

Aus der
Klinik und Poliklinik für Radiologie
Klinikum der Ludwig-Maximilians-Universität München



**Automatische Hirn- und Läsionsvolumetrie bei Multipler Sklerose:
Vergleich von zwei kommerziellen Softwarelösungen**

Dissertation
zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von
Lisa Krause

aus
Quedlinburg

Jahr
2025

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten: Prof. Dr. Mike Notohamiprodjo
Zweites Gutachten: Prof. Dr. Matthias Brendel
Drittes Gutachten: Priv. Doz. Dr. Susanne Schwarzmaier

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 26.06.2025

Inhaltsverzeichnis

1	Zusammenfassung	3
2	Einleitung	4
2.1	Multiple Sklerose	4
2.2	Die Hirnatrophie als Maß der Neurodegeneration	5
2.3	Diagnosestellung mithilfe der Magnetresonanztomografie	6
2.3.1	T2-gewichtete und FLAIR Sequenzen	6
2.3.2	Black Holes und Gadolinium-verstärkte T1-Sequenzen	7
2.3.3	3D-Bildgebung	8
2.4	Hirnvolumetrie mithilfe der MRT	9
2.4.1	Bisherige Hirnvolumetrie	9
2.4.2	KI-basierte Softwarelösungen	10
2.4.3	Deep Learning	13
3	Experimentelle Fragestellung	14
4	Material und Methoden	15
4.1	Patientenkollektiv	15
4.2	MRT und Bildakquirierung	15
4.3	Parameter und Software	17
4.3.1	icobrain ms	17
4.3.2	mdbrain	19
4.4	Statistik	21
5	Ergebnisse	23
5.1	Patientenkollektiv	23
5.2	Volumen und Altersperzentilen des Gesamthirns	23
5.2.1	Native Untersuchungen des Gesamthirns	23
5.2.2	Untersuchungen mit KM des Gesamthirns	26

5.3	Volumen und Altersperzentilen der grauen Substanz	29
5.3.1	Native Untersuchungen der grauen Substanz	29
5.3.2	Untersuchungen mit KM der grauen Substanz	31
5.4	Läsionslast	34
5.4.1	Native Messungen der Läsionslast	34
5.4.2	Kontrastmittelverstärkte Messungen der Läsionslast	35
5.4.3	Messungen ohne Berücksichtigung des KM-Status	36
5.5	Zusammenfassung der Ergebnisse	37
6	Diskussion	39
6.1	Atrophie	39
6.2	Läsionsvolumina	41
6.3	Limitationen der Hirnvolumetrie	41
6.3.1	Limitationen der Hirnvolumetrie generell	41
6.3.2	Limitationen dieser Studie	42
6.4	Schlussfolgerung	43
7	Literaturverzeichnis	44
8	Abkürzungsverzeichnis	52
9	Abbildungsverzeichnis	53
10	Tabellenverzeichnis	54
	Danksagung	55
	Lebenslauf	56
	Publikationen	57
	Eidesstattliche Versicherung	58
	Übereinstimmungserklärung	59

1 Zusammenfassung

Zielsetzung: Die Studie hatte das Ziel, das Hirn- und Läsionsvolumen bei Multiple-Sklerose-Patienten mithilfe von zwei vollautomatisierten, KI-gestützten Software-Lösungen (*icobrain ms* und *mdbrain*) zu quantifizieren und die Übereinstimmung der Ergebnisse zu überprüfen. Die Analyse umfasste die Übereinstimmung des absoluten und relativen Volumens des Gesamthirns und der grauen Substanz, sowie der Läsionslast.

Material und Methoden: Es wurden 86 Patienten mit gesicherter Multipler Sklerose mit einer isotropen T1-MPRAGE-Sequenz und einer isotropen 3D-FLAIR-Sequenz untersucht. Die T1-Sequenzen wurden bei 42 Patienten ohne und bei 44 Patienten mit Kontrastmittel durchgeführt. Es erfolgte eine automatisierte Analyse mit zwei unterschiedlichen Software-Lösungen (*mdbrain* v2.0 (MDB) und *icobrain ms* (ICO)). Das absolute (in ml) und relative (Altersperzentile (PERZ)) Volumen für die gesamte Hirnmasse und graue Substanz wurden berechnet. Das Volumen für periventrikuläre, juxtakortikale, infratentorielle und unspezifische FLAIR-hyperintense Läsionen wurde absolut (in ml) berechnet. Die statistische Auswertung erfolgte mit verbundenen t-Tests für die absoluten Volumen und dem Wilcoxon-Test für die relativen, nicht parametrischen Volumen.

Ergebnisse: Bei nativ durchgeführten Untersuchungen zeigte sich ein signifikanter Unterschied ($p < 0,01$) für das absolute, aber nicht für das relative Gesamthirnvolumen. Das absolute und relative Volumen der grauen Substanz zeigten sich signifikant unterschiedlich.

Bei kontrastverstärkten Untersuchungen zeigte sich kein signifikanter Unterschied für das absolute, aber für das relative Gesamthirnvolumen. Bei der grauen Substanz zeigte sich sowohl das absolute, als auch das relative Volumen signifikant unterschiedlich.

Bei den FLAIR-hyperintensiven Läsionen zeigte sich ein signifikanter Unterschied ($p < 0,01$) nur für infratentorielle Läsionen der kontrastmittelgestützten Messungen, aber nicht für andere Regionen. Die nativen Messungen und ohne Berücksichtigung des Kontrastmittelstatus verblieben ohne signifikanten Unterschied.

Schlussfolgerung: Es zeigen sich bei der automatisierten Auswertung durch zwei unterschiedliche Software-Lösungen signifikante Unterschiede hinsichtlich des absoluten und relativen Volumens des Gesamthirns und der grauen Substanz, sowie für das Volumen von infratentoriellen Läsionen.

2 Einleitung

Die automatische Hirn- und Läsionsvolumetrie bei Multipler Sklerose (MS) hat in den letzten Jahren eine entscheidende Rolle in der klinischen Praxis und der Forschung eingenommen. Die präzise Quantifizierung von Hirnatrophie und Läsionslast ist von entscheidender Bedeutung für die Verlaufskontrolle, Prognosestellung und Behandlungsüberwachung von MS-Patienten. In diesem Zusammenhang haben sich verschiedene kommerzielle Softwarelösungen etabliert, die auf automatisierten Bildanalysealgorithmen basieren und eine schnelle und reproduzierbare Volumetrie ermöglichen. Diese Arbeit stellt einen Vergleich zwischen zwei kommerziellen Softwarelösungen zur automatischen Hirn- und Läsionsvolumetrie bei MS vor und diskutiert deren Vor- und Nachteile sowie deren klinische Anwendbarkeit.

2.1 Multiple Sklerose

Die Multiple Sklerose oder Enzephalitis disseminata ist die häufigste immunvermittelte, entzündliche, demyelinisierende Erkrankung des zentralen Nervensystems und nach der Epilepsie die zweithäufigste neurologische Erkrankung jüngerer Erwachsener (Hartung, 2002). Besonders hohe Prävalenzen treten in den gemäßigten Breiten der nördlichen und südlichen Hemisphäre auf. Häufig betroffen sind vor allem Bewohner von Europa, Russland, dem nördlichen Amerika, Neuseeland und Australien (Mattle und Mumenthaler, 2015), (Browne et al., 2014).

Charakteristisch ist eine multifokale Inflammation mit vermehrtem Einwandern von Makrophagen, Lymphozyten (vorwiegend T-Zellen) und Plasmazellen (Lassmann, 2018). Diese überwinden die Blut-Hirn-Schranke und attackieren das Myelin. Vermehrte Flüssigkeitsansammlung im Gehirngewebe und Verlust der Myelinscheiden mit Untergang der Oligodendrozyten sind die Folgen. Die Demyelinisierung erstreckt sich typischerweise an den lateralen Ventrikeln und juxtakortikalen Marklagern entlang der Venen und Venolen, die von dort ins Hirngewebe ziehen (Giorgio und de Stefano, 2013); (Nakamura et al., 2014); (de Stefano et al., 2014); (Hemond und Bakshi, 2018). Lange ordnete man die Multiple Sklerose als eine Erkrankung der weißen Substanz ein (Naganawa, 2015); (Cortese et al., 2019); (Marciniewicz et al., 2019). Durch Remyelinisierung und reaktiver Proliferation von Astrozyten kommt es zu einer Vernarbung des glialen Gewebes, was der Erkrankung den Namen Sklerose gab (Mattle und Mumenthaler, 2015).

Neuere Studien belegen, dass die Hirnatrophie und der Untergang der grauen Substanz mit

Schädigung der Neuronen, Axone und Synapsen eine ebenfalls große Rolle beim Pathomechanismus der MS spielen (Naganawa, 2015); (Cortese et al., 2019); (Marciniewicz et al., 2019). Betroffen ist besonders der Kortex, die Basalganglien und Hirnstamm, das Kleinhirn und die graue Substanz des Rückenmarks.

Während das Ausmaß der Zellinfiltration, Entzündung und des Ödems vermehrt in der weißen Substanz auftritt, ist die Remyelinisierung führend in der grauen Substanz. Das Ausmaß der Remyelinisierung ist je nach Läsionsort und Patient unterschiedlich ausgeprägt und bedingen die klinischen Symptome (Lassmann, 2018).

2.2 Die Hirnatrophie als Maß der Neurodegeneration

In den letzten zehn Jahren ist die Hirnatrophie einer der wichtigsten Indikatoren der Neurodegeneration und klinischen Fortschreitens der Erkrankung bei MS-Patienten geworden (Marciniewicz et al., 2019). Während der jährliche Hirnverlust in der Normalbevölkerung durch Alterungsprozesse bei circa 0,1 bis 0,3 Prozent liegt, sind bei der MS Minderungen von 0,5 bis 1,35 Prozent anzunehmen (Giorgio und de Stefano, 2013); (Nakamura et al., 2014); (de Stefano et al., 2014). Am Prozess der Atrophie sind unter anderem die Entzündung, die Demyelinisierung, die Neurodegeneration durch axonalen Schaden, sowie der Verlust von Gliazellen und Myelin beteiligt (Giorgio et al., 2008); (Giorgio und de Stefano, 2013). Der Abbau beginnt schon in den frühen Stadien der Erkrankung, wobei hier besonders die graue Substanz betroffen ist. Er schreitet progressiv voran und findet unabhängig vom Verlust der weißen Substanz statt (Hemond und Bakshi, 2018). Die Unterformen der MS sind davon gleichermaßen betroffen (Sastre-Garriga et al., 2020); (Sormani et al., 2017).

Bei der Multiplen Sklerose sind vor allem die Frontal-, Temporal- und Parietallappen von der Volumenminderung betroffen. Auch Strukturen der grauen Substanz, wie der Thalamus, der Hippocampus, die Amygdala und der Nucleus caudatus, werden besonders in Mitleidenschaft gezogen (Giorgio und de Stefano, 2013); (Marciniewicz et al., 2019); (Sinnecker et al., 2018). Die Atrophie allgemein, aber besonders die Atrophie des Thalamus, korreliert mit dem Behinderungsgrad, der physischen und der kognitiven Entwicklung der MS-Patienten. Die Läsionslast scheint hierbei eine untergeordnete Rolle zu spielen (Giorgio et al., 2008); (Sinnecker et al., 2018); (Filippi et al., 2016); (Hemond und Bakshi, 2018); (Sahraian und Radü, 2008); (Sastre-Garriga et al., 2020). Studien zu Folge war der Hirngewebsverlust bei den MS Patienten mit fortschreitender Erkrankung höher, als bei MS Patienten, die klinisch stabil blieben. Auch Patienten mit einem klinisch isolierten Syndrom, die in eine Multiple Sklerose

übergangen, hatten mehr Hirnvolumenverlust als Patienten, die keine MS aus dem klinisch isoliertem Syndrom entwickelten (de Stefano et al., 2014).

In Summe macht dies die Hirnatrophie einen immer bedeutenderen Biomarker der Multiplen Sklerose (Marciniewicz et al., 2019); (Sima et al., 2017).

2.3 Diagnosestellung mithilfe der Magnetresonanztomografie

Um die Diagnose zu stellen und den Verlauf der Erkrankung zu verfolgen, wurde die Magnetresonanztomografie (MRT) dank des guten Gewebekontrastes, der hohen Auflösung und fehlenden Gesundheitsschäden in den letzten 20-30 Jahren als Bildgebung der Wahl etabliert (Giorgio und de Stefano, 2013); (Marciniewicz et al., 2019); (Sastre-Garriga et al., 2020).

Im klinischen Alltag erfolgt die Diagnosestellung auf MRT-Bildern mithilfe der McDonalds Kriterien (Mattle und Mumenthaler, 2015); (McDonald et al., 2001b). Diese in der Routine angewendeten Kriterien machen es leichter, die Diagnose der Multiplen Sklerose zu stellen und sie von ihren Differentialdiagnosen abzugrenzen. Die FLAIR- und T2-gewichtete Sequenz des Gehirns und des Rückenmarks stellen zusammen mit T1-gewichteten Aufnahmen mit und ohne Kontrastmittel die Grundlagen hierfür dar (Thompson et al., 2017); (Hemond und Bakshi, 2018). Die McDonalds-Kriterien fordern neben klinischen Charakteristika unter anderem den Nachweis der typischen Veränderungen in der MRT als räumliche und zeitliche Dissemination. Die räumliche Dissemination in der MRT verlangt den Nachweis von mindestens einer hyperintensiven T2-Läsion in mindestens zwei der typischen Regionen (periventrikulär, juxtakortikal, infratentoriell, spinal, siehe 2.1) (Lukas und Sailer, 2015). Zur Verbesserung der Diagnosestellung empfahl Filippi et al. (2016) den Sehnerv als fünften typischen Ort der Läsionen mit hinzuzufügen. Wegen fehlender Studienlage wurde dies jedoch noch nicht in die McDonalds Kriterien aufgenommen (Filippi et al., 2016).

Die zeitliche Dissemination bei der MRT erfordert entweder den gleichzeitigen Nachweis von Gadolinium anreichernden und nicht anreichernden Läsionen zu einem beliebigen Zeitpunkt, oder die Feststellung einer neuen (Gadolinium anreichernden) T2-Läsion im Vergleich zu einer Referenzaufnahme (Neurologienetz, 2021); (Thompson et al., 2018).

2.3.1 T2-gewichtete und FLAIR Sequenzen

Die typische MS-Läsion stellt sich in der T2-gewichteten MRT-Aufnahme als ovale, anfangs scharf abgrenzbare Struktur mit hoher Signalintensität dar. Später können die Läsionen auch konfluieren. Die Hyperintensität liegt an dem erhöhten Wassergehalt des geschädigten

Gewebes im Gegensatz zum umgebenden Hirnparenchym. Zur Diagnosestellung sind eine oder mehr Hyperintensitäten in mindestens zwei der vier, beziehungsweise fünf typischen MS-Regionen notwendig (Filippi et al., 2016); (Hemond und Bakshi, 2018). Besonders für die Detektion von Läsionen in der hinteren Schädelgrube (Hirnstamm und Kleinhirn) ist die herkömmliche T2-gewichtete Spin-Echo Sequenz führend, da hier Flussartefakte weniger Einfluss nehmen. Die FLAIR-Sequenz ist als stark T2-gewichtete Sequenz mit Unterdrückung des Liquors führend für die Detektion der typischen periventrikulären Läsionen. Auch die Flüssigkeit im Hirngewebe selbst wird unterdrückt, sodass sich ebenfalls der Kontrast innerhalb der grauen Substanz erhöht. Juxtakortikale und kortikale Läsionen können so besser hervorgehoben werden. Mit der FLAIR würde die Läsionslast infratentoriell und im Hirnstamm unterschätzt werden (Rupp, 2016); (Sahraian und Radü, 2008); (Vural et al., 2013); (Hemond und Bakshi, 2018).

Die Hyperintensitäten sind jedoch nicht unbedingt spezifisch für MS. Annähernd jede Änderung der Gewebezusammensetzung kann das Signal hell erscheinen lassen, wie zum Beispiel Entzündungen, Flüssigkeitsansammlung, De-/Remyelinisierung, Gliose, sowie axonaler Verlust (Filippi et al., 2011); (Hemond und Bakshi, 2018). Spezifisch für die MS sind die sogenannten Dawson-Finger. Dies sind periventrikuläre Läsionen, die sich fingerförmig entlang des Marklagers anordnen (Hemond und Bakshi, 2018). Krankhafte Gewebeveränderungen in anderen Regionen, vor allem in der grauen Substanz des Kortex, werden mit der T2-gewichteten Bildgebung oft wegen der Signalähnlichkeit der MS-Läsion zur grauen Substanz nicht erfasst (Sahraian und Radü, 2008). Außerdem ist sie zur Unterscheidung zwischen akuten und chronischen Läsionen nicht imstande. Auch der Grad an Demyelinisierung und axonalem Schaden kann mit der T2-gewichteten Sequenz nicht widerspiegelt werden (Hu et al., 2019).

2.3.2 Black Holes und Gadolinium-verstärkte T1-Sequenzen

Während Strukturen mit einem hohen Fettgehalt, wie zum Beispiel Myelin, in T1-gewichteten Sequenzen hell erscheinen, sind wasserhaltige Strukturen, wie der Kortex, hypointens.

Akute Läsionen erscheinen aufgrund des Ödems im Rahmen der Entzündung in der nativen T1-gewichteten Bildgebung somit dunkel. Chronische, irreversible Gewebeveränderungen, die von axonalem Schaden und Demyelinisierung geprägt sind, bilden sogenannte "Black Holes" in der nativen T1 (Sima et al., 2017). Unter Kontrastmittelgabe kann die entzündliche Aktivität und das Alter der Läsionen aufgedeckt werden. Gadolinium verkürzt die Relaxationszeit und hat

eine hyperintense Erscheinung des aufnehmenden Gewebes zur Folge. Das intravenös verabreichte Kontrastmittel überschreitet die gestörte Blut-Hirn-Schranke akuter Läsionen und tritt ins Gewebe über. Die neuen Läsionen nehmen so vermehrt Kontrastmittel auf und erscheinen in der MRT hell. Die KM-Aufnahme ist dabei meist homogen, kann sich aber auch inhomogen, knotig oder ringförmig darstellen. Dieser Effekt ist stärker ausgeprägt, je höher die Menge an gegebenem Kontrastmittel und je höher die Feldstärke des Tomografen ist. Zudem braucht das Kontrastmittel Zeit in das geschädigte Gewebe einzudringen (Hemond und Bakshi, 2018). Durch Reparatur der Schäden (Blut-Hirn-Schranke, Ödemresorption, Wiederaufbau des Myelins) können sich ehemalige hyperintense Läsionen im Verlauf wieder isointens präsentieren (Hemond und Bakshi, 2018). Dagegen sind chronische Herde mit dauerhaftem myelinem und axonalem Verlust durch die KM-Gabe unbeeinflusst und verbleiben hypointens (Meinl, 2012); (Sahraian und Radü, 2008).

2.3.3 3D-Bildgebung

Die 3D-Bildgebung hat einige Vorteile gegenüber den bisherigen 2D-Bildern (Kitajima et al., 2012). Zum einen werden bei der 3D-Akquirierung isotrope Datensätze des gesamten Schädels generiert. Lücken werden so zwischen den einzelnen Schichten vermieden. Ergänzende Sequenzen sind dadurch in der Regel überflüssig. Es ist mit dem 3D-Datensatz möglich, jede beliebige Schicht zu berechnen und darzustellen. Die erleichterte Rekonstruktion büßt jedoch nicht an Bildqualität ein. Ein besseres Kontrast-zu-Rausch-Verhältnis und geringere Flussartefakte sichern die Qualität. Zum anderen sind durch die 3D-Aufnahmen dünnere Schichtdicken möglich (Nelson et al., 2008); (Kakeda et al., 2012). Dies verspricht die Detektion von kleineren Läsionen. Dank der besseren räumlichen Auflösung mit Voxelvolumen von 0.5 [mm³] im Vergleich zu 1 [mm³] bei 2D-Aufnahmen, werden mehr Herde im gesamten Hirn und Rückenmark detektiert. Dies ermöglicht einen früheren Nachweis der räumlichen und zeitlichen Dissemination der McDonalds Kriterien. Daraus resultiert wiederum eine frühere Therapieeinleitung und bessere Vergleichbarkeit, sowie Verlaufsbeurteilung der Erkrankung. Die genauere Abgrenzung von aktiven Läsionen macht zudem eine frühere Therapieeskalation möglich. Kortikale Herde und Läsionen mit dem für MS spezifischen Zentralvenenzeichen können außerdem einfacher mit den 3D-Bildern detektiert werden. Dies erleichtert die Diagnosestellung und Abgrenzung zu Differential-diagnosen. Mit Hilfe von computerbasierter Verarbeitung der Bilddaten lassen sich zudem Hirn-volumen aus den 3D-Datensätzen bestimmen (Hu et al., 2019).

2.4 Hirnvolumetrie mithilfe der MRT

Die MR-Volumetrie ist ein Verfahren, um den Verlust der Gehirnsubstanz und Läsionsvolumina zu quantifizieren. Auf Grund von großen v.a. absoluten Schwankungen bei der Segmentierung ist die Hirnvolumetrie bisher noch kein Inhalt der McDonalds Kriterien (Sahraian und Radü, 2008); (Sastre-Garriga et al., 2020). Es fehlen außerdem Daten aus längerfristigen Studien, die eine stabile Grundlage zum Vergleich der Hirnvolumenänderungen etablieren lassen (Marciniewicz et al., 2019). Auf Grund der klinischen Relevanz besteht jedoch ein Bedarf an neueren Bildgebungstechniken bei der MS-Bewertung und Überwachung.

Die bisherigen Methoden zur Hirnvolumetrie anhand von MRT-Bildern lassen sich grob in drei Gruppen unterteilen: manuelle, halb- und vollautomatisierte Methoden (Giorgio et al., 2008); (Giorgio und de Stefano, 2013).

2.4.1 Bisherige Hirnvolumetrie

Die Segmentierung des Gehirns wurde lange Zeit händisch durchgeführt. Die Wurzeln der manuellen Methoden lagen in der Messung des Hippocampus und der Amygdala. Spätere Techniken bestimmten das Hirnvolumen durch Abstände und Durchmesser im Gehirn anhand von 2D-Bildern (Giorgio und de Stefano, 2013). So machte man sich bei der MS zum Beispiel den kleinsten Abstand zwischen den beiden Kerngebieten des Nucleus caudatus in Relation zur Hirnweite entlang der selben Linie zu Nutze. Dieser ist bei der Multipler Sklerose im Gegensatz zum gesunden Patienten vergrößert (Bermel et al., 2002). Später kamen auch 3D-Bilder zum Einsatz, in denen mit speziellen Computerprogrammen Umrisse und Strukturen nachgezeichnet wurden (Giorgio und de Stefano, 2013). In Tabelle 1 sind verschiedene manuelle Verfahren aufgelistet. Sie haben den großen Nachteil, dass sie sehr zeitintensiv sind, ausgezeichnete anatomische Kenntnisse voraussetzen und menschlichen Fehlern unterliegen. Die Vergleichbarkeit und Reproduzierbarkeit ist dadurch eingeschränkt (Giorgio und de Stefano, 2013); (Marciniewicz et al., 2019); (Valverde et al., 2015).

Nachdem die 3D-Bildgebung immer mehr zur Routine wurde und so die Datenmenge und Datenkomplexität zunahmen, kam es zur Entwicklung neuerer, (halb)automatischer Softwarelösungen, basierend auf künstlicher Intelligenz (KI). Sie ermöglichten die Auswertung der großen Datenmengen und arbeiteten gleichzeitig schnell und präzise (Akkus, 2017). Anhand computerbasierter Algorithmen können die Volumen der Hirnstrukturen bestimmt werden. Dabei benötigen halbautomatisierte Methoden eine zusätzliche Validierung

der Gehirnvolumenmessungen durch den Benutzer der Softwarelösung. Diese Technik ist durch die händische Korrektur ebenfalls zeitintensiv und nicht ausreichend reproduzierbar. Dahingegen funktionieren vollautomatisierte Algorithmen ohne zusätzliches Eingreifen. Die Messung ist jedoch meist weniger akkurat (Durand-Dubief et al., 2012); (Giorgio und de Stefano, 2013). Einen entscheidenden Vorteil brachten diese Methoden im Gegensatz zu den manuellen und halbautomatisierten Verfahren in der Zeitersparnis, Objektivität und Reproduzierbarkeit (Giorgio et al., 2008); (Marciniewicz et al., 2019).

Manuelle Methoden	Eigenschaften
Bicaudates Verhältnis (BCR = bicaudate ratio)	Kleinster Abstand zwischen den beiden Kerngebieten des Nucleus Caudatus in Relation zur Hirnweite entlang derselben Linie
Gehirnweite	Abstand zwischen zwei Punkten auf der kortikalen Oberfläche, gemessen auf der gleichen Höhe der lateralen Ventrikelbreite auf axialen Schnitten
Bereich des Corpus callosum	Gemessen am mittelsagittalen T1-gewichteten Bild
Abgrenzung von: Hippocampus, Amygdala, entorhinalen Kortex, Parahippocampus, cholinergen Kerne des basalen Vorderhirns	Abgrenzung anhand anatomischer Protokolle
M/P (Mittelhirn-zu-Pons Verhältnis)	Gemessen auf einem mittelsagittalen T1-gewichteten Bild
Breite des mittleren Kleinhirnstiels	Abstand zwischen den oberen und unteren Rändern des mittleren Kleinhirnstiels, begrenzt durch die peripedunculären Liquorräume der pontocerebellären Zisternen
Ventrikelbreite des dritten und der lateralen Ventrikel	Bestimmt entlang einer Ebene, die dem anteroposterioren Mittelpunkt des Ventrikels auf axialen Schnitten entspricht

Tabelle 1: Liste der meistgenutzten manuellen Methoden (Giorgio und de Stefano, 2013); (Marciniewicz et al., 2019).

2.4.2 KI-basierte Softwarelösungen

Das Feld der künstlichen Intelligenz erfährt immer mehr Fortschritte. Jede Technik, die Computer befähigt menschliche Intelligenz nachzuahmen, wird zur künstlichen Intelligenz

gezählt. Diese umfasst beispielsweise Logik, wenn-dann-Szenarien, Entscheidungsbäume und Maschinenlernen (Sarker, 2021). Die in 2.4.2 beschriebenen Algorithmen werden dem Unterfeld des Maschinenlernens zugeordnet (Akkus, 2017); (Sarker, 2021).

Sie quantifizieren das Hirnvolumen meist anhand von 3D T1-gewichteten Bildern (Sinnecker et al., 2018). Beim Akquirieren von 3D-Aufnahmen entstehen große Datensätze, die mit den Softwarelösungen verarbeitet werden und gleichzeitig schnellere Ergebnisse liefern sollen. Künstliche Intelligenz und maschinelles Lernen hat den Computersystemen ermöglicht, genaue Vorhersagen aus großen Datenmengen in kürzester Zeit zu liefern. Dabei müssen Merkmale oft manuell von Experten ausgewählt und extrahiert werden (Sarker, 2021).

In den letzten 10 Jahren fanden verschiedene Softwarelösungen zur Quantifizierung der Atrophie und Läsionslast bei MS Anwendung. Abbildung 1 zeigt eine Reihe von Beispielen der existierenden vollautomatisierten Softwarelösungen. Die unterschiedlichen Methoden sind grob in zwei Untergruppen eingeteilt: die segmentierungs- und die registrierungsbasierten Tools (Sastre-Garriga et al., 2020); (Sinnecker et al., 2018). Eine weitere Einteilung kann anhand von Längsschnitt- und Querschnittmessungen erfolgen (Filippi et al., 2016); (Giorgio et al., 2008).

Die segmentierungsbasierten Techniken kommen vor allem bei Querschnittmessungen zur Anwendung. Sie benötigen nur eine MRT-Aufnahme, um die aktuellen Volumina direkt zu berechnen. Diese Methoden basieren auf der Klassifizierung und Segmentierung der verschiedenen Hirngewebe. Das Ergebnis wird anschließend zum Vergleich mit anderen Individuen oder Zeitpunkten genutzt (Giorgio et al., 2008); (Sinnecker et al., 2018). Die Atrophierate kann indirekt aus den Differenzen der gemessenen Volumina zweier Messungen berechnet werden, ist jedoch ungenauer. BPF, FreeSurfer, SIENAx, SPM, VBM und GIF sind Beispiele der automatischen segmentierungsbasierten Algorithmen (de Stefano et al., 2014); (Durand-Dubief et al., 2012); (Giorgio et al., 2008); (Sinnecker et al., 2018).

Registrierungsbasierte Techniken, wie beispielsweise BBSI, SIENA, SIENAX-MTP, SIENAX- XL, CLADA, und FreeSurfer sind praktikabel für Längsschnittstudien. Sie benutzen Registrierungs-algorithmen, um Volumenänderungen abzuschätzen und ermöglichen einen direkten Vergleich zweier MRT-Bilder des gleichen Individuums zu unterschiedlichen Zeitpunkten. Sie messen so direkt die Atrophierate. Sind serielle Scans einer Person vorhanden, können mit den registrierungsbasierten Techniken genauere Volumenänderungen abgeleitet werden (Giorgio et al., 2008); (Sastre-Garriga et al., 2020).

Laut einer Studie lieferten registrierungsbasierte Techniken akkuratere und sensitivere Er-

gebnisse bei der Messung von Hirnatrophie als segmentierungsbasierte Methoden. Sie sind außerdem weniger beeinflusst von der Bildqualität und Änderungen des Bildgebungssystems als segmentierungsbasierte Techniken (Durand-Dubief et al., 2012); (Giorgio et al., 2008); (Sinnecker et al., 2018).

Tool	Freely available?	Measures
SIENAX	Yes	Global and regional brain volumes for cross-sectional comparisons
SPM/VBM	Yes ^b	Global and regional brain volumes, pixel-to-pixel statistical comparisons between two groups or time points
GIF	Yes	Regional brain volumes for cross-sectional comparisons
Atropos	Yes	Regional brain volumes for cross-sectional and longitudinal comparisons
FreeSurfer	Yes	Cortical thickness, global and regional grey matter and white matter volumes
CIVET	No	Cortical thickness
SIENA	Yes	Percentage brain volume change between two time points
SIENA-XL	No	Grey matter and white matter volumes for longitudinal comparisons
SIENAX-MTP	No	Grey matter and white matter volumes for longitudinal comparisons
BBSI	Yes	Percentage brain volume change between two time points
CLADA	No	Cortical thickness
NeuroQuant (FDA clearance and CE mark received)	No	Global and regional grey matter volumes
Icometrix (FDA clearance and CE mark received)	No	Global and regional grey matter volumes
Biometrix (CE mark received)	No	Global and regional grey matter volumes
Quantib (FDA clearance and CE mark received)	No	Global and regional grey matter volumes

Abbildung 1: Liste einiger vollautomatisierter Softwarelösungen (Sastre-Garriga et al., 2020)

Patientenbezogene und technische Schwankungen bleiben ein ungelöstes Problem, welches die Benutzung und Vergleichbarkeit limitiert (Guo et al., 2019); (Storelli et al., 2018); (Giorgio et al., 2008); (Marciniewicz et al., 2019). Die hier aufgelisteten, auf KI-basierenden Softwarelösungen reichten maximal bis zum Feld des Maschinenlernens. Hierbei sind zur Klassifizierung mehrere aufeinanderfolgende Schritte notwendig, insbesondere Vorverarbeitung, Merkmalsextraktion, kluge Merkmalsauswahl und Lernen (Alzubaidi et al., 2021).

2.4.3 Deep Learning

Verbesserungen in der Hard- und Software eröffneten die neue Möglichkeit des Deep Learnings. Es ist ein Unterfeld des maschinellen Lernens und basiert auf der Idee, dass künstliche neuronale Netzwerke (KNN) in der Lage sind, automatisch Merkmale aus den Daten zu extrahieren und zu lernen, ohne explizit programmiert zu werden (Sarker, 2021); (Li et al., 2024); (Alzubaidi et al., 2021); (Lee et al., 2017). Dabei wird das Netzwerk mit Hilfe von große Mengen an Daten in Form von Ein- und Ausgabedaten trainiert. Das System lernt so komplexe Muster zu erkennen und passt mit jeder zusätzlichen Dateneingabe seine Parameter an. Das System kann anschließend nicht gesehene Daten generalisieren anhand des erlernten Netzwerkes (Sarker, 2021); (Li et al., 2024); (Alzubaidi et al., 2021); (Lee et al., 2017). Dies wird nun auch bei der Segmentierung und Klassifizierung von Bildern angewendet (Akkus, 2017); (Mckinley et al., 2021). Das Deep Learning könnte dabei helfen, Störfaktoren der bisherigen Algorithmen zu umgehen und gleichzeitig nützliche Funktionen in der Quantifizierung von Hirnvolumina und der Läsionslast durch Selbsttraining zu eröffnen. Eine teilweise auf Deep Learning basierte Softwarelösung ist beispielsweise *mdbrain* von mediaire, die wir in unserer Studie mit der häufig genutzten vollautomatischen Softwarelösung *icobrain ms* von icometrix vergleichen.

3 Experimentelle Fragestellung

Ziel dieser Studie war es mit Hilfe von zwei vollautomatisierte, KI-gestützten Softwarelösungen das Hirn- und Läsionsvolumen bei Multiple-Sklerose-Patienten zu quantifizieren und auf die Übereinstimmung der absoluten und relativen Werte, sowie der Läsionsvolumen zu untersuchen. Hierfür sollten MRT-Untersuchungen des Schädels von Patienten mit bekannter Diagnose einer Multiplen Sklerose angefertigt und mithilfe von den zwei kommerziell erhältlichen Softwarelösungen *icobrain ms* und *mdbrain* ausgewertet werden.

Es wurde die Übereinstimmung beider Auswertungen bezüglich der Bestimmung des absoluten und relativen Volumens

- des Gesamthirns,
- der grauen und weißen Substanz,
- sowie der Läsionslast

bei Multipler Sklerose analysiert.

Die Nullhypothese lautete: Die mit den beiden Softwarelösungen gemessenen absoluten und relativen Volumina stimmen überein.

Die Alternativhypothese lautete: Die mit den beiden Softwarelösungen gemessenen absoluten und relativen Volumina stimmen nicht überein.

4 Material und Methoden

4.1 Patientenkollektiv

Für die retrospektive Studie wurden insgesamt 86 Patienten mit bestehender Diagnose einer Multiplen Sklerose herangezogen und anonymisiert.

Die Einschlusskriterien waren:

- Alter über 18 Jahre
- gesicherte MS Diagnose

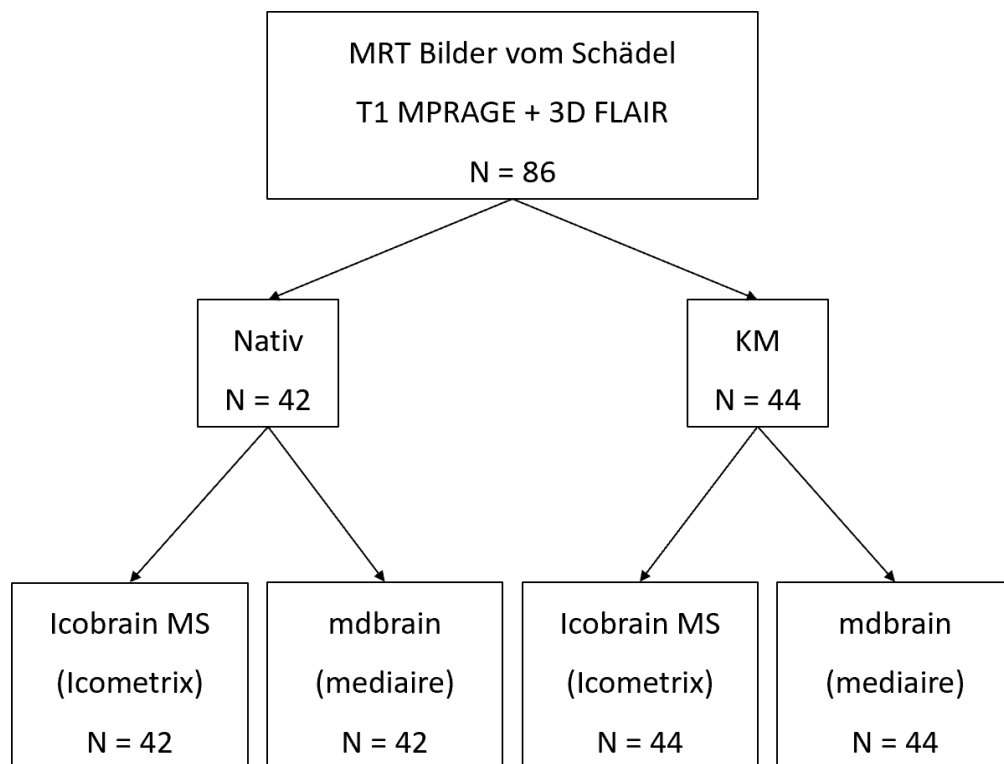
Die Ausschlusskriterien waren:

- Alter unter 18 Jahre
- Voroperationen am Schädel
- schlechte Bildqualität

Das Jahr der Erstdiagnose, die Form der MS sowie abgeschlossene oder laufende Therapien wurden außer Acht gelassen. Die Patienten waren alle an die Marianne-Strauß-Klinik, eine neurologische Fachklinik in Berg, Deutschland. Durch das Anonymisieren der Daten war keine schriftliche Einverständniserklärung der Patienten notwendig. Die Studie wurde gemäß den Vorgaben der Ethik-Kommission der Bayrischen Landesärztekammer durchgeführt.

4.2 MRT und Bildakquirierung

Die Bilderstellung des Kopfes der 86 Patienten erfolgte mit Hilfe des MRT-Scanner Magnetom Skyra von Siemens mit einer Feldstärke von 3,0 T. Es wurden eine isotrope T1- MPRAGE-Sequenz und eine isotrope 3D-FLAIR-Sequenz bei jedem der 86 Patienten akquiriert. Bei 42 Patienten erfolgte die T1-Sequenz ohne, bei 44 Patienten mit Kontrastmittel. In der Abbildung 2 ist die bildliche Darstellung des Aufbaus der Studie dargestellt. N entspricht der Anzahl der Teilnehmer pro Studienarm.

**Abbildung 2:** Studiendesign

In der Tabelle 2 sind die Parameter zur Einstellung der MRT-Sequenzen aufgelistet. TR entspricht der Repetitionszeit, TE der Echozeit und TI der Inversionszeit. Während die Schichtdicke bei beiden Sequenzen identisch ist, unterscheiden sie sich in der Pixelauflösung.

Parameter	T1-MPRAGE	FLAIR
TR [ms]	2.300,00	7.600,00
TE [ms]	2,28	401,00
TI [ms]	950,00	2.400,00
Pixelauflösung [mm]	1,0 x 1,0	1,0 x 1,0
Schichtdicke [mm]	1,0	1,0

Tabelle 2: Eigenschaften der verwendeten Sequenzen

Nach der Bildakquirierung erfolgte zuerst die Anonymisierung und dann der Export der MRT-Bilder von PACS (Picture Archiving and Communication System, ein Bildablage- und Kommunikationssystem) auf einen externen Datenträger. Diese Dateien wurden zum Laden in die Programme von mediaire und icometrix genutzt. Der Bericht wurde nach 10-15 Minuten erstellt und als PDF zum Herunterladen zur Verfügung gestellt.

4.3 Parameter und Software

Das absolute (in ml) und relative (Altersperzentile) Volumen für das Gesamthirn und der grauen Substanz wurden mit Hilfe von zwei unterschiedlichen KI-basierten Softwarelösungen *mdbrain* v2.0 (MDB) und *icobrain ms* (ICO) berechnet. Das Volumen für periventrikuläre, juxtakortikale, infratentorielle und unspezifische FLAIR-hyperintense Läsionen wurde absolut (in ml) berechnet.

4.3.1 icobrain ms

Icobrain ms ist ein kommerziell erhältliches Produkt vom belgischen Unternehmen icometrix, ehemals MSmetrix. Es ist eine vollautomatisierte Softwarelösung zur Bestimmung des Hirnvolumens und der Läsionslast bei MS mithilfe von 3D T1-gewichteten und 3D FLAIR Sequenzen. Dabei können die Aufnahmen mit einem 1,5-T oder einem 3-T Scanner generiert sein. Die Methode erfordert keine Trainingsbilddatenbank von Experten für Läsionsdetektion. Unter Verwendung eines gesunden Hirnatlas teilt icometrix das Gehirn in graue Substanz, weiße Substanz und Liquor ein (Jain et al., 2015). *icobrain ms* quantifiziert sowohl Querschnitts als auch Längsschnittsvolumina. Querschnittsvolumina werden mit einer auf Nifty Seg basierenden Software berechnet. Longitudinale Volumenänderungen sind mit Hilfe der Jacobian Integrationsmethode bestimmt (Sastre-Garriga et al., 2020).

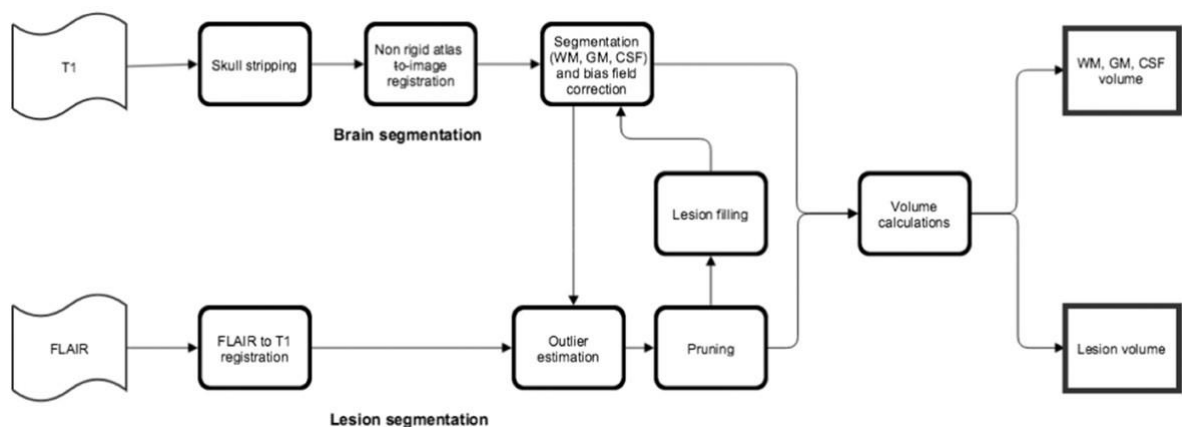


Abbildung 3: Pipeline *icobrain ms*

Außerdem ist die Methode CE-gekennzeichnet und hat die FDA-Zulassung. Dies genehmigt die Anwendung für die klinische Praxis in Europa und den USA (Sastre-Garriga et al., 2020). Der Querschnittsbericht von icometrix unterteilt sich in einen Volumen- und einen Läsionsreport. Auf Ersterem kommen die Volumenmessungen des Ganzhirns und der grauen Substanz

zur Abbildung. Die Volumina sind dabei auf die Schädelgröße genormt. Als Referenz sind Volumina und normative Perzentilen der gesunden Normalbevölkerung angegeben, basierend auf Daten, die vorwiegend von Probanden aus Europa und Nordafrika generiert wurde (Finkelsztejn et al., 2018). Eine Grafik zeigt den alters- und geschlechtsangepassten Durchschnitts- und Normalbereich jeweils für das Ganzhirnvolumen und die graue Substanz (siehe beispielhaft Abbildung 4). Die Patientenmessung ist darauf mit einem schwarzen X gekennzeichnet. Ein Konfidenzintervall gibt die Grenzen an, bei dem mit 90 prozentiger Sicherheit die Unterschiede der Ergebnisse der Test-Retest-Experimente liegen (Icometrix, 2021).

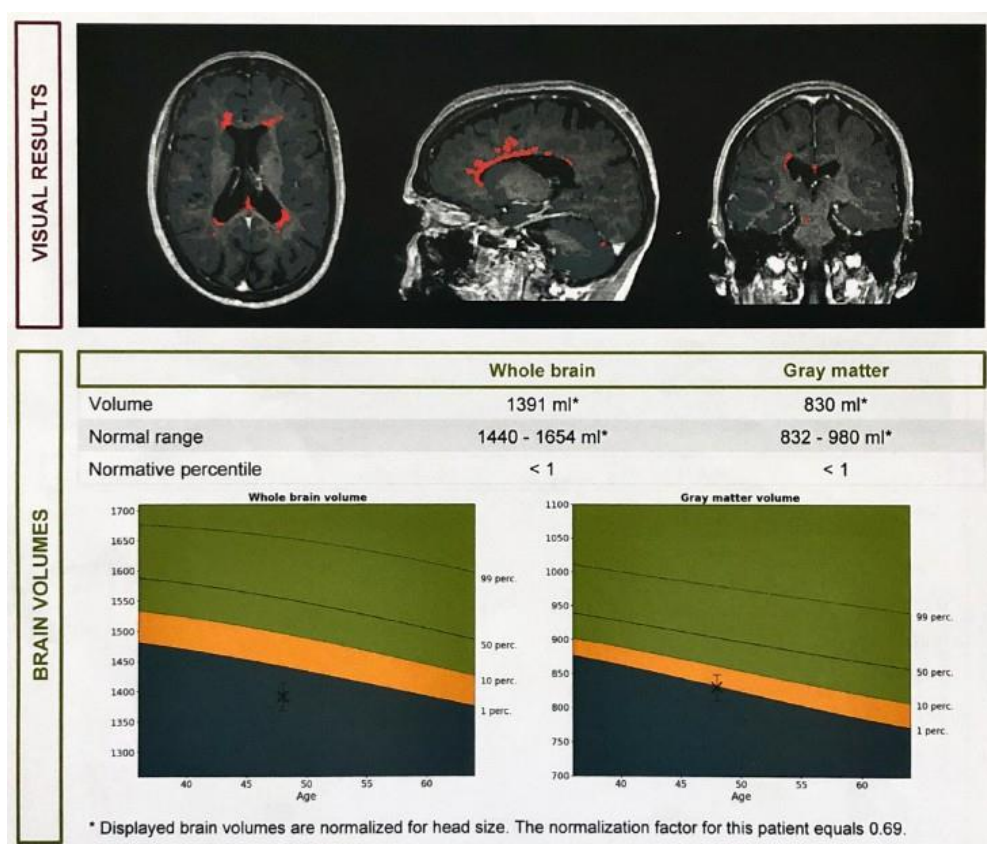


Abbildung 4: Beispiel eines *icobrain ms* Reports, Darstellung der Hirnvolumina (Icometrix, 2021).

Auf dem Ergebnisbericht stellen sich zum Anderen die Hyperintensitäten der weißen Substanz der FLAIR Sequenz dar. Diese werden absolut als gesamte Läsionslast angegeben und zusätzlich den üblichen Regionen der MS im Gehirn zugeordnet (periventrikulär, juxtakortikal, infratentoriell, tiefe weiße Substanz). Sie sind in unterschiedlichen Farben als Balkendiagramm und auf den MRT-Aufnahmen selbst visualisiert. Auch die T1-Hypointensitäten werden in absoluten Werten angegeben. Aktive Läsionen, die in der T1 mit KM hell erscheinen, sind gesondert aufgelistet, sofern eine Sequenz nach Kontrastmittelgabe existiert (Icometrix, 2021).

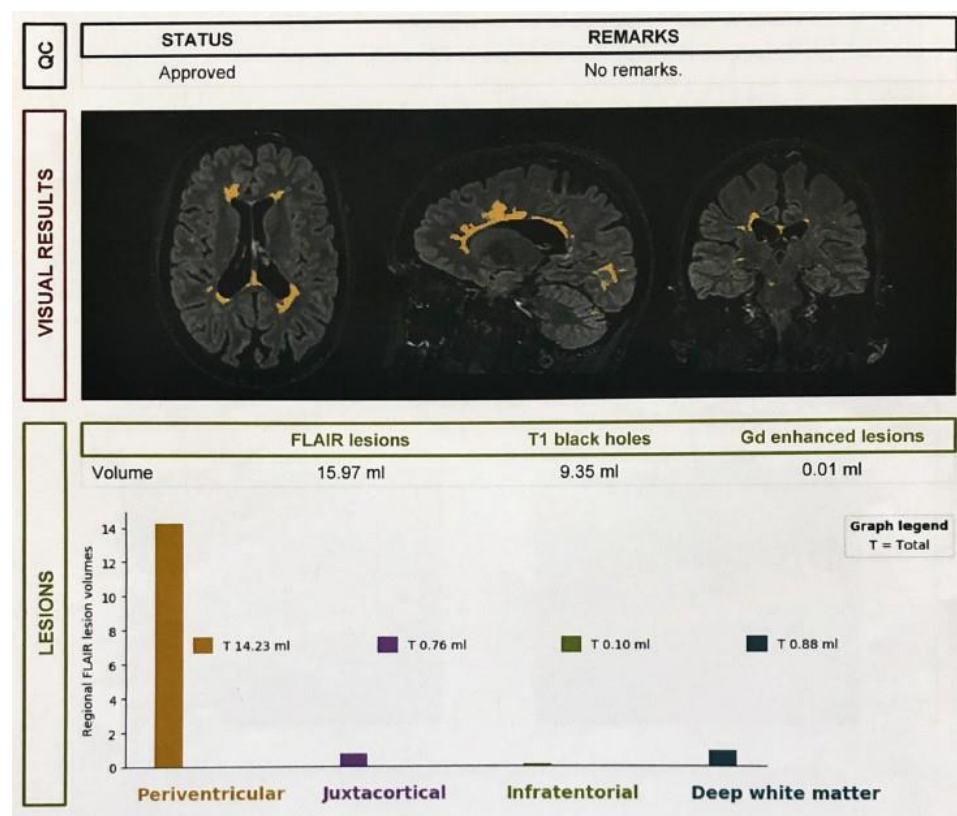


Abbildung 5: Beispiel eines *icobrain ms* Läsionsreports (Icometrix, 2021).

4.3.2 mdbrain

Die 2018 in Berlin gegründete Firma mediaire hat die Softwarelösung *mdbrain* entwickelt. Sie haben es möglich gemacht, das Deep Learning in ein Medizinprodukt mit CE- Zertifizierung zu überführen. Sowohl die Bestimmung der Volumina einzelner Hirngewebe, als auch Läsionscharakteristika lassen sich innerhalb von wenigen Minuten mit der Software ermitteln. Mittlerweile wurden mehrere Versionen herausgebracht. Die verwendete Version 2.0 nutzt bereits Deep Learning U-Net zur Läsionsdetektion. Die Algorithmen werden mit Hilfe von KNNs trainiert und angewendet. Die Läsionsdetektion wird auf 150 vollständig segmentierten Daten trainiert. Für die Hirnvolumetrie benutzt die Firma in der verwendeten Version 2.0 die klassische Volumetrie (skull stripping, atlas registration und compartment segmentation, no Lesion Filtering). In den neueren Versionen stellen die Entwickler auch die Hirnvolumetrie auf Deep Learning mit U-Net um. In den Versionen 3.0 und 4.0 wird der Trainingsdatensatz jeweils erhöht und weitere Zusätze mit aufgenommen (mediaire, 2019).

Der Befundbericht beinhaltet Gesamthirn-, Regional- und Ventrikelvolumina. Visuell werden pathologische Volumenminderungen farblich dargestellt. Dabei sind Werte unterhalb von zwei Standardabweichungen gelb und unterhalb von vier Standardabweichungen rot hervorgehoben.

Das Gesamthirnvolumen, die weiße Substanz und graue Substanz, mit zusätzlicher Unterteilung in Großhirn- und Kleinhirnrinde, kommen als Perzentilenwert und absolutes Volumen zur Darstellung. Als Referenz sind altersentsprechende Normwerte, basierend auf Alter, Geschlecht und Schädelgröße, angegeben. Die regional- und ventrikelvolumetrischen Angaben folgen dem gleichen Prinzip. Regional erfolgt die Unterteilung in Hippocampus, Temporal-, Frontal-, Parietal- und Okzipitallappen, während sich die Ventrikelvolumina in die vier Ventrikel aufgliedern (mediaire, 2019). Ein klassischer Bericht ist Abbildung 6 zu entnehmen.

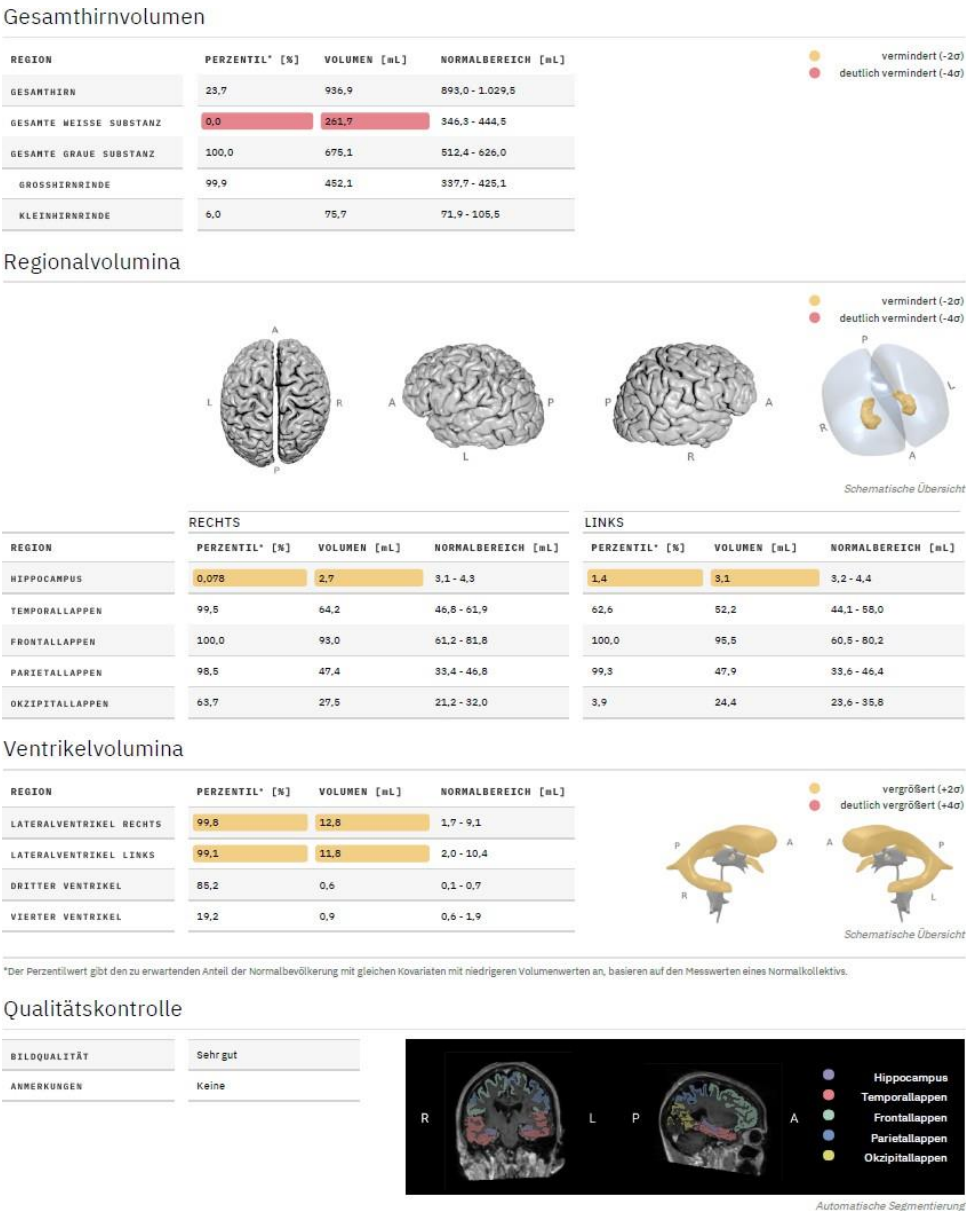


Abbildung 6: Beispiel eines *mdbrain* Volumetriereports. Darstellung von Gesamthirn-, Regional- und Ventrikelvolumina (mediaire, 2019).

Der Läsionsreport (siehe Abbildung 7) liefert eine exemplarische Darstellung der drei Schichten mit den größten Läsionen mit jeweiliger Farbkodierung der MS typischen Region (periventrikulär, juxtakortikal, infratentoriell und übriges Marklager). Des weiteren sind Läsionszahl und Läsionsvolumina der einzelnen Regionen zusätzlich als absolute Werte aufgelistet. Stehen Voraufnahmen zur Verfügung, werden Anzahl und Volumen von aktiven und vergrößerten Läsionen in regionaler Zuordnung separat ausgewiesen und die Auswahl der dargestellten Schichten anhand der Läsionsdynamik ausgewählt (mediaire, 2019).

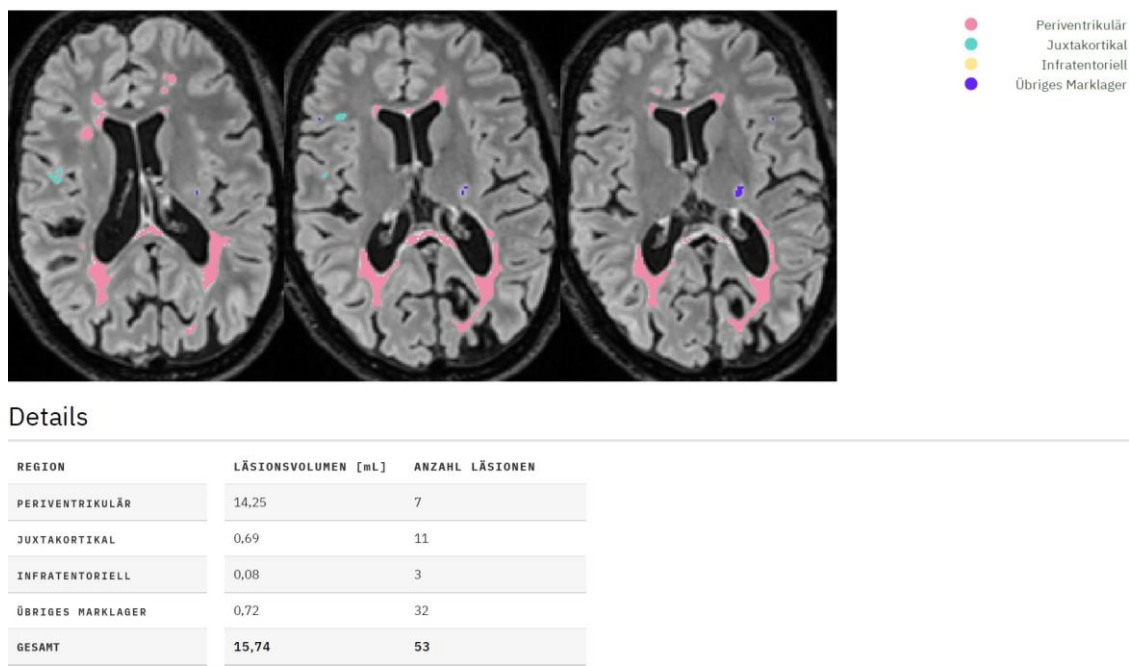


Abbildung 7: Beispiel eines mbrain Läsionsreports (mediaire, 2019).

4.4 Statistik

Von den gemessenen Parametern fand die Berechnung der Mittelwerte und Standardabweichungen statt. Nach vorheriger Prüfung auf Normalverteilung erfolgte die Evaluierung der absoluten Volumina des Gesamthirns und der grauen Substanz mit mehreren verbundenen t-Tests. Boxplots fanden für die graphische Analyse Verwendung. Mithilfe den von Bland und Altman beschriebenen Plots bildeten wir graphisch den Test auf Übereinstimmung der zwei Messwertreihen der unterschiedlichen Softwarelösungen ab. In Rot sind die mittleren Differenzen, in Grün die Übereinstimmungsgrenzen und in Schwarz die Nulllinie der Differenzen abgebildet. Zur Auswertung der relativen, nicht parametrischen Volumina wurde der Wilcoxon-Test angewendet. Wir ermittelten außerdem den Korrelationskoeffizient (r) nach Pearson und Spearman mit Einteilung der Perzentilen nach den Parametern <1%, zwischen 1%

und 5% und >5%. Für die Interpretation des Korrelationskoeffizienten gibt es viele Ansätze. Deskriptoren wie „stark“, „mäßig“ oder „schwach“ werden verwendet, um die Beziehung zu übersetzen. Als Richtschnur können Sie die Tabelle 3 verwenden, um die Stärke der Beziehung anhand des Wertes des Korrelationskoeffizienten zu interpretieren (Schober et al., 2018).

Korrelationskoeffizient	Stärke	Typ
0,7 bis 1,0	Sehr stark	Positiv
0,5 bis 0,7	Stark	Positiv
0,3 bis 0,5	Mäßig	Positiv
0 bis 0,3	Schwach	Positiv
0	Keine	Null
0 bis -0,3	Schwach	Negativ
-0,3 bis -0,5	Mäßig	Negativ
-0,5 bis -0,7	Stark	Negativ
-0,7 bis -1,0	Sehr stark	Negativ

Tabelle 3: Interpretation des Korrelationskoeffizienten

Bei der Betrachtung des Läsionsvolumens wurden die typischen Lokalisationen der Läsionen bei MS berücksichtigt und in periventrikulär, juxtakortikal, infratentoriell und übriges Marklager unterteilt. Die nicht parametrisch verteilten Läsionsvolumina wurden mit dem Wilcoxon-Test auf Signifikanz geprüft. Mit Liniendiagrammen veranschaulichten wir die Differenzen der Läsionsvolumina zwischen den beiden Softwarelösungen.

Es wurde eine statistische Signifikanz mit p-Werten kleiner 0,05 vorausgesetzt. Alle statistischen Tests und Grafiken wurden mit SPSS (IBM SPSS, Inc., Chicago, IL) und Excel Version 2011 (Microsoft, Redmond, WA) durchgeführt und erstellt.

5 Ergebnisse

5.1 Patientenkollektiv

Aus Tabelle 4 ist die deskriptive Statistik des Patientenkollektivs zu entnehmen. Zwei Reports waren nicht verwertbar und wurden von den Daten ausgeschlossen.

Charakteristikum	Anzahl (Prozent)
Kohorte	
Komplette Studienkohorte	86
Subgruppe: native Untersuchung	42 (48,8)
Subgruppe: Kontrastmittel-Gabe	44 (51,2)
Geschlecht	
Frauen	64 (74,4)
Männer	22 (25,6)
Patientenalter (Jahre)	
Mittelwert $\pm$ SD	54 $\pm$ 12
Reichweite	26 - 80

Tabelle 4: Patientenkollektiv

5.2 Volumen und Altersperzentilen des Gesamthirns

5.2.1 Native Untersuchungen des Gesamthirns

Signifikante Unterschiede zwischen beiden Auswertungen mit p-Werten kleiner 0,05 zeigten sich bei der nativen absoluten Volumenmessung des Gesamthirns ($p < 0,001$). Die Korrelation nach Pearson betrug $r = 0,99$. Die deskriptive Statistik ist der Tabelle 5 zu entnehmen.

Statistik	ICO	MDB
Mittelwert [ml]	1057,37	1092,10
Standardabweichung [ml]	140,91	148,88
95% Konfidenzintervall [ml]	1013,46 - 1101,28	1045,70 - 1138,50
Median [ml]	1036,96	1064,70
Minimum [ml]	814,20	837,40
Maximum [ml]	1454,45	1525,30

Tabelle 5: Deskriptive Statistik des Ganzhirnvolumens ohne KM

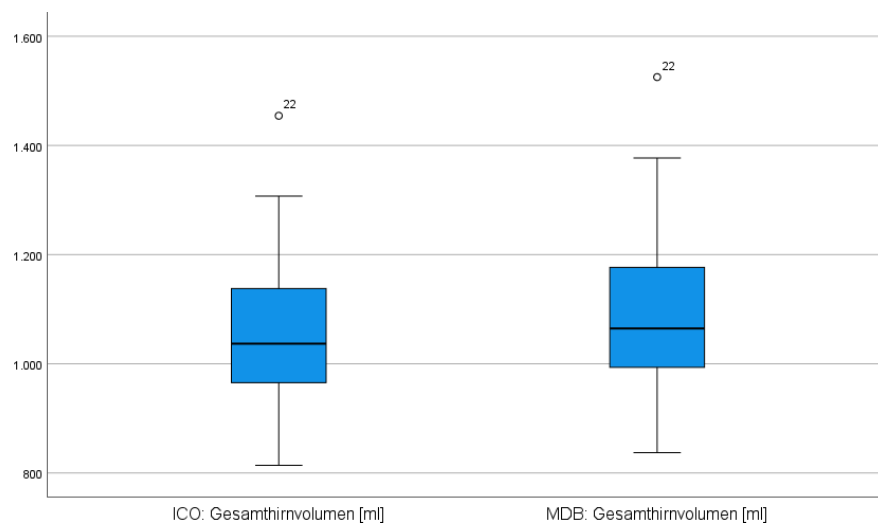


Abbildung 8: Boxplot des Gesamthirnvolumens der nativen Untersuchungen

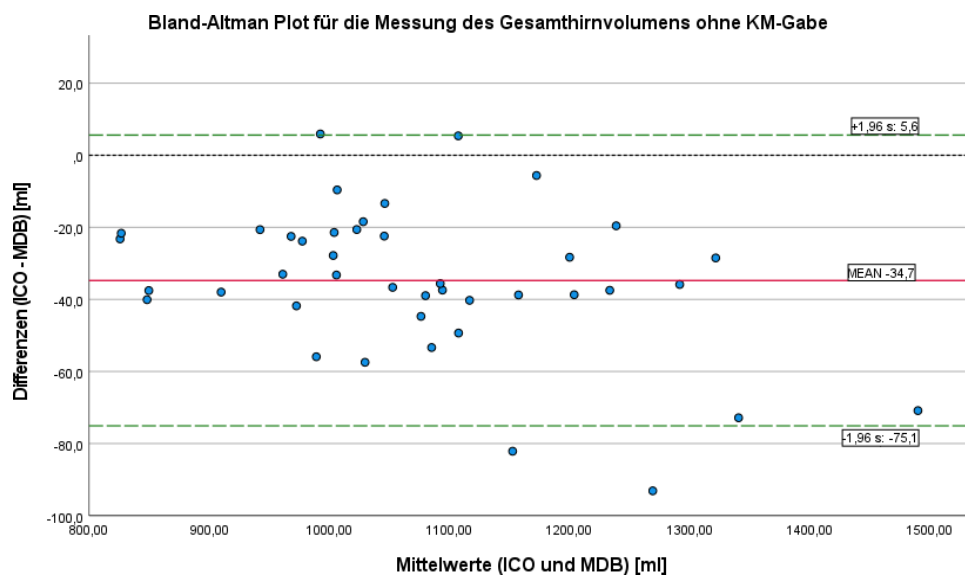
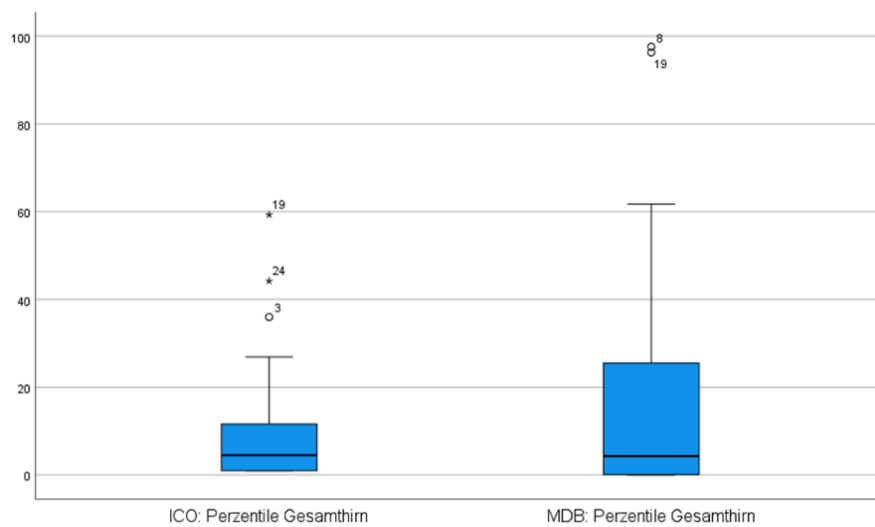
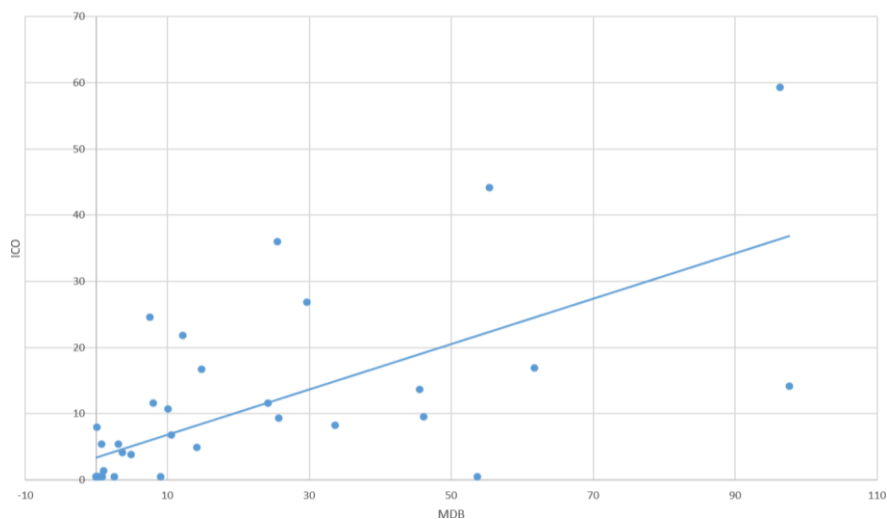


Abbildung 9: Bland Altman Plot des Gesamthirnvolumen ohne KM
Limits of Agreement (grün): 5,6 ml und -75,1 ml. Mittlere Differenz: -34,7 ml.

Die Altersperzentilen unterschieden sich im Wilcoxon-Test nicht signifikant (p-Wert 0,3). Bei der Korrelationsanalyse ergab sich ein Zusammenhang der Werte von 0,67 nach Pearson. Die deskriptive Statistik ist der Tabelle 6 zu entnehmen.

Statistik	ICO	MDB
Mittelwert	9,15	16,67
Standardabweichung	12,94	25,20
Minimum	0,50	0,00
Maximum	59,30	97,60

Tabelle 6: Deskriptive Statistik der Altersperzentilen der nativen Ganzhirnmessungen**Abbildung 10:** Boxplot der Altersperzentilen der nativen Ganzhirnmessungen**Abbildung 11:** Altersperzentilen vom Ganzhirn der nativen Untersuchungen

5.2.2 Untersuchungen mit KM des Gesamthirns

Die deskriptive Statistik der Volumenmessungen des Gesamthirns mit Kontrastmittel wird in Tabelle 7 zusammengefasst. Die absoluten Volumenmessungen waren nicht signifikant unterschiedlich mit einem p-Wert von 0,17. Die Korrelation nach Pearson beträgt $r = 0,95$.

Statistik	ICO	MDB
Mittelwert [ml]	1061,63	1057,39
Standardabweichung [ml]	121,47	122,60
95% Konfidenzintervall [ml]	1028,91 - 1102,84	1020,12 - 1094,66
Median [ml]	1059,93	1047,25
Minimum [ml]	883,32	774,30
Maximum [ml]	1463,73	1431,30

Tabelle 7: Deskriptive Statistik des Ganzhirnvolumens mit KM

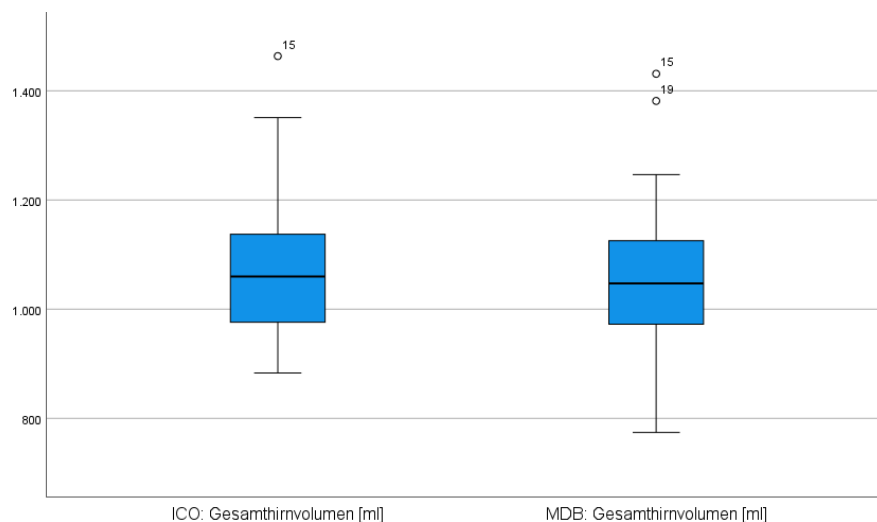


Abbildung 12: Boxplot des Gesamthirnvolumens bei kontrastmittelverstärkten Aufnahmen.

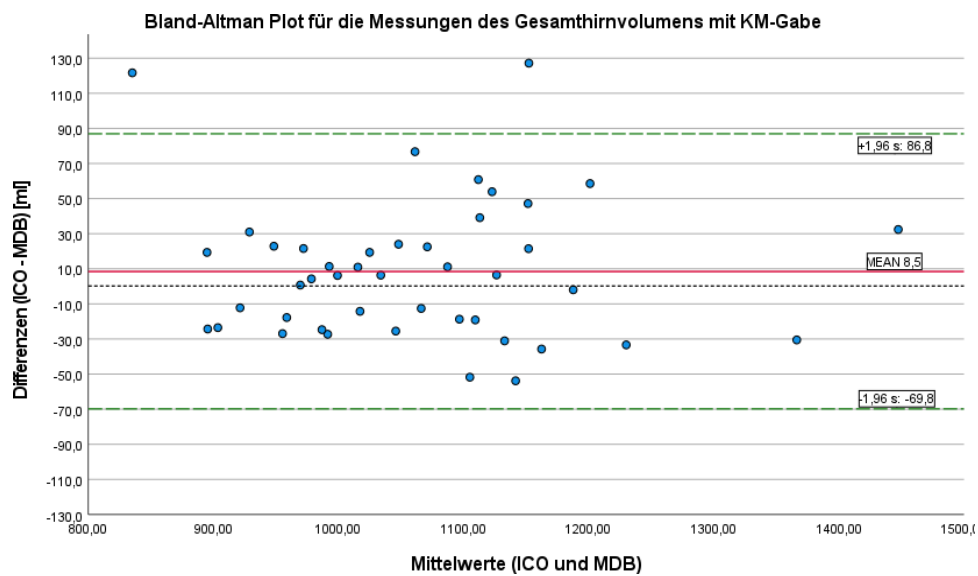


Abbildung 13: Bland-Altman Plot Gesamthirnvolumen mit KM
Limits of Agreement (grün): 86,8 ml und -69,8 ml. Mittlere Differenz (rot): 8,5 ml.

Bei Betrachtung der Altersperzentilen ergab sich ein signifikanter p-Wert von unter 0,05 ($p < 0,001$). Die Korrelationsanalyse zeigte Werte von $r = 0,52$ (Pearson). Die deskriptive Statistik ist der Tabelle 8 zu entnehmen.

Statistik	ICO	MDB
Mittelwert	16,64	53,70
Standardabweichung	21,47	33,31
Minimum	0,50	0,002
Maximum	89,20	99,70

Tabelle 8: Deskriptive Statistik der Altersperzentilen der Ganzhirnmessungen mit Kontrastmittel

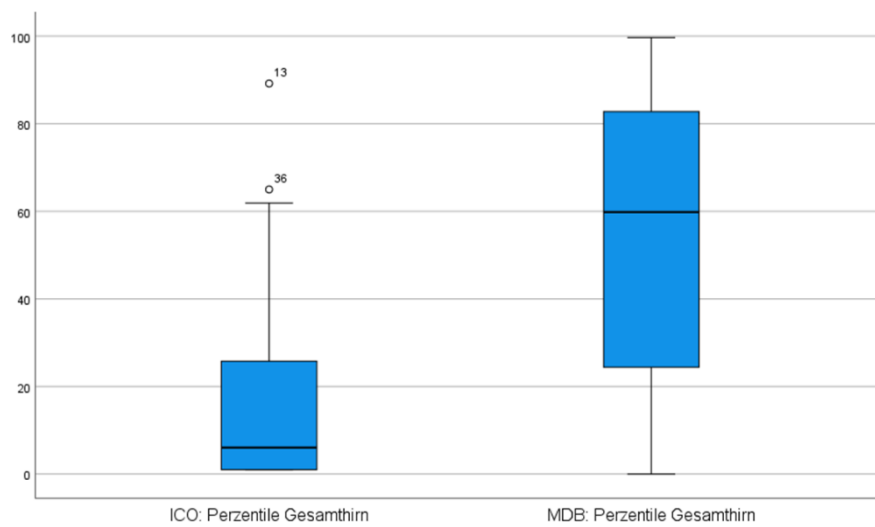


Abbildung 14: Boxplot der Altersperzentilen des Gesamthirnvolumens mit KM-Gabe

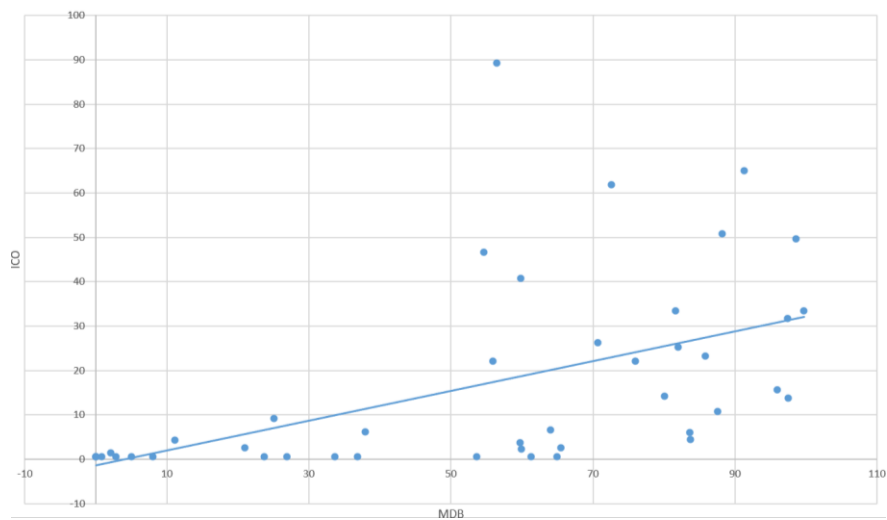


Abbildung 15: Altersperzentilen vom Ganzhirn der Untersuchungen mit KM

5.3 Volumen und Altersperzentilen der grauen Substanz

5.3.1 Native Untersuchungen der grauen Substanz

Die deskriptive Statistik der nativen Untersuchungen der grauen Substanz sind aus der Tabelle 9 zu entnehmen. Der p-Wert nimmt bei der Messung des absoluten Volumens der grauen Substanz signifikante Werte von unter 0,05 an. Die Korrelation nach Pearson beträgt $r = 0,92$.

Statistik	ICO	MDB
Mittelwert [ml]	635,33	682,35
Standardabweichung [ml]	85,24	78,35
95% Konfidenzintervall [ml]	612,30 - 657,69	696,02 - 740,90
Median [ml]	632,43	710,35
Minimum [ml]	503,20	569,50
Maximum [ml]	871,06	940,60

Tabelle 9: Deskriptive Statistik, graue Substanz ohne KM

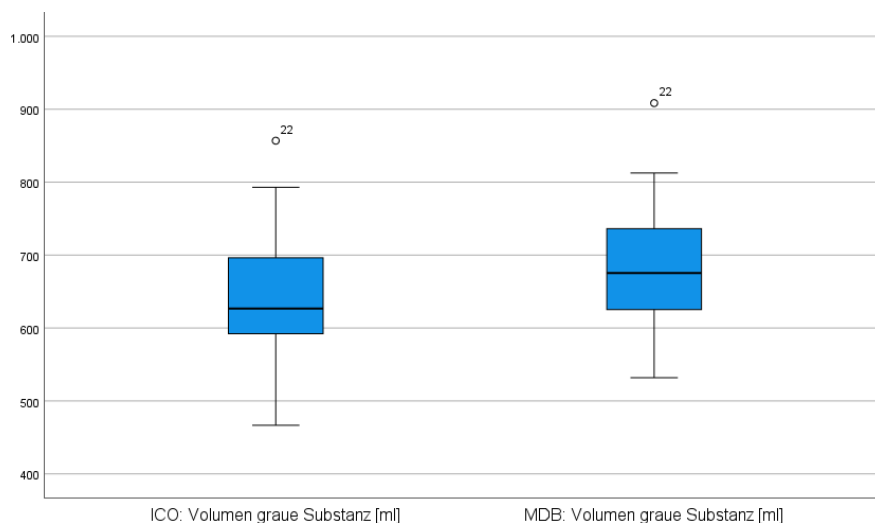


Abbildung 16: Boxplot des Volumens der Grauen Substanz bei nativer Untersuchung

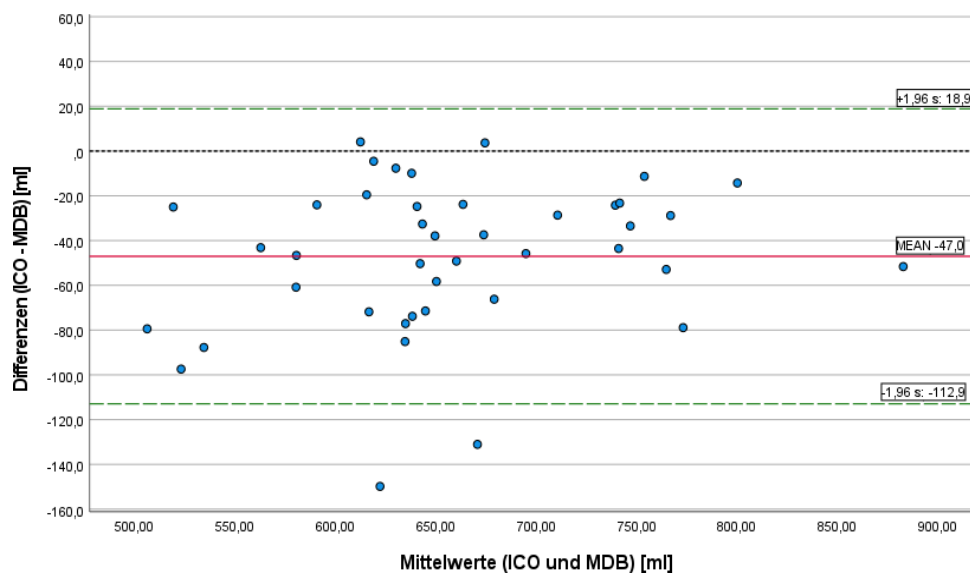


Abbildung 17: Bland-Altman Plot der Messungen der graue Substanz ohne KM
Limits of Agreement (grün): 18,9 ml und -112,9 ml. Mittlere Differenz (rot): -47,0 ml.

Beim Vergleich der Altersperzentilen zeigten sich signifikante p-Werte von unter 0,05. Nach Pearson ergab sich eine Korrelation von 37% der Perzentilenwerte ($r = 0,37$). Die deskriptive Statistik ist der Tabelle 10 zu entnehmen.

Statistik	ICO	MDB
Mittelwert	31,17	50,42
Standardabweichung	26,67	27,77
Minimum	0,50	0,09
Maximum	89,50	97,80

Tabelle 10: Deskriptive Statistik der Altersperzentilen der nativen Messung der grauen Substanz

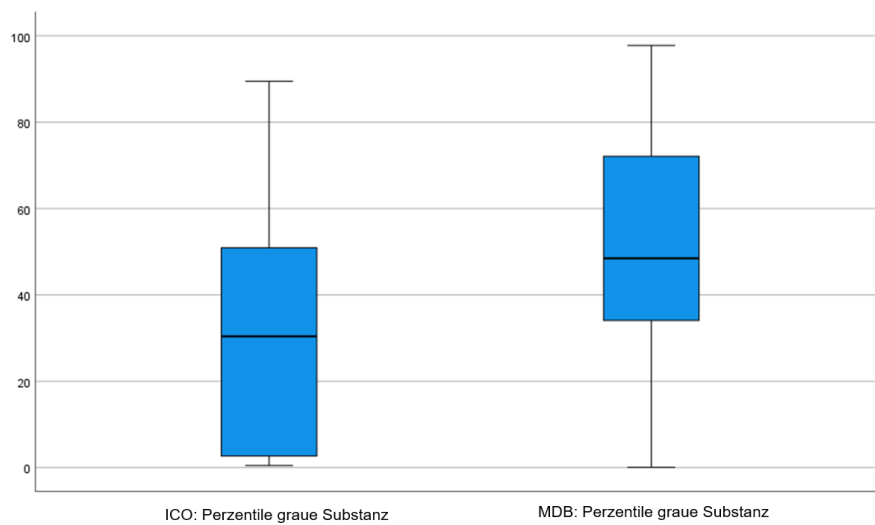


Abbildung 18: Boxplot der Altersperzentilen des Gesamthirns der nativen Untersuchungen

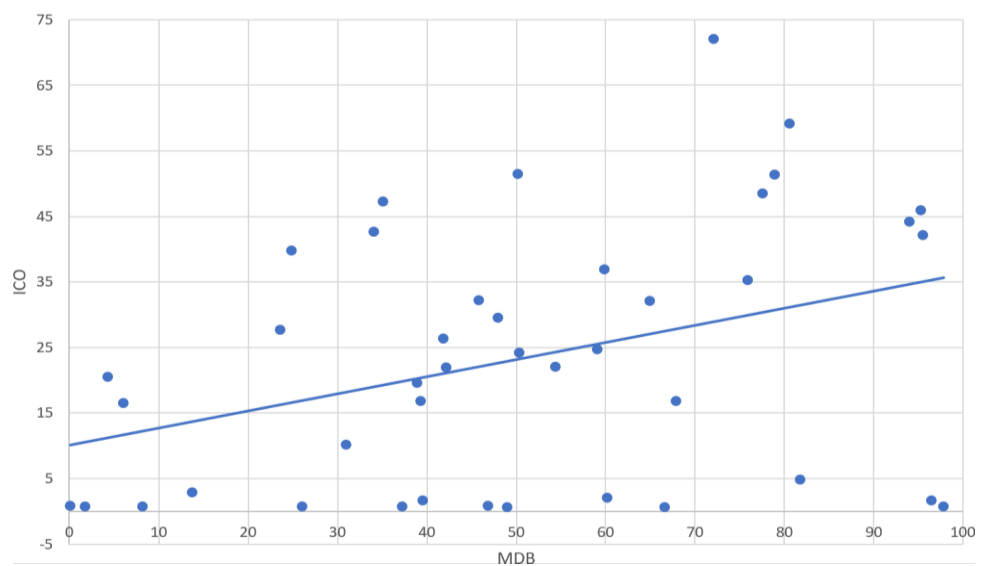


Abbildung 19: Altersperzentilen der grauen Substanz der nativen Untersuchungen

5.3.2 Untersuchungen mit KM der grauen Substanz

Bei der Messung der grauen Substanz mit Kontrastmittel ergab sich ein signifikanter p- Wert unter 0,05 ($<0,001$) bei der Betrachtung der absoluten Volumina. Die Korrelation nach Pearson beträgt $r = 0,87$. Weitere statistische Werte sind der Tabelle 11 zu entnehmen.

Statistik	ICO	MDB
Mittelwert [ml]	635,00	718,50
Standardabweichung [ml]	74,65	73,81
% Konfidenzintervall [ml]	612,30 - 657,69	696,02 - 740,90
Median [ml]	632,43	710,35
Minimum [ml]	503,20	569,50
Maximum [ml]	871,06	940,60

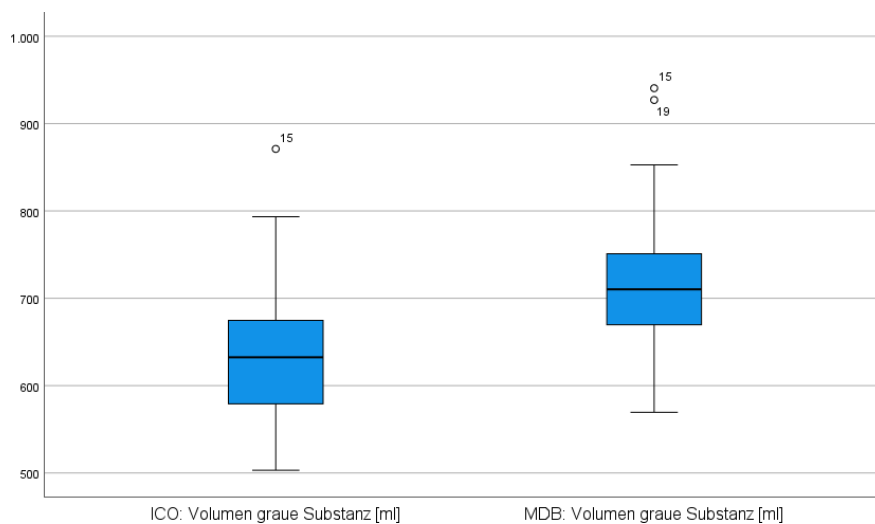
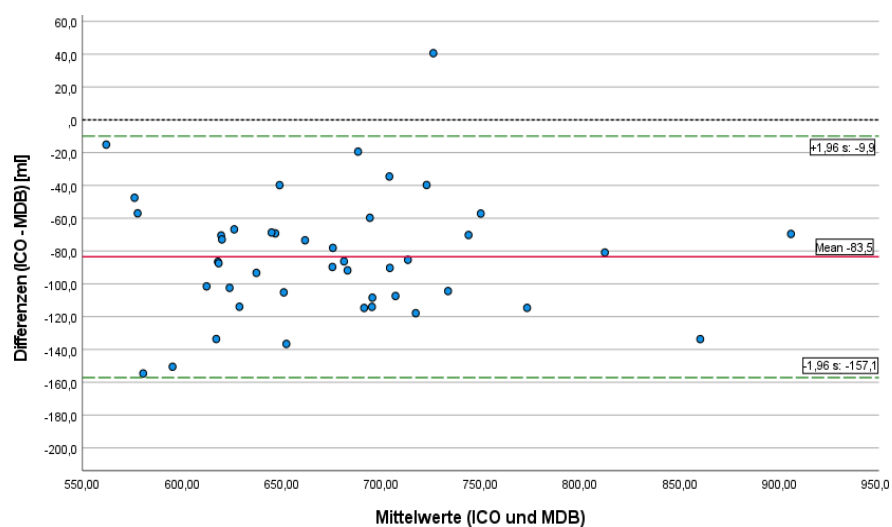
Tabelle 11: Deskriptive Statistik, graue Substanz mit KM**Abbildung 20:** Boxplot des Volumens der grauen Substanz bei KM-Gabe

Abbildung 21: Bland-Altman Plot Volumen graue Substanz mit KM
 Limits of Agreement (grün): -9,9 ml und -157,1 ml. Mittlere Differenz (rot): -83,5 ml.

Der p-Wert der relativen Volumenmessung zeigte sich ebenfalls signifikant unterschiedlich (p-Wert < 0,001). Der Korrelationskoeffizient nach Pearson wies einen Wert von $r = 0,18$ auf.

Statistik	ICO	MDB
Mittelwert	31,92	97,21
Standardabweichung	28,98	12,18
Minimum	0,50	20,50
Maximum	99,90	100,00

Tabelle 12: Deskriptive Statistik der Altersperzentilen der Ganzhirnmessungen mit KM

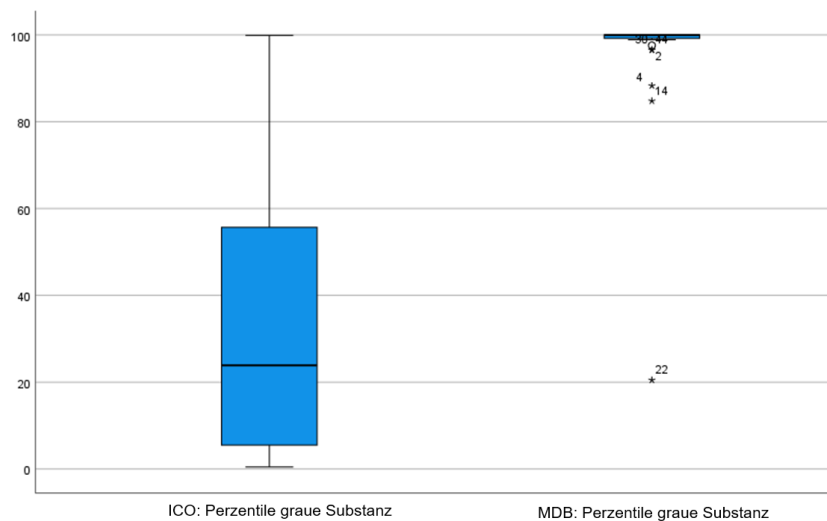


Abbildung 22: Boxplot der Altersperzentilen der grauen Substanz bei KM-Gabe

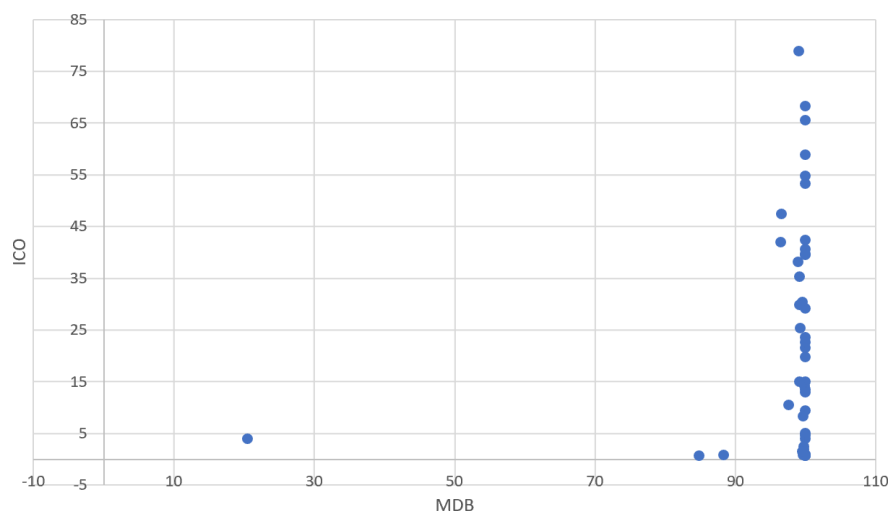


Abbildung 23: Altersperzentilen der grauen Substanz der Untersuchungen mit KM

5.4 Läsionslast

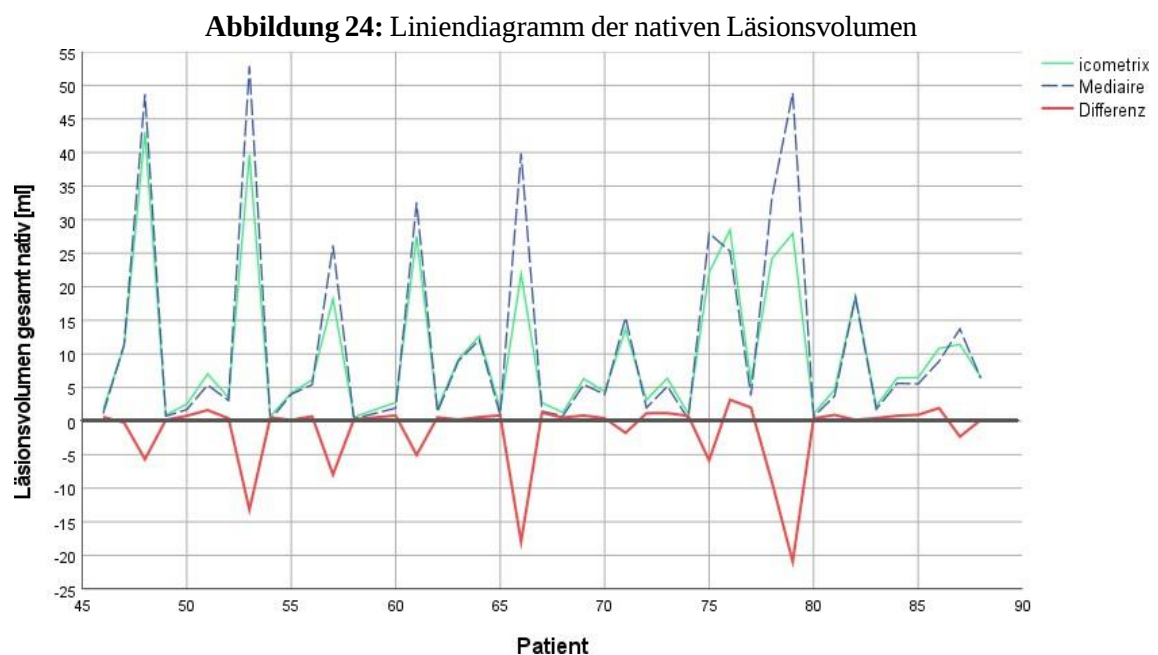
5.4.1 Native Messungen der Läsionslast

Es zeigte sich bei den nativen Messungen eine ähnlich hohe Läsionslast bei beiden Softwarelösungen mit p-Werten stets $>0,05$. Signifikant unterschiedliche Werte konnten nicht nachgewiesen werden. Die gemessenen Mittelwerte mit Standardabweichung und den p-Werten ist in der Tabelle 13 aufgeführt.

Läsionsort	ICO [ml]	MDB [ml]	p-Wert
periventrikulär	$9,41 \pm 10,67$	$10,69 \pm 14,45$	0,23
juxtakortikal	$0,30 \pm 0,28$	$0,44 \pm 0,65$	0,50
infratentoriell	$0,04 \pm 0,08$	$0,07 \pm 0,18$	0,46
übriges Marklager	$0,56 \pm 0,68$	$0,65 \pm 0,76$	0,20
ges. Läsionslast	$10,31 \pm 10,82$	$11,84 \pm 14,85$	0,35

Tabelle 13: Läsionslast nativer Untersuchungen.

Abbildung 24 zeigt die Läsionslast in [ml] der Patienten 45 - 86, die kein Kontrastmittel erhielten. Die Farbzuzuordnung ist der Legende zu entnehmen. In schwarz ist die Nulllinie markiert.



5.4.2 Kontrastmittelverstärkte Messungen der Läsionslast

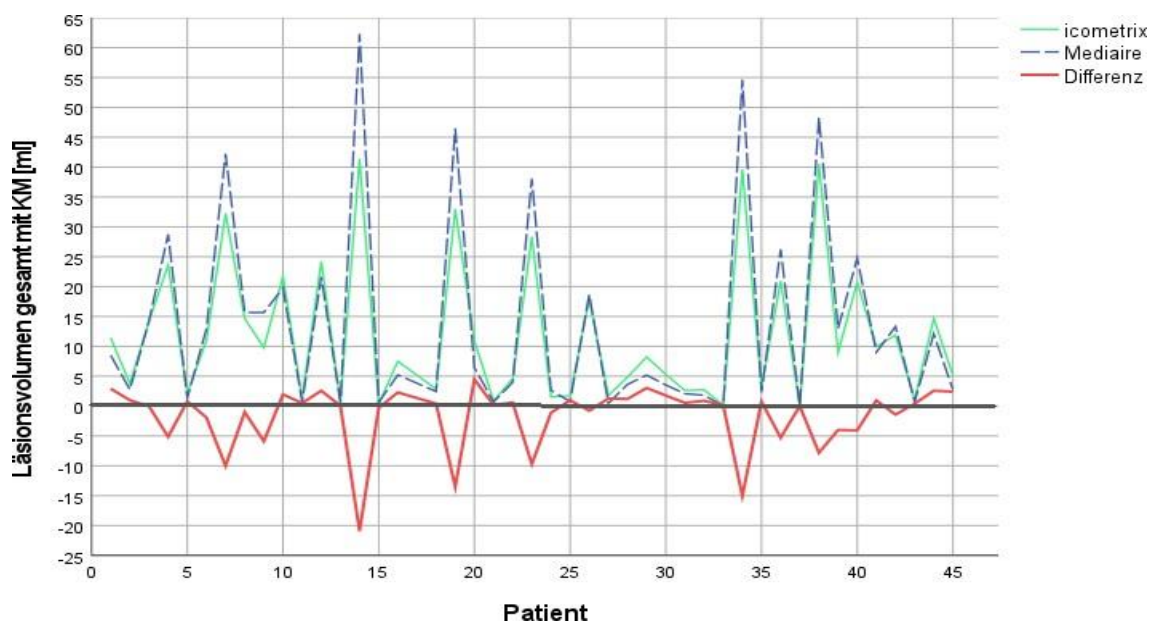
Die Mittelwerte, Standardabweichung und p-Werte der Läsionsmessungen mit KM sind in Tabelle 14 aufgelistet. Bei den kontrastmittelgestützten Messungen zeigte sich ein signifikanter Unterschied bei den infratentoriellen Werten (p-Wert=0,007). Periventrikulär, juxtakortikal und im übrigen Marklager dahingegen, konnten bei p-Werten größer als 0,05 keine signifikanten Diskrepanzen nachgewiesen werden. Auch bei der Betrachtung der gesamten Läsionslast ergab sich keine Signifikanz.

Läsionsort	ICO [ml]	MDB [ml]	p-Wert
periventrikulär	11,92 ± 11,94	13,58 ± 16,38	0,65
juxtakortikal	0,43 ± 0,37	0,47 ± 0,53	0,98
infratentoriell	0,16 ± 0,24	0,11 ± 0,17	0,007
übriges Marklager	1,00 ± 0,69	0,82 ± 0,58	0,06
Ges. Läsionslast	13,58 ± 12,08	14,98 ± 16,59	0,77

Tabelle 14: Läsionslast der KM-verstärkten Untersuchungen.

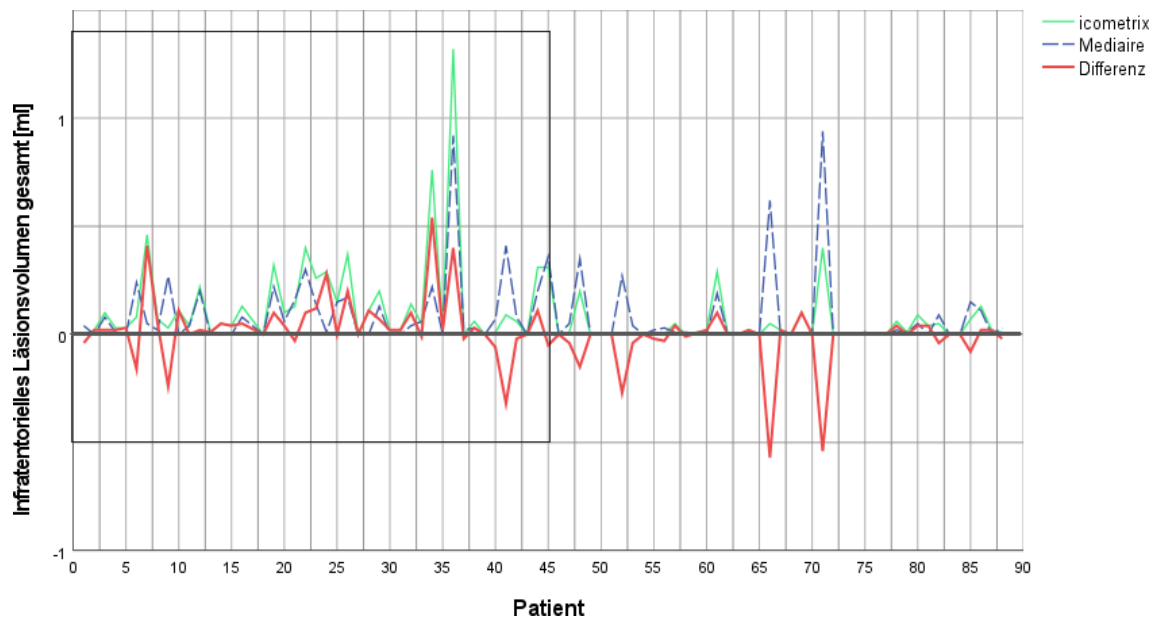
Abbildung 25 gibt die Läsionslast in [ml] der Patienten 1 - 44 wieder, die Kontrastmittel erhielten. Die Farbzuoordnung ist der Legende zu entnehmen. In schwarz ist die Nulllinie markiert.

Abbildung 25: Liniendiagramm der Läsionsvolumen mit KM



In Abbildung 26 ist die infratentorielle Läsionslast in [ml] aller Patienten gezeigt. Die Patienten 1 bis 44 erhielten Kontrastmittel, wohingegen die Patienten 45 bis 86 native Untersuchungen erhielten. Die Grenze der beiden Gruppen ist durch das schwarze Rechteck markiert. Die Farbzuoordnung ist der Legende zu entnehmen. In schwarz ist die Nulllinie markiert.

Abbildung 26: Liniendiagramm des infratentoriellen Läsionsvolumens aller Messungen.



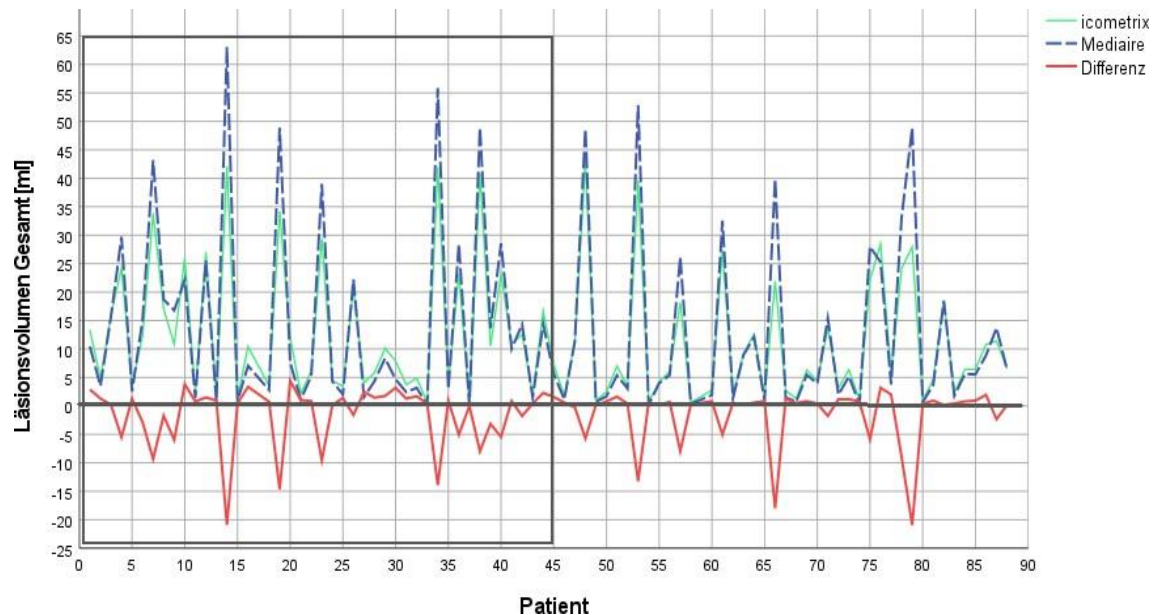
5.4.3 Messungen ohne Berücksichtigung des KM-Status

Lässt man die KM-Gabe außer Acht und betrachtet alle 86 Patientenergebnisse zusammen, so zeigte sich kein signifikanter Unterschied bei den Messungen mit p-Werten stets $>0,05$. Die deskriptive Statistik und die einzelnen p-Werte sind in Tabelle 15 zu finden.

Läsionsort	ICO [ml]	MDB [ml]	p-Wert
periventrikulär	$10,70 \pm 11,34$	$12,17 \pm 15,44$	0,69
juxtakortikal	$0,36 \pm 0,33$	$0,45 \pm 0,59$	0,66
infratentoriell	$0,10 \pm 0,19$	$0,09 \pm 0,17$	0,08
übriges Marklager	$0,78 \pm 0,72$	$0,74 \pm 0,67$	0,40
ges. Läsionslast	$11,90 \pm 11,52$	$13,45 \pm 15,75$	0,74

Tabelle 15: Läsionslast aller Untersuchungen zusammen mit p-Werten $>0,05$

Abbildung 27: Läsionsvolumen Gesamt



Abgebildet ist in Abbildung 27 die gesamte Läsionslast in [ml] aller Patienten. Die Patienten 1 bis 44 erhielten Kontrastmittel, wohingegen die Patienten 45 bis 86 native Untersuchungen erhielten. Die Grenze der beiden Gruppen ist durch das schwarze Rechteck markiert. Die Farbzusordnung ist der Legende zu entnehmen. In schwarz ist die Nulllinie markiert.

5.5 Zusammenfassung der Ergebnisse

Beim Vergleich der beiden Softwarelösungen *mdbrain* und *icobrain ms* zeigten sich z.T. signifikante Unterschiede in verschiedenen Volumenmessungen. Bei der Messung des Gesamthirnvolumen erwiesen sich die absoluten Volumina bei den nativen Untersuchungen signifikant unterschiedlich. Bei Betrachtung der grauen Substanz zeigten sich sowohl unter Kontrastmittelgabe als auch nativ deutliche Unterschiede bei den absoluten Messungen. Bei den nativen Volumenmessungen des Gesamthirns und der grauen Substanz war auffällig, dass die Messungen von *mdbrain* im Gegensatz zu *icometrix* ein signifikant höheres Volumen ausgaben. Bei den kontrastmittelgestützten Untersuchungen wurden bei *mdbrain* und *icobrain ms* ähnlich hohe Ganzhirnvolumina gemessen bei gleichzeitig sehr starker Korrelation der Werte ($r = 0,95$). Die kontrastmittelgestützten Volumina der grauen Substanz unterschieden sich jedoch signifikant. Hier maß *mdbrain* erneut mehr Volumen als *icobrain ms*.

Deutliche Unterschiede der Altersperzentilen ergaben sich bei der Messung des Gesamthirns mit KM bei einer Korrelation von $r = 0,52$, sowie bei der grauen Substanz mit und ohne KM.

Diese Werte zeigten jedoch nur eine schwache bis mäßige Korrelation ($r = 0,37$ der nativen Messungen und $0,18$ der KM-Messungen). Dabei fiel auf, dass *mdbrain* vor allem bei Kontrastmittelgabe vermehrt hohe Perzentilenwerte auswies. Besonders auffallend hoch waren die Perzentilen der grauen Substanz mit KM von *mdbrain* im Gegensatz zu *icobrain ms*.

Die Bewertung der Läsionslast zwischen den Softwarelösungen verblieben ohne signifikanten Unterschied, insofern der Kontrastmittelstatus unberücksichtigt blieb. Wurden die Messungen der nativen und kontrastmittelgestützten Untersuchung getrennt analysiert, so ergaben sich lediglich signifikante Werte bei den mit Kontrastmittel durchgeführten infratentoriellen Messungen.

Auffällig war bei den nativen Messungen, dass *mdbrain* vor allem bei höherer Läsionslast tendenziell ein höheres Volumen ausgab als *icobrain ms*. Statistisch waren die Unterschiede sehr gering, ein signifikanter Unterschied konnte nicht nachgewiesen werden. Bei Betrachtung der gesamten Läsionslast mit Kontrastmittelgabe fielen, wie bei den nativen Messungen, tendenziell niedrigere Volumina bei *icobrain ms* auf. Besonders betroffen waren wieder höhere Läsionslasten. Bei nur geringer Ausprägung der Herde schien der Unterschied weniger bemerkbar zu sein.

In Bezug auf die einzelnen Läsionsorte waren bei Kontrastmittelgabe die infratentoriellen Läsionsvolumina signifikant unterschiedlich. *icobrain ms* gab hier teilweise signifikant höhere Volumina aus als *mdbrain*. Bei der Berechnung der Messungen ohne Berücksichtigung des Kontrastmittelstatus, ergab sich keine Signifikanz.

6 Diskussion

6.1 Atrophie

Die absoluten Messwerte zeigen zwischen den beiden verwendeten Softwarelösungen bis auf die kontrastmittelgestützten Ganzhirnmessungen einen signifikanten Unterschied, allerdings stets eine sehr hohe Korrelation.

Die Berechnung der Altersperzentilen zeigt einen signifikanten Unterschied, außer bei den nativen Ganzhirnmessungen, jedoch nur bei den Ganzhirnmessungen sind die Korrelationen hoch. Bei den Messungen der grauen Substanz korrelieren die Werte der Altersperzentilen der beiden Softwarelösungen nur schwach bis gar nicht.

Die Ergebnisse unserer Studie sind wichtig, weil die Ergebnisse von zwei Softwarelösungen zwar sehr hoch korrelieren, d.h. also die gleichen Veränderungen anzeigen, jedoch die absoluten Werte nicht vergleichbar sind. Die Bestimmung eines hohen oder niedrigen Hirnvolumens kann also mit beiden Lösungen durchgeführt werden, allerdings können diese nicht miteinander verglichen werden, welches einen wichtigen Einfluss auf den klinischen Alltag hat, da z.B. beim Wechsel der Institution bei der Verlaufskontrolle eine Nachberechnung der Voruntersuchungen erfolgen müsste.

Die Ergebnisse der Studie von Lee et al. (2021) (Lee et al., 2021) sind sehr ähnlich zu unseren Ergebnissen. Auch diese verglichen zwei unterschiedliche Softwaremethoden (NeuroQuant und InBrain) und zeigten ebenfalls, dass sich signifikante absolute Unterschiede bei der Bestimmung der Hirnvolumina ergaben, jedoch die Korrelation sehr gut war - außer für das Pallidum. Im Gegensatz zu Lee et al. haben wir auch die für die klinische Einschätzung relevanten Altersperzentilen berechnet.

In einer weiteren Studie haben Zaki et al. (2022) (Zaki et al., 2022) zwei Softwarepakete mit künstlicher Intelligenz verwendet zum Vergleich normativer Gehirnvolumetrie und untersucht, ob sie die Demenzdiagnostik im klinischen Kontext unterschiedlich beeinflussen. Übereinstimmend mit unseren Ergebnissen zeigte sich ein signifikanter Unterschied zwischen den absoluten Volumina und eine schwache Korrelation für normative regionale Volumendaten, die für die Diagnosestellung verwendet wurden. Im Gegensatz zu unserer Studie wurde auch die Atrophieform bestimmt und verglichen. Wir haben jedoch MS-Patienten untersucht, bei denen nicht eine primäre Neurodegeneration wie bei der Demenz im Fokus steht, sondern die Volumenverminderung im Rahmen der Grunderkrankung.

In der Studie von Vinke et al. (2019) (Vinke et al., 2019) wurden Querschnittsstichproben aus

drei Referenzpopulationen verwendet, um die Perzentilenkurven der subkortikalen Volumina zu schätzen und zu vergleichen. Die Referenzpopulationen umfassten die Rotterdam-Studie, die United Kingdom Biobank (UKBB) und normale Kontrollen aus der Alzheimer's Disease Neuroimaging Initiative (ADNI). Trotz unterschiedlicher Basispopulationen und Bildgebungsparameter wie Feldstärke und Scannertyp zeigten die Ergebnisse, dass die normativen Daten der subkortikalen Volumina aus den verschiedenen Kohorten weitgehend austauschbar sind. Der Unterschied zu unserer Studie ist, dass wir die Einordnung der Perzentilen für das Ganzhirn und die grauen Substanz machten, nicht aber für subkortikale Strukturen.

Eine weitere Studie, die den Vergleich von Altersperzentilen untersuchte, ist die Studie von Finkelsztejn et al. (2018) (Finkelsztejn et al., 2018). Hier wurden MRT Scans von gesunden, brasilianischen Probanden mit den normativen Daten speziell von MSmetrix verglichen, um die Anwendbarkeit des Normkollektivs (basierend auf europäischen und nordafrikanischen Daten) auf eine lateinamerikanische Population zu untersuchen. Es zeigte sich keine signifikanter Unterschied zwischen den in den verschiedenen Untergruppen erzielten Perzentilen (Alter, Geschlecht), weshalb die MSmetrix-Datenbank mit Normalpersonen in verschiedenen Populationen verwendet werden kann. Im Gegensatz zu Finkelsztejn et al. (2018) zeigten sich unsere Ergebnisse bis auf die nativen Ganzhirnmessungen signifikant unterschiedlich. Besonders die Werte der Altersperzentilen der grauen Substanz korrelierten zudem nur schwach, was möglicherweise auf eine unterschiedliche Normalpopulation zurückzuführen ist. Den Einfluss von Kontrastmitteln auf die Ergebnisse der Hirnvolumetrie hat eine Studie von Lie et al. (2022) für die Softwarelösung FreeSurfer adressiert (Lie et al., 2022). Hier zeigte sich in einem intraindividuellen Szenario zwar eine sehr gute Korrelation zwischen den Ergebnissen mit und ohne Kontrastmittel, allerdings doch signifikante, absolute Unterschiede. Für unsere verwendeten Softwarelösungen existieren publizierte Daten bzgl. der Empfindlichkeit des Algorithmus auf appliziertes Kontrastmittel, jedoch empfiehlt der Hersteller mediate die Verwendung ihrer Softwarelösung nur auf nativen Bildern. Tatsächlich zeigte sich bei dieser Lösung auch ein signifikanter Unterschied in dem Hirnvolumen zwischen der nativ und mit KM untersuchten Population - wobei allerdings kein intraindividueller Vergleich vorliegt.

6.2 Läsionsvolumina

Bei den FLAIR-hyperintensiven Läsionen zeigte sich ein signifikanter Unterschied ($p < 0,01$) nur für infratentorielle Läsionen der Kontrastmittelgestützten Messungen, aber nicht für andere Regionen. Die nativen Messungen und die Messungen ohne Berücksichtigung des Kontrastmittelstatus verblieben ohne signifikanten Unterschied.

Die Ergebnisse unserer Studie hinsichtlich der Läsionsvolumina sind wichtig, da die Werte beider Softwarelösungen vergleichbar sind. Die Bestimmung der Läsionslast kann mit beiden Softwarelösungen bestimmt werden. Zur Bestimmung der Läsionslast kann auf Kontrastmittel verzichtet werden.

Ähnlich zu unserer Studie verglich die Gruppe von Fyllingen et al. (2016) drei Softwarepakete (BrainVoyagerTM QX, ITK-Snap und 3D Slicer) für die semi-automatische Segmentierung von Glioblastomen in MRT-Scans (Fyllingen et al., 2016). Die Ergebnisse zeigen eine hohe Übereinstimmung innerhalb und zwischen den Softwarepaketen. Die verwendeten Softwaretools sind jedoch halb-automatische Softwarelösungen, während die Softwarepakete unserer Studie bereits vollautomatisiert und somit komplett Untersucher unabhängig, sowie zeitsparender sind.

In einer Studie von Kollmann et al. (2020) wurde das Volumen von Vestibularschwannomen bei Neurofibromatose Typ 2 mit drei Bildanalyse-Tools (BrainLab, ITK-Snap und OsiriX) hinsichtlich ihrer Genauigkeit und Reproduzierbarkeit verglichen (Kollmann et al., 2020). Die Ergebnisse zeigten, dass alle Lösungen, ähnlich wie bei unserer Studie, eine geringe Variabilität aufwiesen. Die Volumenschätzungen unterschieden sich jedoch zwischen den Tools bei den tatsächlichen Tumoren. Dabei zu beachten ist jedoch die nur sehr geringe eingeschlossene Patientenzahl mit nur 8 Patienten. Anders als bei unserer Studie wurden zusätzlich volumetrische Messungen an Phantomen mit vordefinierten Volumina durchgeführt. Es zeigte sich, dass alle drei Lösungen signifikant vom tatsächlichen Volumen der Phantome abwichen, jedoch nicht untereinander (OsiriX und BrainLab).

6.3 Limitationen der Hirnvolumetrie

6.3.1 Limitationen der Hirnvolumetrie generell

Ein zentraler Aspekt ist das Fehlen eines etablierten Goldstandards für die Hirnvolumetrie, was die Validierung und Vergleichbarkeit der Messungen erschwert.

Zusätzlich müssen technische und personenbezogene Schwankungen beachtet werden, die zu

Messgenauigkeit führen können, wie etwa Updates der MRT-Scanner, kleinste Veränderungen in der Kopfposition oder Bewegungsartefakte. Patientenbezogene Faktoren wie physiologische Variablen (Alter, Geschlecht, Schädelform, Wasserhaushalt), die Uhrzeit der Untersuchung sowie Lebensgewohnheiten (Rauchen, Substanzmissbrauch, chronische Erkrankungen) können ebenfalls Einfluss auf die Messungen haben. (Durand-Dubief et al., 2012); (Duning et al., 2005); (Giorgio und de Stefano, 2013); (Nakamura et al., 2014); (Sastre-Garriga et al., 2020), (Sinnecker et al., 2018); (Enzinger et al., 2005); (Marciniewicz et al., 2019).

Inbesondere bei MS-Patienten kann die Krankheitsaktivität durch akute Entzündungen und Ödembildung vorübergehend zu einer Zunahme des Hirnvolumens führen. Auch gliale Zellproliferation und Gliose können zu falschen Annahmen einer Volumenzunahme führen (Marciniewicz et al., 2019); (Nakamura et al., 2014).

Die begrenzte Anzahl an kommerziellen MS-QReport-Lösungen mit nachgewiesener klinischer Validierung und Evaluierung durch klinische Endnutzer stellt eine weitere Limitation dar. Dies unterstreicht die Notwendigkeit einer systematischen Validierung und Bewertung solcher Berichtswerkzeuge, um ihre klinische Anwendung und Akzeptanz zu verbessern. (Mendelsohn et al., 2023).

Zusammenfassend ist es entscheidend, Verlaufsuntersuchungen auf demselben MRT-Gerät durchzuführen, um die Reproduzierbarkeit der Messungen sicherzustellen und Schwankungen zu minimieren. Die Zusammenarbeit zwischen klinischen Einrichtungen und Herstellern sowie klare regulatorische Rahmenbedingungen sind notwendig, um die Weiterentwicklung und Integration solcher Werkzeuge in klinische Praktiken voranzutreiben.

6.3.2 Limitationen dieser Studie

Die vorliegende Studie hat bedeutende Erkenntnisse geliefert, jedoch sind einige Limitationen zu berücksichtigen, die die Interpretation der Ergebnisse beeinflussen könnten.

Das Fehlen eines etablierten Referenzstandards stellt ein zentrales Problem dar und war auch in dieser Studie eine Einschränkung (Filippi et al., 2016); (Sastre-Garriga et al., 2020). Eine potenzielle Lösung könnte die Verwendung von Phantomgehirnen sein, die beispielsweise mithilfe moderner 3D-Drucktechnologie mit präzisen Volumina hergestellt werden könnten. Diese könnten dann mittels MRT und den untersuchten Softwarelösungen analysiert werden, ähnlich wie in der Studie von Kollmann et al. (2020) (Kollmann et al., 2020).

Eine weitere Limitation der vorliegenden Studie ist die begrenzte Stichprobengröße. Eine Erhöhung der Teilnehmerzahl wäre wünschenswert, um die Zuverlässigkeit der Ergebnisse zu stärken und eine bessere Generalisierbarkeit der Befunde zu ermöglichen.

6.4 Schlussfolgerung

Die vorliegenden Untersuchungen zeigen, dass beide Softwarelösungen zuverlässige Ergebnisse bei der Segmentierung und Volumenbestimmung des Gehirns und der grauen Substanz liefern, unabhängig davon, ob native oder kontrastmittelgestützte MRT-Aufnahmen verwendet werden. Für konsistente Ergebnisse bei Kontrolluntersuchungen ist die Nutzung derselben Software unerlässlich, und es sollten die allgemeine Empfehlungen zur Reduktion von Schwankungen beachtet werden.

Die Altersperzentilen bieten eine grobe Orientierung innerhalb der Normalbevölkerung, wobei eine Überarbeitung des Normkollektivs zur Verbesserung der Standardisierung der Algorithmen und Erweiterung der Trainingsdatensätze sinnvoll sein könnte.

Beide Anwendungen erweisen sich als verlässlich in der Bestimmung der Läsionslast, wobei der Kontrastmittel-Status keinen signifikanten Einfluss zeigt. Zur Bestimmung infratentorieller Läsionen sind jedoch weitere Studien notwendig, um die Vorteile zwischen kontrastmittelgestützten und nativen Messungen sowie unterschiedlichen Softwarelösungen zu klären. Ob Deep Learning im Vergleich zu herkömmlichem maschinellen Lernen bei der Bestimmung der Atrophierate und Läsionslast einen Vorteil bietet, bleibt aufgrund des fehlenden Goldstandards unklar. Es wird jedoch erwartet, dass durch den Einsatz fortschrittlicher KI-Algorithmen die Genauigkeit und Zuverlässigkeit der Hirnvolumetrie weiter verbessert werden kann. Dies könnte zu präziseren Diagnosen neurologischer Erkrankungen und zur Validierung der Effektivität von Therapien beitragen, wie zum Beispiel in der Studie von Andravizou et al. (2019) verdeutlicht (Andravizou et al., 2019).

Basierend auf dieser Studie empfiehlt es sich, die vorliegenden MRT-Bilder mit den aktuellen Softwarelösungen von *mdbrain* (Version 3 und 4) zu untersuchen und mit den bisherigen Ergebnissen zu vergleichen. Bemerkenswert ist, dass bisher keine Deep Learning-basierte Anwendung zur Bestimmung der Hirnatrophierate verwendet wurde. Die neuesten Versionen von *mdbrain* bieten verbesserte neuronale Netzwerke und erweiterte Trainingsdatensätze, insbesondere für die Analyse der Läsionslast. Ein Vergleich mit anderen Anbietern, die ebenfalls auf Deep Learning setzen, könnte zusätzliche Erkenntnisse liefern.

Zusammengefasst zeigen sich bei der automatisierten Auswertung durch zwei unterschiedliche Software-Lösungen signifikante Unterschiede hinsichtlich des absoluten und relativen Volumens des Gesamthirns und der grauen Substanz, sowie für das Volumen von infratentoriellen Läsionen.

7 Literaturverzeichnis

- Akkus Z., Galimzianova, A., Hoogi, A., Rubin D., & Erickson J. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of digital imaging*, 30(4):449–459, 2017. ISSN 1618-727X. doi: 10.1007/s10278-017-9983-4.
- Alzubaidi L., Zhang J., Humaidi A., Al-Dujaili A., Duan Y., Al-Shamma O., Santamaría J., Fadhel M., Al-Amidie M., & Farhan L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1):1–74, 2021. ISSN 2196-1115. doi: 10.1186/s40537-021-00444-8.
- Andravizou A., Dardiotis E., Artemiadis A., Sokratous M., Siokas V., Tsouris Z., Aloizou A., Nikolaidis I., Bakirtzis C., Tsivgoulis G., Deretzi G., Grigoriadis N., Bogdanos D., & Hadjigeorgiou G. Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options. *Auto- immunity highlights*, 10(1):7, 2019. ISSN 2038- 0305. doi: 10.1186/s13317-019-0117-5.
- Bermel R., Bakshi R., Tjoa C., Puli S., & Jacobs L. Bicaudate ratio as a magnetic resonance imaging marker of brain atrophy in multiple sclerosis. *Archives of neurology*, 59(2):275–280, 2002. ISSN 0003-9942. doi: 10.1001/archneur.59.2.275.
- Bernstein M., King K., & Zhou X. Handbook of MRI pulse sequences. *Academic Press, Amsterdam*, 2004. ISBN 9780080533124. URL <http://www.sciencedirect.com/science/book/9780120928613>.
- Browne P., Chandraratna C., Angood C., Tremlett H., Baker C., Taylor B., Thompson A. Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology*, 83(11):1022–1024, 2014. doi: 10.1212/WNL.0000000000000768.
- Cortese R., Collorone S., Ciccarelli O., & Toosy A.. Advances in brain imaging in multiple sclerosis. *Therapeutic advances in neurological disorders*, 12:1756286419859722, 2019. ISSN 1756-2856. doi: 10.1177/1756286419859722.
- Cosmus T. & Parizh M. *Advances in Whole-Body MRI Magnets*. IEEE, 2011.
- De Stefano N, Airas L, Grigoriadis N, Mattle H, O' Riordan J, Oreja-Guevara C, Sellebjerg F, Stankoff B, Walczak A, Wiendl H, Kieseier B. *Clinical relevance of brain volume measures in multiple sclerosis*. 2014; 28(2):147-56. doi: 10.1007/s40263-014-0140-z.

- Duning T, Kloska S, Steinsträter O, Kugel H, Heindel W, Knecht S. Dehydration confounds the assessment of brain atrophy. *Neurology*, 64(3):548–550, 2005. doi: 10.1212/01.WNL.0000150542.16969.CC.
- Durand-Dubief F, Belaroussi B, Armspach JP, Dufour M, Roggerone S, Vukusic S, Hannoun S, Sappey Marinier D, Confavreux C, Cotton F. Reliability of longitudinal brain volume loss measurements between 2 sites in patients with multiple sclerosis: comparison of 7 quantification techniques. *American Journal of Neuroradiology*, 2012 Nov;33(10):1918-1924. DOI: 10.3174/ajnr.a3107.
- Enzinger C, Fazekas F, Matthews P, Ropele S, Schmidt H, Smith S & Schmidt R. Risk factors for progression of brain atrophy in aging: six-year follow-up of normal subjects. *Neurology*, 64(10):1704–1711, 2005. doi: 10.1212/01.WNL.0000161871.83614.BB.
- Filippi M, Rocca M. *MRI evidence for multiple sclerosis as a diffuse disease of the central nervous system*. J Neurol. 2005 Nov; 252 Suppl 5:v16-24. doi:10.1007/s00415-005-5004-5.
- Filippi M, Rocca M, De Stefano N, Enzinger C, Fisher E, Horsfield M, Pelletier D, Comi G. Magnetic Resonance Techniques in Multiple Sclerosis: The Present and the Future. *Arch Neurol*. 2011;68(12):1514–1520. doi:10.1001/archneurol.2011.914
- Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, Rovira A, Sastre-Garriga J, Tintorè M, Frederiksen JL, Gasperini C, Palace J, Reich DS, Banwell B, Montalban X, Barkhof F; MAGNIMS Study Group. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol*. 2016 Mar;15(3):292-303. doi: 10.1016/S1474-4422(15)00393-2.
- Filippi M, Preziosa P, Banwell B, Barkhof F, Ciccarelli O, de Stefano N, Geurts J, Paul F, Reich D, Toosy A, Traboulsee A, Wattjes M, Yousry T, Gass A, Lubetzki C, Weinshenker B, Rocca M. *Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines*. A journal of neurology, 142(7):1858–1875, 2019. ISSN 0006-8950. doi: 10.1093/brain/awz144.
- Finkelsztejn A, Fragoso Y, Bastos E, Duarte J, Varela J, Houbrechts R, Sima D, van Hecke W. Intercontinental validation of brain volume measurements using MSmetrix. *The neuroradiology journal*, 31(2):147–149, 2018. ISSN 1971-4009. doi: 10.1177/1971400917721521.

- Fyllingen E, Stensjøen A, Berntsen E, Solheim O, & Reinertsen I. Glioblastoma Segmentation: Comparison of Three Different Software Packages. *PloS one*, 11(10):e0164891, 2016. doi: 10.1371/journal.pone.0164891.
- García-Lorenzo, D., Francis, S.J., Narayanan, S., Arnold, D., & Collins, D.L. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical image analysis*, 17 1, 1-18, 2013. doi.org/10.1016/j.media.2012.09.004
- Giorgio A., Battaglini M., Smith S., & de Stefano N. Brain atrophy assessment in multiple sclerosis: importance and limitations. *Neuroimaging clinics of North America*, 18(4):675–86, xi, 2008. ISSN 1052-5149. doi: 10.1016/j.nic. 2008.06.007.
- Giorgio A, De Stefano N. Clinical use of brain volumetry. *J Magn Reson Imaging*. 2013 Jan;37(1):1-14. doi: 10.1002/jmri.23671. PMID: 23255412.
- Guo C., Ferreira D., Fink K., Westman E., & Granberg T. Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. *European radiology*, 29(3):1355–1364, 2019. doi: 10.1007/s00330-018-5710-x.
- Hartung H. P. Entmarkungskrankheiten. In Johannes Jörg, editor, *Neurologische Therapie*, pages 479–500. Springer Berlin Heidelberg, 2002. ISBN 978- 3-540-67246-3.
- Heinen R, Bouvy W, Mendrik A, Viergever M, Biessels G, de Bresser J. Robustness of Automated Methods for Brain Volume Measurements across Different MRI Field Strengths. *PloS one*, 11(10):e0165719, 2016. doi: 10.1371/journal.pone.0165719.
- Hemond C. & Bakshi R. Magnetic Resonance Imaging in Multiple Sclerosis. *Cold Spring Harbor perspectives in medicine*, 8(5), 2018. doi: 10.1101/cshperspect. a028969.
- Hidalgo-Tobon S. *Theory of gradient coil design methods for magnetic resonance imaging*. Wiley Subscription Services, Inc., A Wiley Company, Hoboken, 2010.
- Hoseinipourasl M, Zandkarimi M, Abdolmohammadi J, Sharifi K, Miraki S. Evaluation the FLAIR Sensitivity and DWI Post-inject in Comparison with Delayed Enhancement T1w for Better Detection of Active MS Lesions. *Journal of biomedical physics & engineering*, 8(4):365–374, 2018. ISSN 2251-7200.

- Hu X, Rajendran I, Lapointe E, Tam R, Li D, Traboulsee A, Rauscher A. Three-dimensional MRI sequences in MS diagnosis and research. *Multiple sclerosis (Houndmills, Basingstoke, England)*, 25(13): 1700–1709, 2019. doi: 10.1177/1352458519848100.
- Icometrix. icobrain-ms | icometrix, 2021. URL <https://icometrix.com/services/icobrain-ms>.
- Ismailova K., Poudel P., Parlesak A., Frederiksen P., & Heitmann B. Vitamin D in early life and later risk of multiple sclerosis-A systematic review, meta-analysis. *PloS one*, 14(8):e0221645, 2019. doi: 10.1371/journal.pone.0221645.
- Jain S, Sima DM, Ribbens A, Cambron M, Maertens A, Van Hecke W, De Mey J, Barkhof F, Steenwijk MD, Daams M, Maes F, Van Huffel S, Vrenken H, Smeets D. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. *Neuroimage Clin.* 2015 May 16;8:367-75. doi: 10.1016/j.nicl.2015.05.003.
- Jörg J., editor. *Neurologische Therapie*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002. ISBN 978-3-540-67246-3. doi: 10.1007/978-3-642-56398-0.
- Kakeda S., Korogi Y., Hiai Y., Ohnari N., Sato T., & Hirai T. Pitfalls of 3D FLAIR brain imaging: a prospective comparison with 2D FLAIR. *Academic radiology*, 19(10):1225–1232, 2012. doi: 10.1016/j.acra.2012.04.017.
- Kitajima M., Hirai T., Shigematsu Y., Uetani H., Iwashita K., Morita K., Komi M., & Yamashita Y. Comparison of 3D FLAIR, 2D FLAIR, and 2D T2-weighted MR imaging of brain stem anatomy. *AJNR. American journal of neuroradiology*, 33(5): 922–927, 2012. doi: 10.3174/ajnr.A2874.
- Kollmann P., Mautner V., Koeppen J., Wenzel R., Friedman J., Salamon J., Farschtschi S. MRI based volumetric measurements of vestibular schwannomas in patients with neurofibromatosis type 2: comparison of three different software tools. *Scientific reports*, 10(1):11541, 2020. doi: 10.1038/s41598-020-68489-y.
- Ladd M., Quick H., & Bock M. *Magnetresonanztomographie und -spektroskopie: In: Schlegel W., Karger C., Jäkel O. (eds) Medizinische Physik*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN 3662548003. doi: 10.1007/978-3-662-54801-1
- Lassmann H. Multiple Sclerosis Pathology. *Cold Spring Harbor perspectives in medicine*, 8(3), 2018. doi: 10.1101/cshperspect.a028936.

- Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol.* 2017 Jul-Aug;18(4):570-584. doi: 10.3348/kjr.2017.18.4.570.
- Lee J., Oh S., Sun Chung M., Park J., Moon Y., Jeon H., & Moon W. Clinically Available Software for Automatic Brain Volumetry: Comparisons of Volume Measurements and Validation of Intermethod Reliability. *Korean journal of radiology*, 22(3):405–414, 2021. ISSN 1229-6929. doi: 10.3348/kjr.2020.0518.
- Li J, Jiang P, An Q, Wang GG, Kong HF. Medical image identification methods: A review. *Computers in biology and medicine.* 2024 Feb;169:107777. doi: 10.1016/j.combiomed.2023.107777.
- Lie I., Kerklingh E., Wesnes E., van Nederpelt D., Brouwer I., Torkildsen Ø., Myhr K., Barkhof F., Bø L., & Vrenken H.. The effect of gadolinium-based contrast-agents on automated brain atrophy measurements by FreeSurfer in patients with multiple sclerosis. *European radiology*, 32(5): 3576–3587, 2022. doi: 10.1007/s00330-021-08405-8.
- Lukas C, Sailer M. Magnetresonanztomografie bei Patienten mit Multipler Sklerose: Relevanz in der Diagnose und Verlaufsbeurteilung. *Fortschritte der Neurologie- Psychiatrie*, 83(9):522–32; quiz 533–4, 2015. doi: 10.1055/s-0035-1553712.
- Marciniewicz E., Podgórski P., Sasiadek M., Bladowska J. The role of MR volumetry in brain atrophy assessment in multiple sclerosis: A review of the literature. *Advances in clinical and experimental medicine : official organ Wroclaw Medical University*, 28(7):989–999, 2019. ISSN 1899-5276. doi: 10.17219/acem/94137.
- Mattle H. und Mumenthaler M. *Kurzlehrbuch Neurologie*. Georg Thieme Verlag, Stuttgart and New York, 4., vollständig überarbeitete und aktualisierte Auflage, 2015. ISBN 9783131429346.
- McDonald W. I., Compston A., Edan G., Goodkin D., Hartung H., Lublin F., McFarland F., Paty D., Polman C., Reingold S., Sandberg-Wollheim M., Sibley W., Thompson A., van den Noort S., Weinshenker B., & Wolinsky J. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of neurology*, 50(1):121–127, 2001a. ISSN 0364-5134. doi: 10.1002/ana.1032.

- McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, Sandberg-Wollheim M, Sibley W, Thompson A, van den Noort S, Weinshenker BY, Wolinsky JS. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*. 2001 Jul;50(1):121-7. doi: 10.1002/ana.1032.
- McKinley R, Wepfer R, Aschwanden F, Grunder L, Muri R, Rummel C, Verma R, Weisstanner C, Reyes M, Salmen A, Chan A, Wagner F, Wiest R. Simultaneous lesion and brain segmentation in multiple sclerosis using deep neural networks. *Sci Rep*. 2021 Jan 13;11(1):1087. doi: 10.1038/s41598-020-79925-4.
- mediaire. Mediaire - mdbrain, 2019. URL <https://mediaire.de/product/mdbrain/>.
- Meinl I. Therapie der Multiplen Sklerose mit Natalizumab: Klinische Wirksamkeit und Veränderung der zirkulierenden Immunzellsubtypen bei einem hochaktiven Patientenkollektiv. 2012. DOI: 10.5282/edoc.14263
- Mendelsohn Z., Pemberton H., Gray J., Goodkin O., Carrasco F., Scheel M., Nawabi J., & Barkhof F. Commercial volumetric MRI reporting tools in multiple sclerosis: a systematic review of the evidence. *Neuroradiology*, 65(1):5–24, 2023. doi: 10.1007/s00234-022-03074-w.
- mh/ktg. neuroRAD 2019 – Hirnvolumetrie mit KI-basierten Tools, 2019. URL <https://radiologie.bayer.de/aktuelles/kongressberichte/neurorad-koeln/neurorad-2019-hirnvolumetrie-mit-ki-basierten-tools>.
- Naganawa S. The Technical and Clinical Features of 3D-FLAIR in Neuroimaging. *Magn Reson Med Sci*. 2015;14(2):93-106. doi: 10.2463/mrms.2014-0132.
- Nakamura K., Brown R., Araujo D., Narayanan S., & Arnold D. Correlation between brain volume change and T2 relaxation time induced by dehydration and rehydration: implications for monitoring atrophy in clinical studies. *NeuroImage. Clinical*, 6:166–170, 2014. doi: 10.1016/j.nicl.2014.08.014.
- Nelson F., Poonawalla A., Hou P., Wolinsky J., & Narayana P. 3D MPRAGE improves classification of cortical lesions in multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)*, 14(9):1214–1219, 2008. doi: 10.1177/1352458508094644.
- Notohamiprodjo, M; Lummel, Nina; Baum, T, Röttinger, M; Krause, L & Kleiner, I. *Vergleich*

- von zwei Programmen zur automatisierten Quantifizierung des Hirn- und Läsionsvolumens bei Multiple-Sklerose-Patienten. *RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*. 2021. Doi: 10.1055/s-0041-1723201.
- Rupp R. *Läsionssuche bei Patienten mit Multipler Sklerose: Sensitivität verschiedener MRT-Sequenztypen*. Regensburg, 2016.
- Sahraian M. und Radü E. *MRI atlas of MS lesions*. Springer, Berlin and Heidelberg, 2008. ISBN 9783540713715. doi: 10.1007/978-3-540-71372-2.
- Sarker I. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(6):1–20, 2021. ISSN 2661-8907. doi: 10.1007/s42979-021-00815-1.
- Sastre-Garriga J., Pareto D., Battaglini M., Rocca M., Ciccarelli O., Enzinger C., Wuerfel J., Sormani M., Barkhof F., Yousry T., de Stefano N., Tintoré M., Filippi M., Gasperini C., Kappos L., Río J., Frederiksen J., Palace J., Vrenken H., Montalban X., & Rovira A. MAGNIMS consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. *Nature reviews. Neurology*, 16(3):171–182, 2020. doi: 10.1038/s41582-020-0314-x.
- Schober P, Boer C, Schwarte L. Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia and analgesia*, 126(5):1763–1768, 2018. doi: 10.1213/ANE.0000000000002864.
- Sima D., Loeckx D., Smeets D., Jain S., Parizel P., & van Hecke W. Use Case I: Imaging Biomarkers in Neurological Disease. Focus on Multiple Sclerosis. In Luis Martí-Bonmatí and Angel Alberich-Bayarri, editors, *Imaging Biomarkers*, pages 169–180. Springer International Publishing, Cham, 2017. doi: 10.1007/978-3-319-43504-6{\textunderscore}15.
- Sinnecker T, Granziera C, Wuerfel J, Schlaeger R. Future Brain and Spinal Cord Volumetric Imaging in the Clinic for Monitoring Treatment Response in MS. *Curr Treat Options Neurol*. 2018 Apr 20;20(6):17. doi: 10.1007/s11940-018-0504-7. PMID: 29679165.
- Sormani M., Kappos L., Radue E., Cohen J., Barkhof F., Sprenger T., Meier D., Häring D., Tomic D., & de Stefano N. Defining brain volume cutoffs to identify clinically relevant atrophy in RRMS. *Multiple sclerosis (Houndmills, Basingstoke, England)*, 23(5):656–664, 2017. doi: 10.1177/1352458516659550.
- Storelli L., Rocca M., Pagani E., van Hecke W., Horsfield M., de Stefano N., Rovira A., Sastre-

Garriga J., Palace J., Sima D., Smeets D., & Filippi M. Measurement of Whole-Brain and Gray Matter Atrophy in Multiple Sclerosis: Assessment with MR Imaging. *Radiology*, 288(2):554–564, 2018. doi: 10.1148/radiol.2018172468.

Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, Correale J, Fazekas F, Filippi M, Freedman MS, Fujihara K, Galetta SL, Hartung HP, Kappos L, Lublin FD, Marrie RA, Miller AE, Miller DH, Montalban X, Mowry EM, Sorensen PS, Tintoré M, Traboulsee AL, Trojano M, Uitdehaag BMJ, Vukusic S, Waubant E, Weinshenker BG, Reingold SC, Cohen JA. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria *The Lancet. Neurology*. 2018 Feb;17(2):162-173. doi: 10.1016/S1474-4422(17)30470-2.

Toledano-Massiah S., Sayadi A., de Boer R., Gelderblom J., Mahdjoub R., Gerber S., Zuber M., Zins M., & Hodel J. Accuracy of the Compressed Sensing Accelerated 3D-FLAIR Sequence for the Detection of MS Plaques at 3T. *AJNR. American journal of neuroradiology*, 39(3):454–458, 2018. doi: 10.3174/ajnr.A5517.

Valverde S, Oliver A, Roura E, Pareto D, Vilanova JC, Ramió-Torrentà L, Sastre-Garriga J, Montalban X, Rovira À, Lladó X. Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling. *Neuroimage Clin*. 2015 Oct 28;9:640-7. doi: 10.1016/j.nicl.2015.10.012.

Vinke E., Huizinga W., Bergtholdt M., Adams H., Steketee R., Papma J., de Jong F., Niessen W., Ikram A., Wenzel F., Vernooij M. Normative brain volumetry derived from different reference populations: impact on single-subject diagnostic assessment in dementia. *Neurobiology of aging*, 84:9–16, 2019. doi: 10.1016/j.neurobiolaging.2019.07.008.

Vural G., Keklikoğlu H., Temel S., Deniz O., Ercan K. Comparison of double inversion recovery and conventional magnetic resonance brain imaging in patients with multiple sclerosis and relations with disease disability. *The neuroradiology journal*, 26 (2):133–142, 2013. ISSN 1971-4009. doi: 10.1177/197140091302600201.

Zaki L., Vernooij M., Smits M., Tolman C., Papma J., Visser J., & Steketee R. Comparing two artificial intelligence software packages for normative brain volumetry in memory clinic imaging. *Neuroradiology*, 64(7):1359–1366, 2022. doi: 10.1007/s00234-022-02898-w.

8 Abkürzungsverzeichnis

FLAIR	fluid attended inversion recovery
FSL	FMRI Software Library
ICO	icobrain ms
KI	Künstliche Intelligenz
KM	Kontrastmittel
KNN	Künstliche neuronale Netze
MDB	mdbrain
MPRAGE	magnetization-prepared rapid gradient-echo
MRT	Magnetresonanztomographie
MS	Multiple Sklerose
PACS	Picture Archiving and Communication System
PERZ	Altersperzentile
R	Korrelationskoeffizient
SD	standard deviation, Standardabweichung
SPM	Statistical Parametric Mapping
T	Tesla
TE	Echozeit, „time echo“
TI	Inversionszeit
TR	Repetitionszeit, „time repetition“
ZNS	Zentralnervensystem

9 Abbildungsverzeichnis

1:	Liste einiger vollautomatisierter Softwarelösungen	12
2:	Studiendesign	16
3:	Pipeline <i>icobrain ms</i>	17
4:	Beispiel eines <i>icobrain ms</i> Reports, Volumenreport (Icometrix, 2021).....	18
5:	Beispiel eines <i>icobrain ms</i> Läsionsreports (Icometrix, 2021).....	19
6:	Beispiel eines <i>mdbrain</i> Volumetriereports. Darstellung von Gesamthirn-, Regional- und Ventrikelvolumina (mediaire, 2019).	20
7:	Beispiel eines <i>mdbrain</i> Läsionsreports (mediaire, 2019).....	21
8:	Boxplot des Gesamthirnvolumens der nativen Untersuchungen	24
9:	Bland-Altman Plot des Gesamthirnvolumen ohne KM	24
10:	Boxplot der Altersperzentilen der nativen Ganzhirnmessungen	25
11:	Altersperzentilen vom Ganzhirn der nativen Untersuchungen.....	25
12:	Boxplot des Gesamthirnvolumens bei kontrastmittelverstärkten Aufnahmen.....	26
13:	Bland-Altman Plot Gesamthirnvolumen mit KM	27
14:	Boxplot der Altersperzentilen des Gesamthirnvolumens mit KM-Gabe.....	28
15:	Altersperzentilen vom Ganzhirn der Untersuchungen mit KM.....	28
16:	Boxplot des Volumens der Grauen Substanz bei nativer Untersuchung	29
17:	Bland-Altman Plot der Messungen der graue Substanz ohne KM.....	30
18:	Boxplot der Altersperzentilen des Gesamthirns der nativen Untersuchungen.....	31
19:	Altersperzentilen der grauen Substanz der nativen Untersuchungen	31
20:	Boxplot des Volumens der grauen Substanz bei KM-Gabe	32
21:	Bland-Altman Plot Volumen graue Substanz mit KM.....	32
22:	Boxplot der Altersperzentilen der grauen Substanz bei KM-Gabe	33
23:	Altersperzentilen der grauen Substanz der Untersuchungen mit KM	33
24:	Liniendiagramm der nativen Läsionsvolumen.....	34
25:	Liniendiagramm der Läsionsvolumen mit KM.....	35
26:	Liniendiagramm des infratentoriellen Läsionsvolumens aller Messungen.....	36
27:	Läsionsvolumen Gesamt.....	37

10 Tabellenverzeichnis

1:	Liste der meistgenutzten manuellen Methoden.	10
2:	Eigenschaften der verwendeten Sequenzen.	16
3:	Interpretation des Korrelationskoeffizienten.	22
4:	Patientenkollektiv	23
5:	Deskriptive Statistik des Ganzhirnvolumens ohne KM.	23
6:	Deskriptive Statistik der Altersperzentilen der nativen Ganzhirnmessungen	25
7:	Deskriptive Statistik des Ganzhirnvolumens mit KM	26
8:	Deskriptive Statistik der Altersperzentilen der Ganzhirnmessungen mit Kon- trastmittel	27
9:	Deskriptive Statistik, graue Substanz ohne KM.	29
10:	Deskriptive Statistik der Altersperzentilen der nativen Messung der grauen Substanz.	30
11:	Deskriptive Statistik, graue Substanz mit KM	32
12:	Deskriptive Statistik der Altersperzentilen der nativen Ganzhirnmessungen	33
13:	Läsionslast nativer Untersuchungen.	34
14:	Läsionslast der KM-verstärkten Untersuchungen.	35
15:	Läsionslast aller Untersuchungen zusammen mit p-Werten >0,05.	36

Danksagung

Bei meiner Studentenwerkstelle bin ich durch Prof. Dr. Notohamiprodjo auf das Thema meiner Doktorarbeit gestoßen. Meine Mutter als gelernte MTRA wies mich in die Magnetresonanztomografie ein und rasch bestand das technische und medizinische Interesse für das Thema.

Ich möchte mich an dieser Stelle herzlich bei meinem Doktorvater Prof. Dr. Mike Notohamiprodjo bedanken für die gute und freundschaftliche Betreuung. Er gab mir den roten Faden für meine Arbeit und ließ mir gleichzeitig viel Freiraum. Ich bin dankbar für die viele Zeit, die du, Mike, für mich aufgebracht hast und dass du mir die Chance gegeben hast, diesen Abschnitt in meinem Leben zu vollenden. Auch danke ich meinen Zweit- und Drittbetreuern für die Zeit und die Mühe meine Doktorarbeit zu begleiten.

Ein herzliches Dankeschön an DIE RADIOLOGIE, die mir die Magnetresonanztomografen und Speichermedien zur Erstellung der Bilder und Nutzung der Softwarelösungen zur Verfügung stellte. Danke auch an die Gründer der Softwarelösungen mediaire und icometrix, die mir die Versionen der genutzten Softwarelösungen zur Verfügung stellten.

Eine weitere Danksagung geht an meine wundervolle Familie, die mir immer in den Hoch- und Tiefphasen zur Seite stand, mich antrieb, wenn die Energie knapp war und meine noch so kleinen Erfolge als Anlass zum Feiern nahm. Ihr habt mir immer aufgezeigt, dass ich es schaffen kann, hier her zu gelangen. Ein besonderer Dank geht hierbei an meine Mama, du warst mein engster Wegbegleiter. Danke, dass du mir deine Schulter gabst und mir ein wahrer Freund warst auf diesem Weg.

Danke auch an Christoph, der mich ein Stück auf diesem Weg begleitet hat. Dank deiner Hilfe hat das Geschriebene seine Form bekommen. Danke für deine Freundschaft in dieser Zeit.

„Alle Hindernisse und Schwierigkeiten sind Stufen, auf denen wir in die Höhe steigen.“

Friedrich Wilhelm Nietzsche.

Das Schreiben der Arbeit hat mich sehr bereichert, mich aber auch an meine Grenzen gebracht. Meine Persönlichkeit wurde hierdurch ein Stück weit geformt und ich bin auf meinem Karriereweg weiter gekommen, wofür ich sehr dankbar bin. Ich blicke stolz auf diesen Abschnitt meines Lebens und freue mich auf die neuen Türen, die nun nach Abschluss der Doktorarbeit auf mich warten.

Publikationen

Notohamiprodjo, M; Lummel, Nina; Baum, T, Röttinger, M; Krause, L & Kleiner, I. *Vergleich von zwei Programmen zur automatisierten Quantifizierung des Hirn- und Läsionsvolumens bei Multiple-Sklerose-Patienten*. RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren, 2021. Doi: 10.1055/s-0041-1723201.



Eidesstattliche Versicherung

Krause, Lisa

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

Automatische Hirn- und Läsionsvolumetrie bei Multipler Sklerose: Vergleich von zwei kommerziellen Softwarelösungen

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Dachau, 19.10.2024

Ort, Datum

Lisa Krause

Unterschrift Lisa Krause



Erklärung zur Übereinstimmung der gebundenen Ausgabe der Dissertation mit der elektronischen Fassung

Krause, Lisa

Name, Vorname

Hiermit erkläre ich, dass die elektronische Version der eingereichten Dissertation mit dem Titel:

**Automatische Hirn- und Läsionsvolumetrie bei Multipler Sklerose: Vergleich von zwei
kommerziellen Softwarelösungen**

in Inhalt und Formatierung mit den gedruckten und gebundenen Exemplaren übereinstimmt.

Dachau, 30.12.2024

Ort, Datum

Lisa Krause

Unterschrift Lisa Krause