer und physiologischer Hirnatrophie wichtig und von klinischer Relevanz, da somit auch Rückschlüsse auf die Progression der Beeinträchtigung der PatientInnen gezogen werden können (Sormani, Arnold and De Stefano, 2014).

Wie bereits erwähnt, kann mit den vorliegenden Daten jedoch nicht allgemein bestätigt werden, dass eine solche Differenzierung mit freiem Auge möglich ist. Um diese Frage zu beantworten, wäre ein Vergleich der Auswertungen von mehreren unterschiedlichen Personen notwendig.

4.3 Visuelles Rating im klinischen Alltag

In Bezug auf die klinischen Daten der Studienpopulation unterschied sich statistisch signifikant nur die Dauer des Scanintervalls zwischen Baseline und Follow-up Untersuchung zwischen richtig zugeordneten (44 (33.75, 58) Monate), falsch zugeordneten (38 (23.5, 38.5) Monate) und nicht zuordenbaren (30 (24.5, 43.5) Monate) MRT-Scans.

Die Ursache liegt vermutlich darin, dass je länger der Scanintervall desto höher ist auch die Abnahme des Hirnvolumens. Daher war es für den Rater bzw. für die Raterin bei Scan-Paaren mit längerem Zeitintervall zwischen den beiden Untersuchungen einfacher, den Unterschied zwischen Baseline und Follow-up Scan zu erkennen. Im Umkehrschluss bedeutet das, dass für eine visuell verlässliche Beurteilung ein ausreichend langer Untersuchungszeitraum essenziell ist.

Die PatientInnen der richtig zugeordneten Scans waren im Durchschnitt bei Erkrankungsbeginn und bei Durchführung des Baseline Scans älter als PatientInnen, deren Scans nicht zuordenbar waren. Auch die Krankheitsdauer bei der Baseline Untersuchung war im Schnitt höher als bei den anderen PatientInnen. Da die Abnahme des Hirnvolumens unabhängig vom MS-Subtyp mit der Länge der Krankheitsdauer kontinuierlich ansteigt, hat dies vermutlich die Zuordnung zum Untersuchungszeitpunkt vereinfacht (De Stefano *et al.*, 2010). Allerdings war der Unterschied zwischen den Zuordnungen weder für das Alter noch für die Krankheitsdauer statistisch signifikant.

Bisher gibt es keine vergleichbaren Studien zur visuellen Atrophiemessung bei MS PatientInnen anhand von MRT-Scans. Dennoch wird die Verwendung derartiger Ratingsysteme in der Forschung immer wieder diskutiert; beispielsweise für die Quantifizierung von Läsionen der weißen Substanz (Fazekas *et al.*, 2002; Prins *et al.*, 2004; Olsson *et al.*, 2013) oder für die Atrophiemessung bei AlzheimerpatientInnen (Harper *et al.*, 2015, 2016).

In den letzten Jahren wurde intensiv an der Entwicklung von automatisierten computergestützten Techniken zur Messung von Hirnvolumen geforscht. Durch diese neuen Verfahren können auch Volumina von verschiedenen Hirnarealen genauestens bestimmt werden. Allerdings ist die Sensitivität dieser Quantifizierungsmethoden je nach verwendetem Algorithmus unterschiedlich und wird teilweise stark von äußeren Faktoren wie dem verwendeten MRT-Scan Gerät, der Scan-Prozedur oder von den physiologischen Gegebenheiten der untersuchten Person selbst beeinflusst (Durand-Dubief *et al.*, 2012; Biberacher *et al.*, 2016; Storelli *et al.*, 2018; Zivadinov *et al.*, 2018).

Das visuelle Rating hingegen stellt eine einfachere, schnellere und billigere Methode zur Quantifizierung von Hirnvolumen dar, und wäre daher für einen Einsatz im klinischen Alltag sehr gut geeignet. Im Gegensatz zur klinischen Studie ist eine genaue quantitative Erfassung des Hirnvolumens nicht unbedingt notwendig, stattdessen könnte das Erkennen von pathologischer Hirnatrophie und von Veränderungen über die Zeit ausreichend sein, um Krankheitsverlauf und Therapieerfolg zu verfolgen.

Visuelle Rating Verfahren können zudem von jedem Neurologen und jeder Neurologin schnell erlernt und in jeder Klinik durchgeführt werden, ohne dass dafür ein eigenes Software-Programm notwendig ist. Dies könnten somit auch für die Radiologie Entlastungen bringen (Harper *et al.*, 2015).

Durchgeführt von entsprechend erfahrenen Ratern und Raterinnen ist eine visuelle Auswertung zudem unabhängiger vom verwendeten MRT-Gerät, von der Qualität der MRT-Bilder, von Schwankungen im Bildkontrast und Änderungen der Angulation. Vor allem in der Beurteilung von artefakt-behafteten Bildern, die in der klinischen Routine vorkommen können, ist das menschlichen Auge besser als der Computer (Fazekas *et al.*, 2002).

Der größte Nachteil von visuellen Rating-Skalen ist die Abhängigkeit von der individuellen Erfahrung der auswertenden Person selbst. Die Art der verwendeten Rating-Skala beeinflusst zusätzlich das Ergebnis (Fazekas *et al.*, 2002; Harper *et al.*, 2015). Gut strukturierte Rating-Skalen und bestmöglich trainierte RaterInnen könnten diesem Problem

entgegenwirken. Allerdings fehlen bisher einheitliche Konzepte zur Entwicklung und Verwendung von visuellen Rating-Systemen.

4.4 Limitationen

Wie bereits erwähnt, ist der wichtigste limitierende Faktor dieser Arbeit die geringe Anzahl der RaterInnen. Um das Rating-System auf dessen Reliabilität zu prüfen, wäre die Auswertung von mehr als nur zwei Ratern bzw. Raterinnen notwendig gewesen.

Zudem wurden die computerbasierten Messdaten für das Hirnvolumen nur mit den Programmen SIENA/SIENAX erhoben und diese als einzig richtige Referenz verwendet. Die Werten wurden nicht mit den Ergebnissen von anderen Software-Programmen verglichen

Andere Ursachen, die zu Schwankungen des Hirnvolumens führen können, wie beispielsweise Entzündungen und Pseudoatrophie oder Komorbiditäten der ProbandenInnen, wurden ebenfalls nicht berücksichtigt. Diese würden wahrscheinlich sämtliche Methoden zur Quantifizierung von Hirnvolumen betreffen.

4.5 Zusammenfassung

In Zusammenschau aller Ergebnisse kann diese Diplomarbeit als ein erster Hinweis dafür gesehen werden, dass eine visuelle, semiquantitative Rating Skala neben computerbasierten Methoden eine weitere Möglichkeit für die Erfassung von Veränderungen des Hirnvolumens bei MS-PatientInnen anhand von seriellen MRT-Untersuchungen darstellt.

Abnahme von Hirnvolumen ist bis zu einem bestimmten Schwellenwert mit freiem Auge sichtbar und mit Hilfe von Skalen und Vergleichsbildern quantifizierbar. Anhand eines solchen Cut-offs wäre zumindest die Unterscheidung zwischen pathologischer und physiologischer Hirnatrophie möglich.

Es bedarf allerdings noch weiterer Studien v.a. bezüglich des Interrater-Agreements, um eine genauere Aussage darüber treffen zu können.

5 Literaturverzeichnis

- Aktas, O. *et al.* (2018) 'Diagnose der Multiplen Sklerose: Revision der McDonald-Kriterien 2017', *Der Nervenarzt*, 89(12), pp. 1344–1354. doi: 10.1007/s00115-018-0550-0.
- Alonso, A. and Hernán, M. A. (2008) 'Temporal trends in the incidence of multiple sclerosis', *Neurology*, 71(2), pp. 129–135. doi: 10.1212/01.wnl.0000316802.35974.34.
- Altman Douglas G. (1990) *Practical Statistics for Medical Research*.
- Amiri, H. *et al.* (2018) 'Urgent challenges in quantification and interpretation of brain grey matter atrophy in individual MS patients using MRI', *NeuroImage: Clinical*. Elsevier, 19(April), pp. 466–475. doi: 10.1016/j.nicl.2018.04.023.
- Ascherio, A. *et al.* (2014) 'Vitamin D as an early predictor of multiple sclerosis activity and progression.', *JAMA neurology*, 71(3), pp. 306–314. doi: 10.1001/jamaneurol.2013.5993.
- Ascherio, A. and Munger, K. L. (2007a) 'Environmental risk factors for multiple sclerosis. Part I: The role of infection', *Annals of Neurology*, 61(4), pp. 288–299. doi: 10.1002/ana.21117.
- Ascherio, A. and Munger, K. L. (2007b) 'Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors', *Annals of Neurology*, 61(6), pp. 504–513. doi: 10.1002/ana.21141.
- Banwell, B. *et al.* (2007) 'Multiple sclerosis in children: clinical diagnosis, therapeutic strategies, and future directions', *The Lancet Neurology*, 6(10), pp. 887–902. doi: 10.1016/S1474-4422(07)70242-9.
- Barkhof, F. *et al.* (2009) 'Imaging outcomes for neuroprotection and repair in multiple sclerosis trials', *Nature Reviews Neurology*, 5(5), pp. 256–266. doi: 10.1038/nrneurol.2009.41.
- Barnes, J. *et al.* (2010) 'Head size, age and gender adjustment in MRI studies: A necessary nuisance?', *NeuroImage*. Elsevier Inc., 53(4), pp. 1244–1255. doi: 10.1016/j.neuroimage.2010.06.025.
- Battaglini, M. *et al.* (2009) 'Voxel-wise assessment of progression of regional brain atrophy in relapsing-remitting multiple sclerosis.', *Journal of the neurological sciences*. Elsevier B.V., 282(1–2), pp. 55–60. doi: 10.1016/j.jns.2009.02.322.
- Baumhackl, U. (2014) *Multiple Sklerose: Prävalenz & Therapie im 12-Jahres-Vergleich in Österreich*. 1. Auflage. Facultas.
- Beadnall, H. N. *et al.* (2019) 'Comparing longitudinal brain atrophy measurement techniques in a real-world multiple sclerosis clinical practice cohort: towards clinical integration?', *Therapeutic Advances in Neurological Disorders*, 12, p. 175628641882346. doi: 10.1177/1756286418823462.
- Beiske, A. G. *et al.* (2004) 'Pain and sensory complaints in multiple sclerosis', *European Journal of Neurology*, 11(7), pp. 479–482. doi: 10.1111/j.1468-1331.2004.00815.x.
- Beiske, A. G. *et al.* (2008) 'Depression and anxiety amongst multiple sclerosis patients', *European Journal of Neurology*, 15(3), pp. 239–245. doi: 10.1111/j.1468-1331.2007.02041.x.
- Biberacher, V. *et al.* (2016) 'Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis', *NeuroImage*. Elsevier Inc., 142, pp. 188–197. doi: 10.1016/j.neuroimage.2016.07.035.
- Böcker, W. (2012) *Pathologie*. Edited by 5. Auflage. Urban & Fischer Verlag/Elsevier GmbH.
- Bodini, B. *et al.* (2013) 'Corpus callosum damage predicts disability progression and cognitive dysfunction in primary-progressive MS after five years', *Human Brain Mapping*, 34(5), pp. 1163–1172. doi: 10.1002/hbm.21499.

- Bresciani, L. *et al.* (2005) 'Visual assessment of medial temporal atrophy on MR films in Alzheimer's disease: comparison with volumetry.', *Aging clinical and experimental research*, 17(1), pp. 8–13. doi: 2804 [pii].
- Calabrese, M. *et al.* (2013) 'The changing clinical course of multiple sclerosis: A matter of gray matter', *Annals of Neurology*, 74(1), pp. 76–83. doi: 10.1002/ana.23882.
- Calabrese, M. *et al.* (2015) 'Regional distribution and evolution of gray matter damage in different populations of multiple sclerosis patients', *PLoS ONE*, 10(8), pp. 1–12. doi: 10.1371/journal.pone.0135428.
- Calabrese, M. and Gallo, P. (2009) 'Magnetic resonance evidence of cortical onset of multiple sclerosis', *Multiple Sclerosis Journal*, 15(8), pp. 933–941. doi: 10.1177/1352458509106510.
- Caramanos, Z. *et al.* (2010) 'Gradient distortions in MRI: Characterizing and correcting for their effects on SIENA-generated measures of brain volume change', *NeuroImage*. Elsevier Inc., 49(2), pp. 1601–1611. doi: 10.1016/j.neuroimage.2009.08.008.
- Cepok, S. *et al.* (2005) 'Identification of Epstein-Barr virus proteins as putative targets of the immune response in multiple sclerosis.', *The Journal of clinical investigation*, 115(5), pp. 1352–60. doi: 10.1172/JCI23661.
- Chandran, S. *et al.* (2008) 'Myelin repair: the role of stem and precursor cells in multiple sclerosis.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(1489), pp. 171–83. doi: 10.1098/rstb.2006.2019.
- Chard, D. *et al.* (2002) 'Brain metabolite changes in cortical grey and normal-appearing white matter in clinically early relapsing-remitting multiple sclerosis.', *Brain : a journal of neurology*, 125(Pt 10), pp. 2342–2352. doi: 10.1002/ana.1123.
- Charil, A. *et al.* (2007) 'Focal cortical atrophy in multiple sclerosis: Relation to lesion load and disability', *NeuroImage*, 34(2), pp. 509–517. doi: 10.1016/j.neuroimage.2006.10.006.
- Compston, A. and Coles, A. (2008) 'Multiple sclerosis', *The Lancet*. Elsevier Ltd, 372(9648), pp. 1502–1517. doi: 10.1016/S0140-6736(08)61620-7.
- Correale, J. and Gaitán, M. I. (2015) 'Multiple sclerosis and environmental factors: the role of vitamin D, parasites, and Epstein-Barr virus infection', *Acta Neurologica Scandinavica*, 132(S199), pp. 46–55. doi: 10.1111/ane.12431.
- Dalton, C. M. *et al.* (2004) 'Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes', *Brain*, 127(5), pp. 1101–1107. doi: 10.1093/brain/awh126.
- Dargahi, N. *et al.* (2017) 'Multiple Sclerosis: Immunopathology and Treatment Update', *Brain Sciences*, 7(12), p. 78. doi: 10.3390/brainsci7070078.
- Dendrou, C. a., Fugger, L. and Friese, M. a. (2015) 'Immunopathology of multiple sclerosis', *Nature Reviews Immunology*. Nature Publishing Group, 15(9), pp. 545–558. doi: 10.1038/nri3871.
- Deutsche Gesellschaft für Neurologie (2015) 'Diagnose und Therapie der Multiplen Sklerose', pp. 1–52. Available at: <http://www.dgn.org/leitlinien/2333-ll-31-2012-diagnose-und-therapie-der-multiplen-sklerose>.
- Duning, T. *et al.* (2005) 'Dehydration confounds the assessment of brain atrophy', *Neurology*, 64(3), pp. 548–550. doi: 10.1212/01.WNL.0000150542.16969.CC.
- Durand-Dubief, F. *et al.* (2012) 'Reliability of Longitudinal Brain Volume Loss Measurements between 2 Sites in Patients with Multiple Sclerosis: Comparison of 7 Quantification Techniques', *American Journal of Neuroradiology*, 33(10), pp. 1918–1924. doi: 10.3174/ajnr.A3107.
- Eshaghi, A. *et al.* (2014) 'Temporal and spatial evolution of grey matter atrophy in primary progressive multiple sclerosis', *NeuroImage*. Elsevier Inc., 86, pp. 257–264. doi: 10.1016/j.neuroimage.2013.09.059.
- Evans, C. *et al.* (2013) 'Incidence and prevalence of multiple sclerosis in the americas: A

systematic review', *Neuroepidemiology*, 40(3), pp. 195–210. doi: 10.1159/000342779.

Fazekas, F. *et al.* (2002) 'CT and MRI rating of white matter changes.', *Journal of neural transmission. Supplementum*, 13(62), pp. 41–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12456048>.

Fazekas, F. *et al.* (2013) 'MRI in acute cerebral ischemia of the young: The Stroke in Young Fabry Patients (sifap1) Study', *Neurology*, 81(22), pp. 1914–1921. doi: 10.1212/01.wnl.0000436611.28210.ec.

Filippi, M. *et al.* (2013) 'Gray matter damage predicts the accumulation of disability 13 years later in MS', *Neurology*, 81(20), pp. 1759–1767. doi: 10.1212/01.wnl.0000435551.90824.d0.

Filippi, M. *et al.* (2016) 'MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines', *The Lancet Neurology*, 15(3), pp. 292–303. doi: 10.1016/S1474-4422(15)00393-2.

Fisher, E. *et al.* (2008) 'Gray matter atrophy in multiple sclerosis: A longitudinal study', *Annals of Neurology*, 64(3), pp. 255–265. doi: 10.1002/ana.21436.

Flachenecker, P., Henze, T. and Zettl, U. K. (2014) 'Spasticity in patients with multiple sclerosis - clinical characteristics, treatment and quality of life', *Acta Neurologica Scandinavica*, 129(3), pp. 154–162. doi: 10.1111/ane.12202.

Galeazzi, G. M. *et al.* (2005) 'Psychiatric disorders and depression in multiple sclerosis outpatients: Impact of disability and interferon beta therapy', *Neurological Sciences*, 26(4), pp. 255–262. doi: 10.1007/s10072-005-0468-8.

Giovannoni, G. *et al.* (2015) 'Is it time to target no evident disease activity (NEDA) in multiple sclerosis?', *Multiple Sclerosis and Related Disorders*, 4(4), pp. 329–333. doi: 10.1016/j.msard.2015.04.006.

Harper, L. *et al.* (2015) 'Using visual rating to diagnose dementia: a critical evaluation of MRI atrophy scales.', *Journal of neurology, neurosurgery, and psychiatry*, 86(11), pp. 1225–33. doi: 10.1136/jnnp-2014-310090.

Harper, L. *et al.* (2016) 'MRI visual rating scales in the diagnosis of dementia: evaluation in 184 post-mortem confirmed cases', *Brain*, 139(4), pp. 1211–1225. doi: 10.1093/brain/aww005.

Hasan, K. M. *et al.* (2011) 'Multimodal Quantitative Magnetic Resonance Imaging of Thalamic Development and Aging across the Human Lifespan: Implications to Neurodegeneration in Multiple Sclerosis', *The Journal of Neuroscience*, 31(46), pp. 16826–16832. doi: 10.1523/JNEUROSCI.4184-11.2011.

Healy, B. *et al.* (2009) 'Sample size requirements for treatment effects using gray matter, white matter and whole brain volume in relapsing-remitting multiple sclerosis.', *Journal of neurology, neurosurgery, and psychiatry*, 80(11), pp. 1218–23. doi: 10.1136/jnnp.2008.154732.

Hohlfeld, R. and Wekerle, H. (2004) 'Autoimmune concepts of multiple sclerosis as a basis for selective immunotherapy: from pipe dreams to (therapeutic) pipelines', *Proc Natl Acad Sci U S A*, 101 Suppl, pp. 14599–14606. doi: 10.1073/pnas.0404874101.

International Multiple Sclerosis Genetics Consortium *et al.* (2007) 'Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study', *New England Journal of Medicine*, 357(9), pp. 851–862. doi: 10.1056/NEJMoa073493.

Jacobsen, C. *et al.* (2014) 'Brain atrophy and disability progression in multiple sclerosis patients: a 10-year follow-up study', *J Neurol Neurosurg Psychiatry*, 85, pp. 1109–1115. doi: 10.1136/jnnp-2013-306906.

Jokinen, H. *et al.* (2012) 'Brain atrophy accelerates cognitive decline in cerebral small vessel disease: The LADIS study', *Neurology*, 78(22), pp. 1785–1792. doi: 10.1212/WNL.0b013e3182583070.

Kappos, L. *et al.* (2016) 'Inclusion of brain volume loss in a revised measure of "no

- evidence of disease activity” (NEDA-4) in relapsing–remitting multiple sclerosis’, *Multiple Sclerosis Journal*, 22(10), pp. 1297–1305. doi: 10.1177/1352458515616701.
- van Kessel, K. and Moss-Morris, R. (2006) ‘Understanding multiple sclerosis fatigue: A synthesis of biological and psychological factors’, *Journal of Psychosomatic Research*, 61(5), pp. 583–585. doi: 10.1016/j.jpsychores.2006.03.006.
- Kingwell, E. *et al.* (2013) ‘Incidence and prevalence of multiple sclerosis in Europe: a systematic review.’, *BMC neurology*, 13, p. 128. doi: 10.1186/1471-2377-13-128.
- Kisic Tepavcevic, D. *et al.* (2017) ‘Bladder dysfunction in multiple sclerosis: a 6-year follow-up study’, *Acta Neurologica Belgica*. Springer International Publishing, 117(1), pp. 83–90. doi: 10.1007/s13760-016-0741-z.
- Koch, M. *et al.* (2007) ‘Tremor in multiple sclerosis’, *Journal of Neurology*, 254(2), pp. 133–145. doi: 10.1007/s00415-006-0296-7.
- Kurtzke, J. F. (1983) ‘Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS)’, *Neurology*, 33(11), pp. 1444–1444. doi: 10.1212/WNL.33.11.1444.
- Lansley, J. *et al.* (2013) ‘Localized grey matter atrophy in multiple sclerosis: A meta-analysis of voxel-based morphometry studies and associations with functional disability’, *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd, 37(5), pp. 819–830. doi: 10.1016/j.neubiorev.2013.03.006.
- Lavorgna, L. *et al.* (2014) ‘Clinical and magnetic resonance imaging predictors of disease progression in multiple sclerosis: a nine-year follow-up study’, *Multiple Sclerosis Journal*, 20(2), pp. 220–226. doi: 10.1177/1352458513494958.
- Lill, C. M. (2014) ‘Recent advances and future challenges in the genetics of multiple sclerosis’, *Frontiers in Neurology*, 5 JUL(July), pp. 1–4. doi: 10.3389/fneur.2014.00130.
- Lublin, F. D. *et al.* (2014) ‘Defining the clinical course of multiple sclerosis: The 2013 revisions’, *Neurology*, 83(3), pp. 278–286. doi: 10.1212/WNL.0000000000000560.
- Lünemann, J. D. *et al.* (2006) ‘Increased frequency and broadened specificity of latent EBV nuclear antigen-1-specific T cells in multiple sclerosis’, *Brain*, 129(6), pp. 1493–1506. doi: 10.1093/brain/awl067.
- Mahad, D. H., Trapp, B. D. and Lassmann, H. (2015) ‘Pathological mechanisms in progressive multiple sclerosis’, *The Lancet Neurology*. Elsevier Ltd, 14(2), pp. 183–193. doi: 10.1016/S1474-4422(14)70256-X.
- van der Mei, I. A. F. *et al.* (2007) ‘Vitamin D levels in people with multiple sclerosis and community controls in Tasmania, Australia.’, *Journal of neurology*, 254(5), pp. 581–90. doi: 10.1007/s00415-006-0315-8.
- Miller, D. *et al.* (2008) ‘Differential diagnosis of suspected multiple sclerosis: a consensus approach’, *Multiple Sclerosis Journal*, 14(9), pp. 1157–1174. doi: 10.1177/1352458508096878.
- Miller, D. H., Chard, D. T. and Ciccarelli, O. (2012) ‘Clinically isolated syndromes’, *The Lancet Neurology*. Elsevier Ltd, 11(2), pp. 157–169. doi: 10.1016/S1474-4422(11)70274-5.
- Minagar, A. *et al.* (2004) ‘Pathogenesis of Brain and Spinal Cord Atrophy in Multiple Sclerosis’, *Journal of Neuroimaging*, 14(3), pp. 5–10. doi: 10.1177/1051228404266263.
- Minagar, A. *et al.* (2013) ‘The thalamus and multiple sclerosis: modern views on pathologic, imaging, and clinical aspects.’, *Neurology*, 80(2), pp. 210–9. doi: 10.1212/WNL.0b013e31827b910b.
- Mühlau, M. *et al.* (2013) ‘White-matter lesions drive deep gray-matter atrophy in early multiple sclerosis: support from structural MRI.’, *Multiple sclerosis (Houndmills, Basingstoke, England)*, 19(11), pp. 1485–92. doi: 10.1177/1352458513478673.
- Nakamura, K. *et al.* (2015) ‘Diurnal fluctuations in brain volume: Statistical analyses of MRI from large populations’, *NeuroImage*. Elsevier Inc., 118, pp. 126–132. doi:

10.1016/j.neuroimage.2015.05.077.

Neacsu, V. *et al.* (2008) 'Agreement between different input image types in brain atrophy measurement in multiple sclerosis using SIENAX and SIENA', *Journal of Magnetic Resonance Imaging*, 28(3), pp. 559–565. doi: 10.1002/jmri.21501.

O’Gorman, C. *et al.* (2013) 'Modelling Genetic Susceptibility to Multiple Sclerosis with Family Data', *Neuroepidemiology*, 40(1), pp. 1–12. doi: 10.1159/000341902.

Okuda, D. T. *et al.* (2009) 'Incidental MRI anomalies suggestive of multiple sclerosis: The radiologically isolated syndrome', *Neurology*, 72(9), pp. 800–805. doi: 10.1212/01.wnl.0000335764.14513.1a.

Olsson, E. *et al.* (2013) 'White matter lesion assessment in patients with cognitive impairment and healthy controls: Reliability comparisons between visual rating, a manual, and an automatic volumetrical MRI method - The gothenburg MCI study', *Journal of Aging Research*, 2013. doi: 10.1155/2013/198471.

Ozakbas, S., Ormeci, B. and Idiman, E. (2005) 'Utilization of the multiple sclerosis functional composite in follow-up: Relationship to disease phenotype, disability and treatment strategies', *Journal of the Neurological Sciences*, 232(1–2), pp. 65–69. doi: 10.1016/j.jns.2005.01.008.

Pantoni, L. *et al.* (2005) 'Impact of Age-Related Cerebral White Matter Changes on the Transition to Disability – The LADIS Study: Rationale, Design and Methodology', *Neuroepidemiology*, 24(1–2), pp. 51–62. doi: 10.1159/000081050.

Pareto, D. *et al.* (2016) 'Lesion filling effect in regional brain volume estimations: a study in multiple sclerosis patients with low lesion load', *Neuroradiology*. *Neuroradiology*, 58(5), pp. 467–474. doi: 10.1007/s00234-016-1654-5.

Pérez-Miralles, F. *et al.* (2013) 'Clinical impact of early brain atrophy in clinically isolated syndromes', *Multiple Sclerosis Journal*, 19(14), pp. 1878–1886. doi: 10.1177/1352458513488231.

Pichler, A. *et al.* (2016) 'Combined analysis of global and compartmental brain volume changes in early multiple sclerosis in clinical practice', *Multiple Sclerosis Journal*, 22(3), pp. 340–346. doi: 10.1177/1352458515593405.

Polman, C. H. *et al.* (2005) 'Diagnostic criteria for multiple sclerosis: 2005 Revisions to the “McDonald Criteria”', *Annals of Neurology*. doi: 10.1002/ana.206703.

Popescu, V. *et al.* (2013) 'Brain atrophy and lesion load predict long term disability in multiple sclerosis', *Journal of Neurology, Neurosurgery & Psychiatry*, 84(10), pp. 1082–1091. doi: 10.1136/jnnp-2012-304094.

Popescu, V. (2015) *Brain atrophy in multiple sclerosis : quantifi cation, pathological substrate and clinical relevance*. Edited by Universiteit Amsterdam.

Popescu, V. *et al.* (2015) 'What drives MRI-measured cortical atrophy in multiple sclerosis?', *Multiple sclerosis (Houndmills, Basingstoke, England)*, 21(10), pp. 1280–90. doi: 10.1177/1352458514562440.

Poser, C. M. and Brinar, V. V (2001) 'Diagnostic criteria for multiple sclerosis.', *Clinical neurology and neurosurgery*, 103(1), pp. 1–11. doi: 10.1016/j.clineuro.2004.02.004.

Prins, D. *et al.* (2004) 'Measuring progression of cerebral white matter lesions on MRI', *Neurology*, pp. 1533–1539.

Reuter, M. *et al.* (2015) 'Head motion during MRI acquisition reduces gray matter volume and thickness estimates', *NeuroImage*. Elsevier Inc., 107, pp. 107–115. doi: 10.1016/j.neuroimage.2014.12.006.

Rieckmann, P. (2009) 'Concepts of induction and escalation therapy in multiple sclerosis', *Journal of the Neurological Sciences*. Elsevier B.V., 277(SUPPL. 1), pp. S42–S45. doi: 10.1016/S0022-510X(09)70012-7.

Rocca, M. A. *et al.* (2010) 'Thalamic damage and long-term progression of disability in multiple sclerosis.', *Radiology*, 257(2), pp. 463–9. doi: 10.1148/radiol.10100326.

- Rocca, M. A. *et al.* (2017) 'Brain MRI atrophy quantification in MS', *Neurology*, 88(4), pp. 403–413. doi: 10.1212/WNL.0000000000003542.
- Salhofer-Polanyi, S. *et al.* (2017) 'Epidemiology of Multiple Sclerosis in Austria', *Neuroepidemiology*, 49(1–2), pp. 40–44. doi: 10.1159/000479696.
- Sastre-Garriga, J. *et al.* (2005) 'Metabolite changes in normal-appearing gray and white matter are linked with disability in early primary progressive multiple sclerosis.', *Archives of Neurology*, 62(4), pp. 569–73. doi: 10.1001/archneur.62.4.569.
- Sawcer, S., Franklin, R. J. M. and Ban, M. (2014) 'Multiple sclerosis genetics', *The Lancet Neurology*. Elsevier Ltd, 13(7), pp. 700–709. doi: 10.1016/S1474-4422(14)70041-9.
- Schmidt, R. M. (2015) *Multiple Sklerose*. Elsevier. doi: 10.1016/C2013-0-22879-8.
- Sicotte, N. L. *et al.* (2008) 'Regional hippocampal atrophy in multiple sclerosis', *Brain*, 131(4), pp. 1134–1141. doi: 10.1093/brain/awn030.
- Siegert, R. J. and Abernethy, D. A. (2005) 'Depression in multiple sclerosis: a review.', *Journal of neurology, neurosurgery, and psychiatry*, 76(4), pp. 469–475. doi: 10.1136/jnnp.2004.054635.
- Siffrin, V. *et al.* (2010) 'Multiple sclerosis – candidate mechanisms underlying CNS atrophy', *Trends in Neurosciences*. Elsevier Ltd, 33(4), pp. 202–210. doi: 10.1016/j.tins.2010.01.002.
- Simon, J. H. *et al.* (1999) 'A longitudinal study of brain atrophy in relapsing multiple sclerosis', *Neurology*, 53(1), pp. 139–148. doi: 10.1212/WNL.53.1.139.
- Simpson Jr., S. *et al.* (2011) 'Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis', *J Neurol Neurosurg Psychiatry*, 82(10), pp. 1132–1141. doi: jnnp.2011.240432 [pii]r10.1136/jnnp.2011.240432.
- Smith, S. M. *et al.* (2002) 'Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis', *NeuroImage*, 17(1), pp. 479–489. doi: 10.1006/nimg.2002.1040.
- Smith, S. M. *et al.* (2004) 'Advances in functional and structural MR image analysis and implementation as FSL', *NeuroImage*, 23(SUPPL. 1), pp. 208–219. doi: 10.1016/j.neuroimage.2004.07.051.
- Sormani, M. P., Arnold, D. L. and De Stefano, N. (2014) 'Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis', *Annals of Neurology*, 75(1), pp. 43–49. doi: 10.1136/jnnp.2008.154732.
- Steenwijk, M. D. *et al.* (2015) 'Unraveling the relationship between regional gray matter atrophy and pathology in connected white matter tracts in long-standing multiple sclerosis', *Human Brain Mapping*, 36(5), pp. 1796–1807. doi: 10.1002/hbm.22738.
- De Stefano, N. *et al.* (2010) 'Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes.', *Neurology*, 74(23), pp. 1868–76. doi: 10.1212/WNL.0b013e3181e24136.
- De Stefano, N. *et al.* (2014) 'Clinical relevance of brain volume measures in multiple sclerosis', *CNS Drugs*, 28(2), pp. 147–156. doi: 10.1007/s40263-014-0140-z.
- De Stefano, N. *et al.* (2016) 'Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis.', *Journal of neurology, neurosurgery, and psychiatry*, 87(1), pp. 93–9. doi: 10.1136/jnnp-2014-309903.
- Storelli, L. *et al.* (2018) 'Measurement of Whole-Brain and Gray Matter Atrophy in Multiple Sclerosis: Assessment with MR Imaging', *Radiology*, 288(2), pp. 554–564. doi: 10.1148/radiol.2018172468.
- Thacker, E. L., Mirzaei, F. and Ascherio, A. (2006) 'Infectious mononucleosis and risk for multiple sclerosis: A meta-analysis', *Annals of Neurology*, 59(3), pp. 499–503. doi: 10.1002/ana.20820.
- Thompson, A. J. *et al.* (2018) 'Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria', *The Lancet Neurology*, 17(2), pp. 162–173. doi: 10.1016/S1474-

4422(17)30470-2.

Thompson, A. J. and Baneke, P. (2013) *Multiple Sclerosis International Federation (MSIF) Design and editorial support by Summers Editorial & Design Graphics by Nutmeg Productions Printed by Modern Colour Solutions*. Available at:

<https://www.msif.org/wp-content/uploads/2014/09/Atlas-of-MS.pdf> (Accessed: 28 February 2016).

Tintore, M. *et al.* (2015) 'Defining high, medium and low impact prognostic factors for developing multiple sclerosis', *Brain*, 138(7), pp. 1863–1874. doi: 10.1093/brain/awv105.

Trapp, B. D. *et al.* (1998) 'Axonal Transection in the Lesions of Multiple Sclerosis', *New England Journal of Medicine*, 338(5), pp. 278–285. doi: 10.1056/NEJM199801293380502.

Uher, T. *et al.* (2016) 'Combining clinical and magnetic resonance imaging markers enhances prediction of 12-year disability in multiple sclerosis.', *Multiple sclerosis (Houndmills, Basingstoke, England)*, pp. 1–11. doi: 10.1177/1352458516642314.

Vidal-Jordana, A. *et al.* (2015) 'Treating relapsing–remitting multiple sclerosis: therapy effects on brain atrophy', *Journal of Neurology*. Springer Berlin Heidelberg, 262(12), pp. 2617–2626. doi: 10.1007/s00415-015-7798-0.

Vollmer, T. *et al.* (2015) 'The natural history of brain volume loss among patients with multiple sclerosis: a systematic literature review and meta-analysis', *Journal of the neurological sciences*. Elsevier B.V., 357(1–2), pp. 8–18. doi: 10.1016/j.jns.2015.07.014.

Wahlund, L.-O. (2000) 'Visual rating and volumetry of the medial temporal lobe on magnetic resonance imaging in dementia: a comparative study', *Journal of Neurology, Neurosurgery & Psychiatry*, 69(5), pp. 630–635. doi: 10.1136/jnnp.69.5.630.

Wingerchuk, D. M. and Carter, J. L. (2014) 'Multiple sclerosis: Current and emerging disease-modifying therapies and treatment strategies', *Mayo Clinic Proceedings*. Elsevier Inc, 89(2), pp. 225–240. doi: 10.1016/j.mayocp.2013.11.002.

Young, C. A. (2011) 'Factors predisposing to the development of multiple sclerosis', *QJM*, 104(5), pp. 383–386. doi: 10.1093/qjmed/hcr012.

Zhang, P. *et al.* (2016) 'The risk of smoking on multiple sclerosis: a meta-analysis based on 20,626 cases from case-control and cohort studies', *PeerJ*, 4, p. e1797. doi: 10.7717/peerj.1797.

Zhou, Y. *et al.* (2014) 'The potential role of epigenetic modifications in the heritability of multiple sclerosis.', *Multiple sclerosis (Houndmills, Basingstoke, England)*, 20, pp. 135–40. doi: 10.1177/1352458514520911.

Zivadinov, R. *et al.* (2008) 'Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis.', *Neurology*, 71(2), pp. 136–44. doi: 10.1212/01.wnl.0000316810.01120.05.

Zivadinov, R. *et al.* (2013) 'Evolution of cortical and thalamus atrophy and disability progression in early relapsing-remitting MS during 5 years', *American Journal of Neuroradiology*, 34(10), pp. 1931–1939. doi: 10.3174/ajnr.A3503.

Zivadinov, R. *et al.* (2018) 'Feasibility of Brain Atrophy Measurement in Clinical Routine without Prior Standardization of the MRI Protocol: Results from MS-MRIUS, a Longitudinal Observational, Multicenter Real-World Outcome Study in Patients with Relapsing-Remitting MS', *American Journal of Neuroradiology*, 39(2), pp. 289–295. doi: 10.3174/ajnr.A5442.

6 Anhang

6.1 Rating Sheet

Visuelles Atrophie-Rating

Patient: ID- (ersten 5 Zahlen)

Rater: Johanna Leitgeb Christian Enzinger

Beurteilt am:

Zuordnung der Scans zu einem Untersuchungszeitpunkt:

Welcher Scan entspricht der Baseline Untersuchung: ID- (s.oben) _

Welcher Scan entspricht der Follow-Up Untersuchung: ID- (s.oben) _

Sicherheitsgrad der Zuordnung:

möglich(er Unterschied) sicher(er Unterschied) nicht möglich, kein Unterschied

Bild-Qualität:

Baseline Untersuchung: gut ausreichend nicht beurteilbar

Follow-Up Untersuchung: gut ausreichend nicht beurteilbar

Rating der Ventrikel-Weite:

Baseline Untersuchung: V1 V2 V3 V4 V5 V6 V7 V8

Follow-Up Untersuchung: V1 V2 V3 V4 V5 V6 V7 V8

Rating der sulkalen Weite:

Baseline Untersuchung: S1 S2 S3 S4 S5 S6 S7 S8

Follow-Up Untersuchung: S1 S2 S3 S4 S5 S6 S7 S8

Visuell globale Beurteilung:

Globale Änderung der Ventrikelweite: keine gering mäßig ausgeprägt

Globale Änderung der sulkalen Weite: keine gering mäßig ausgeprägt

Lineare Maße (nach J Simon et al, Neurology. 1999 Jul 13;53(1):139-48):

III-Ventrikel Weite : Baseline Follow-up.....

Parenchym entlang dieser Linie: Baseline Follow-up.....

Weite Seitenventrikel: Baseline Follow-up.....

Parenchym entlang dieser Linie: Baseline Follow-up

Bemerkung:

6.2 SIFAP-Rating Skala

