

SPINAL CORD VOLUME QUANTIFICATION AND CLINICAL APPLICATION IN MULTIPLE SCLEROSIS

Inaugural dissertation

to

be awarded the degree of

Dr. sc. med. in the Department of Clinical Research

presented at

the Faculty of Medicine
of the University of Basel

by

Charidimos Tsagkas

from Athens, Greece

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch

Basel, 2019

Approved by the Faculty of Medicine

On application of

Prof. Ludwig Kappos, University Hospital Basel,

Prof. Philippe Claude Cattin, University of Basel

PD Dr. med. Katrin Parmar, University Hospital Basel

PD Dr. M. P. Wattjes, Medical School Hannover, Germany

Basel, 26th August 2019

Prof. Dr. Primo Schär

Dean

To my friends and family

Contents

Acknowledgements.....	3
Summary.....	5
1. Introduction.....	7
1.1. Magnetic Resonance Imaging of the Spinal Cord: Current State and Challenges.....	7
1.2. Spinal Cord Volume Quantification.....	10
1.2.1. Image Segmentation.....	10
1.2.3. SC Segmentation	11
1.3. Multiple Sclerosis and Spinal Cord Involvement.....	13
1.3.1. Epidemiology and Risk Factors.....	13
1.3.2. Pathogenesis.....	13
1.3.3. Pathology.....	14
1.3.4. Spinal Cord Involvement.....	16
2. Gaps in research and aims of this PhD thesis.....	19
3. Manuscripts.....	22
3.1. Reliable and fast volumetry of the lumbar spinal cord using cord image analyser (Cordial)....	22
3.2. Automatic Spinal Cord Gray Matter Quantification: A Novel Approach	35
3.3. Spinal Cord Volume Loss: A Marker of Disease Progression in Multiple Sclerosis.....	57
3.4. Preferential Spinal Cord Volume Loss in Primary Progressive Multiple Sclerosis.....	81
4. Discussion and Outlook.....	99
4.1. Spinal Cord Volume Quantification.....	99
4.2. Spinal Cord Volume Loss in Multiple Sclerosis.....	101
4.3. Outlook and Future Research.....	103
4.3.1. Lumbar Spinal Cord Segmentation.....	103
4.3.2. Spinal Cord Grey Matter Segmentation.....	103

4.3.3. Future Volumetric Studies in Multiple Sclerosis.....	104
5. Contributions by the PhD student.....	105
6. Conclusion and closing remarks.....	106
7. References.....	107
8. Curriculum Vitae.....	121

Acknowledgements

First of all, I would like to thank Prof. Ludwig Kappos, head physician of the Department of Neurology at the University Hospital Basel and first supervisor of my PhD. studies. He gave me the opportunity to explore my academic skills and opened the door to clinical research within the field of multiple sclerosis, which has always been my dream already since my medical studies in Athens. Further I am very grateful for his words of encouragement and inspiration within these last 3 years, which pushed me to work even harder.

Next, I would like to thank Prof. Philippe Claude Cattin, also part of the PhD. committee as a Co-referee, for his feedback and the excellent advice and input not only in engineering questions. I would like to express my special appreciation and thanks to my supervisor PD Dr. med. Katrin Parmar, who demonstrated admirable endurance and burning scientific interest – despite personal life transitions - throughout the course of my PhD. She provided me with motivational as well as scientific and psychological support and led me safely through multiple challenges all these years. Finally, I would like to thank PD Dr. Mike P. Wattjes from MH Hannover, for joining the PhD. committee as an external expert. Special thanks go to Dr. Stefano Magon and Dr. Laura Gaetano for their help, guidance and teaching as well as our fruitful collaboration during the past years.

One of the most important things I have learnt during the past years is that clinical research requires teamwork. Therefore, I would like to warmly thank all our collaborating partners both from the Department of Biomedical Engineering, the Division of Radiological Physics of the Department of Radiology and Nuclear Medicine at the University of Basel and MIAC AG. In particular, I would like to thank Dr. Simon Pezold, whose contributions were essential for all studies and always provided us with fast and reliable solutions; the brilliant M.Sc. Antal Horváth for his work on the method development for the spinal cord grey matter segmentation and M.Sc. Simon Andermatt for his contribution on the same part; Dr. Michael Amann for his significant help and contribution in the method development and analysis in all of our studies; Prof. Dr. phil. Oliver Bieri Paravicini, head of the Division of Radiological Physics of the Department of Radiology and his postdoctoral researcher Dr. rer. nat. Matthias Weigel for the essential work in the development of new sequences enabling the visualization of the spinal cord grey matter; Tanja Haas for her exceptional talents with the MRI scanners and her always positive attitude, enthusiasm and help in our projects; Dr. Jens Wuerfel for his support, advice and ideas in our collaboration. I am also grateful to all coauthors in our studies –who have not been mentioned so far- for their productive feedback and ideas, namely Dr. Yvonne Naegelin, Dr. Ulrike Bonati, Prof. Dr. med. Dirk Fischer, Julia Reinhard and PD Dr. med. Arne Fischmann. Finally, I am truly very thankful to the brilliant friend and colleague M.Sc. Anna Altermatt not only for her valuable and essential work in two of our studies, but also for always being there for me at the nearby desk, ready to discuss and help with any scientific, professional or personal problems I encountered from the first day since I came to Basel. Also special thanks to M.Sc. Priska Zuber and M. Sc. Emanuel Geiter for their help and friendly advice.

Finally, I would like to thank a number of people not directly involved in this scientific work, who however supported me through small and big distances in this toilsome 3-year studies and contributed enormously to

my work–life balance. Hence, I would like to express my gratitude to Alexandros Polymeris, Anneza Panagiotou, Evgenia Dodoula, Giorgos Kanavakis, Tasos Petrou, my brother Giannis Tsagkas, Sotiris Manikas, Giannis Papoulakos, Youla Lymperaki, Stavroula Mendrinou, Eleni Papageorgiou, Fiona Nordemann, Ivan Bozin, Milan Prnjic, Vincent Haenen and many others. Last but not least, I would like to thank my parents for their immense support way before these PhD studies started.

Summary

Magnetic resonance imaging of the spinal cord is a valuable part of the diagnostic work-up in patients with multiple sclerosis and other neurological disorders. Currently, mainly signal intensity changes within the cord in MR-images are considered in the clinical management of disorders of the central nervous system. However, cross-sectional or longitudinal measurements of spinal cord volume may deliver additional valuable information. Hence, the overall goal of this doctoral thesis was twofold: i) to clinically validate methods for quantification of spinal cord volume and spinal cord compartments, which are suitable for longitudinal assessment of large patient cohorts and clinical practice and ii) to evaluate spinal cord volume as a potential valuable biomarker and provide new insights into the role of spinal cord damage in multiple sclerosis.

The first part focuses on the validation of quantification methods for spinal cord volume and includes two projects. While several MRI-based approaches of semi- and fully automatic techniques for volumetric spinal cord measurements have been proposed, up to now no gold standard exists and only a few methods have been validated and/or evaluated on patient follow-up scans to demonstrate their applicability in longitudinal settings. One of the latter segmentation methods was recently developed in-house and showed excellent reliability for cervical cord segmentation (Cordial, the cord image analyzer). In a first project, we extended its applicability to the lumbar cord, since no other software has been tested so far within this anatomical region of interest. On a well-selected dataset of 10 healthy controls (scanned in a scan-rescan fashion) we were able to show that - even within this technically challenging region - this segmentation algorithm provides excellent inter- and intra-session reproducibility showing high potential for application in longitudinal trials (Manuscript 1, published in *European Radiology*). In a second project, we aimed at obtaining volumetric information on particular compartments of the spinal cord such as the cord grey and white matter, since recent studies in multiple sclerosis provided evidence that measuring spinal cord grey matter volume changes may be a better biomarker for disease progression than quantifying cord white matter pathology or even volumetric brain measures. We therefore implemented a novel imaging approach, the averaged magnetization inversion recovery acquisitions sequence, for better grey and white matter visualization within the cord and scanned 24 healthy controls in a scan-rescan fashion. Further we applied an innovative fully automatic variational segmentation algorithm with a shape prior modified for 3D data with a slice similarity prior to segment the spinal cord grey and white matter. This pipeline allowed for highly accurate and reproducible grey and white matter segmentation within the cord. In view of its features, our automatic segmentation method seems promising for further application in both cross-sectional and longitudinal and in large multi-center studies (Manuscript 2, accepted *American Journal of Neuroradiology*).

The second goal of this thesis was the clinical application of the above-mentioned methods for the evaluation of spinal cord volume changes as a potential biomarker in multiple sclerosis patients. For this purpose, we quantified spinal cord volume change in a large cohort of 243 multiple sclerosis patients, followed over a period of 6 years with annual clinical and MRI examinations. Spinal cord volume proved to be a strong

predictor of physical disability and disease progression, indicating that it may be a suitable marker for monitoring disease activity and severity in all disease types but especially in progressive multiple sclerosis. Spinal cord volume also proved to be the only MRI metric to strongly explain the clinical progression over time as opposed to brain atrophy and lesion measures (Manuscript 3+4 published in Neurology and Multiple Sclerosis Journal respectively).

1. Introduction

1.1. Magnetic Resonance Imaging of the Spinal Cord: Current State and Challenges

Magnetic resonance imaging (MRI) is the method of choice in investigating disorders of the spinal cord (SC) in a fairly quick and non-invasive fashion. Based on the magnetic properties of the hydrogen atoms, abundant in the human body in fat and water, various MRI contrasts may be generated and deliver anatomical detail, information on structural composition and tissue function¹⁻³. For that purpose, a number of different MR-sequences and contrasts are used including relaxation time-weighted (T1, T2, T2*), proton-density weighted, magnetization prepared rapid gradient echo, fast-spin echo, phase-sensitive and short tau inversion recovery techniques⁴.

However, in contrast to brain MRI, the environment of the SC presents additional challenges for MRI methods^{1-3,5}. Greatest challenge is the inhomogeneous magnetic field across the SC due to the different magnetic properties of the surrounding tissues, e.g. cerebrospinal fluid (CSF), fat, vertebral bones, and air-tissue interfaces. This may lead to image distortions and a loss of spatial resolution. Further, the SC is a fairly thin and curved structure with a maximal antero-posterior diameter of $8.3 \pm 1.6\text{mm}$ at the C1 level and maximal latero-lateral diameter of $13.3 \pm 2.2\text{mm}$ at the C5 level⁶⁻¹³, which results in partial volume effects (a mix of tissues with different relaxation properties in one voxel) at the SC/CSF borders^{1-3,5}. Another challenge is the cord's physiological movement in the spinal canal as a result of cardiac-induced pulsatile CSF motion, respiratory motion, and swallowing resulting in MRI motion artifacts¹⁴⁻¹⁶. Contact of the SC with some point of the osseous canal, which partially eliminates contrast between the SC and its surroundings further hampers assessment. Furthermore, osteophytes of the spinal column can cause focal changes in CSF flow dynamics within the spinal canal causing so called CSF flow artifacts. Finally, Gibbs truncation artifacts are very common in SC MRI resulting in high signal in the center and dark edges of the SC.

Another limitation is the currently insufficient contrast between SC grey matter (GM), white matter (WM) and CSF on conventional SC MRI. Only recently, advanced MRI sequences were able to overcome those obstacles providing images with sufficient signal- and contrast-to-noise ratios between SC compartments in 3 Tesla MRI machines. First attempts utilized a 3D-fast field echo sequence (T2w)¹⁷, a 3D multiple echo data image combination sequence (T2*w)¹⁸, a 2D phase sensitive inversion recovery sequence (T1w)¹⁹ and most recently a modified look-locker inversion recovery sequence with single-shot balanced steady state free precession (T2w)²⁰ as well as an averaged magnetization inversion recovery acquisitions (AMIRA) sequence (T2w)²¹ (Figure 1). In particular, AMIRA acquires eight images of remarkable tissue contrast variation between SC GM and WM CSF in clinically feasible acquisition times. Averaging the first images enhances

the contrast-to-noise ratio between GM and WM, whereas averaging the last images considerably improves the WM/CSF contrast-to-noise ratio.

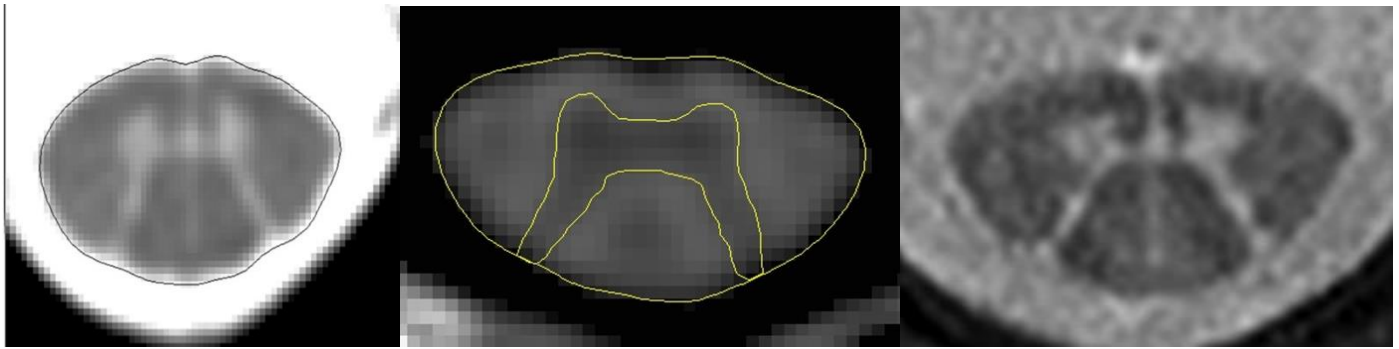


Figure 1. SC GM and WM contrast acquired with: fast field echo (left) ¹⁷, phase sensitive inversion recovery (middle) ¹⁹, averaged magnetization inversion recovery acquisitions (right) ²¹ sequences.

Despite anatomical and methodological challenges, SC MRI is essential in clinical routine and can be generally assessed in two ways: qualitatively or quantitatively. Qualitative SC MRI assessment involves neuroradiologists searching for MRI signal intensity changes within the SC in order to distinguish normal SC tissue from focal intramedullary pathology such as demyelination, edema, or inflammation e.g. in multiple sclerosis (MS) lesions (Figure 2). This is currently the sole use of SC MRI in clinical settings. However, some of the disorders affecting the SC do not present with focal abnormalities in the sense of hypo- or hyperintense lesions on MRI but are rather characterized by neurodegeneration of various aetiologies leading to neuronal loss and shrinkage, Wallerian degeneration and axonal loss (e.g. spinal muscular atrophy, amyotrophic lateral sclerosis). Others do manifest with MRI intensity changes indicating inflammatory and demyelinating lesions, which -however- do not represent the entire underlying SC pathology and often do not serve as reliable biomarkers (e.g. MS, human-T-cell lymphotropic virus type-1 (HTLV-1) associated myelopathy). Nevertheless, the aftermath of those neurodegenerative and demyelinating processes is tissue shrinkage and can be assessed in vivo on MRI as SC volume loss. Hence, cross-sectional or longitudinal quantitative measurements of SC volume and/or cross-sectional area indirectly deliver additional valuable information regarding mechanisms of neuropathology, that cannot be appreciated with the naked eye in qualitative assessments of SC MRI. Nonetheless, this is not part of clinical routine for the time being.

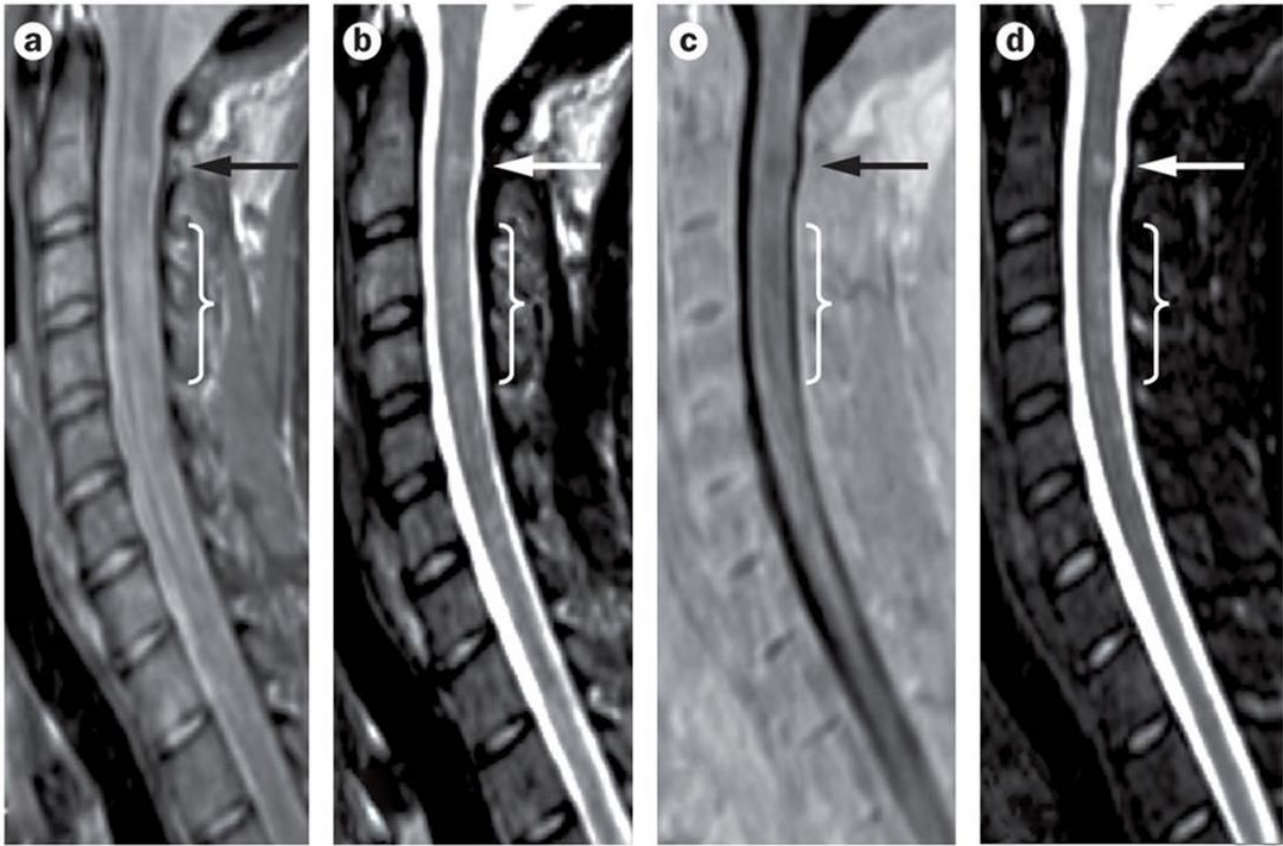


Figure 2. Multiple sclerosis SC lesions in sagittal images acquired with: a) Proton-density-weighted, b) T2-weighted, c) phase-sensitive inversion recovery and d) short tau inversion recovery imaging ⁵.

1.2. Spinal Cord Volume Quantification

1.2.1. Image Segmentation

Generally speaking, image segmentation is the process of dividing an image into different parts, aiming to define specific regions, whose boundaries separate image parts that display distinct features. This is especially interesting in the medical field in order to be able to quantify certain image properties for clinical or scientific goals. In general, this can be performed in two ways: manually or computer-based.

Manual image segmentations are performed by human observers, who may outline the objects or regions of interest intuitively or after completing some specific training. However, this approach is associated with 2 major limitations. The first one is that the task requires a considerable amount of time depending on the region of interest. This especially applies to medical imaging where the generation of swift diagnostic answers in a high number of patients with the least number of medical staff possible involved is necessary for logistic reasons. The second limitation is much more important, since it concerns the more systematic problem of reproducibility, which is essential in every diagnostic approach to be applied in scientific studies and/or clinical routine. It is very likely that two observers (a.k.a. raters) manually segmenting the same image will not produce exactly the same result, thus the inter-observer (or inter-rater) reproducibility is limited. Moreover, even the same observer will most probably create different segmentations should it be required to segment an image twice, thus the intra-observer (or intra-rater) reliability is limited. Adding other limitations related to medical images (such as MR-imaging) to the variability of measurements extracted out of manual segmentations can result in a considerable variation. This is important when taking into consideration that many image-extracted measurements are for example meant to be deployed for the purpose of patient monitoring in a longitudinal fashion, where small changes occur over time.

In the light of the above-mentioned limitations of manual image segmentations, computer-based image segmentation could release the medical staff from a highly time-consuming task and increase the reproducibility of those measurements. Computer-based image segmentations are, however, confronted with a new series of challenges. As a result, a plethora of computer-based image segmentation techniques have been proposed. In summary, those methods can be divided in 3 categories: a) basic or ad hoc ²², b) machine learning ²³ and c) energy-based techniques ^{24–26}. A complete review and analysis of those techniques is beyond the scope of this thesis.

1.2.2. SC Segmentation

Much like any other medical image, manual segmentation of SC MRI is a very time-consuming and rater-dependent task. Implementing image segmentation techniques to quantify the SC volume and cross-sectional areas have been introduced since 1996²⁷. Despite that, computer-based SC segmentation remains demanding in part due to limitations hampering SC MR-imaging (as mentioned in 1.1). A big number of semi- and fully automatic techniques have been proposed²⁸. The most important include active contours of surfaces^{29–32}, level sets³³, partial volume modeling³⁴, gradient vector flow³⁵, atlas-based approaches^{36,37}, and tubular deformable models³⁸ with variable required user interaction from manual identification of the SC centerline³¹ to the identification of multiple³⁰ or single^{29,37} anatomical landmarks, with completely automated approaches presented only recently (Figure 3)³⁹. In the past, SC atrophy was usually determined by assessing the cross-sectional area of the cervical cord, usually at the C2/C3 level, which has been shown to correlate with clinical measures, although reproducibility was limited and depending on data quality as well as repositioning^{27,40,41}. However, until now only a few of those methods have been validated and/or evaluated on patient follow-up data to demonstrate the applicability in longitudinal trial settings with up to two years follow up time^{27,31,39,42–45}. Most recently, a new method (*Cordial*, the cord image analyzer) was developed with longitudinal studies in mind, which provides measurements of SC sections with fixed length and location, while it relies on natural landmarks on the SC itself. Thus it allows for a reliable measurement of localized SC volume and its comparison in the same subject over time⁴⁶ (Figure 4).

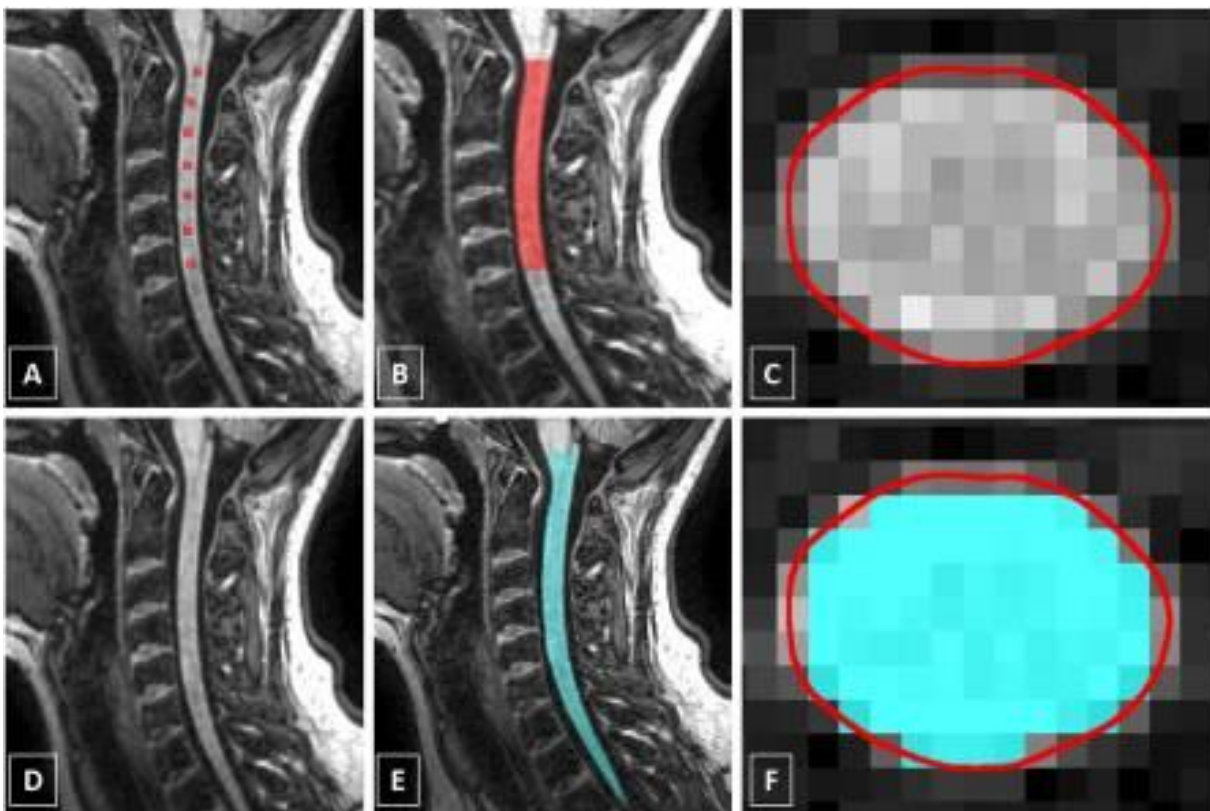


Figure 3. A-C shows a SC segmentation performed with an active surface method, while D-F demonstrates a fully automatic SC segmentation.³⁹

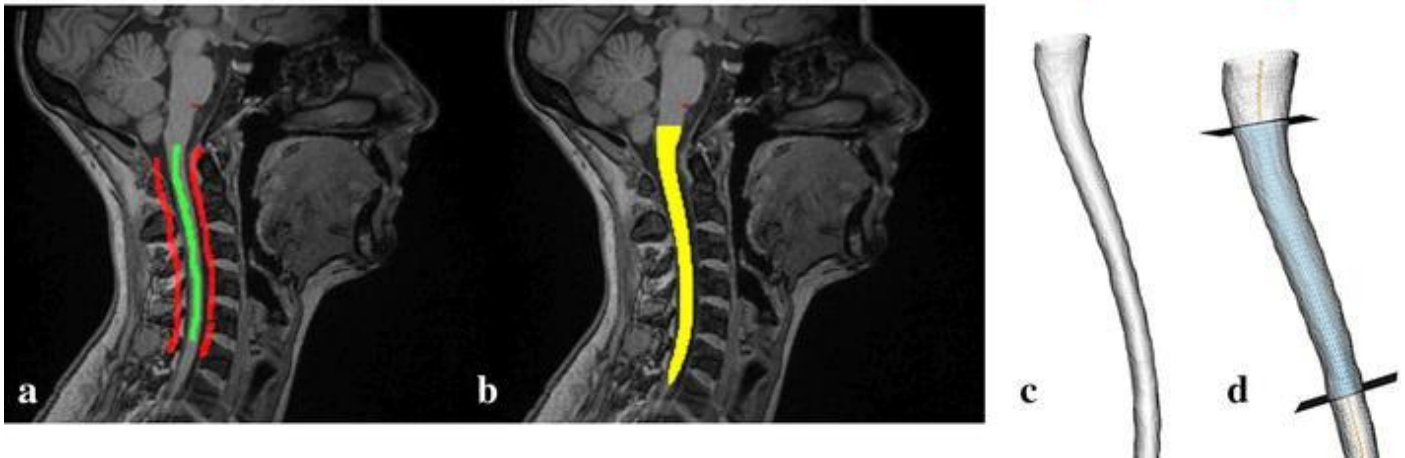


Figure 4. SC segmentation with Cordial, the cord image analyzer. a) In a first step, a subset of SC voxels (green) and background voxels (e.g., cerebrospinal fluid, vertebral bodies; red) are selected manually. In addition, an anatomical landmark—the medullopontine sulcus—is defined manually (red dot). b) Presegmentation result of the SC (yellow). c) SC surface reconstruction. d) Definition of the cutting planes for SC volume calculation ⁴⁶.

Most of the above mentioned cord volume quantification approaches focused on cervical and mid thoracic levels of the SC ^{27–34,36–39,42–46}. In comparison to the cervical and thoracic SC, the lumbar SC poses additional challenges because of its location in an environment with a different bone, soft tissue and air composition resulting in image distortion and lower signal intensity due to magnetic field inhomogeneities. In addition, the lumbar SC is surrounded by multiple spinal roots exiting the spinal canal, which can lead to reduced SC/CSF contrast due to partial volume effects and misclassifications of the SC contour.

Moreover, accurate SC GM segmentation also remains a challenge. While all the above-mentioned segmentation algorithms focus on the whole cord, only recently, Yiannakas et al. (2012) demonstrated the feasibility to distinguish between the WM and GM by performing manual segmentation ¹⁷. As a result of the improvement in image quality and post-processing techniques, the first fully automatic SC GM segmentation methods were also established in the past few years ^{47–50}. A gold standard pipeline for accurate and reproducible SC GM measurements is, however, still a matter of ongoing discussion.

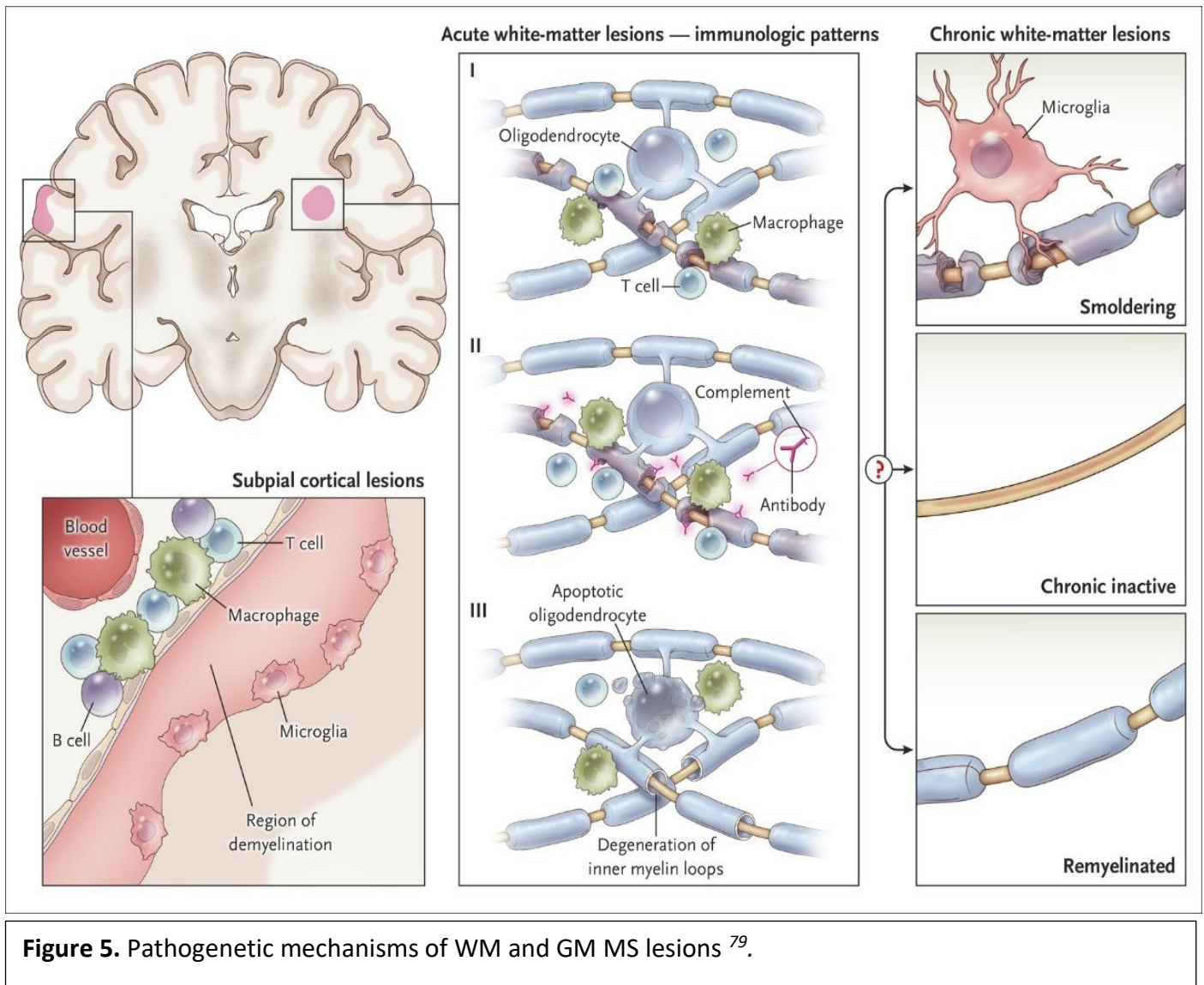
1.3. Multiple Sclerosis and Spinal Cord Involvement

1.3.1. Epidemiology and risk factors

MS is a chronic inflammatory and neurodegenerative disease of the optic nerves, SC, and brain, affecting more than 2-3 million people worldwide⁵¹⁻⁵⁴. It predominantly affects young adults, and has a huge impact functionally, on quality of life, and financially whereas healthcare costs are considerable and rise with increasing disability⁵¹⁻⁵⁵. Age of onset is usually between 20 and 50 years and women are more often affected than men, while the female to male sex ratio seems to be increasing in the last 50 years and is currently estimated to be around 3.2:1⁵⁶. Despite the fact that a specific etiologic trigger has not been identified, various genetic and environmental risk factors, e.g. female sex, carriers of a HLA DRB1*1501 haplotype, viral infections, low vitamin D levels, smoking, and others, have been shown to increase the risk of acquiring the disease⁵⁷. Once multiple sclerosis diagnosis is confirmed, older age, male sex, and higher disability at baseline are associated with a worse prognosis⁵⁸.

1.3.2. Pathogenesis

MS is generally accepted to be an autoimmune demyelinating disease, although the reason why immune responses against CNS antigens are initiated and maintained is currently unclear. Furthermore, there is also an ongoing debate about whether the root cause of MS is intrinsic or extrinsic to the CNS. However, both the adaptive and the innate immune system have been described in numerous studies to participate in tissue damage in MS (Figure 5). Helper (CD4+) and cytotoxic (CD8+) T cells⁵⁹, B-lineage cells and respective antibody production are involved⁶⁰. However, no specific autoantigens interacting with T and B cells have been so far confirmed^{61,62}. On the other side, blood-borne macrophages⁶³, microglial cells are (although unclear if their role is pathogenic, protective or both)^{54,64,65}, astrocytes⁶⁶ and blood-brain barrier leakage are part of the pathogenic processes in MS⁶⁷.



1.3.3. Pathology

The hallmarks of MS pathology are inflammation, immune cell infiltration, axonal or neuronal loss, demyelination, and astrocytic gliosis ^{52,54}. MS patients demonstrate characteristic lesions, which are disseminated in multiple regions of the CNS, including brain WM and GM, brain stem, SC, and optic nerve. Besides the latter, temporal distribution represents the second cornerstone of the diagnosis of the disease ⁶⁸ (Figure 6).

Evidence from MRI and pathological assessment (biopsies and autopsies) indicates that early “active” WM lesions demonstrate different histopathological patterns of demyelination (described as type I-III) and evolve over the course of months ⁶⁰. However, studies suggest that a single immune-mediated mechanism and thus a specific “lesion type” dominates in each MS patient ⁶⁹. Acute MS plaques show activation of astrocytes and microglia and sometimes caspase-independent oligodendrocyte apoptosis ⁶⁰. After the initial phase, the inflammation of a given lesion may resolve or persist leading to a so-called “smoldering lesion”.

Demyelinating lesions do not only occur in the WM but can be also located at the WM/GM junction (also known as juxtacortical lesions) or even arise entirely in the GM. Brain cortical lesions are shown to be less inflammatory than their WM counterparts and have substantially less permeability of the blood–brain barrier⁷⁰. Approximately half of all brain cortical lesions are perivascular, whereas the rest appear adjacent to the pial surface of the brain^{71–73}. In the early stages these lesions are active and correspond to leptomeningeal inflammatory aggregates which most likely become inactive during the course of the disease⁷³. Leptomeningeal inflammation has been shown to also organise into self-sustaining lymphoid-like follicles in secondary progressive MS⁷⁴.

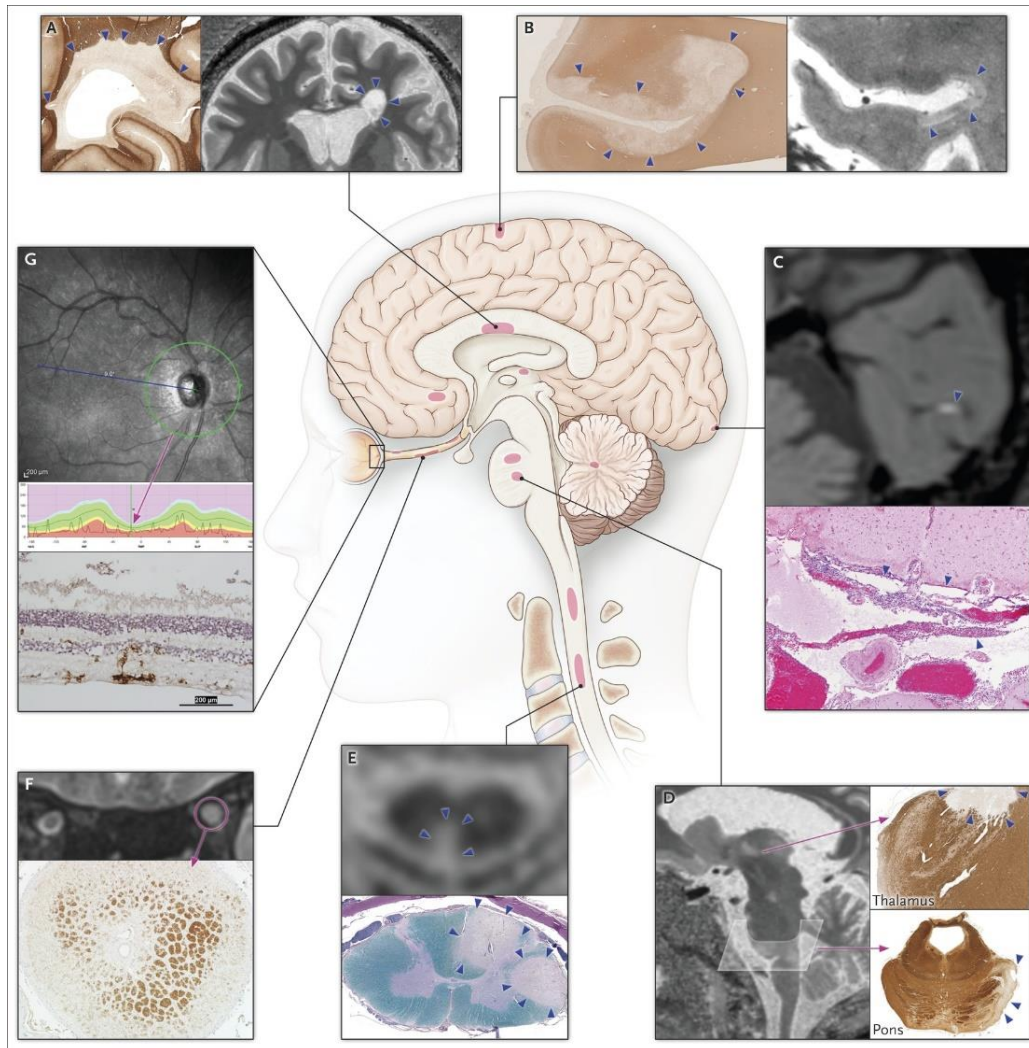


Figure 6. Imaging and pathological examples of MS lesion location in A) the periventricular white matter lesion, B) the subpial cortex, C) the leptomeninges, D) thalamus and pons, E) the spinal cord, F) the optic nerve, G) the retina⁷⁹.

Finally, axonal transections especially in acute MS lesions as well as slow neuronal and axonal degeneration in chronically demyelinated lesions lead to irreversible neuronal injury and permanent clinical disability^{75,76}. Chronic neurodegeneration is thought to be a result of impaired axonal transport, mitochondrial dysfunction, the loss of myelin trophic support -leading to progressive swelling and cytoskeletal disorganization- and increased energy demands related to the upregulation of ion channels⁷⁶.

1.3.4. Spinal Cord Involvement

SC abnormalities have been observed in up to 83% of MS patients, with 60% of them occurring in the cervical region^{77,78}. Patients with progressive MS often demonstrate confluent SC lesions, which appear as diffuse MRI abnormalities^{77,78}. SC lesions mostly arise in the SC circumference, affecting predominantly the dorsal WM tracts and the lateral corticospinal tracts and may not respect the SC WM/GM borders^{5,77,78} (Figure 8 and 9). SC lesions have a central diagnostic role and the presence of a single SC lesion is predictive of conversion to clinically definite MS^{79,80}. SC involvement in CIS is also associated with an increased risk of future disability⁸⁰.

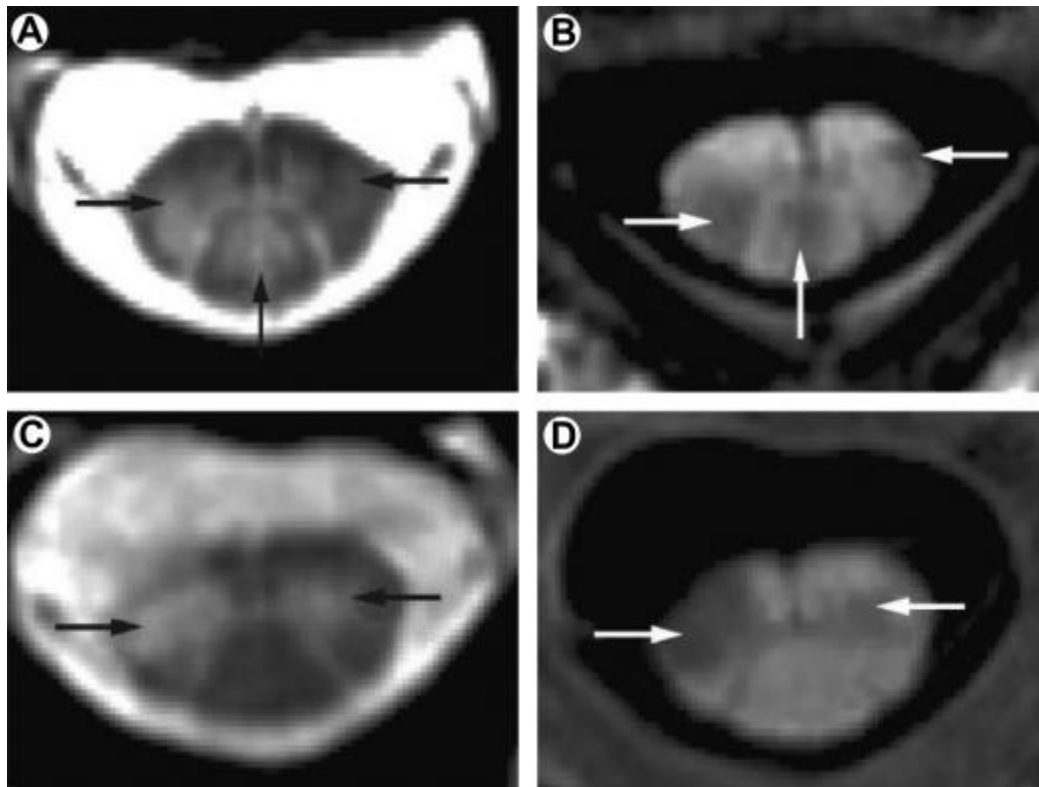
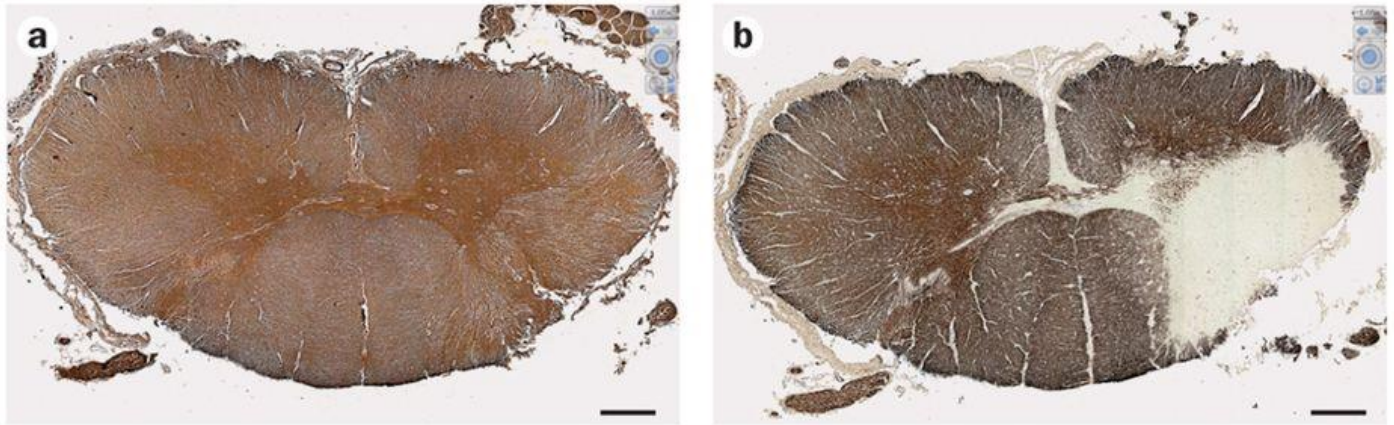


Figure 7. Axial sections of cervical spinal cord (SC) lesions in multiple sclerosis. (A) Hyperintense signal abnormalities on axial fast field echo (FFE) sequences demonstrate lesions in the lateral and posterior columns. (B) Hypointense signal abnormalities seen on phase sensitive inversion recovery (PSIR) sequence imaging indicate lesions in the lateral and posterior columns. (C, D) Lateral column lesions extending from the SC white matter into the gray matter are seen on (C) FFE sequence and (D) PSIR sequence⁵.

MS is, however, characterized by diffuse damage of the SC. In post mortem studies, axonal density in MS was found to be lower by around 60% across all levels and affect all fibers regardless of diameter, whereas loss of corticospinal tract fibers may be disproportional to the rest of the SC⁸¹. Demyelination seems to affect 24–48% of the GM and 11–13% of the WM, with no significant differences across levels. Finally, a significant association was detected between focal demyelination and decreased axonal density.



Nature Reviews | Neurology

Figure 8. *Transverse, paraffin-embedded sections (thickness: 10 μ m) sampled from the same level (the sixth cervical nerve) of the spinal cord from a patient with secondary progressive multiple sclerosis. a | SMI-31 (a phosphorylated neurofilament) immunostaining shows the anatomical organization of the grey matter (the darker core with a butterfly-like shape) and white matter in the spinal cord. b | Myelin basic protein immunostaining shows a large lesion of both grey and white matter. Scale bar = 1 mm⁵.*

Despite the abundance of focal demyelinating lesions in MS patients, this disorder is generally characterized by a weak correlation between lesion load and clinical disability also known as clinical-radiological paradox^{40,82}. This paradox is also seen for SC lesional burden. One explanation for that are the repair and compensatory mechanisms known to arise following MS-associated CNS injury such as the redistribution and increase of the axonal sodium channels^{83,84}, remyelination⁸⁵, recovery of function due to secondary to cortical plasticity⁸⁶ and long-term potentiation of synaptic transmission⁸⁷. Another explanation is that -instead of focal inflammatory lesions- a diffuse neurodegenerative mechanism is responsible -for the greater part- for the patient's neurologic deficits.

Given the abovementioned limitations using SC lesion metrics as biomarker in MS, alternative methods that have the ability to quantify MS-related SC injury were sought and the main interest was focused on SC atrophy. SC atrophy is thought to be the aftermath of neuronal loss and shrinkage, axonal loss and demyelination^{81,88,89}, which seems at best weakly related to focal brain and SC MS lesions^{88,90}. Therefore, in vivo measures of SC atrophy assessed on MRI, e.g. SC volume and SC cross-sectional area, reflect a diffuse process. Fortunately, those metrics consistently demonstrate strong correlations with clinical disability^{44,45,91–93}.

SC atrophy mainly SC volume loss have been shown to occur across all stages of MS (although there are conflicting data regarding SC atrophy in clinical isolated syndrome^{44,45,91–93}) and has been shown to be more extensive in progressive forms of the disease in cross-sectional studies^{44,94,95}. Former studies demonstrate controversial results concerning differences in short-term longitudinal SC atrophy assessments between MS subgroups. While RRMS display clusters of atrophy primarily localized to the posterior SC, whereas progressive MS patients seemed to have more generalized volume loss confined in the posterior and lateral columns⁹⁶.

SC GM cross-sectional area changes were also investigated in two recent cross-sectional studies. Schlaeger and colleagues observed evidence of GM atrophy of the cervical and thoracic SC compared to healthy controls (Figure 10)^{95,97}. These studies also demonstrated marked GM atrophy in progressive MS subtypes compared to RRMS^{95,97}. More importantly, SC GM atrophy was found to be the most significant correlate of clinical disability over brain metrics such as GM and WM^{95,97}.

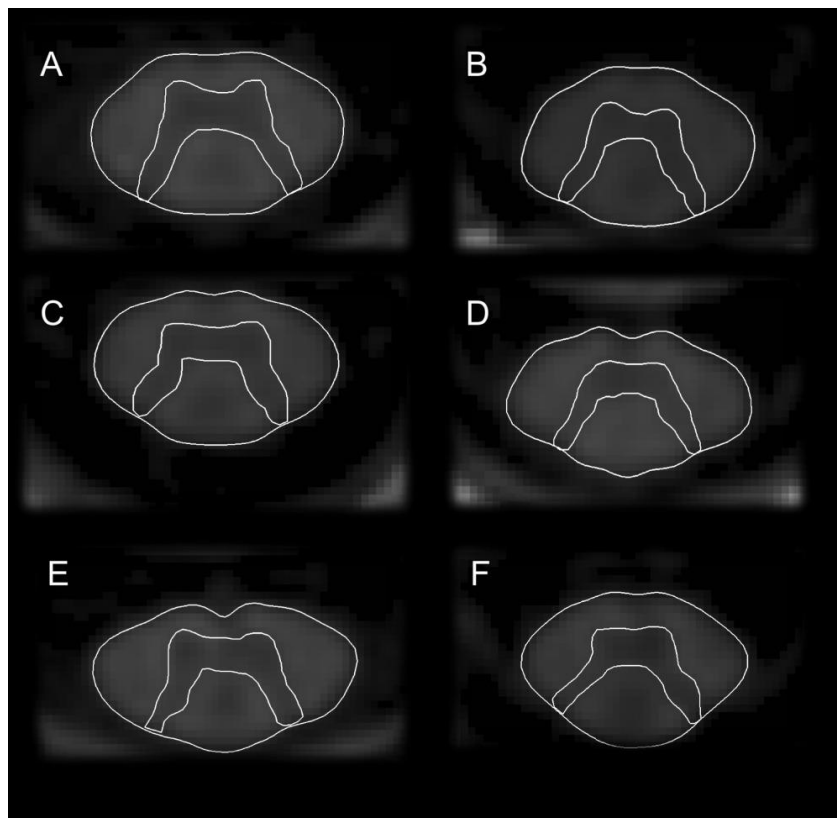


Figure 9. A–F) Axial 2-dimensional phase-sensitive inversion images at the C2/C3 disk level of multiple sclerosis patients⁹⁵.

Despite the interest raised by cross-sectional and short-term follow-up volumetric SC studies, there is still lack in larger scale longer-term longitudinal studies on SC volume loss in MS. This is partly due to the technical difficulties discussed before as well as time and financial restrictions for an additional MRI of the SC. As a result, there is a knowledge gap on the dynamic changes of SC measures and their association with the patient’s clinical picture over time.

2. Gaps in research and aims of this PhD thesis

The overall aims of this thesis were to: i) develop and clinically validate methods for volume quantification of the SC and its compartments (GM/WM) suitable for longitudinal assessment and ii) evaluate spinal cord volume as a potential valuable biomarker and provide new insights into the contribution of SC pathology to the evolution multiple sclerosis.

Manuscript 1:

As a first step, we aimed to expand the main field of application of our new in-house developed quantification software Cordial (Cord image analyzer) for quantification of the lumbar cord. The lumbar cord is of great importance for the sensorimotor function of the legs, sexual function, and bladder and bowel control. Moreover, volumetric measurements of the lumbar cord may depend on the integrity of all of those structures and serve as a potential biomarker e.g. for inflammatory or neurodegenerative disorders (e.g. HTLV-1-associated myelopathy, MS, neuromyelitis optica, amyotrophic lateral sclerosis, etc.). Such a segmentation method for lumbar SC volume would be a major step forward in a number of diseases with affection of the lumbar SC in terms of disease monitoring, and development and evaluation of new treatment options.

In comparison to the cervical and thoracic SC, the lumbar SC poses additional challenges (e.g. environment of different bone, soft tissue and air composition leading to magnetic field inhomogeneities; multiple spinal roots in close vicinity prone to partial volume effects) and therefore has not been in the focus of method development/validation yet. Most recently, the quantification software Cordial was presented with longitudinal studies in mind, which provides measurements of SC sections with fixed length and location, while it relies on natural landmarks on the SC itself. Thus, it allows for a reliable measurement of localized SC volume and its comparison in the same subject over time. Although this method was validated for the cervical SC, its features allow a potential application in other SC segments.

Hence, “Manuscript 1” of this thesis had the aim to verify the accuracy and reproducibility of Cordial for quantification of lumbar SC volume. Our objective was to evaluate the performance of this SC volume quantification algorithm in MR-images of healthy controls scanned repeatedly and then post-processed by multiple “raters”.

Manuscript 2:

In the past, quantification of the spinal cord predominantly focused on the structural changes as a whole. However, histopathological studies show changes of SC GM independently of the surrounding WM and vice versa in a large spectrum of neurological conditions. The open question is to what extent each individual

compartment is affected and how their respective longitudinal course looks like in vivo. To answer this, not only advanced MR imaging methods, but also an automatic segmentation method providing accurate and reproducible measurements of each individual SC compartment is needed.

Fortunately, a novel MRI technique, the Averaged Magnetization Inversion Recovery Acquisitions (AMIRA) sequence was developed ²¹. This technique overcomes the major requirements for reliable quantification of SC tissue changes over time. It shows remarkable tissue contrast between GM and WM and has a high in-plane resolution, which is necessary for a structure with such a small diameter. Despite the latter it is fast with low sensitivity to motion.

Next to the limitations for of high-quality imaging methods of the SC, accurate SC GM/WM segmentation also remains a challenge. The above-mentioned in-plane resolution of around 0.5 mm is barely enough to visualize the SC's butterfly-shaped GM structure. Taking into account, that mean SC volume changes in disorders such as multiple sclerosis range between 0.5-2.2% per year, it becomes apparent that new accurate and reproducible methods for the development of such a biomarker are a difficult task. Hence, a gold standard pipeline for accurate and reproducible SC GM measurements is still a matter of ongoing discussion.

“Manuscript 2” as part of this thesis had the objective to provide a reliable pipeline for the quantification of SC GM and WM. To do so, this work analyzed the accuracy and reproducibility of the quantification of SC GM and WM with a fully automatic post-processing approach in SC images acquired with a recently proposed imaging protocol (including the AMIRA sequence). Our objective was to test the performance of this SC GM quantification algorithm in “scan-rescan” MR-images of healthy controls. The SC images were segmented manually multiple times by multiple “raters” and then compared with the automatic method.

Manuscript 3 & 4:

In MS, SC atrophy is common and is thought to be a hallmark of neurodegeneration. It represents a number of pathologic processes of the SC such as neuronal loss and shrinkage, axonal loss and demyelination. Despite the fact that reproducible measurements of the spinal cord cross-sectional area have been possible for more than 20 years and the interest raised so far by cross-sectional and short-term follow-up volumetric SC studies, there is a lack of larger scale longer-term longitudinal studies on SC volume loss in this disorder. This is partly due to the technical difficulties discussed before as well as time and financial restrictions for an additional MRI of the SC in a routine setting. As a result, there is a knowledge gap on the dynamic changes of SC measures, their association with the patient's clinical symptoms over time and the predictive value of this potential biomarker. Studies so far have shown controversial result to whether MS disease subtypes differ in terms of SC atrophy.

Hence, “Manuscript 3 & 4” investigated the cervical SCV for atrophic changes over 6 years in an existing large cohort of overall 243 MS patients. In particular, the first study (Manuscript 3) analyzed relapse-onset

MS patients only, while the second study (Manuscript 4) focused mainly on the PPMS group. One objective of those studies was to demonstrate SCV loss occurring over time in these patients and if evident, focus on between-group differences. Moreover, the relationship between SCV loss and clinical outcomes as well as the relative contribution of SCV loss together with brain atrophy and lesion measurements to the deterioration of physical disability were examined. In a last step, these studies verified the hypothesis that SC atrophy may be able to predict clinical outcomes later on.

3. Manuscripts

3.1. Reliable and fast volumetry of the lumbar spinal cord using cord image analyser (Cordial)

Charidimos Tsagkas, MD¹⁺²; Anna Altermatt, M.Sc.²⁺³; Ulrike Bonati, MD⁴; Simon Pezold, Ph.D.³; Julia Reinhard, Ph.D.⁵; Michael Amann, Ph.D.¹⁺²⁺⁵; Philippe Cattin, Ph.D.³; Jens Wuerfel, MD²⁺³; Dirk Fischer, MD⁴; Katrin Parmar,* MD¹; Arne Fischmann,* MD⁵⁺⁶

*equally contributing last authors

1. Department of Neurology, University of Basel Hospital, Petersgraben 4, CH-4031 Basel
2. Medical Image Analysis Center (MIAC AG), Basel, Mittlere Strasse 83, CH - 4031 Basel
3. Center for medical Image Analysis & Navigation (CIAN), Department of Biomedical Engineering, University of Basel, Gewerbestrasse 14 ,CH-4123 Allschwil
4. Division of Neuropediatrics, University of Basel Children's Hospital, Spitalstrasse 33, CH-4056 Basel
5. Division of Diagnostic and Interventional Neuroradiology, Department of Radiology, University of Basel Hospital, Petersgraben 4, CH-4031 Basel
6. Division of Neuroradiology, Hirslanden Klinik St. Anna, St. Anna-Strasse 32, CH-6006 Luzern

Publication: European Radiology. April 2018. DOI: 10.1007/s00330-018-5431-1

Abstract

Objective: To validate the precision and accuracy of the semi-automated cord image analyzer (*Cordial*) for lumbar spinal cord (SC) volumetry in 3D T1w MRI data of healthy controls (HC).

Materials and Methods: 40 3D T1w images of 10 HC (w/m: 6/4; age range: 18-41y) were acquired at one 3T-scanner in two MRI sessions (time interval 14.9 ± 6.1 d). Each subject was scanned twice per session, allowing determination of test-retest reliability both in back-to-back (intra-session) and scan-rescan images (inter-session). *Cordial* was applied for lumbar cord segmentation twice per image by two raters allowing for assessment of intra- and inter-rater reliability and compared to a manual gold standard.

Results: While manually segmented volumes were larger (mean: 2028 ± 245 mm³ vs *Cordial*: 1636 ± 300 mm³, $p<0.001$), accuracy assessments between manually and semi-automatically segmented images showed a mean Dice-coefficient of 0.88 ± 0.05 . Calculation of within-subject coefficients of variation (COV) demonstrated high intra-session (1.22-1.86%), inter-session (1.26-1.84%), as well as intra-rater (1.73-1.83%) reproducibility. No significant difference was shown between intra- and inter-session reproducibility as well as between intra-rater reliabilities. Although inter-rater reproducibility (COV: 2.87%) was slightly lower compared to all other reproducibility measures, between rater consistency was very strong (intraclass correlation coefficient: 0.974).

Conclusion: While under-estimating the lumbar SCV, *Cordial* yet provides excellent inter- and intra-session reproducibility showing high potential for application in longitudinal trials.

Keywords: spinal cord; volumetry; semi-automated segmentation; magnetic resonance imaging; imaging biomarker

Key points:

- Lumbar spinal cord segmentation using the semi-automated cord image analyzer (*Cordial*) is feasible.
- Lumbar spinal cord was defined as a 40mm spinal cord segment located 60 mm above the conus medullaris.
- *Cordial* provides excellent inter- and intra-session reproducibility in the region of the lumbar spinal cord.
- *Cordial* shows high potential for application in longitudinal trials.

Abbreviations and acronyms: central nervous system (CNS); coefficient of variation (COV); cerebrospinal fluid (CSF); field-of-view (FOV); healthy controls (HC); intra-class correlation coefficient (ICC); magnetization-prepared rapid gradient-echo (MPRAGE); spinal cord (SC); spinal cord volume (SCV); standard deviation (SD); echo time (TE); repetition time (TR); volumetric interpolated breath-hold examination (VIBE)

Introduction

The spinal cord (SC) is affected by a number of inflammatory, hereditary and degenerative diseases of the central nervous system (CNS). Pathological changes in this region may cause significant disabling deficits including sensorimotor as well as bowel and bladder dysfunction. MR imaging of the SC plays a crucial role in the diagnostic process, depicting *in vivo* pathological processes within and surrounding the SC. Atrophy of the cord reflects loss or damage of myelin, axons and neurons and has already been proven to be a valuable biomarker correlating well with physical disability⁵.

Despite fast technical developments within the past decades, the anatomical properties of the SC (a relatively thin, long and mobile structure) still hamper acquisition of high-quality MR images and reliable quantification of SC metrics^{5,98–100}. Manual segmentation of the cord is not only time-consuming, but also rater-dependent, and therefore both, reproducibility and accuracy are compromised. Several semi- and fully-automated segmentation techniques have been deployed to overcome these limitations^{26,29–38,46,101–103}. However, most of these approaches focused on cervical and mid thoracic levels of the SC^{26,29–31,33,34,36,37,46,101–104}. Measurements of the lumbar cord volume are still limited due to its anatomical structure and surroundings in the human body rendering it susceptible to image artifacts as well as relatively low tissue contrast³. In comparison to the cervical and thoracic SC, the lumbar SC poses additional challenges. It is located in an environment with a different bone, soft tissue and air composition resulting in image distortion and lower signal intensity due to magnetic field inhomogeneities. In addition, the lumbar SC is surrounded by multiple spinal roots exiting the spinal canal, which can lead to reduced SC-cerebrospinal fluid (CSF) contrast due to partial volume effects and misclassifications of the SC contour.

Nevertheless, the lumbar cord is of great importance for the sensorimotor function of the legs, sexual function, and bladder and bowel control. It contains not only both afferent and efferent fibers to the lower extremities, bladder and bowel, but also nuclei for those pathways (anterior and posterior horns, sacral sympathetic nuclei, Onuf's nuclei, etc.). Volumetric measurements of the lumbar cord may depend on the integrity of all of those structures and serve as a potential biomarker for disorders such as spinal cord injury, spinal muscular atrophy, HTLV-1-associated myelopathy, tumors, multiple sclerosis, neuromyelitis optica, amyotrophic lateral sclerosis, etc.)^{2,3}.

In this study, we validated the precision of the semi-automated cord segmentation tool *Cordial* (cord image analyzer) in terms of its reliability and reproducibility in segmenting the lumbar SC of healthy controls (HC) in one 3T-scanner^{26,46,101}. This approach is able to provide SC volume (SCV) measurements of fixed length and location in relation to the conus medullaris. We hypothesize that *Cordial* could deliver reliable lumbar cord volume quantifications, providing efficient and reliable analyses for studies with longitudinal design.

Materials and methods

Subjects & MRI acquisition

10 HC (6 women; mean age \pm standard deviation (SD): 28.7 ± 7.3 ; age range 18-41 years) were scanned on one 3 Tesla whole-body MR scanner (Prisma, Siemens Medical, Erlangen, Germany) with a T1w fat-suppressed volumetric interpolated breath-hold examination (VIBE) sequence. The sequence parameters were: repetition time (TR) of 7 ms, echo time (TE) of 2.46 ms, Matrix 256×256 and field-of-view (FOV) of $256 \times 256 \text{ mm}^2$, 160 transverse slices with an isotropic resolution of $1 \times 1 \times 1 \text{ mm}^3$, covering 9.8 cm of the lumbar spine, located between the 11th thoracic vertebral body and 3rd lumbar vertebral body, acquisition time 5:16 min. Sagittal images were reconstructed from the axial VIBE images. To ascertain reproducibility the sequence was centered on the superior endplate of the most caudal vertebral segment containing the SC.

Each subject underwent two MRI sessions. In each MRI session subjects were scanned twice (back-to-back) with repositioning between scans. The second session took place with a mean time interval of 14.9 ± 6.1 days, allowing for scan-rescan tests. All subjects gave written consent. The experimental procedures conformed to the Declaration of Helsinki and the study protocol was approved by the local ethics committee.

Lumbar SC segmentation

Segmentations were performed in a manual (as a gold standard) and semi-automated fashion.

The manual segmentations of the spinal cord were performed on the first scan of the first MR-imaging session of each HC by two expert raters (C.T.; A.A.) using ITK-Snap (Version 3.6.0). The latter were visually inspected and, if necessary, adjusted by two experienced radiologists (J.W.; A.F.). A consensus reading between all four raters (C.T.; A.A.; J.W.; A.F.) resulted in a final single manual spinal cord segmentation per scan, which was then deployed as the gold standard and used for the comparison with the automated segmentations (described below).

In a second step all scans were segmented using the semi-automated tool *Cordial*, described in detail by Amann et al. ^{46,101,102}. Briefly, the method comprises three separate steps: pre-segmentation, segmentation refinement, and volumetric measurement. While the pre-segmentation requires user interaction of about 2–5 min. per scan, the subsequent steps are fully automated. During pre-segmentation an anatomical marker was set at the tip of the conus medullaris, defined as the most distal point on the lowest axial image in which spinal cord tissue was identifiable with confirmation on sagittal images (on average between level L1 and L2). The segmentation of the caudal part of the lumbar cord including the conus medullaris (region below lower blue border on Figure 18) was limited by artifacts (low contrast between the SC and the surrounding structures such as spinal nerves, spinal nerve roots, vertebrae etc.). These parts were therefore excluded from the calculation of the volume. The segmentation of the upper part was restricted from the FOV-limits. In the light

of the aforesaid, the measured volume (hereinafter referred to as SCV) was located 60 mm above the previously fixed marker at the conus medullaris and comprised a segment of 40mm along the SC center line (for details see Figure 10). After completion of the pre-segmentation and the segmentation refinement the MR-images were visually inspected for quality. All images were segmented twice by two experienced raters (C.T; A.A.), with an interval of at least 7 days between the two evaluations (runs).

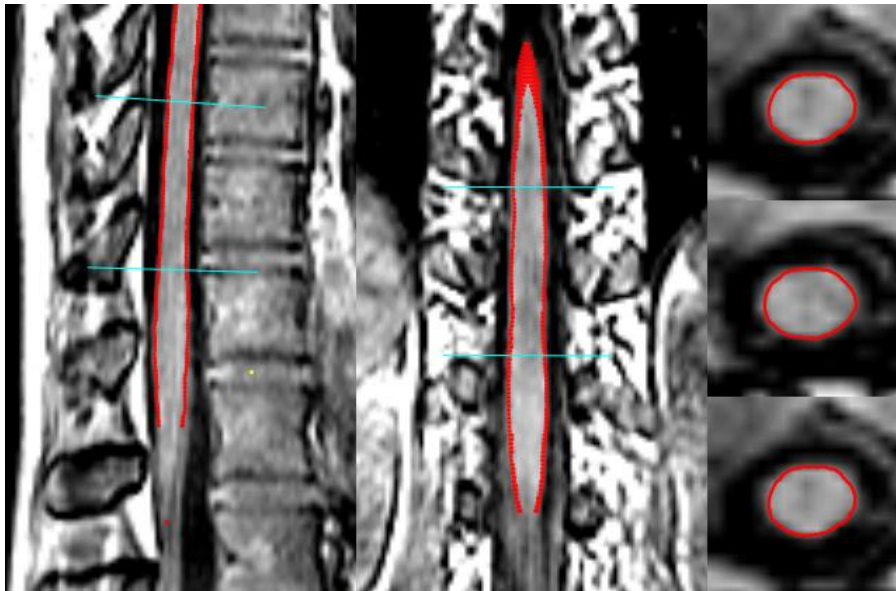


Figure 10. Final lumbar spinal cord (SC) segmentation (red lines) with Cordial of a representative subject. Blue lines mark the region of maximum feasible SC volume over all subjects (starting 60mm above the manual marked tip of the conus medullaris (red dot) over a defined cord centre line distance of 40mm). Low image and segmentation quality restrict measurement beyond the blue lines in the majority of scans.

Statistical Analysis

All analyses were performed using R Version 3.2.3 (<http://r-project.org/>).

The reproducibility of SCV in terms of intra-session and inter-session, as well as intra- and inter-rater reliability of *Cordial* was assessed using the coefficient of variation (COV) independently for each subject. 1) Paired back-to-back acquisitions allowed for computation of intra-session variability (reflecting noise and segmentation errors), whereas 2) paired scan-rescan acquisitions were used for computation of inter-session variability (depicting biological changes (e.g. hydration status) and/or differences due to repositioning of the subject). 3) Paired measurements of the two runs by the same rater were used for intra-rater reproducibility, which corresponds to variability depending on the segmentation method and image quality. 4) Paired measurements of single images segmented by different raters were deployed for inter-rater reliability. The latter corresponds to variability depending on the application of the method by different raters (emerging during the pre-segmentation step of the method).

For assessment of the respective COV the standard deviation of these paired differences for each subject was defined as:

$$SD_n = \sqrt{\sum (\Delta SCV_n^i - \overline{\Delta SCV}_n)^2 / (m_n - 1)},$$

where ΔSCV_n^i and $\overline{\Delta SCV}_n$, the SCV difference of the i th paired measurements and the mean SCV difference of the n th subject, over the m pairs. Subsequently the respective intra- and inter-session as well as intra- and inter-rater COV_s of each individual patient was computed in the form of a percentage as follows:

$$COV_n = \frac{SD_n}{\overline{SCV}_n} * 100,$$

where \overline{SCV}_n the arithmetic mean of the SCV values used for the calculation of the respective n th subject's paired differences. Comparisons between intra- and inter-session variability and between raters were performed through 2-paired t-tests, with Bonferroni corrections for multiple comparisons where applicable.

As a relative measure of reliability which measures the contribution of between-rater variance to total variance we calculated the intra-class correlation coefficient (ICC), using a two-way linear mixed effect model with random intercepts allowing for random between subject variability and different raters as a fixed effect. ICC is capable of measuring the ability of a method to detect differences between subjects consistently, and typically ranges between 0–1, with values close to 1 indicating high reliability.

Further, in order to measure *Cordial*'s accuracy, Dice-coefficients as well as symmetric Hausdorff distances and symmetric mean surface distances between the manually and semi-automatically segmented images were calculated. Manually segmented cord volumes were compared to *Cordial*'s results using 2-paired t-tests.

Association between SCV, sex and age were investigated using analysis of covariance (ANCOVA) with type III calculation of the sum of squares.

Results

All 40 lumbar cord segmentations of 10 HC were accepted after quality control for image quality and artifacts for both raters and runs (Figure 11).

For all HC, mean \pm SD SCV was 1652 \pm 262mm³ as calculated through 4 MRI scans, segmented by two raters, twice per rater (16 measurements per subject). SCVs and COVs for different raters and runs as well as comparisons between different types of reliabilities are displayed in Table 1 and demonstrated in Figure 12. All COVs were less than 3%. Mean intra-session, inter-session and intra-rater reproducibility as measured by COV were in the range of 1.22-1.86%, 1.22-1.84% and 1.73-1.83% respectively. Intra- and inter-session reliability as well as intra-rater reliability between the two raters did not significantly differ as shown by paired two sample t-tests. Inter-rater reproducibility (mean: 2.87%) was lower than all other reproducibility measures. When using a two-way linear mixed effect model with raters as fixed effect, ICC was 0.974. This indicates a very strong consistency of lumbar SC measurements between the two raters.

Using ANCOVA with the SCV as dependent variable, men ($F(1,7)=6.14$, $r=0.68$) had significantly larger volumes compared to women (men: $1986\pm 77\text{mm}^3$, women: $1489\pm 53\text{mm}^3$, $p<0.05$), while age ($F(1,7)=0.95$, $r=0.35$) had no significant effect.

Accuracy assessments showed a mean Dice-coefficient between manually and semi-automatically segmented images of 0.88 ± 0.05 . The mean symmetric Hausdorff distance was 1.55 ± 0.34 mm and the mean symmetric mean surface distance was 0.47 ± 0.19 mm. The manually segmented volumes (mean: $2028\pm 245\text{mm}^3$) were significantly larger ($p<0.001$) in comparison to the results of *Cordial* (mean: 1636 ± 300 mm^3).

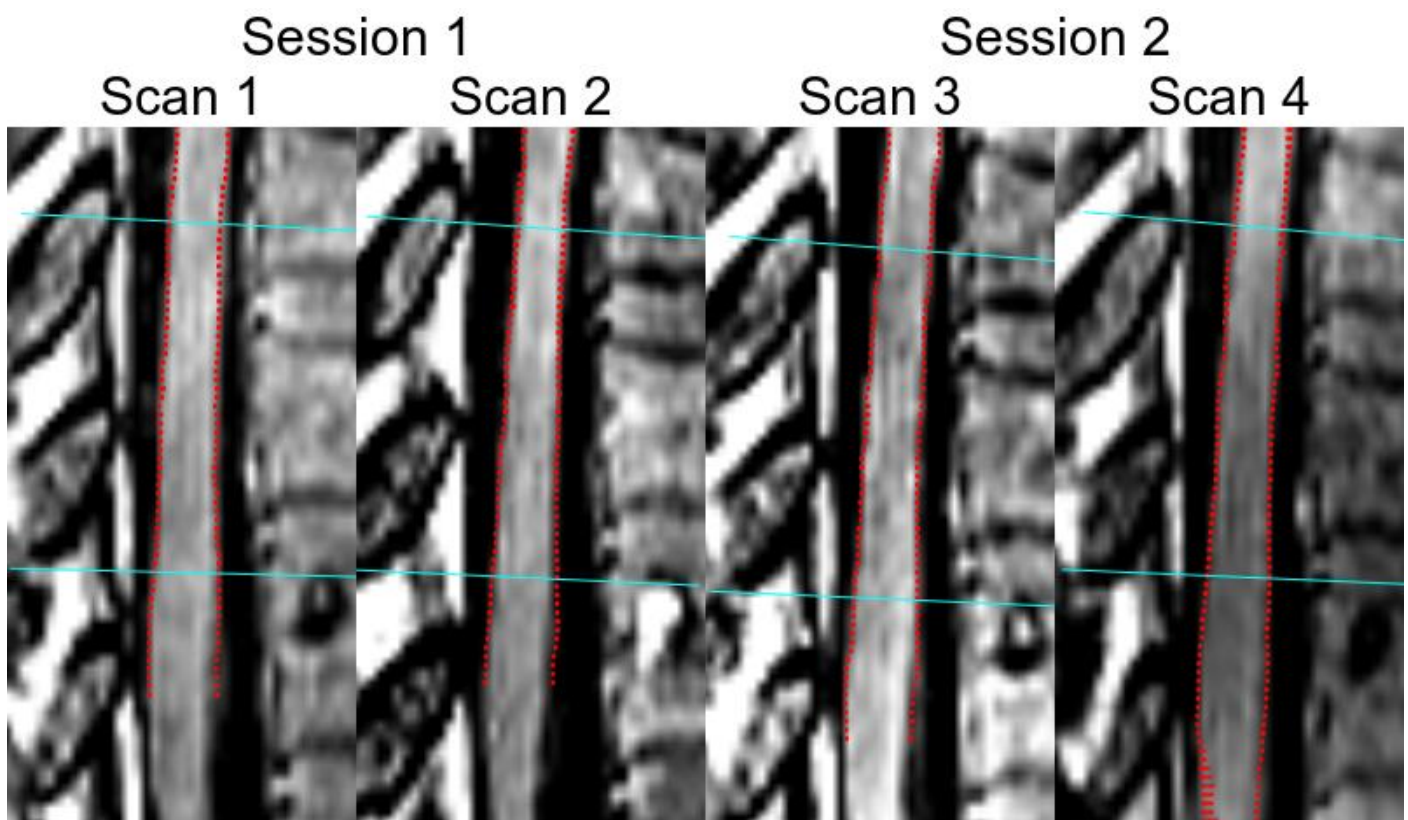


Figure 11. Successful spinal cord segmentations of the four scans of one representative healthy control. Scan 1 and 2 were performed in the 1st session, whereas scan 3 and 4 in the 2nd session.

Table 1. Spinal cord volumes and coefficients of variation of both raters and runs.

Rater	Run	SCV (mm ³)	Intra- session COV (%)	Inter-session COV (%)	Intra-rater COV (%)	Inter-rater COV (%)	p-value	
							<i>Intra-session vs. Inter-session</i>	<i>Intra-rater Between Raters</i>
<i>Rater 1</i>	<i>Run 1</i>	1620±270	1.22±0.95	1.26±0.60	1.73±1.43	2.87±1.20	ns	ns
	<i>Run 2</i>	1653±261	1.50±1.33	1.42±0.97			ns	
<i>Rater 2</i>	<i>Run 1</i>	1673±261	1.63±0.97	1.53±0.90	1.83±0.91		ns	
	<i>Run 2</i>	1661±264	1.86±1.12	1.84±0.91			ns	

Abbreviations: SCV= spinal cord volume; COV= coefficient of variation; ns= not significant

All values are shown as follows: mean ± standard deviation. P-values were obtained through paired 2 sample t-tests. Comparisons between the intra-session and inter-session reliability were conducted between COVs of a single run by a single rater. In comparisons between the inter-rater and intra-rater reliability of Cordial, Bonferroni correction was applied.

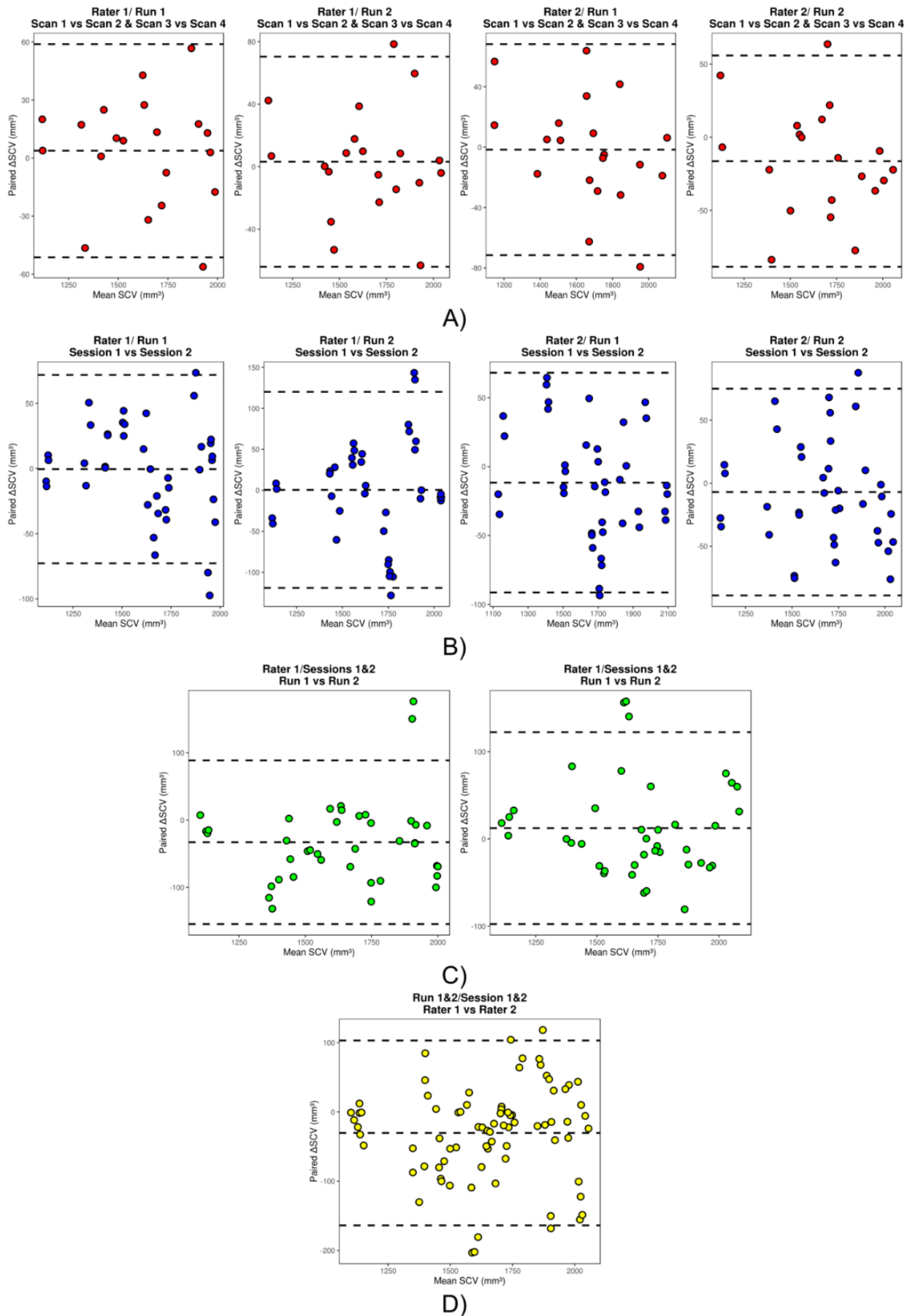


Figure 12. Bland-Altman plots. Intra-session A), Inter-session B), Intra-rater C) and Inter-rater D) paired differences (y-axis) plotted against mean SCV (x-axis) of the respective scans. Dashed lines indicate mean

difference and limits of agreement (mean difference ± 1.96 standard deviation of the paired difference). SCV = spinal cord volume (x-axis), ΔSCV = spinal cord volume differences (y-axis).

Discussion

While various SC segmentation techniques have been applied in the cervical and thoracic cord ^{26,29–31,33,34,36,37,46,101–104}, little is known regarding volumetry of the lumbar cord ¹⁰⁵. In this study, we successfully deployed the semi-automatic SC segmentation tool *Cordial* ⁴⁶ to measure the lumbar SCV in healthy controls producing robust results.

Cordial showed similar intra- and inter-session variability with a mean COV range of 1.22-1.86%, indicating high reliability, precision and consistency of the SCV measurements of back-to-back and scan-rescan MRI acquisitions. Repositioning did not seem to have significant effect on the variation of the measurements, making *Cordial* an appropriate choice for longitudinal studies. Intra-rater reproducibility showed comparable quality with a mean COV range of 1.73-1.83%, which was similar for both raters.

Of all results inter-rater reliability was lowest (mean COV of 2.87%), although the measurements of the 2 raters were highly consistent. Correct manual positioning of our anatomical marker at the tip of the conus medullaris was challenging even for experienced radiologists. Due to suboptimal contrast between the cord, surrounding CSF and spinal nerves, the exact positioning of the tip of the conus medullaris was at times difficult to discern. A difference of just one or two slices results in increased between-measurement variances. This remains a problematic issue for a fully manual or semi-automated approaches such as *Cordial* and always has to be taken into account. It could be overcome by using the same reader for all images of a study, however this would be impractical in a clinical setting. In addition, in terms of SC-CSF contrast the use of others sequences, e.g. 3D magnetization-prepared rapid gradient-echo (MPRAGE) or combination of T1- and T2-weighted contrasts might be better suited than T1w fat suppressed VIBE alone.

We saw the known effect of male sex on CNS structures also in the spinal cord ^{103,106,107}, showing larger volumes in men than women. The missing aging effect is not surprising and explained by the young age range of our subjects as previously shown in cervical and thoracic SC volumes ¹⁰³ as well as the small sample size. However, conclusions about the association of age and sex are limited by our small sample size.

Cordial demonstrated high accuracy in comparison to a consensus reading manual segmentation, deployed as a gold standard. Nevertheless, the segmented spinal cord volume measured with *Cordial* underestimated the spinal cord volume compared to the manually segmented spinal cord. The conservative spinal cord segmentation approach of *Cordial* is also evident in the axial slices of Figure 13. This should be taken into consideration in future analyses when comparing *Cordial* with other segmentation methods. Likewise, longitudinal studies using *Cordial* should focus only on the relative changes of the SCV.

In comparison to the application of *Cordial* in 3D magnetization-prepared rapid gradient-echo and T2-weighted 3D turbo spin echo images of the upper cervical SC in a previous study¹⁰², our study demonstrated a marginally lower reliability, reproducibility as well as accuracy which most probably reflects the lower image quality of the lumbar SC MR-imaging. The different environment and surrounding structures of the lumbar SC result in magnetic field inhomogeneities and lower SC-CSF contrast. Moreover, we attempted to validate *Cordial* in 3D T2-weighted SPACE images to compare the software's reliability in a T2-weighted contrast. Unfortunately, *Cordial* failed to deliver high quality segmentations of the lumbar SC in this data set (Figure 13). We believe that the spinal nerve roots in these heavily T2-weighted images generate a stronger signal compared to T1-weighted images, making them isointense to the spinal cord and therefore posing difficulties for an accurate segmentation. Integration of multiple contrasts in the segmentation algorithm in the future may improve precision of *Cordial* within this region. As an important limitation of our study, it should be mentioned, that *Cordial* was not able to deliver a reliable volumetry of the conus medullaris, which is a relevant spinal cord segment for a number of spinal cord disorders. In our view, this is, at least partly, due to the low SC-CSF contrast rendering it difficult to separate conus tissue from surrounding spinal nerve roots.

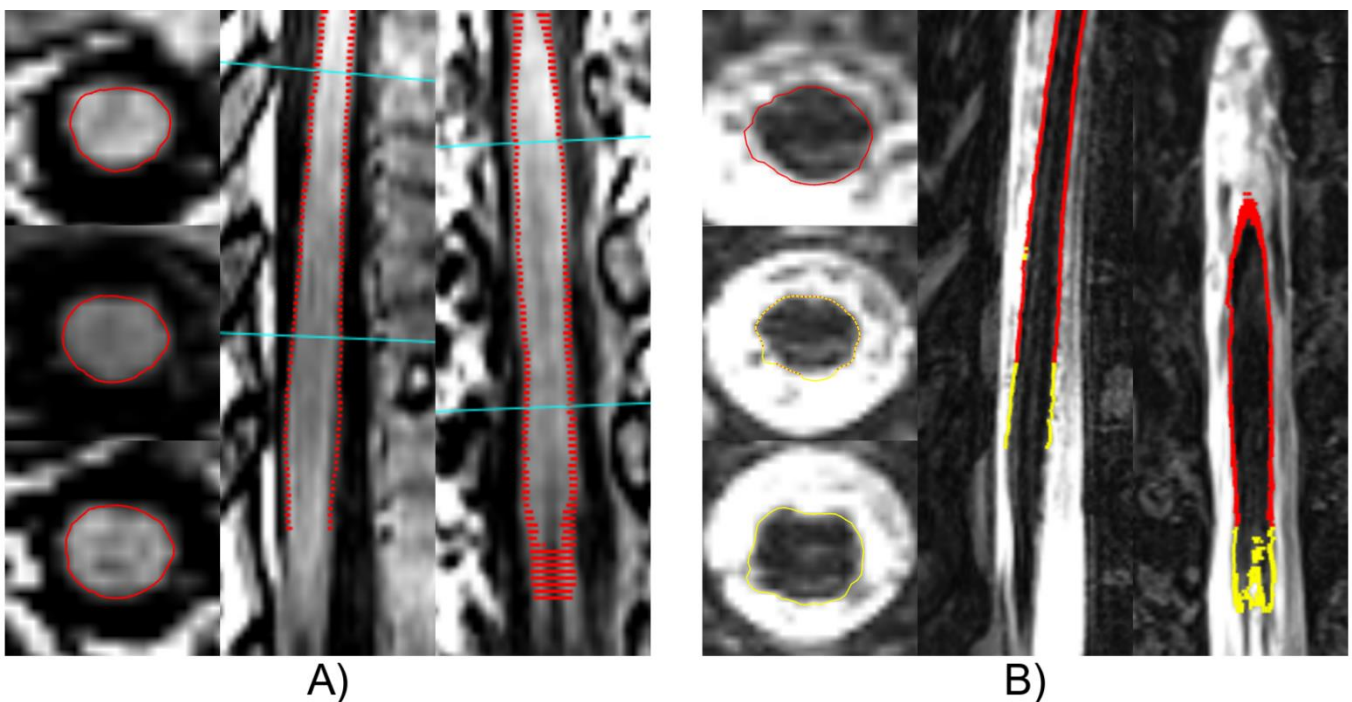


Figure 13. Lumbar spinal cord (SC) segmentation (red lines) using *Cordial* in a representative subject. A) Successful segmentation on 3D T1-weighted fat-suppressed VIBE images. As demonstrated in the sagittal image, *Cordial* aborts the segmentation in the region of the conus medullaris. Of note, *Cordial* shows a rather conservative segmentation with its outer boundary slightly within the cord's surface. B) *Cordial* failed to deliver high quality segmentations of the whole lumbar SC in 3D T2-weighted SPACE images of the same subject due to artifacts based on the spinal nerve roots exits and nerve roots adjacent to the spinal cord within this region.

Lastly, *Cordial* is to this point limited to whole-cord measurements and is currently not able to perform any other measurements such as grey and white matter measurements.

In conclusion, *Cordial* results in the lumbar cord were highly reliable and reproducible. *Cordial* is therefore suitable for the use in longitudinal trials. The COV values calculated in this study can be used for effect size calculations for SCV for studies using the same processing pipeline.

Acknowledgments:

We would like to thank Tanja Haas and Pascal Kuster for MRI data acquisition and data management. Most of all, we are grateful to the healthy controls for participating in the study. Data acquisition was funded by F. Hoffman La Roche. F. Hoffman La Roche did not have any additional role in the study design, data collection, analysis, interpretation of data, writing of the report and decision to submit the paper for publication.

3.2. Automatic Spinal Cord Gray Matter Quantification: A Novel Approach

*C. Tsagkas¹⁺²⁺³, *A. Horvath⁴, A. Altermatt²⁺³⁺⁴, S. Pezold⁴, M. Weigel²⁺⁴⁺⁵, T. Haas⁵, M. Amann¹⁺³⁺⁶, L. Kappos¹⁺², T. Sprenger¹⁺⁷, O. Bieri⁴⁺⁵, P. Cattin⁴, K. Parmar¹⁺²

* equally contributing first authors

1. Neurologic Clinic and Policlinic, Departments of Medicine, Biomedical Engineering and Clinical Research, University Hospital Basel, University of Basel, Basel, CH
2. Translational Imaging in Neurology (ThINK) Basel, Department of Medicine and Biomedical Engineering, University Hospital Basel, University of Basel, Basel, Switzerland
3. Medical Image Analysis Center (MIAC AG), Basel, CH
4. Department of Biomedical Engineering, University of Basel, Allschwil, CH
5. Division of Radiological Physics, Department of Radiology, University Hospital Basel, University of Basel, Basel, CH
6. Division of Diagnostic and Interventional Neuroradiology, Department of Radiology, University Hospital Basel, University of Basel, Basel, CH
7. Department of Neurology, DKD HELIOS Klinik Wiesbaden, Germany

Grant support:

Swiss National Science Foundation (Grant number: 320030_156860)

Foundation for Sponsorship of Gastroenterological and General Clinical Research as well as of Medical Imaging (Application ID 02/2015)

Publication: American Journal of Neuroradiology. September 2019. DOI: <https://doi.org/10.3174/ajnr.A6157>

Abstract

Background and purpose: To assess the reproducibility and accuracy of cervical spinal cord (SC) gray matter (GM) and white matter (WM) cross-sectional area measurements using averaged magnetization inversion recovery acquisition (AMIRA) images and a fully-automatic post-processing segmentation algorithm.

Materials and Methods: The cervical SC of 24 healthy subjects (14 female, age 40 ± 11 years) was scanned in a test-retest fashion on a 3T MRI-system. Twelve axial AMIRA slices were acquired over a 48mm cord segment. GM and WM were both manually segmented by two experienced readers and compared to an automatic variational segmentation algorithm with a shape prior modified for 3D data with a slice similarity prior. Precision and accuracy of the automatic method were evaluated using coefficients of variation (CV) and Dice similarity coefficients (DSC).

Results: Mean GM area was $17.20\pm 2.28\text{mm}^2$ and mean WM area $72.71\pm 7.55\text{mm}^2$ using the automatic method. Reproducibility was high for both methods, while being better for the automatic approach (all mean automatic CVs $\leq 4.77\%$, all differences $p < 0.001$). The accuracy of the automatic method, as compared to the manual reference standard, was excellent (mean DSC: 0.86 ± 0.04 for GM and 0.90 ± 0.03 for WM). The automatic approach demonstrated similar CVs between intra- and inter-session reproducibility as well as among all acquired SC slices.

Conclusion: Our novel approach including the AMIRA-sequence and a fully-automated post-processing segmentation algorithm demonstrated an accurate and reproducible SC GM and WM segmentation. This pipeline is promising for both the exploration of longitudinal structural GM changes and application in clinical settings in disorders affecting the SC.

Abbreviations

SC=spinal cord; magnetic resonance imaging=MRI; gray matter=GM; white matter=WM; cerebral spinal fluid=CSF; averaged magnetization inversion recovery acquisitions=AMIRA; 3 dimensional=3D; total spinal cord=TSC; coefficients of variation=CV, dice similarity coefficients=DSC; hausdorff distances=HD; multivariate one-way analysis of variance=MANOVA.

Introduction

The human spinal cord (SC) can be affected by numerous neurologic disorders of variable pathophysiology (e.g. genetic, inflammatory and demyelinating, degenerative etc.)^{108,109} and magnetic resonance imaging (MRI) is a valuable part of the diagnostic work-up in patients with suspected intramedullar pathology^{2,5}. SC gray matter (GM) and white matter (WM) can be involved to various extent not only between different SC disorders but also between patients suffering from the same disease (e.g. multiple sclerosis, amyotrophic lateral sclerosis)^{95,110}. Hence, quantification of SC compartments may add to our understanding of SC pathology^{95,110} and hopefully help in the management of individual patients in the future.

However, the SC presents additional challenges for MRI. The SC is surrounded by a number of different tissue types including cerebral spinal fluid (CSF), bone and air. This creates significant signal inhomogeneities along this thin, elongated structure^{1-3,5}. As a result, “conventional” SC MRI was – until recently – not able to differentiate sufficiently between SC GM, WM and CSF. In the last years, first attempts towards this differentiation using a series of acquisition approaches were made¹⁷⁻²⁰. More recently an averaged magnetization inversion recovery acquisitions (AMIRA) sequence was proposed, delivering a notable SC GM/WM contrast while maintaining short acquisition times at the same time²¹. The latter is especially important for imaging small-sized structures (like the SC GM/WM) in patients with disabilities having a short time window where they can lie still.

Moreover, accurate SC GM segmentation remains challenging. First, manual approaches demonstrated the feasibility to distinguish between WM and GM¹⁷. However, manual approaches require a considerable amount of time, are prone to error and demonstrate significant inter-observer and intra-observer variability. As a result of improvements in image quality and post-processing techniques, the first fully automatic SC GM segmentation methods were established in the past few years⁴⁷⁻⁵⁰. These methods have deployed atlas-based GM segmentation algorithms, which may however lead to misestimations or segmentation errors especially in case of pathology, image artifacts or large between-individual anatomical variations^{111,112}. A non-invasive broadly accepted reference standard for accurate and reproducible SC GM measurements is still a matter of ongoing discussion.

In this study, we validate a fully automatic method for SC GM and WM segmentation in terms of its reproducibility and accuracy in segmenting the cervical SC of healthy controls against a manual segmentation. The proposed approach utilized a variational segmentation algorithm with a shape prior¹¹³ modified for 3 dimensional (3D) data with a slice similarity prior on AMIRA images.

Materials and Methods

Subjects & MRI acquisition

24 healthy subjects (14 female, age 40 ± 11 years) were scanned in a test-retest fashion on a 3T whole-body MRI system (Magnetom Prisma, Siemens). All subjects gave written consent. Experimental procedures conformed to the Declaration of Helsinki and the study protocol was approved by the local ethics committee. We acquired 12 axial AMIRA images ²¹ (FOV = $128 \times 128 \text{mm}^2$, slice thickness 8mm, 4mm slice overlap, in-plane resolution = $0.67 \times 0.67 \text{mm}^2$, $TE_{bSSFP} = 2.14$ ms, $TR_{bSSFP} = 5.13$ ms, signal averaging = 1, acquisition time = 51s per slice) over a 48mm cervical SC segment, approximately extending from the C2-C5 vertebral level²¹. The most rostrally acquired slice was placed with its lower surface adjacent to the most rostral surface of the C2/C3 intervertebral disc. For precise positioning of each individual slice and its orthogonal angulation to the course of the SC, a strongly T2 weighted TSE with high contrast between CSF and SC was used as a reference. For each slice, the AMIRA approach acquired eight images of considerably different tissue contrast between GM, WM and CSF with $TI_{\text{eff}} = [97.1; 158.7; 220.2; 281.8; 343.3; 404.9; 466.5; 528.0]$ ms. Averaging the first five images enhances the GM/WM CNR, whereas averaging the last three images clearly improves the WM/CSF CNR (Figure 1). For more details on the AMIRA sequence, please see Weigel et al. 2018 ²¹.

Each subject was scanned three times in one MRI session. The first two scans were performed in a back-to-back fashion without repositioning to allow for intra-session comparisons. The third scan was conducted after patient repositioning to allow for inter-session comparisons.

All scans underwent basic pre-processing including 2D and 3D correction for field inhomogeneities using the scanner software prior to segmentation. To minimize numerical errors of the validation metrics, we performed a 5-fold in-slice upsampling of the slices using the Lanczos-3 interpolation kernel.

Each subject was scanned three times in one MRI session. The first two scans were performed in a back-to-back fashion without repositioning to allow for intra-session comparisons. The third scan was conducted after patient repositioning to allow for inter-session comparisons.

SC segmentation

As proposed in a previous study ¹¹³, a variational segmentation approach based on the continuous min-cut max-flow framework was used, which includes total variation regularization to segment WM and GM. The min-cut max-flow's capacity functions are modeled using edge, region, and prior information as well as an appearance model built from manual segmentations. Aiming for high accuracy, the proposed approach prefers the actual image's intensities and tries to include prior information as little as possible, which regularizes for higher precision. Compared to the previous study ¹¹³, we added a slice similarity prior ¹⁰², included all inversion images of one slice (Figure 1) into the calculation of the max-flow's capacity functions and improved the GM and WM principal component appearance models with posterior appearance models ¹¹⁴.

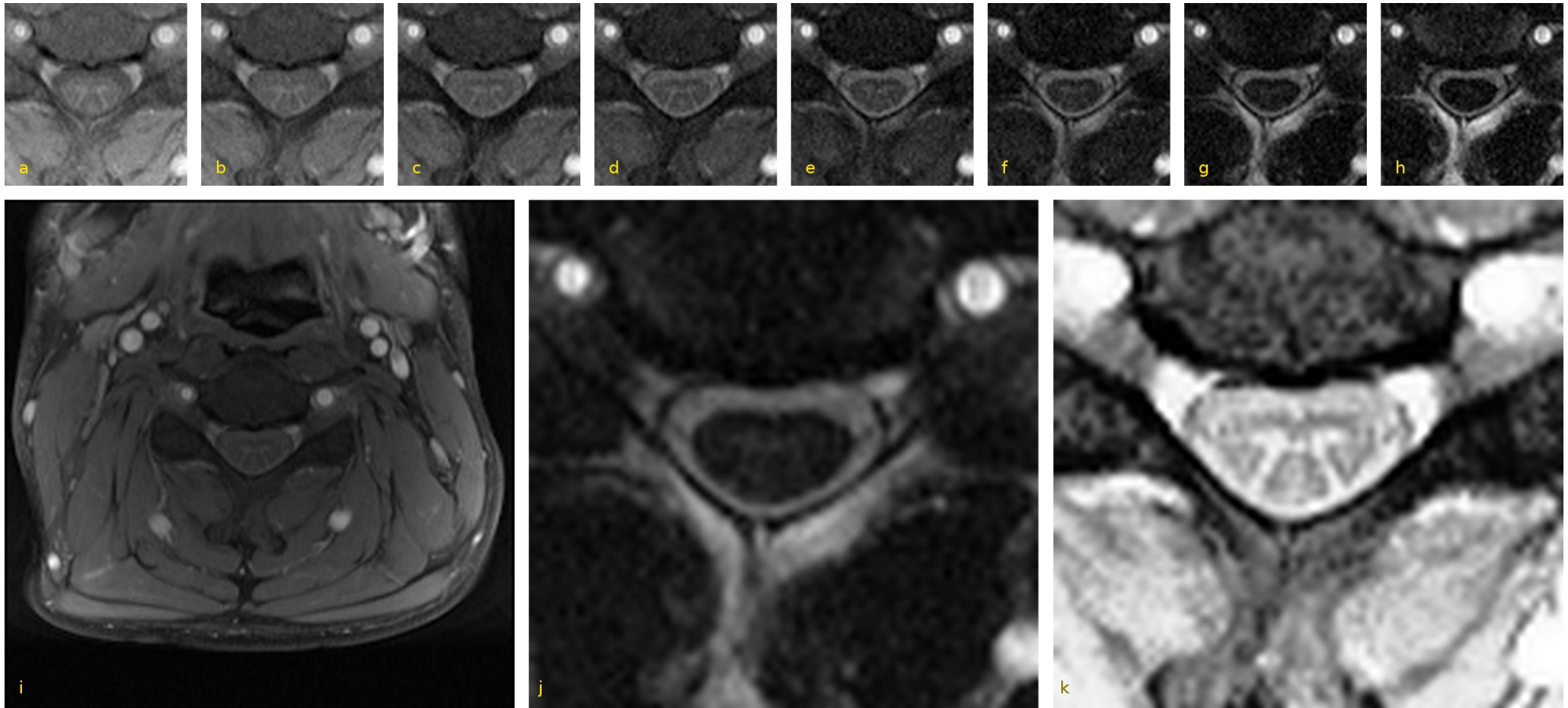


Figure 14. Exemplary axial AMIRA slice of one representative volunteer at the C4 level. a-h) Eight images of different tissue contrast acquired by the AMIRA sequence, shown in chronological order from lowest to highest inversion time. i) Average image from a→e in full view, which delivers a high contrast-to-noise-ratio for GM/WM. j) Average image from f→h, which delivers a high contrast-to-noise ratio for SC/CSF. k) Same average image as in i) but histogram-equalized and zoomed.

As a first step of the algorithm to align the 12 slices, the images are center cropped and slice-wise successively co-registered rostral to caudal using translations in pixel size steps to prevent further interpolation. Then, the algorithm automatically locates and delineates the ring-shaped CSF from its surroundings and extracts the cross-sectional SC surface. Finally, it uses the previously segmented SC surface as a mask for GM/WM differentiation. An illustration of the algorithm is shown in Figure 15. Segmentations were achieved in a leave-one-subject-out cross-validation, that is: with the currently segmented subject being left out in the used appearance model.

The segmentation algorithm was implemented in MATLAB. Processing time on an Intel Xeon CPU E5-2620 v3 @ 2.40GHz is around one minute for each segmentation step (CSF-SC and WM-GM segmentation), and less than 8 GB of RAM are used to segment a stack of 12 slices. Code is available on [github link].

Two experienced raters (C.T.; A.A.) were involved in the manual segmentations. Both raters had more than 4 years experience in neuroimaging research including SC volumetric studies. In a first step, this was conducted on the average of the last three AMIRA images for total SC cross-sectional area. Using the already delineated total SC masks, manual segmentations of the GM and WM cross-sectional areas were then performed on the average of the first five AMIRA images (Supplementary Figure 1). C.T. segmented all images once. These results were further applied as “manual reference standard”. C.T. also conducted a second “run” of 60 randomly selected slices to assess intra-rater comparisons. This second “run” was conducted with slightly different contrast adjustments than the first in order to evaluate the robustness of intra-rater manual segmentation. A.A. segmented all images of the first scan of all 24 healthy controls to allow for inter-rater comparisons.

In order to evaluate the performance of our method on SC slices, in which the fully automatic approach failed (in total 12% of acquired slices, see also “Results”), we applied a semi-automatic approach as follows: the SC/CSF boundaries were segmented manually (manual reference standard) and segmentation of the GM and WM was then performed using the fully automatic approach described above given the manual total SC masks. In order to compare our automatic method with currently available algorithms, we tested the iterative non-local STAPLE (iNLS)¹¹⁵ on our AMIRA images using the algorithm in the SCFusion_Demo package*. Asman et al. use atlases consisting of SC GM-WM contrast images and SC-GM-WM manual reference standard segmentations, which are rigidly registered to the target slice and fused together with the most fitting manual reference standard segmentation as an estimation of the targeted segmentation. We built our own atlases and tested iNLS in a leave-one-subject-out fashion.

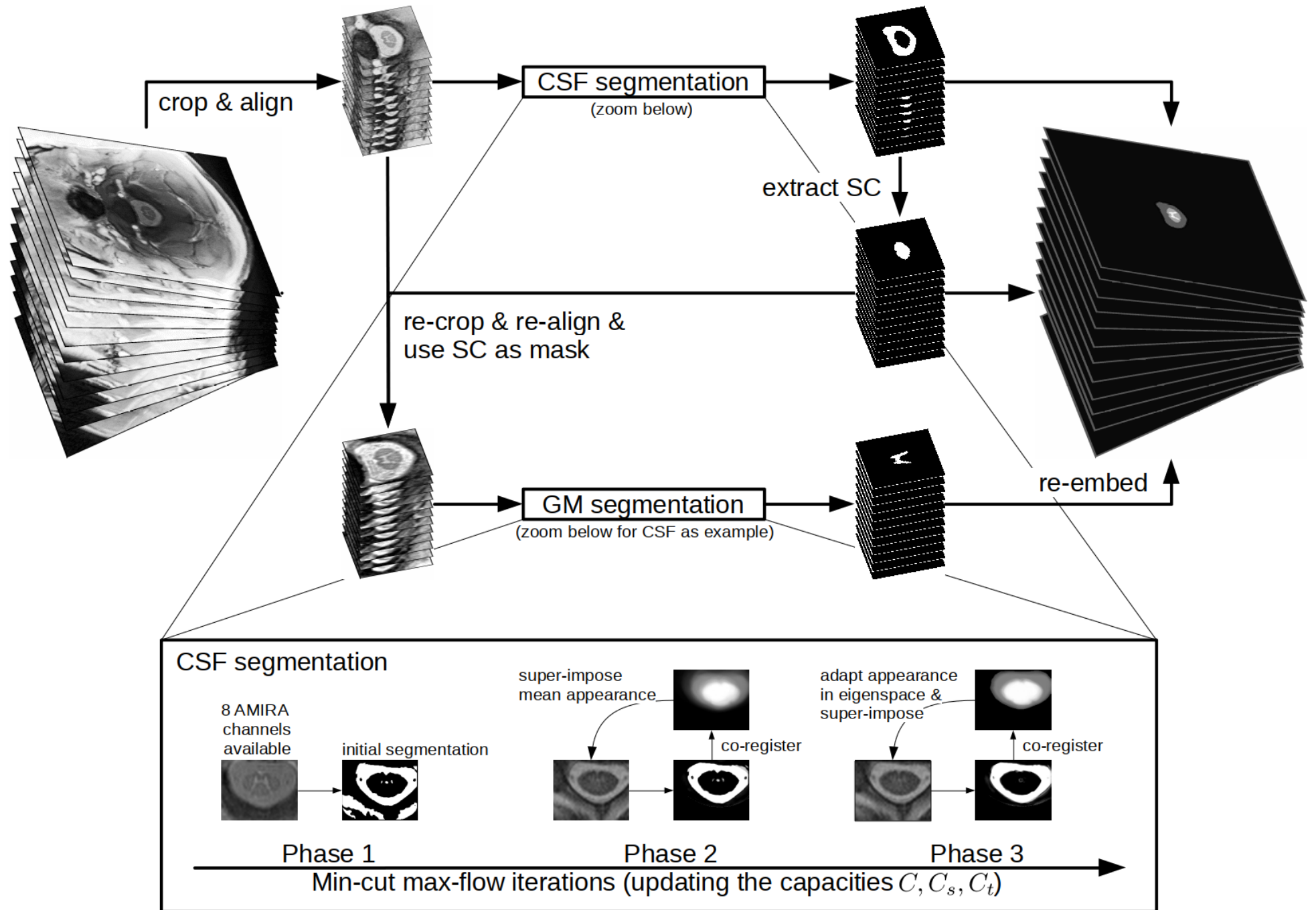
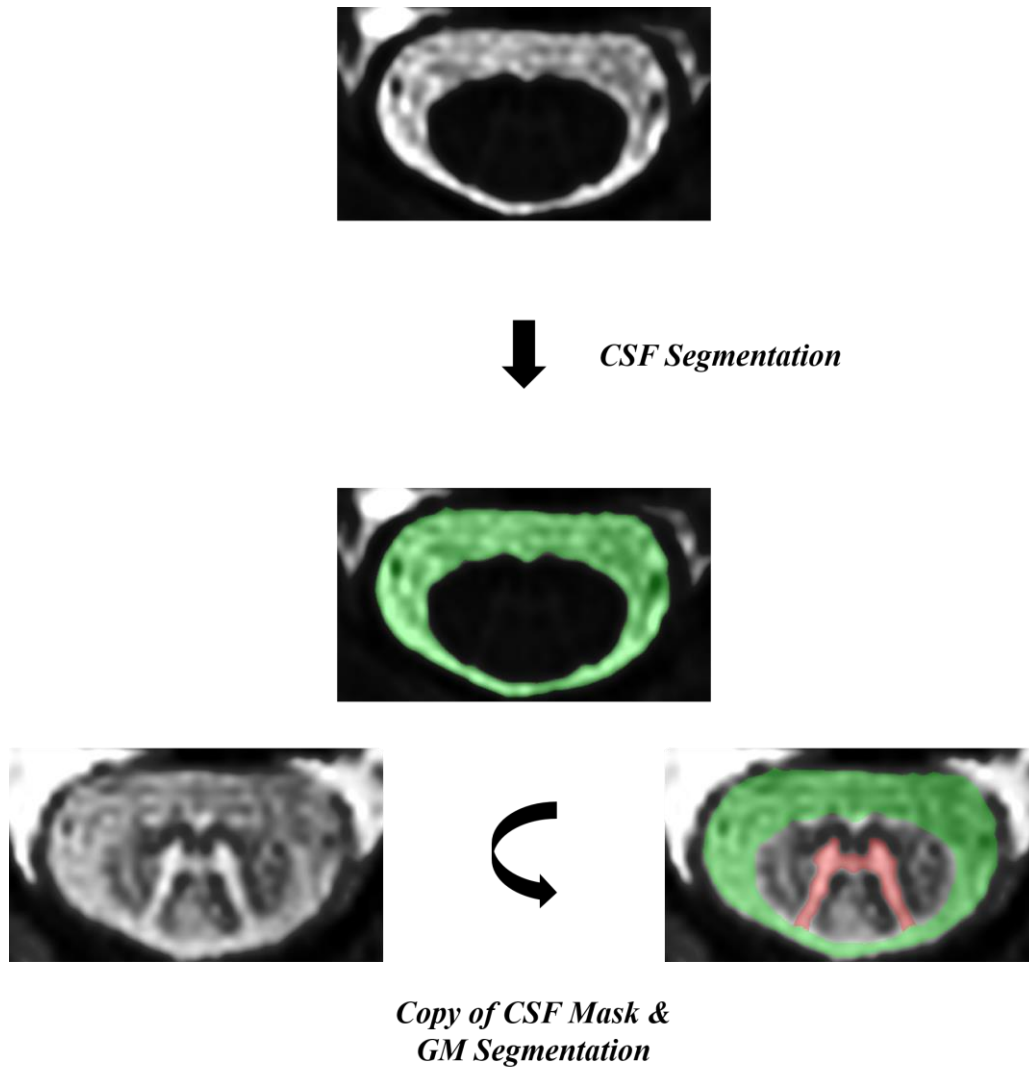


Figure 15. *Flow-chart of the automatic segmentation pipeline. As a first step of the algorithm to align the 12 slices, the images are center-cropped and slice-wise successively co-registered rostral to caudal using translations in pixel size steps to prevent further interpolation. Then, the algorithm automatically locates and delineates the ring-shaped CSF from its surroundings and extracts the cross-sectional SC surface. Finally, it uses the previously segmented SC surface as a mask for GM/WM differentiation. The iterative steps of CSF segmentation are shown as a zoomed-in view. GM segmentation uses essentially the same steps and is thus not shown in detail.*



Supplementary figure 1. *Manual segmentations of acquired axial AMIRA slices. In a first step, the CSF was manually segmented using the average of the last three AMIRA images. The CSF “mask” was then overlaid on the average of the first five AMIRA images. Subsequently, the SC GM was also manually segmented.*

Statistical Analysis

Intra- and inter-session as well as intra- and inter-rater reproducibility of the two approaches were evaluated using coefficients of variation (CV), Dice similarity coefficients (DSC) and Hausdorff distances (HD). The accuracy of the automatic method compared to the manual reference standard was evaluated using DSC and HD. CVs between two masks A and B were calculated with the following formula:

$$CV = \sqrt{2} \times \frac{|A - B|}{|A + B|} \%$$

DSCs were calculated as follows:

$$DSC = 2 \times \frac{|GT \cap AM|}{|GT| + |AM|}$$

HDs were calculated as follows:

$$d(X \rightarrow Y) = \max(d_i^{X \rightarrow Y}, i = 1 \dots N_X)$$

$$HD = \max(d(GT \rightarrow AM), d(AM \rightarrow GT))$$

where d is the Euclidean distance between voxel x and y .

Because of non-normally distributed data, we performed a square root transformation of the CV, a cubic transformation of the DSC and a logarithmic transformation of the HD before conducting all t-test analyses and MANOVA. 2-paired t-tests were performed for the following comparisons after controlling for normal data distribution: 1) manual segmentation versus automatic method reproducibility, 2) manual segmentation versus automatic method total SC, WM and GM cross-sectional area. For the automatic method, differences in measures of reproducibility and accuracy a) between intra-session and inter-session b) between GM, WM and total SC and c) between the axial slices' levels (1→12) were investigated using multivariate one-way analysis of variance (MANOVA). Additional Tukey's post-hoc-tests were conducted, where applicable.

Results

In total, 864 slices were acquired from 24 volunteers with 12 slices per scan, performed a total of three times for each subject. Out of 864 acquired axial SC slices, nine were excluded from further analysis, because of severe imaging artifacts. The automatic method successfully segmented 88% (752 slices) of all remaining slices. Because of imaging artifacts, localization problems, or posterior gaps of the CSF, 8% of all slice-wise SC segmentations and 4% GM segmentations would have needed further manual interventions and thus were excluded from the reproducibility analysis.

Cross-sectional SC measurements

Mean total SC area was $89.98 \pm 7.88\text{mm}^2$, mean WM area $72.71 \pm 7.55\text{mm}^2$ and mean GM area $17.20 \pm 2.28\text{mm}^2$ as measured by the automatic method. Compared to the manual reference standard, the automatic method delivered significantly higher total SC and WM area as well as significantly lower GM area (86.88 ± 11.87 , 69.18 ± 10.16 and 17.77 ± 3.05 respectively, all $p < 0.001$). Cross-sectional areas per slice of automatic method are shown in Figure 16.

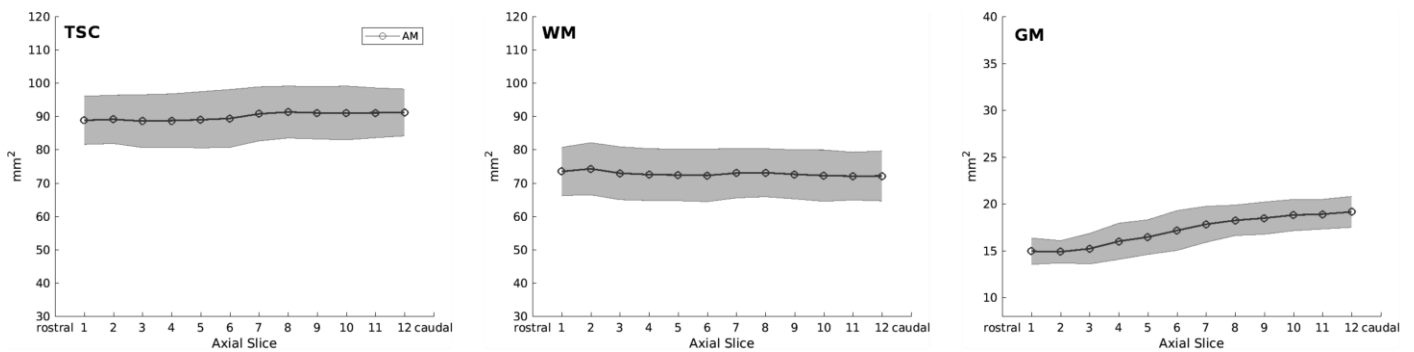


Figure 16. Cross-sectional areas of total spinal cord (TSC), white matter (WM) and grey matter (GM) per axial slice, as measured by automatic segmentations. Notice the slight increase of TSC and the marked GM cross-sectional area increase caudally, which corresponds to the cervical SC enlargement. The light grey area depicts the limits of \pm one standard deviation.

Reproducibility

Measurements of intra- and inter-session as well as intra- and inter-rater reproducibility are displayed in Table 2. Reproducibility of SC GM and WM is also depicted per slice in Figure 17. All mean CVs of the automatic method were $\leq 4.77\%$ and mean DSC ≥ 0.88 between scans and raters. The latter was significantly better than for the manual segmentation (all $p < 0.001$).

Using MANOVA with DSC, HD and CV as multivariate outcome, a significant difference between intra- and inter-session reproducibility for total SC, WM and GM using the automatic method was shown ($p < 0.001$ for all 3 models). However, CVs differed only for WM and total SC ($p < 0.05$ and $p < 0.001$ respectively), but not for GM. In our automatic method, intra- and inter-session reproducibility was significantly decreasing in the order total SC \rightarrow WM \rightarrow GM (all $p < 0.001$), as shown by MANOVA and post-hoc tests. No difference was found in intra- and inter-session reproducibility between slices for GM, but a significant decrease was found for WM (both $p < 0.001$) and total SC ($p < 0.05$ and $p < 0.001$ respectively) in caudal slices, as shown by MANOVA. However, CVs were similar for all SC metrics between all slices.

Accuracy

Measurements of accuracy of our fully automatic method compared to a manual reference standard are shown in detail in Table 3 and are also displayed per slice in Figure 18. The automatic method showed a mean DSC ≥ 0.86 in all SC metrics. Accuracy was significantly decreasing in the order total SC \rightarrow WM \rightarrow GM (all $p < 0.001$), as shown by MANOVA and post-hoc tests. In MANOVA, accuracy was lower for GM ($p < 0.05$) and total SC ($p < 0.001$) in caudal slices, but not for WM. However, DSC was similar between acquired slices for total SC.

Measurements of accuracy of the initially discarded SC slices (12% of all acquired AMIRA slices) analyzed in a semi-automatic fashion are also shown in Table 3. The semi-automatic approach showed a mean DSC of ≥ 0.83 in both GM and WM. When compared to the fully automatic approach on the initially non-discarded SC slices, a statistically significant accuracy decrease was observed in the semi-automatic approach (both $p < 0.001$).

Comparison with the iterative non-local STAPLE algorithm

In comparison to the original study 23 performed on T2* MR-images, the application of the iterative non-local STAPLE algorithm in our AMIRA images showed a higher accuracy. Mean DSC and HD for the total SC, GM and WM were as follows: total SC: mean DSC 0.93 ± 0.03 (median 0.94), mean HD 0.96 ± 0.39 mm (median 0.84mm); GM: mean DSC 0.80 ± 0.06 (median 0.82), mean HD 1.09 ± 0.42 mm (median 1.04mm); WM: mean DSC 0.87 ± 0.04 (median 0.88), mean HD 0.98 ± 0.37 (median 0.89mm).²³ Moreover, our proposed automatic method had higher accuracy for all total SC, WM and GM compared to the iterative non-local STAPLE algorithm in our AMIRA images (all $p < 0.001$).

Table 2. Grey matter, white matter and total spinal cord reproducibility measurements.

Parameter	GM		WM		TSC		MAS vs AM	
	MAS	AM	MAS	AM	MAS	AM		
<i>Intra-session</i> <i>n=240</i>	CV (%)	5.54 ± 4.14	4.10 ± 3.48	3.69 ± 3.23	2.54 ± 2.69	2.99 ± 2.73	1.64 ± 1.63	GM: p<0.001 WM: p<0.001
	DSC	0.86 ± 0.03	0.89 ± 0.03	0.93 ± 0.02	0.95 ± 0.02	0.97 ± 0.01	0.98 ± 0.01	GM: p<0.001 WM: p<0.001
	HD (mm)	0.64 ± 0.22	0.63 ± 0.39	0.54 ± 0.13	0.54 ± 0.24	0.44 ± 0.17	0.32 ± 0.15	GM: p=0.77 WM: p=0.80
<i>Inter-session</i> <i>n=454</i>	CV (%)	6.34 ± 4.66	4.77 ± 4.71	4.68 ± 3.75	2.95 ± 2.58	3.75 ± 3.13	2.21 ± 1.93	GM: p<0.001 WM: p<0.001
	DSC	0.85 ± 0.03	0.88 ± 0.04	0.92 ± 0.02	0.94 ± 0.02	0.97 ± 0.02	0.98 ± 0.01	GM: p<0.001 WM: p<0.001
	HD (mm)	0.68 ± 0.24	0.71 ± 0.46	0.58 ± 0.16	0.59 ± 0.28	0.50 ± 0.19	0.37 ± 0.14	GM: p=0.10 WM: p=0.45
<i>Intra-rater</i> <i>n=60</i>	CV (%)	19.18 ± 12.16	0 ± 0	4.14 ± 3.34	0 ± 0	1.17 ± 1.92	0 ± 0	n.a.
	DSC	0.85 ± 0.07	1 ± 0	0.96 ± 0.02	1 ± 0	0.99 ± 0.01	1 ± 0	n.a.
	HD (mm)	0.62 ± 0.30	0 ± 0	0.44 ± 0.15	0 ± 0	0.20 ± 0.13	0 ± 0	n.a.
<i>Inter-rater</i> <i>n=283</i>	CV (%)	12.03 ± 8.87	n.a.	8.06 ± 6.15	n.a.	6.50 ± 5.17	n.a.	n.a.
	DSC	0.85 ± 0.05	n.a.	0.90 ± 0.04	n.a.	0.95 ± 0.03	n.a.	n.a.
	HD (mm)	0.75 ± 0.32	n.a.	0.75 ± 0.27	n.a.	0.68 ± 0.34	n.a.	n.a.

Abbreviations: MAS = manual segmentations; AM = automatic segmentations; GM = grey matter; WM = white matter; TSC = total spinal cord; CV = coefficient of variation; DSC = Dice coefficient; ns = not significant; n signifies the number of MRI axial slices taken into account for each parameter; n.a. = not applicable.

All values are shown in the form: mean \pm standard deviation. P -values were obtained through paired 2 sample t -tests ($\alpha = 0.05$), after performing a square root transformation of the CV, a cubic transformation of the DSC and a logarithmic transformation of the HD.

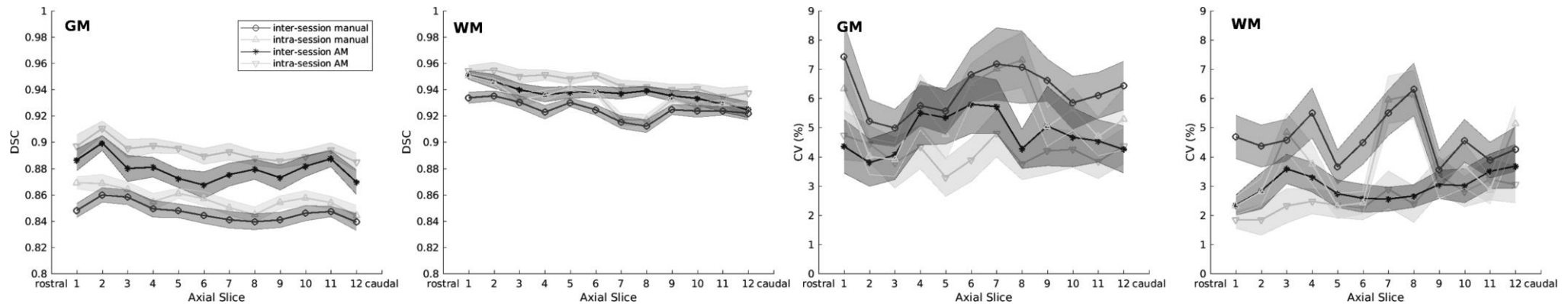


Figure 17. Comparison between the reproducibility of manual and automatic measurements (AM) of spinal cord grey matter (GM) and white matter (WM) per axial slice. Intra-session and inter-session reproducibility is assessed in terms of Dice coefficients (DSC; graphic 1+2) and coefficients of variation (CV; graphic 3+4). Manual and automatic inter-session reproducibility is shown dark grey, whereas manual and automatic intra-session reproducibility is shown in light grey. Error bars display mean values ± 0.2 standard deviations.

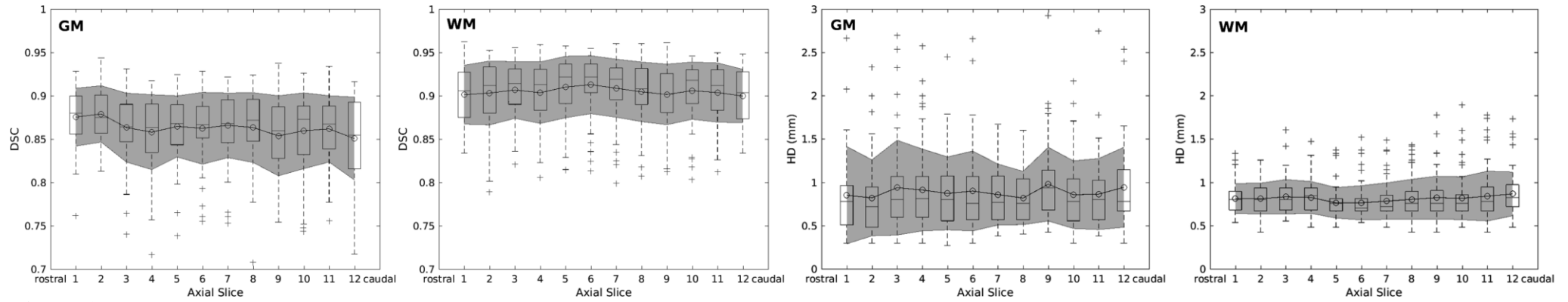


Figure 18. Accuracy measurements in terms of Dice coefficients (DSC; graphic 1+2) and Hausdorff distances (HD; graphic 3+4) of white matter (WM) and grey matter (GM) per slice. Overlaid boxplots display median values as well as 25th and 75th value percentiles. Grey areas depict the mean standard error values \pm one standard deviation.

Table 3. Table 2. SC GM, WM and TSC accuracy of automatic and semi-automatic segmentations against the manual reference standard segmentations.

	Parameter	GM	WM	TSC
AM (on 88% of acquired slices, non-discarded from initial analysis)	DSC	0.86 ± 0.04 (0.87)	0.90 ± 0.03 (0.91)	0.95 ± 0.03 (0.95)
	HD (mm)	0.90 ± 0.44 (0.72)	0.82 ± 0.22 (0.75)	0.64 ± 0.27 (0.67)
SAM (on 12% of acquired slices, discarded from initial analysis)	DSC	0.83 ± 0.04 (0.84)	0.96 ± 0.01 (0.96)	-
	HD (mm)	1.11 ± 0.55 (0.93)	0.64 ± 0.15 (0.67)	-
AM (on non-discarded samples) and SAM (on discarded slices) mixed (100% of acquired slices)	DSC	0.86 ± 0.04 (0.86)	0.91 ± 0.04 (0.92)	0.96 ± 0.03 (0.96)
	HD (mm)	0.91 ± 0.46 (0.81)	0.80 ± 0.22 (0.75)	0.60 ± 0.29 (0.55)

Abbreviations: GM=grey matter; WM= white matter cross-section; TSC= spinal cord, AM = automatic segmentations; SAM = semi-automatic segmentation; DSC = Dice coefficient; HD = Hausdorff distances

All values are shown in the form: mean \pm standard deviation (median).

Discussion

Visualization of the SC GM in MRI was hampered by technical difficulties until recently ³. Despite technological advancements, segmentation of SC compartments remains a challenge ¹¹⁶. In this work, we successfully deployed the novel MRI-approach AMIRA and a fully automatic variational segmentation algorithm with a shape prior modified for 3D data with a slice similarity prior to demonstrate a fully automated approach for segmentation of SC GM and WM.

In contrast to brain MRI, the environment of the SC presents additional challenges for MRI methods and inherently to SC segmentation. Greatest challenges include magnetic field inhomogeneities across the SC, cord curvature, shape and size, contact of the SC and the osseous canal, osteophytes causing focal changes in CSF flow dynamics within canal, motion artifacts, Gibbs artifacts, partial volume effects, and B1 inhomogeneity ¹⁴⁻¹⁶. The used AMIRA sequence ¹³ is based on a 2D approach that is, generally, less motion-sensitive than a 3D sequence. It utilizes a relatively short acquisition time of 51s per slice, which leads to a reduction of motion artifacts and is especially suitable for disabled patients with limited abilities in lying still (e.g. due to spasticity). The AMIRA approach makes use of a balanced steady state free precession readout, which is inherently of low flow sensitivity or "inherently flow compensated". ²¹The inversion recovery preparation is global and non-selective; hence, it does not pose an issue for CSF flow sensitivity either. Furthermore, since the SC has a small cross section of roughly 1.3cm by 0.7cm and our slices were located close to the isocenter, effects of B1 inhomogeneity do not play a significant role for the present AMIRA acquisitions. The even smaller size of the SC GM depicts additional difficulties for MRI methods, requiring submillimeter in-plane resolutions, especially for the needs of morphometry. Visualization and segmentation of the SC GM and WM is hampered by the similar relaxation times of the two SC compartments limiting the use of conventional SC MR-imaging to this matter. Finally, the complex "butterfly" shape of the SC GM makes the segmentation of the structure a rather hard task for computer-based segmentation methods. The AMIRA approach was able to produce SC images with a high GM/WM contrast in all participants. This was achieved in clinically feasible acquisition times (10.2min for a 48mm cervical SC segment). Out of 864 slices, only 9 were excluded due to image artifacts mainly resulting from magnetic field inhomogeneities produced by bone structures (clavicles, scapulae, humeri, ribs etc.) as well as due to aliasing and motion artifacts. Although, these artifacts occurred in a rather small percentage of the acquired images (1%), this should be taken into account in future applications of the AMIRA approach.

A further argument in favor of the use of AMIRA for the purpose of SC GM and WM quantification is the fact that our pipeline was able to not only deliver higher accuracy measures compared to a previous study ¹¹⁶ demonstrating results from various MRI sequences and segmentation algorithms, but also the better accuracy performance of an established algorithm on AMIRA compared to T2* MR-images (see also below). This may be an indirect indication for superior quality of AMIRA compared to other sequences used so far for spinal cord GM and WM quantification. Nevertheless, due to the fact that AMIRA has a non-isotropic resolution,

our MRI acquisition may have been more prone to partial volume effects, despite our slices being angulated individually in an orthogonal way to the course of the SC.

The proposed automatic segmentation method showed excellent precision in terms of inter- and intra-session reproducibility and was superior to the manual segmentation performed by experienced raters for all SC metrics, as measured by both CV and DSC. Our automatic method was also superior in terms of HD for total SC, although it did not differ with regard to SC WM and GM. At the same time, accuracy of the automatic method was high for total SC, GM and WM, as measured by both DSC and HD. Comparing the present data to results of the SC GM segmentation challenge ²⁴, we achieve a superior mean GM DSC of 0.86 vs 0.80 performed by the “deepseg” algorithm in the SC GM segmentation challenge dataset. This could be potentially explained by the high quality of the AMIRA images and/or the use of a multi-center dataset within the challenge with results from various MRI sequences and segmentation algorithms. Application of the previously published iterative non-local STAPLE algorithm on our AMIRA images showed higher accuracy than the original work of Asman et al. ¹¹⁵ (SC GM: median DSC of 0.82 vs 0.75, median HD 1.04mm vs 2.5mm), which was performed on T2* weighted 3D gradient echo images. While our atlases were constructed from a pool of around 800 samples, Asman et al. had around 2000 available slices. Thus, the better accuracy seen here can be explained in terms of a possible higher image quality in AMIRA images compared to T2* weighted 3D gradient echo images, however a direct comparison of MR sequences within the same subjects was not performed. The proposed algorithm's shallow architecture with only a few parameters may make it less prone to overfitting to the training set, compared to a state-of-the-art deep neural network. However, a direct comparison of our method with the recently presented deep learning methods ¹¹⁷ was not possible in this study.

Precision and accuracy of our automatic method was decreasing in the order total SC→WM→GM. This may be caused by the accordingly decreasing size of WM and GM compared to total SC, since small differences may be translated into a larger variance. Moreover, the more complex geometry of the GM and WM compared to total SC may be more prone to misclassification errors. Finally, despite the good image quality, signal contrast was stronger for SC/CSF compared to GM/WM, which in turn could have partly contributed to differences in total SC and GM segmentation. Moreover, a slightly lower reproducibility and accuracy of our measurements in more caudally acquired slices could also be identified, which may reflect decrease in contrast intensity and “noise” increase in AMIRA images acquired closer to the lungs and surrounded by overall greater body mass (thorax, shoulders and arms) compared to the more rostral cervical SC.

Our automatic method also displayed significantly lower intra-session than inter-session variability for all SC metrics. However, GM intra- and inter-session CVs were similar with mean values ranging between 4.10–4.77%. Accordingly, our method demonstrated similar mean intra- and inter-session CVs between 2.54-2.95% for WM. We therefore conclude that patient repositioning only slightly influences GM and WM area measurements, providing evidence for the suitability of our automatic segmentation method in longitudinal settings.

In our work, minimal contrast adjustment differences in our manual segmentation led to a marked decrease of reproducibility, especially in GM area quantifications, as shown in the manual intra- and inter-rater measurements (mean CV up to 19.18%). Since the proposed method is fully automatic and requires no user-software interaction, it is devoid of additional variation produced by intra- and inter-rater variability. Therefore, our method provides significant advantages in big datasets or multi-center studies and as mentioned above may also be valuable in the longitudinal evaluation of individual patients (e.g. multiple sclerosis patients).

Compared to the manual reference standard, the automatic method slightly overestimated total SC and WM area, while underestimating the GM area. This might be due to different intensity-thresholding in the manual segmentation compared to automatic method. The caudal GM area increase shown in Figure 16 can be explained by the cervical SC enlargement's increased volume of motor cells in the GM ventral horns, which innervate the upper limb muscles.

Although a fully automatic segmentation was not feasible on 12% of acquired SC slices, a semi-automatic approach with manual total SC segmentation and fully automatic GM and WM segmentation could be performed on those slices. This approach also showed high accuracy measurements with mean DSC ≥ 0.83 in both GM and WM. However, compared to the fully automatic method on the initially non-discarded slices, a slight accuracy decrease was observed, which could be interpreted in terms of a lower image quality of those AMIRA acquisitions. Nevertheless, these results demonstrate a relative robustness of our automatic approach even in MR-images of suboptimal quality, which are a rather common phenomenon in clinical routine.

The present work focuses on SC GM and WM segmentation using AMIRA images of healthy controls. Nevertheless, the motivation of our research is to deploy this method in patient data (e.g. multiple sclerosis patients) for the development of a potential widely applied MRI biomarker. Exemplary segmentation of multiple sclerosis patient data (not shown in detail here) showed that lesion appearance was similar to GM and therefore challenged the algorithm where lesions did not respect the GM boundaries (see Figure 19). In future work we intend to adjust the current method and/or apply different approaches (e.g. deep learning methods ²⁸) to address the current limitations.

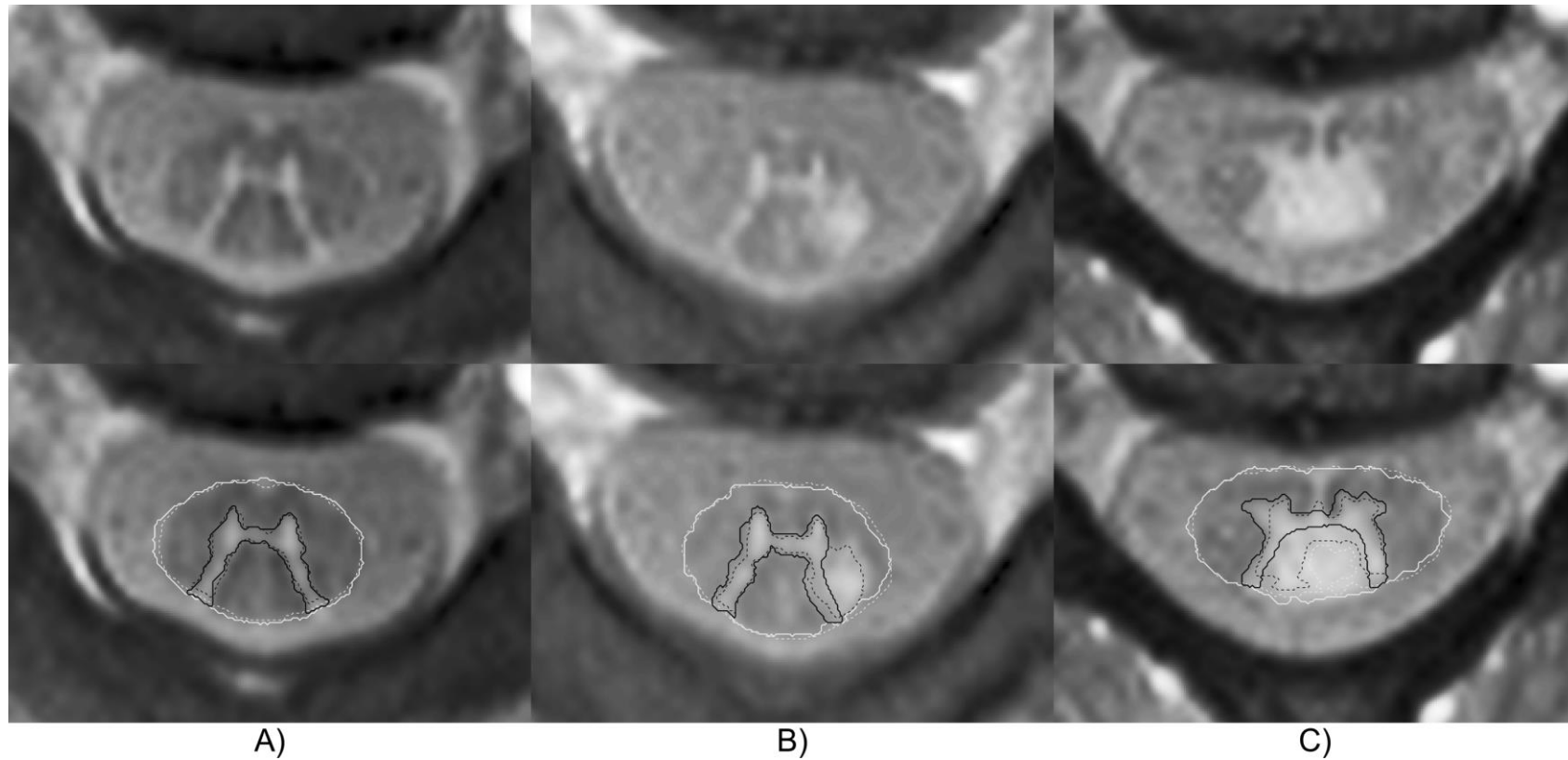


Figure 19. Examples of segmentations of representative multiple sclerosis (MS) patients. Thick continuous line: manual reference standard, dashed line: automatic segmentation. A) 54-year old female MS patient. Rostral cervical spinal cord (SC) slices of the C1/C2 level without focal lesions. Automatic segmentation highly corresponds to the manual reference standard. B) 32-year old male MS patient. Rostral cervical SC slice of the C2 level with a focal posterolateral lesion fused with the left posterior grey matter (GM) horn. Automatic segmentation misclassifies the focal lesion as SC GM. C) 33-year old female MS patient. Cervical SC slice of the C3/C4 level with a focal posterior lesion fusing with the posterior SC GM horns and the central SC GM commissure. Automatic segmentation misclassifies the focal lesion as SC GM and CSF.

Conclusion

In conclusion, the AMIRA sequence presented as a time-efficient and reproducible MR imaging approach within the cervical cord. Our fully automatic segmentation method for SC GM and WM demonstrated further high reproducibility and accuracy. We were able to show that a shallow algorithm produces state-of-the-art GM-WM segmentation results on the AMIRA data. It is therefore suitable in large longitudinal studies investigating upper cervical SC volumes. Reproducibility measures of this work could be further used for effect size calculations of SC compartment metrics for studies using the same processing approach. In future work we will address the use of deep learning approaches, as demonstrated in recent studies ¹¹⁷.

Acknowledgements

We would like to thank the Swiss National Science Foundation (Grant number: 320030_156860) and Foundation for Sponsorship of Gastroenterological and General Clinical Research as well as of Medical Imaging (Application ID 02/2015) for the financial support. The study was approved by the "Ethikkommission Nordwest- und Zentralschweiz EKNZ".

3.3. Spinal Cord Volume Loss: A Marker of Disease Progression in Multiple Sclerosis

Charidimos Tsagkas, MD¹⁺², Stefano Magon, Ph.D.¹⁺², Laura Gaetano, Ph.D.¹⁺², Simon Pezold, Ph.D.³, Yvonne Naegelin, MD¹, Michael Amann, Ph.D.¹⁺²⁺⁴, Christoph Stippich, MD⁴, Philippe Cattin, Ph.D.³, Jens Wuerfel, MD², Oliver Bieri, Ph.D.⁵, Till Sprenger, MD¹⁺⁶, Ludwig Kappos, MD¹, Katrin Parmar, MD¹

1. Department of Neurology, University Hospital Basel, University of Basel, CH
2. Medical Image Analysis Center (MIAC AG), Basel, CH
3. Department of Biomedical Engineering, University of Basel, CH
4. Division of Diagnostic and Interventional Neuroradiology, Department of Radiology, University Hospital Basel, University of Basel, CH
5. Division of Radiological Physics, Department of Radiology, University Hospital Basel, University of Basel, CH
6. Department of Neurology, DKD HELIOS Klinik Wiesbaden, Germany

Author Contributions

CT: Study concept and design, image analysis, statistical analysis and interpretation of the data, draft, revision and final approval of manuscript.

SM: Study concept and design, statistical analysis and interpretation of the data, revision and final approval of manuscript.

LG: Image analysis, statistical analysis, final approval of manuscript.

SP: Methods development and data analysis, final approval of manuscript.

YN: Acquisition of clinical data, final approval of manuscript.

MA: Acquisition of data, Image analysis, final approval of manuscript.

CS: Final approval of manuscript.

PC: Methods development and data analysis, final approval of manuscript.

OB: Final approval of manuscript.

JW: Final approval of manuscript.

TS: Study concept and design, interpretation of the data, draft, revision and final approval of manuscript.

LK: Study concept and design, acquisition of clinical data, interpretation of the data, revision and final approval of manuscript.

KP: Study concept and design, acquisition of clinical data, image analysis, statistical analysis and interpretation of the data, draft, revision and final approval of manuscript.

Conflicts of Interest

C. Tsagkas, S. Pezold, Y. Naegelin, M. Amann, C. Stippich, P. Cattin, O. Bieri: Nothing to report.

S. Magon: Travel support from Biogen.

L. Gaetano: Novartis advisory board.

J. Wuerfel: CEO of MIAC AG, Basel, Switzerland; speaker honoraria (Bayer, Biogen, Novartis, Teva); advisory boards and research grants (Biogen, Novartis); supported by the German Ministry of Science (BMBF/KKNMS) and German Ministry of Economy (BMWi).

T. Sprenger: The current and previous employers of TS have received compensation for his serving on scientific advisory boards or speaking fees from Novartis, ATI, Electrocore, Sanofi Genzyme, Actelion, Jansen, Teva, Mitsubishi Pharma Europe and Biogen Idec.

L. Kappos: Author's institution (University Hospital Basel) has received in last 3 years and used exclusively for research support: steering committee, advisory board, and consultancy fees (Actelion, Addex, Bayer HealthCare, Biogen Idec, Biotica, Genzyme, Lilly, Merck, Mitsubishi, Novartis, Ono Pharma, Pfizer, Receptos, Sanofi, Santhera, Siemens, Teva, UCB, XenoPort); speaker fees (Bayer HealthCare, Biogen Idec, Merck, Novartis, Sanofi, Teva); support of educational activities (Bayer HealthCare, Biogen, CSL Behring, Genzyme, Merck, Novartis, Sanofi, Teva); royalties (Neurostatus Systems GmbH); grants (Bayer HealthCare, Biogen Idec, European Union, Merck, Novartis, Roche Research Foundation, Swiss MS Society, Swiss National Research Foundation).

K. Parmar: Travel support from Novartis Switzerland unrelated to this work.

Abstract

Objective: Cross-sectional studies have shown that spinal cord volume (SCV) loss is related to disease severity in multiple sclerosis (MS). However, long-term data are lacking. Aim was to evaluate SCV loss as a biomarker of disease progression in comparison to other MRI-measurements in a large cohort of relapse-onset MS-patients with 6-year follow-up.

Methods: The upper-cervical SCV, the total brain volume (TBV) and the brain T2 lesion volume (T2LV) were measured annually in 231 MS-patients (180 relapsing-remitting (RRMS) and 51 secondary progressive (SPMS)) over 6 years on 3D T1-weighted MPRAGE images. Expanded disability status scale (EDSS) and relapses were recorded at every follow-up.

Results: Patients with SPMS had lower baseline SCV ($p < 0.01$), but no accelerated SCV loss compared to those with RRMS. Clinical relapses were found to predict SCV loss over time ($p < 0.05$) in RRMS. Further, SCV loss, but not TBV and T2LV, was a strong predictor of EDSS worsening over time ($p < 0.05$). The mean annual rate of SCV loss was the strongest MRI-predictor for the mean annual EDSS change of both RRMS and SPMS separately, while correlating stronger in SPMS. Every 1% increase of the annual SCV loss rate was associated with an extra 28% risk-increase of disease progression in the following year in both groups.

Conclusion: SCV loss over time relates to the number of clinical relapses in RRMS, but does overall not differ between RRMS and SPMS. SCV proved to be a strong predictor of physical disability and disease progression, indicating that SCV may be a suitable marker for monitoring disease activity and severity.

Abbreviations

ANCOVA: analysis of covariance

EDSS: expanded disability status scale

LMER: linear mixed effects regression

MLR: multiple linear regression

MPRAGE: magnetization-prepared rapid gradient-echo images

MS: multiple sclerosis

RRMS: relapsing-remitting multiple sclerosis

SC: spinal cord

SCV: spinal cord volume

SPMS: secondary progressive multiple sclerosis

TBV: total brain volume

T1w: T1-weighted

T2LV: T2 lesion volume

T25fwt: timed 25-foot walk test

Introduction

Multiple sclerosis (MS) is an inflammatory disease of the brain and the spinal cord (SC) that leads to demyelination and neurodegeneration. Atrophy is considered to be the consequence of neurodegeneration in MS and can be measured in vivo using magnetic resonance imaging (MRI) as a reduction of central nervous tissue volume^{88,118–121}.

SC abnormalities have been observed in up to 83% of MS patients, with 60% occurring in the cervical region^{5,77,78}. SC lesions are of diagnostic as well as prognostic importance in MS⁷⁹. However, previous studies were inconsistent regarding the correlation between lesional SC abnormalities and MS-related disability^{45,78,88}. In contrast, SC volume (SCV) or cross-sectional area measurements have indicated a stronger and more consistent relationship to disability in MS^{5,27,44,45,95,97,121–123}. Indeed, SCV loss has been shown to be more extensive in progressive forms of the disease^{44,94,95}. It mainly reflects a diffuse process, which seems at best weakly related to focal brain and SC MS lesions^{88,90}.

Despite the interest raised by cross-sectional and short-term follow-up volumetric SC studies, there is a lack of larger scale longer-term longitudinal studies on SCV loss in MS. This is mainly because the SC is infrequently included in imaging protocols, because of cost, time restrictions, technical difficulties in acquiring high-quality MR-images and in quantifying SC metrics in a reliable and time-efficient way^{3,124}. As a result, there is a knowledge gap on the dynamic changes of SC measures and their association with the patient's clinical picture over time.

We hypothesized that SCV loss could be strongly associated with clinical changes over time, contributing stronger to disease progression than brain MRI measurements. Aim of this study was to evaluate SCV loss in a large cohort of relapsing-onset MS patients, which have been followed over a period of 6 years with annual clinical and MRI exams. Our goal was to evaluate SCV loss as a feasible biomarker in MS in comparison to brain MRI metrics and assess potential between-group differences in a fast, clinically easy applicable fashion using high-resolution brain and not SC scans.

Methods

Study design and participants

Clinical and MRI data of relapsing-remitting (RR) and secondary-progressive (SP) MS patients of an ongoing large scale cohort study⁷⁷ (240 relapse onset MS patients in total) from a single centre (MS Center, University Hospital, Basel) were analysed. Patients were followed over a maximum of 6 years (7 annual time points).

The diagnosis of MS was made in accordance with international panel established criteria¹²⁵. The local ethics committee approved the study.

Procedures

All patients underwent a standardized neurological examination including the Expanded Disability Status Scale (EDSS; www.neurostatus.org) by trained and certified examiners and timed 25-foot walk test (T25fwt) annually. Definite clinical disease progression was defined according to the following conventions: 1a) an increase of 1 point in EDSS if the baseline EDSS score was ≤ 5.5 OR, 1b) an increase of 0.5 point in EDSS if the baseline EDSS score was > 5.5 AND 2) no relapse in the last 12 months. The number of relapses during the 12 months prior to every follow-up was noted with every clinical evaluation, as well as their sum from baseline to each follow-up.

All MRI scans were performed on the same 1.5T Magnetom Avanto MR scanner (Siemens Medical Solutions, Erlangen, Germany). Morphological analyses were performed on high-resolution three-dimensional T1-weighted (T1w) magnetization-prepared rapid gradient-echo (MPRAGE) brain MRI scans acquired in sagittal orientation (TR/TI/TE = 2080/1100/3.0 ms; $\alpha = 15^\circ$, 160 slices, resolution: $0.98 \times 0.98 \times 1 \text{ mm}^3$), covering the upper cervical SC.

SCV analysis was performed using an established semi-automatic software (CORDIAL), which allows a fast and reliable segmentation and volumetry of the SC with minimal user interaction^{46,102}. The segmentation was carried out over a 35mm long SC segment, starting 27mm below the cisterna pontis, which corresponds roughly to the SCV between the Foramen magnum and the C2/C3 intervertebral disc (Figure e-1). Segmentations were visually inspected for quality and excluded from further statistical analysis in case of segmentation errors.

Total brain volume (TBV) was computed for each patient from the T1w-images with the fully automated tool “Structural Image Evaluation, using Normalization, of Atrophy” for cross-sectional studies (SIENAX version 2.6)¹²⁶. The SIENAX volume-correction factor was used for normalizing the TBV regarding variations of head size. All analyses were performed on these corrected volumes.

All brain white matter lesions (T2LV) were segmented by trained expert observers according to the standard operating procedures used at the local institution for the analysis of clinical phase II and phase III trial as described before¹²⁷.

Statistical analysis

The mean annual SCV loss rate, the mean annual TBV loss rate and the mean rate of annual T2LV increase over 6 years was determined for every patient as the average of the annualized changes between all available time points. In order to approximate a normal distribution, logarithmic (EDSS) and inverse transformations (T25fwT) were performed. The annual EDSS and annual T25fwT changes were calculated based on the transformed scores. See Appendix 1 (incl. Figure e-2) for exact calculation formulas of mean annual rates and changes.

Comparisons of baseline demographic factors, clinical measurements and number of follow-ups between subtypes were made using Welch's and Pearson's chi-squared test with Yate's continuity correction. Between-group differences in regard to baseline MRI measures and annual rates were performed using analyses of covariance (ANCOVA), while correcting for age, sex and disease duration.

Hierarchical multiple linear regression (MLR) analyses were performed to investigate the associations between annual rates of MRI-metrics and annual changes of clinical scores, in a backwards stepwise fashion. Independent variables were entered blockwise keeping the following sequence: first demographics and clinical factors, then SCV and finally brain metrics. Baseline MRI and clinical measures were always entered into the model as correction factors to the respective annual rates, irrespective of whether they reached levels of statistical significance. All other independent variables that did not reach levels of statistical significance were not included in the final model.

Linear mixed effects regression (LMER) analyses were deployed to explore the longitudinal correlations between clinical and MRI measures in a forward stepwise fashion, using a "random intercept" and a "random slope" to allow for within-subject and between-subject variance. Each factor was tested both for its contribution to the fit's intercept as well as to the fit's slope. The fit's intercept corresponds to the average of the dependent variable, whereas the fit's slope to the change of the dependent variable over time. Again, independent variables were entered blockwise in the above-mentioned sequence (see MLR). All independent variables without statistical significance were excluded from the final model.

In order to evaluate the prediction capabilities of MRI on disease progression and time-to-disease progression, a Cox analysis in a backwards stepwise fashion was performed. In patients showing disease progression the annual rates were recalculated for the time period between baseline and time of progression. A 10-fold cross-validation was performed and the corrected concordance probability index (c-index) was assessed. The contribution of each factor in the final model was assessed using the proportion of chi-squared.

All statistical analyses were performed using R Version 3.2.3 (<https://www.r-project.org/>).

Results

Out of 1180 available MRI datasets of 240 relapse onset patients, 95 datasets (8%; including the complete MRI series of 9 patients) were excluded from further analysis due to segmentation errors or image artefacts. This resulted in a total of 231 MS patients analysed (180 RRMS and 51 SPMS), which were followed on average over 5.1 ± 1.99 years (Figure e-3). Demographics and basic clinical characteristics are described in Table 4. Corrected baseline MRI metrics, annual rates, and between group comparisons are shown in Table 5.

Table 4. Demographics and basic clinical characteristics of patients with MS

Characteristics	Overall	RRMS	SPMS	p-value
Number of patients	231	180	51	
Baseline Age (y)				<0.001
<i>Mean ± SD</i>	44.7 ± 11.2	41.5 ± 10.1	55.4 ± 7.6	
<i>Range</i>	19 – 67	19 – 65	38 – 67	
Sex (Female/Male)	160 / 71	133 / 47	27 / 24	<0.01
Baseline Disease Duration (y)				<0.001
<i>Mean ± SD</i>	13.1 ± 9.2	11.4 ± 8.4	19.1 ± 9.7	
<i>Range</i>	0 – 47	0 – 40	1 – 47	
Baseline EDSS				<0.001
<i>Median</i>	3.0	2.5	4.5	
<i>Range</i>	0 – 7.5	0 – 7.5	1.5 – 7.5	
Baseline T25fwt z-score				<0.001
<i>Mean ± SD</i>	0 ± 1	0.26 ± 0.86	-0.98 ± 0.90	
<i>Range</i>	-2.69 – 3.90	-2.69 – 3.90	- 2.67 – 1.02	
Number of Follow-ups				0.38
<i>Mean ± SD</i>	5.10 ± 1.99	5.16 ± 1.99	4.88 ± 2.00	
<i>Range</i>	1 – 7	1 – 7	1 – 7	

SD=Standard Deviation; EDSS=Expanded Disability Status Scale; T25fwt=Timed 25-Foot Walk Test; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS

Between-group comparisons were performed using Welch's two sample t-test and Pearson's chi-squared test with Yate's continuity correction where appropriate.

Table 5. *Adjusted MRI measures of patients with MS*

MRI measure	Overall	RRMS	SPMS	p-value
Baseline SCV (mm³)				
<i>Mean ± SD</i>	2380 ± 167	2416 ± 148	2253 ± 170	<0.001
<i>Range</i>	1883 – 2724	2121 – 2724	1883 – 2499	
Annual SCV Rate (%/y)				
<i>Mean ± SD</i>	-0.43 ± 0.20	-0.38 ± 0.17	-0.62 ± 0.17	0.24
<i>Range</i>	-0.95 – 0.04	-0.73 – 0.04	-0.95 – (-0.31)	
Baseline TBV (cm³)				
<i>Mean ± SD</i>	1489 ± 52	1498 ± 46	1433 ± 40	<0.001
<i>Range</i>	1340 – 1595	1355 – 1595	1340 – 1543	
Annual TBV Rate (%/y)				
<i>Mean ± SD</i>	-0.43 ± 0.12	-0.41 ± 0.11	-0.50 ± 0.13	0.41
<i>Range</i>	-0.75 – (-0.17)	-0.69 – (-0.17)	-0.75 – 0.17	
Baseline T2LV (mm³)				
<i>Mean ± SD</i>	6287 ± 1720	5959 ± 1515	7437 ± 1910	0.18
<i>Range</i>	2647 – 12340	2647 – 10250	7437 – 12340	
Annual T2LV rate (%/y)				
<i>Mean ± SD</i>	3.48 ± 2.53	4.34 ± 1.95	0.41 ± 1.91	<0.05
<i>Range</i>	-4.86 – 8.07	-2.90 – 8.07	-4.86 – 4.61	

SCV= Spinal Cord Volume; SD=Standard Deviation; TBV= Total Brain Volume; T2LV= Brain T2 Lesion Volume; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS

Baseline SCV, TBV and T2LV were adjusted for age, disease duration and gender. Annual SCV rate, annual TBV rate and annual T2LV rate were adjusted for age, disease duration, gender and baseline SCV/TBV/T2LV respectively. Adjusted values and significance of difference between RRMS and SPMS were obtained through analysis of covariance (ANCOVA).

SCV changes

In a first series of statistical analyses, SCV and its change over time was evaluated with respect to demographic, clinical and brain MRI metrics using LMER. The analyses revealed that men had larger average volumes than females (Males: $B= 289$, $p<0.001$), and a significantly faster SCV loss over time (Males: $B= -6.24$, $p<0.01$). Age ($B= 3.36$, $p<0.01$) and disease duration ($B= -7.94$, $p<0.001$) correlated with the average SCV but not with the SCV loss over time. Moreover, TBV ($B= 6 \times 10^{-4}$, $p<0.001$) and T2LV ($B= 2 \times 10^{-3}$, $p<0.05$) correlated with the average SCV, but not with the SCV loss over time. SPMS on average had lower SC volumes (SPMS: $B= -175$, $p<0.01$) than RRMS but not an accelerated SCV loss over time. Overall, the full model accounted for 98.7% of SCV variance, while the fixed effects alone accounted for 32.7%.

The sum of clinical relapses in RRMS patients was associated with SCV loss over time ($B= -1.06$, $p<0.05$), but not with the average SCV. Disease duration ($B= -9.58$, $p<0.001$) was associated with the average SCV but not with the SCV loss over time. Age did not correlate with the SCV. The final model accounted for 98.7% of SCV variance, while the fixed effects alone accounted for 22.3% of the SCV variance.

SCV changes and disability

In a second step, we were interested in whether measures of disability (EDSS and T25fwT) are associated with SCV and brain MRI metrics and their changes over time.

EDSS

In the LMER analyses, SCV was inversely associated with the average EDSS ($B= -4 \times 10^{-4}$, $p<0.001$) and the EDSS worsening over time ($B= 2.1 \times 10^{-5}$, $p<0.05$). TBV and T2LV were also correlated with the average EDSS ($B= -2.2 \times 10^{-7}$, $p<0.05$; $B= 8.8 \times 10^{-6}$, $p<0.001$ respectively), but not with the EDSS worsening over time. Finally, sex (Males: $B= 0.13$, $p<0.05$), age ($B= 5.6 \times 10^{-3}$, $p<0.001$), disease duration ($B= 2.8 \times 10^{-3}$, $p<0.01$), and disease subtype (SPMS: $B= 0.35$, $p<0.001$) correlated with the average EDSS, but not with the EDSS worsening over time. The final model accounted for 87.7% of EDSS variance, while the fixed effects alone accounted for 42.6%.

Separate MLR analyses were performed to specifically investigate the extent of correlation between EDSS and SCV in each disease type. In RRMS the annual SC rate was the only predictive MRI factor, while age and baseline EDSS contributed significantly. The final model demonstrated a weak prediction of the annual EDSS change variance, with an adjusted- R^2 of 19.2%. Similar to RRMS the annual SC rate was the most important predicting factor for the annual EDSS change in SPMS, while again age and baseline EDSS also contributed

significantly. The final model demonstrated, however, a moderate prediction of the annual change in EDSS, with an adjusted- R^2 of 53.8%. The models are described in detail in Table 6.

Table 6. Multivariate analysis of clinical measures by disease type.

Group	Final Model	Variable	β	p-value	ΔR^2	
RRMS	Annual EDSS change ~ Age + Baseline EDSS + Baseline SCV + Annual SCV Rate Adj-R ² =19.19% P<0.001	Age	0.25	<0.001	17.38%, p<0.001	
		Sex	-	n.s.		
		Disease Duration	-	n.s.		
		Baseline EDSS	-0.51	<0.001		
		Baseline SCV	-0.10	n.s.	1.81%, p<0.05	
		Annual SCV Rate	-0.16	<0.05		
		Baseline TBV	-	n.s.		-
		Annual TBV Rate	-	n.s.		-
		Baseline T2LV	-	n.s.		-
		Annual T2LV Rate	-	n.s.		-
SP	Annual T25fwt change ~ Baseline T25fwt + Baseline SCV + Annual SCV Rate + Baseline TBV Adj-R ² =13.87% P<0.001	Age	-	n.s.	0.33%, p=0.23	
		Sex	-	n.s.		
		Disease Duration	-	n.s.		
		Baseline T25fwt	-0.08	n.s.		
		Baseline SCV	0.07	n.s.	11.51%, p<0.001	
		Annual SCV Rate	0.35	<0.001		
		Baseline TBV	0.18	<0.05		2.03%, p<0.05
		Annual TBV Rate	-	n.s.		-
		Baseline T2LV	-	n.s.		-
		Annual T2LV Rate	-	n.s.		-
S		Age	-0.27	<0.05	18.25%, p<0.001	

	<p><i>Annual EDSS change ~ Age + Baseline EDSS + Baseline SCV + Annual SCV Rate</i></p> <p><i>Adj-R² = 53.83 %</i></p> <p><i>P < 0.001</i></p>	Sex	-	n.s.	35.58%, p < 0.001	
		Disease Duration	-	n.s.		
		Baseline EDSS	-0.23	<0.05		
		Baseline SCV	0.03	n.s.		
		Annual SCV Rate	-0.60	<0.001		
		Baseline TBV	-	-		-
		Annual TBV Rate	-	-		-
		Baseline T2LV	-	-		-
		Annual T2LV Rate	-	-		-
	<p><i>Annual T25fwt change ~ Baseline T25fwt + Baseline SCV + Annual SCV Rate</i></p> <p><i>Adj-R² = 49.89 %</i></p> <p><i>P < 0.001</i></p>	Age	-	n.s.	8.38%, p < 0.01	
		Sex	-	n.s.		
		Disease Duration	-	n.s.		
		Baseline T25fwt	0.06	n.s.		
		Baseline SCV	-0.26	<0.05		
		Annual SCV Rate	0.65	<0.001		41.51%, p < 0.001
		Baseline TBV	-	n.s.		-
		Annual TBV Rate	-	n.s.		-
		Baseline T2LV	-	n.s.		-
Annual T2LV Rate	-	n.s.	-			

SCV=Spinal Cord Volume, TBV=Total Brain Volume, T2LV=Brain T2 Lesion Volume; EDSS=Expanded Disability Status Scale; T25fwt=Timed 25-Foot Walk Test; RRMS=Relapsing Remitting MS, SPMS=Secondary Progressive MS; B=Standardised Regression Coefficients; ΔR^2 =Adjusted R-Squared Difference; N.S.=Not Significant

Timed 25 foot walk test

In the LMER analyses, SCV as well as TBV was inversely associated with the average T25fwt ($B= 3.7 \times 10^{-5}$, $p<0.001$; $B= 6.8 \times 10^{-8}$, $p<0.01$ respectively), but not with its worsening over time. T2LV correlated with both, the average T25fwt ($B= -1.4 \times 10^{-6}$, $p<0.01$) and its worsening over time ($B= -2.2 \times 10^{-7}$, $p<0.05$). Age ($B= -1 \times 10^{-3}$, $p<0.001$), disease duration ($B= -6.2 \times 10^{-4}$, $p<0.001$), and disease subtype (SPMS: $B= -5.7 \times 10^{-2}$, $p<0.001$), but not sex, correlated with the average T25fwt. Only sex (Males: $B= -2.9 \times 10^{-3}$, $p<0.01$) and disease type (SPMS: $B= -4.5 \times 10^{-3}$, $p<0.05$) correlated with the T25fwt worsening over time. The final model accounted for 94.2%, while the fixed effects alone accounted for 40% of the T25fwt variance.

Separate MLR analyses were performed to specifically investigate the extent of correlation between T25fwt and SCV in each disease type. In the RRMS group, the annual SC rate was the most important predicting factor, while baseline TBV also contributed significantly. The final model demonstrated a weak prediction of the annual T25fwt change variance, with an adjusted- R^2 of 13.9%. In terms of the annual T25fwt changes in SPMS patients, the annual SC rate was the most important predicting factor, while baseline SCV also contributed significantly. The final model demonstrated a moderate prediction of the annual change in T25fwt, with an adjusted- R^2 of 49.9%. The models are described in detail in Table 6.

Prediction of disease progression

Last, the predictive capabilities of MRI metrics on disease progression and time-to-disease progression were evaluated. A Cox analysis showed that disease type, baseline EDSS, annual rate of SCV loss and baseline TBV were significant predictors of disease progression. The annual rate of SCV loss explained 34% of the final model's chi-squared and was the strongest MRI measure. Its hazard ratio was 0.72 (95% confidence interval [CI]: 0.61 – 0.84, $p<0.001$). Every 1% increase of the annual rate of SCV loss was associated with an extra 28% risk increase to develop disease progression in the following year. The c-index resulting from a 10-fold cross validation of the cohort was 70%, showing a moderate predictive power of the model. Details are displayed in Table 7.

Table 7. *Cox analysis of Disease Progression*

Final Model	Variable	HR	(95% CI)	Proportion of Chi-Square	p-value
(Disease Progression, Time to Disease Progression) ~ Disease Type +Baseline EDSS + Baseline SCV + Annual SCV Rate + Baseline TBV Concordance=0.717 (se=0.04) Wald test=51.53, p<0.001 Score (logrank) test=54.16, p<0.001 Total Chi-Square=51.54, p<0.001 C-Index=0.70	Age	-	-	-	n.s.
	Sex	-	-	-	n.s.
	Disease Duration	-	-	-	n.s.
	Disease Type (SPMS)	4.62	2.46 - 8.67	43.9%	<0.001
	Baseline EDSS	0.29	0.15 - 0.58	24.3%	<0.001
	Baseline SCV	1.0008	1.0 - 1.002	6.7%	n.s.
	Annual SCV Rate	0.72	0.61 – 0.84	34%	<0.001
	Baseline TBV	1.0000	1.0000 – 1.0000	18%	<0.01
	Annual TBV Rate	-	-	-	n.s.
	Baseline T2LV	-	-	-	n.s.
Annual T2LV	-	-	-	n.s.	

aSCVLR=annualised SC volume loss rate; B=regression coefficients; HR=Hazard Ratios; 95% CI=95% Confidence Intervals; c-index=10-fold crossvalidation corrected concordance probability index; n.s.=Not Significant

Discussion

This study is the first longitudinal longer-term analysis of SCV loss in MS investigating the upper cervical SCV in a large cohort of patients with relapse-onset MS for up to 6 years of follow-up. The temporal profile of SC tissue loss is illustrated with its association to clinical changes and its potential to monitor the clinical course of the disease. For that purpose we deployed a fast, clinically easy applicable pipeline using high-resolution brain scans.

The SC segmentation was performed in a mainly automatic fashion with minimal user interaction, thus avoiding bias from manual segmentations or manual corrections. Our approach using the upper cervical SCV instead of the upper SC area is thought to be less susceptible to possible bias from focal atrophy due to lesions and possesses high test-retest reliability ¹⁰².

In line with previous cross-sectional studies ^{27,44,45,94,95,97,118,121–123}, patients with SPMS had lower SCV compared to RRMS. However, there was no evidence of accelerated SCV loss over time in SPMS compared with RRMS. The observed between-group differences in cross-sectional settings can be explained by longer disease duration and higher age of the SPMS patients (Figure 20), arguing against a theoretical preferential SCV loss in this group as suggested by previous smaller scale and shorter duration studies ^{27,44,45,94,95,97,118,121–123}. Similar findings were present regarding TBV loss when compared between groups in the present cohort. This is in line with a large study by De Stefano et al. showing that there is no difference in the rate of brain volume loss in earlier versus more advanced MS phenotypes on a group level ¹¹⁹. Overall, these results indicate that the rate of tissue loss in the central nervous system is in general comparable in relapse-onset MS phenotypes.

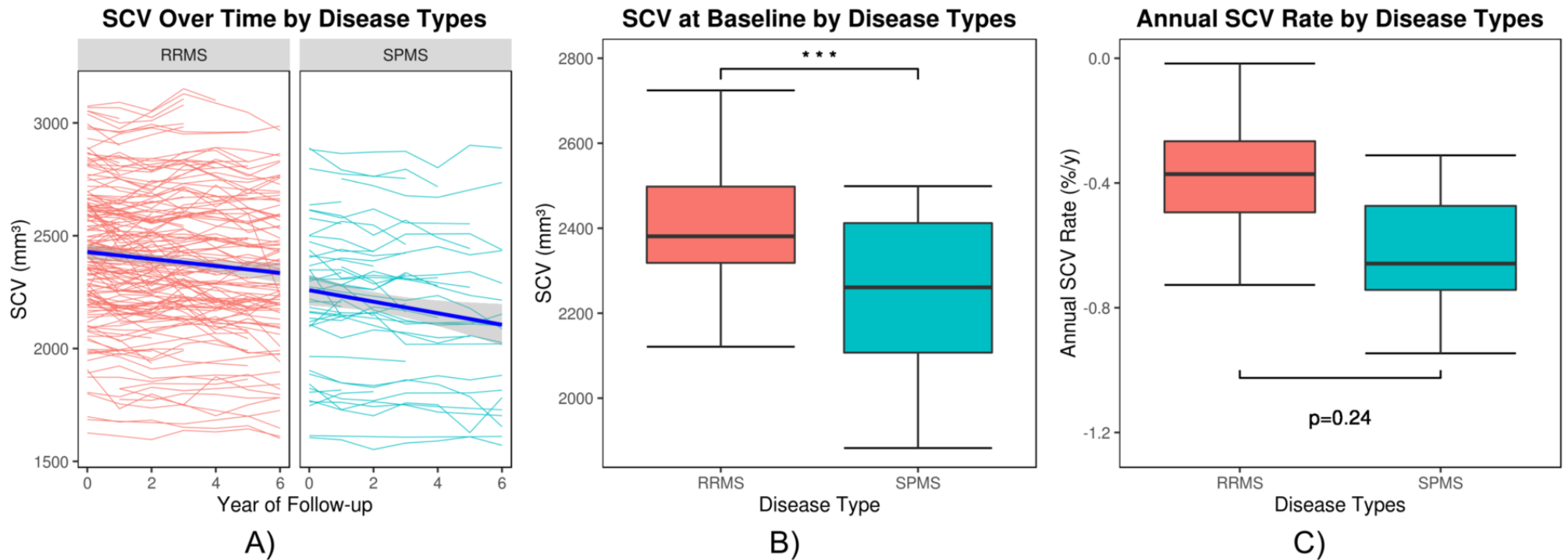


Figure 20. SCV and its change over time separated by disease type: RRMS (red) and SPMS (turquoise) patients.

(A) The fine lines represent the raw measurements of SCV for each patient over 6 years of follow-up, while the thicker blue lines depict the respective “trends” of SC volume loss in the two disease types. Notice, that despite a significant difference of SCV at baseline (B), the annual SCV rate did not differ significantly between disease types, after adjusting for age, sex and disease duration (C). The error bars (whiskers) represent the mean +/- the standard deviation.

In terms of T2LV, RRMS patients showed faster accumulation of T2LV, reflecting the inflammatory activity dominating in this phase of the disease. However, SCV loss over time seemed to be largely independent of change in TBV and T2LV.

As expected, men had larger SCV at baseline, however, exhibited more pronounced SCV loss over time, which is in line with the clinical observation that male sex is an indicator of poorer prognosis¹²⁸. Yet, our statistical analyses predicting the EDSS and disease progression could not confirm a faster clinical worsening of male patients compared to females. This may be explained by the follow up span which might not have been long enough. Neither age nor disease duration seemed to have an effect on SCV loss over time, suggesting a relatively steady SC tissue loss in both patient groups.

In RRMS, an increased cumulative number of clinical relapses was associated with faster SCV loss over time. One may speculate that with every new clinical relapse in a RRMS patient, some degenerative cascade is activated accelerating SCV loss, an effect that seems to persist way beyond the initial acute inflammatory phase of the relapse. The data further suggest that the accumulation of clinical relapses shapes - at least partially - a patient's profile in terms of SCV loss. It also implies that the reduction of clinical relapses -which constitutes a fundamental therapeutic strategy of RRMS - should beneficially affect this process¹²⁹. In the LMER, SCV proved to be the only MRI metric to strongly explain the clinical progression over time as measured by speed of EDSS worsening for the whole cohort. TBV and T2LV were associated with the average patients' disability, but did not independently relate to EDSS worsening over time, when added to a statistical model that included SCV. Regarding the T25ftw, SCV and TBV correlated with the patients walking speed although they failed to correlate with its worsening over time. Instead, the T2LV was a predictor of these changes, probably driven by the RRMS group.

Regardless of a similar SCV loss pace in both patient groups, association of the annual rate of SCV loss with clinical scores was remarkably higher for patients with SPMS than for the RRMS group in all analyses. This shows that despite the assumption of same underlying causal pathomechanisms^{83,129}, the clinical impact in SPMS is more pronounced. This can be explained by higher age and accumulation of damage over time in SPMS patients, entailing chronic immune activation, increased oxidative stress-related damage, loss of trophic support and exhaustion of repair and compensatory mechanisms^{83,84}. This allows axonal and myelin damage to be "translated" in a much more straightforward way into clinical deficits once the threshold of neuronal injury and/or repair has been exceeded. On the other hand, it is possible to hypothesize that RRMS patients still have sufficient reserves of cortical adaptation, remyelination, axonal repair and neuroprotection, which allow them on one hand to maintain or re-establish the functionality of neuronal tissue and on the other hand "mask" the produced axonal loss through neuroplasticity occurring at higher cortical centres^{237,265,266}. Considering that, in future works additional measures of neuroprotective mechanisms such as quantitative MRI measurements (e.g. magnetization transfer imaging, diffusion tensor imaging, and proton magnetic resonance spectroscopy) may shed light into this interesting aspect of the disease^{83,130,131}. The annual SCV loss rate was to a great extent the only MRI factor in predicting clinical worsening, while the annual TBV and

T2LV did not significantly contribute. Predicting disease progression using a cox analysis, it was shown that the annual rate of SCV loss was the strongest MRI measure and a 1% rate increase was associated with a 28% risk increase to develop disease progression in the following year. Further SPMS patients were shown to be more susceptible to disease progression. This is also adding to the evidence that SCV could possibly be a better disease marker compared to brain volume loss and lesion load in the brain.

This study has a number of limitations. It aimed to analyse follow-up data of a group of patients with MS in a retrospective fashion. Some patients were lost to follow-up during the study, leading to incomplete data sets and potential bias. While the SPMS patient group was smaller than the RRMS, the two groups were followed-up for a similar amount of time. The latter at least does not support a preferential loss of patients with progression. Furthermore, the lack of a representative control group of healthy subjects made it impossible to assess MS-related SC volume loss as compared to SCV decline due to normal aging. However, it is likely that the observed SCV loss is mainly a disease-related effect, as in the normal population the degree of SCV loss seems to be smaller¹⁰³. Unfortunately, no study has so far assessed the effect of age on SCV loss in healthy controls longitudinally. In this study, the potential effect of disease modifying drugs (DMD's) on inflammation, and ultimately on SCV variation was not included in our analysis. In our cohort 68% of patients were treated with DMD's at baseline including primarily first line injectables (63%). While injectables also show an effect on brain atrophy we believe that this effect is rather negligible¹³². Finally, since the study had no T2w sequence covering the area of volume measurement over all time points analysed here, the impact of SC lesions was not assessed.

In conclusion, this study indicates that the measurement of SCV loss represents a reliable imaging marker for the monitoring of disease activity and progression in MS. SCV loss was shown to be directly affected by inflammatory events (relapses), supporting at least a partial role of inflammation in driving neurodegeneration in MS. Finally, the study provides evidence of a dissociation in the clinical consequences of SC volume loss in RRMS versus SPMS. That said, SCV could be a useful endpoint in clinical trials of therapeutic agents that are aiming at the degenerative process in MS, a much wanted unmet need.

Acknowledgements

C.T. was financially supported from the Swiss National Science Foundation. K.P. is holding a grant from the Baasch-Medicus Foundation Switzerland.

We are very grateful to all participants and staff involved in the GeneMSA cohort study, in particular Alain Thoeni.

Supplementary Material

Appendix 1

The annualized percentage changes (APC) in the volume (V) of the spinal cord, the total brain mass, and the T2-weighted lesions were calculated for every patient (x) with n available follow ups at given time points $[t_1, t_2, \dots, t_n]$ (in years) with corresponding fractional annual time intervals $\Delta t_i := t_{i+1} - t_i$ (for $i = 1, 2, \dots, n - 1$) using

$$V_{APC}^{(x)}(i) := \frac{V^{(x)}(t_{i+1})/V^{(x)}(t_i) - 1}{\Delta t_i} \times 100$$

and subsequently the patient's mean annualized percentage volume change

$$\langle V_{APC}^{(x)} \rangle = \frac{1}{n-1} \sum_{i=1}^{n-1} V_{APC}^{(x)}(i)$$

In complete analogy, the annualized changes (AC) of the logarithmically-scaled expanded disability status scale (EDSS) and the inversely-scaled timed 25-foot walk test (T25fwt) were derived using

$$EDSS_{AC}^{(x)}(i) := \frac{EDSS^{(x)}(t_{i+1}) - EDSS^{(x)}(t_i)}{\Delta t_i}$$

$$T25fwt_{AC}^{(x)}(i) := \frac{T25fwt^{(x)}(t_{i+1}) - T25fwt^{(x)}(t_i)}{\Delta t_i}$$

and subsequently the patient's EDSS and T25fwt mean annualized change

$$\langle EDSS_{AC}^{(x)} \rangle = \frac{1}{n-1} \sum_{i=1}^{n-1} EDSS_{AC}^{(x)}(i)$$

$$\langle T25fwt_{AC}^{(x)} \rangle = \frac{1}{n-1} \sum_{i=1}^{n-1} T25fwt_{AC}^{(x)}(i)$$

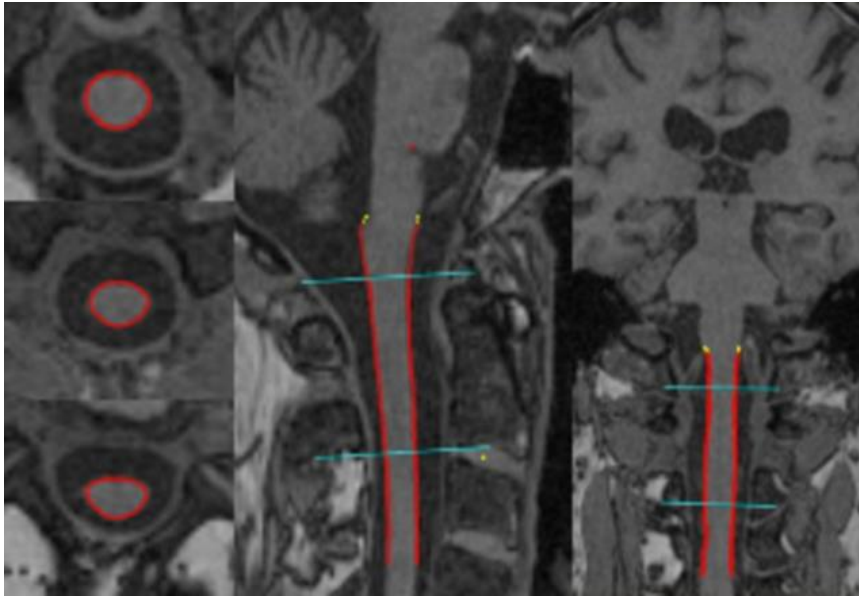


Figure e-1. Cervical cord segmentation using CORDIAL. Display of the fine-segmentation in red in axial (left), sagittal (middle) and coronal view (right). The spinal cord volume is measured between the blue lines.

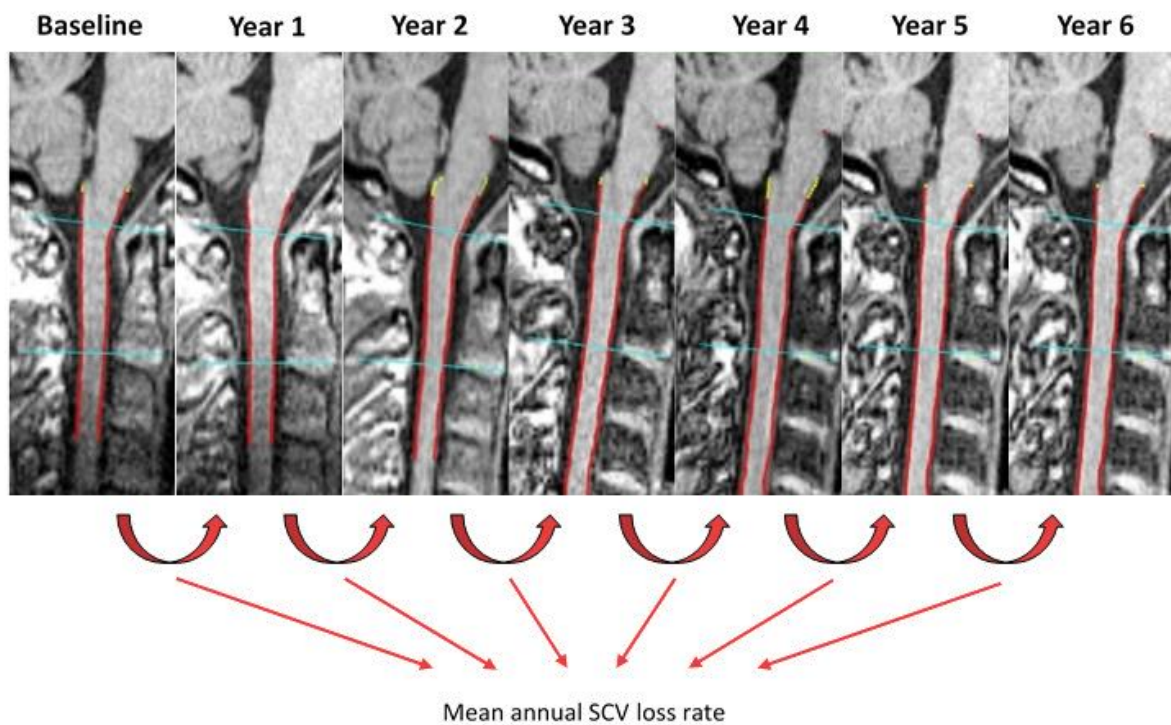


Figure e-2. Segmented MR images of one representative RRMS patient. The blue lines on each image depict the segmented volume, which corresponds to the upper cervical cord between the foramen magnum and the C2/C3 intervertebral disc. The mean annual SCV loss rate for each patient was calculated using the subsequent SC volume values from a follow-up to the next.

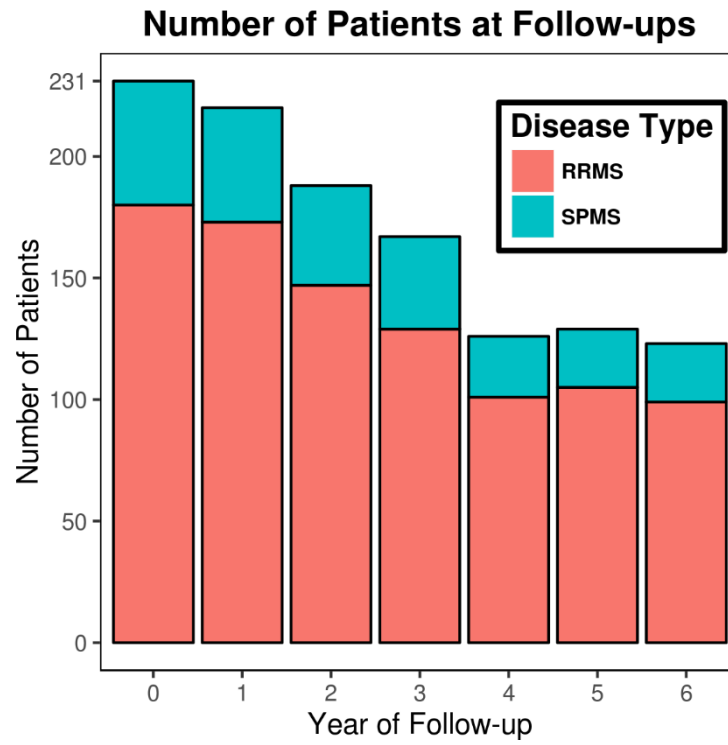


Figure e-3. Number of patients participating in each follow-up. RRMS: relapsing remitting MS; SPMS; secondary progressive MS.

3.4. Preferential Spinal Cord Volume Loss in Primary Progressive Multiple Sclerosis

Charidimos Tsagkas, MD¹⁺², Stefano Magon, Ph.D.¹⁺², Laura Gaetano, Ph.D.¹⁺², Simon Pezold, Ph.D.³, Yvonne Naegelin, MD¹, Michael Amann, Ph.D.¹⁺²⁺⁴, Prof. Christoph Stippich, MD⁴, Prof. Philippe Cattin, Ph.D.³, Jens Wuerfel, MD²⁺³, Prof. Oliver Bieri, Ph.D.⁵, Prof. Till Sprenger, MD¹⁺⁶, Prof. Ludwig Kappos, MD¹, Katrin Parmar, MD¹

1. Department of Neurology, University Hospital Basel, University of Basel, CH
2. Medical Image Analysis Center (MIAC AG), Basel, CH
3. Department of Biomedical Engineering, University of Basel, CH
4. Division of Diagnostic and Interventional Neuroradiology, Department of Radiology, University Hospital Basel, University of Basel, CH
5. Division of Radiological Physics, Department of Radiology, University Hospital Basel, University of Basel, CH
6. Department of Neurology, DKD HELIOS Klinik Wiesbaden, Germany

Abstract

Background: Little is known on longer-term changes of spinal cord volume (SCV) in primary progressive multiple sclerosis (PPMS).

Objective: Longitudinal evaluation of SCV loss in PPMS and its correlation to clinical outcomes, compared to relapse-onset MS subtypes.

Methods: A total of 60 MS age-, sex- and disease duration-matched patients (12 PPMS, each 24 relapsing-remitting (RRMS) and secondary progressive MS (SPMS)) were analysed annually over 6 years of follow-up. The upper cervical SCV was measured on 3D T1-weighted MPRAGE images using a semi-automatic software (CORDIAL), along with the total brain volume (TBV), brain T2 lesion volume (T2LV) and Expanded Disability Status Scale (EDSS).

Results: PPMS showed faster SCV loss over time than RRMS ($p < 0.01$) and by trend ($p = 0.066$) compared to SPMS. In contrast to relapse-onset MS, in PPMS SCV loss progressed independently of TBV and T2LV changes. Moreover, in PPMS, SCV was the only MRI measurement associated with EDSS increase over time ($p < 0.01$), as opposed to RRMS and SPMS.

Conclusions: SCV loss is a strong predictor of clinical outcomes in PPMS and was shown to be faster and independent of brain MRI metrics when compared to relapse-onset MS.

Introduction

Primary progressive multiple sclerosis (PPMS) is clinically characterised by sustained neurologic disability progression without early relapses or remissions¹²⁵. Other distinctive features are an older onset age, a relatively equal sex incidence and the most common presentation being a progressive paraparesis rather than visual or sensory disturbances¹³³. The spinal cord (SC) is therefore an important locus of pathology in PPMS, with a frequent occurrence of SC lesions and diffuse SC abnormalities, even in the very onset of the disease^{125,134,135}.

In contrast to brain magnetic resonance imaging (MRI) metrics, studies have shown a strong clinico-radiological correlation between SC atrophy measurements and clinical disability at all stages of the disease^{44,45,134,136–139}. In a recent study SC atrophy together with the number of asymptomatic SC lesions predicted the development of MS disability 5 years after a non-spinal clinical isolated syndrome⁸⁰. SC atrophy was further found to be an early distinctive feature between relapsing and progressive disease types (in cross-sectional analyses), predominantly seen in the grey matter, and being more pronounced in progressive MS with superior correlation to clinical disability compared to brain MRI measures^{95,140}.⁴ Reduction of SC cross-sectional area was found to occur largely independent from brain atrophy and lesion load in PPMS^{5,44,134,136,138,140,141}. Qualitative MRI studies demonstrate irreversible axonal injury and reactive gliosis in the SC of PPMS¹⁴² and even in absence of extensive SC atrophy, the structural integrity of neurons, demyelination and abnormalities in the glutamatergic pathways in the cervical cord of early PPMS show strong associations with physical disability¹¹⁸.

Former studies demonstrate controversial results concerning between-group differences not only in cross-sectional but also short-term longitudinal SC atrophy assessments. While Lukas et al.⁴⁴ could not differentiate disease types by 24-month SC atrophy rates, Cawley et al.¹⁴³ demonstrated a significant reduction of cervical cross-sectional area over 1 year in PPMS over secondary progressive MS (SPMS) patients suggesting SC atrophy as a suitable marker for clinical trials. Others, however, showed also an evident reduction of cervical cross-sectional area in the very onset of SPMS, showing a great need for more data supporting the role of spinal cord damage in PPMS¹⁴⁴.

The aim of this study was to analyse long-term changes of spinal cord volume over 6 years with focus on primary progressive MS in comparison to matched relapse-onset MS patients. We hypothesized that in PPMS the upper cervical cord volume would be the predilection site of volume loss over time when compared to whole brain and brain lesion volumes. Further the relationship between MRI measures and physical disability were investigated in a longitudinal fashion.

Methods

Study design and participants

Clinical and MRI data of 6 years of follow-up (7 annual time points) of 12 PPMS, 24 RRMS, and 24 secondary progressive (SP) MS patients of an ongoing large scale cohort study⁷⁷ (incl. 253 MS patients in total) from a single centre (MS Centre, University Hospital, Basel) were selected retrospectively. Selection of patients was based on propensity score matching, as described in the statistical analysis section. Patients' diagnosis of MS was made in accordance with international panel established criteria¹²⁵. The local ethics committee approved the study.

Procedures

All patients received annual standardized neurological examination including the Expanded Disability Status Scale (EDSS; www.neurostatus.org) and timed 25-foot walk test (T25fwt).

All MRI scans were performed on a single 1.5 T Magnetom Avanto MR scanner (Siemens Medical Solutions, Erlangen, Germany). Morphological analyses were performed on high-resolution three-dimensional T1-weighted (T1w) magnetization-prepared rapid gradient-echo images (MPRAGE) brain scans acquired in sagittal plane (TR/TI/TE = 2080/1100/3.0 ms; flip angle = 15°, 160 slices, resolution: 0.98x0.98x1mm³), covering the upper cervical SC.

SC volume (SCV) was estimated using an established semi-automatic software (CORDIAL), which allows a fast and reliable segmentation and volumetry of the spinal cord with minimal user interaction^{46,102}. The segmentation was carried out over a 35mm long SC segment, starting 27mm below the cisterna pontis, which corresponds roughly to the SC volume between the Foramen magnum and the C2/C3 intervertebral disc (Figure 21). Segmentations were visually inspected for quality and excluded from further statistical analysis in case of segmentation errors.

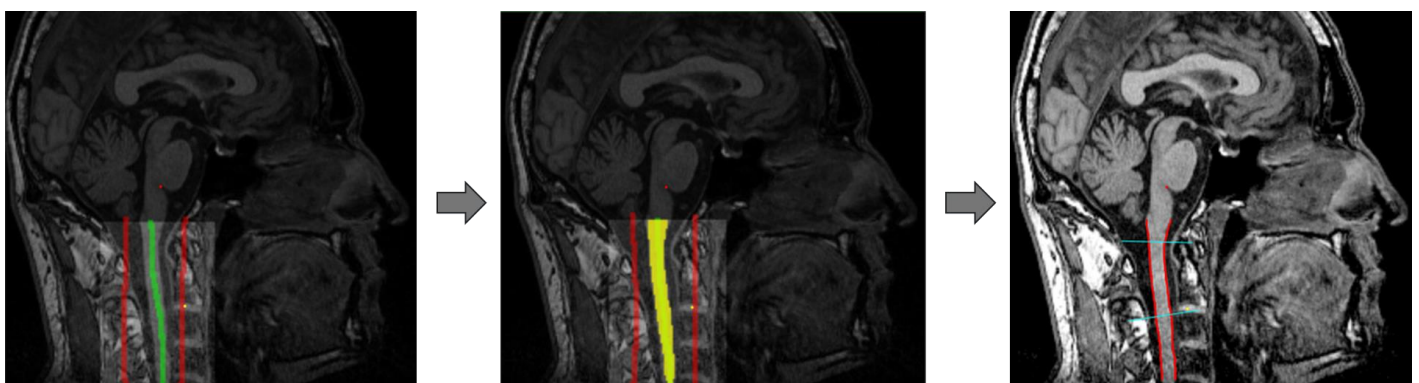


Figure 21. In the pre-segmentation step (left), the user selects manually a subset of voxels belonging to the SC (green) versus background (e.g., cerebrospinal fluid, vertebral bodies; red) and additionally an

anatomical landmark at the medullopontine sulcus (red dot). Subsequently, the segmentation (middle) is performed in a fully automatic fashion (yellow). In a last step (right), a volume is calculated in a defined distance of the anatomical landmark. For details see also Amann et al. 2016.

Total brain volume (TBV) was computed for each patient from the T1w images with the fully automated tool “Structural Image Evaluation, using Normalization, of Atrophy” for cross-sectional studies (SIENAX version 2.6) ¹²⁶. The SIENAX volume-correction factor was used for normalizing the TBV regarding variations of head size. All analyses were performed on these corrected volumes.

All brain white matter lesions (T2LV) were segmented by trained experts according to the standard operating procedures used at the local institution for the analysis of clinical phase II and phase III trials as described before ¹²⁷.

Statistical analysis

Baseline covariates, including gender, age and disease duration, were used to calculate propensity scores for each individual patient. These propensity scores were derived from a logistic regression model, in which the primary progressive disease type was the outcome variable and baseline characteristics formed the explanatory variables. RRMS and SPMS were matched with PPMS patients, based on high similarity of propensity scores, on a 2:1 basis for each group. Success of matching was assessed using analysis of variance (ANOVA) and post hoc analyses using Bonferroni correction. All further statistical analyses were performed on the matched data.

Mean annual SCV loss rate, mean annual TBV loss rate and mean annual T2LV change rate over 6 years were determined for every patient as the average change between all available time points. In order to approximate a normal distribution, logarithmic (EDSS) and inverse transformations (T25fwt) were performed.

Comparisons of baseline demographics, clinical measurements and number of follow-ups between subtypes were made using Welch’s Two Sample t-Tests and Pearson’s Chi-squared test with Yate’s continuity correction. Between-group differences in regard to baseline MRI measures and annual rates were performed using analyses of covariance (ANCOVA), while correcting for age, sex and disease duration.

Linear mixed effect regression (LMER) analyses were deployed to explore the longitudinal associations between clinical and MRI measures in a forward stepwise fashion, using a “random intercept” and a “random slope” to allow for within-subject and between-subject variance. Each variable was tested both for its contribution to the fits’ intercept as well as to the fits’ slope. In all regression LMER analyses with a clinical measurement as outcome, the independent variables were evaluated in four blocks (demographics and clinical

factors-TBV-SCV-T2LV). Multiple comparisons in post hoc analyses were performed using Bonferroni correction.

All statistical analyses were performed using R Version 3.2.3 (<https://www.r-project.org/>).

Results

Out of 253 MS patients (12 PPMS, 180 RRMS, 51 SPMS), 60 patients (all 12 PPMS, 24 RRMS, 24 SPMS) were included in the study after propensity score matching of the RRMS and SPMS to the PPMS patients. The selected cohort was monitored for a mean number of 5.15 ± 1.91 annual follow-ups (Figure 22). Out of 304 available MRI datasets, 26 (9%) were excluded from further analysis due to segmentation or image artefacts. These exclusions were evenly distributed among different disease types of our cohort; none of the selected patients had to be fully excluded. Baseline demographics and clinical characteristics are described in Table 8. Corrected baseline MRI metrics, annual rates, and between-group comparisons are displayed in Table 9 and Figure 23.

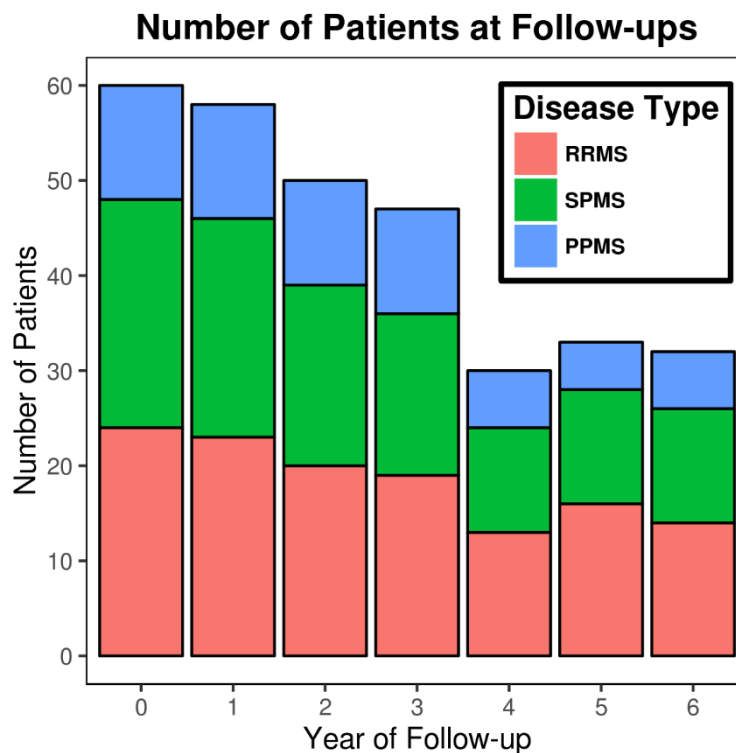


Figure 22. Number of patients participating in each follow-up. RRMS: relapsing remitting MS; SPMS: secondary progressive MS; PPMS: primary progressive MS.

Table 8. Demographics and basic clinical characteristics of patients with MS

Characteristics	Overall	PPMS	RRMS	SPMS	p-value
Number of patients	60	12	24	24	
Baseline Age (y)					n.s.
<i>Mean ± SD</i>	48.9 ± 8.2	46.7 ± 6.6	48.2 ± 9.6	50.6 ± 7.4	
<i>Range</i>	34 – 65	34 – 58	34 – 65	38 – 64	
Sex (Female/Male)	32 / 28	6 / 6	15 / 9	11 / 13	n.s.
Baseline Disease Duration (y)					n.s.
<i>Mean ± SD</i>	10.2 ± 6.6	8.6 ± 7.1	8.9 ± 5.9	12.3 ± 6.6	
<i>Range</i>	1 – 27	1 – 27	2 – 20	1 – 25	
Baseline EDSS					***, †††
<i>Median</i>	3.5	4.625	2.25	4.0	
<i>Range</i>	0 – 7.5	2.5 – 7.5	0 – 4.0	1.5 – 7.5	
Baseline T25fwt z-score					**
<i>Mean ± SD</i>	0 ± 1	-0.22 ± 1.24	0.52 ± 0.74	-0.44 ± 0.90	
<i>Range</i>	-2.07 – 1.74	-2.07 – 1.74	-0.61 – 1.68	-2.07 – 1.28	
Number of Follow-ups					n.s.
<i>Mean ± SD</i>	5.15 ± 1.91	5.25 ± 1.48	5.33 ± 2.04	4.92 ± 2.02	
<i>Range</i>	1 – 7	2 – 7	1 – 7	2 – 7	

SD=Standard Deviation; EDSS=Expanded Disability Status Scale; T25fwt=Timed 25-Foot Walk Test; PPMS=Primary Progressive MS; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS;

n.s.=not significant for any comparisons between PPMS, RRMS and SPMS

PPMS versus RRMS: † ≤ 0.05, †† ≤ 0.01, ††† ≤ 0.001.

PPMS versus SPMS: ‡ ≤ 0.05, ‡‡ ≤ 0.01, ‡‡‡ ≤ 0.001.

RRMS versus SPMS: * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 .

Between-group comparisons were performed using Welch's two sample t-test and Pearson's chi-squared test with Yate's continuity correction where appropriate.

Table 9. Corrected MRI measures of patients with MS

MRI measure	Overall	PPMS	RRMS	SPMS	p-value
Baseline SCV (mm³)					n.s.
<i>Mean ± SD</i>	2379 ± 144	2247 ± 116	2465 ± 110	2359 ± 133	(RRMS vs PPMS: 0.08)
<i>Range</i>	2108 – 2655	2108 – 2435	2308 – 2655	2133 – 2565	
Annual SCV Rate (%)					
<i>Mean ± SD</i>	-0.91 ± 0.64	-1.89 ± 0.34	-0.35 ± 0.33	-0.98 ± 0.27	†
<i>Range</i>	-2.38 – 0.28	-2.38 – (-1.1)	-0.78 – 0.33	-1.58 – (-0.55)	
Baseline TBV (cm³)					n.s.
<i>Mean ± SD</i>	1468 ± 46	1473 ± 43	1483 ± 46	1451 ± 42	
<i>Range</i>	1378 – 1562	1378 – 1535	1390 – 1562	1396 – 1561	
Annual TBV Rate (%)					
<i>Mean ± SD</i>	-0.45 ± 0.18	-0.44 ± 0.14	-0.33 ± 0.12	-0.56 ± 0.18	n.s.
<i>Range</i>	-0.85 – (-0.05)	-0.62 – (-0.15)	-0.54 – (-0.05)	-0.85 – (-0.21)	
Baseline T2LV (mm³)					n.s.
<i>Mean ± SD</i>	6121 ± 3618	5514 ± 3686	4852 ± 3184	7693 ± 3534	
<i>Range</i>	590 – 15270	1894 – 15270	1091 – 11800	590 – 14480	
Annual T2LV rate (%)					
<i>Mean ± SD</i>	1.66 ± 0.84	2.14 ± 0.62	1.97 ± 0.66	1.13 ± 0.83	n.s.
<i>Range</i>	-0.32 – 3.21	1.21 – 3.21	0.79 – 3.04	-0.32 – 2.68	

SCV= Spinal Cord Volume; SD=Standard Deviation; TBV= Total Brain Volume; T2LV= Brain T2 Lesion Volume; PPMS=Primary Progressive MS; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS;

n.s.=not significant for any comparisons between PPMS, RRMS and SPMS

PPMS versus RRMS: † ≤ 0.05, †† ≤ 0.01, ††† ≤ 0.001.

PPMS versus SPMS: ‡ ≤ 0.05, ‡‡ ≤ 0.01, ‡‡‡ ≤ 0.001.

RRMS versus SPMS: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001.

Baseline SCV, TBV and T2LV were corrected for age, disease duration and sex. Annual SCV rate, annual TBV rate and annual T2LV rate were corrected for age, disease duration, sex and baseline SCV/TBV/T2LV respectively. Corrected values and significance of difference between RRMS and SPMS were obtained through analysis of covariance (ANCOVA).

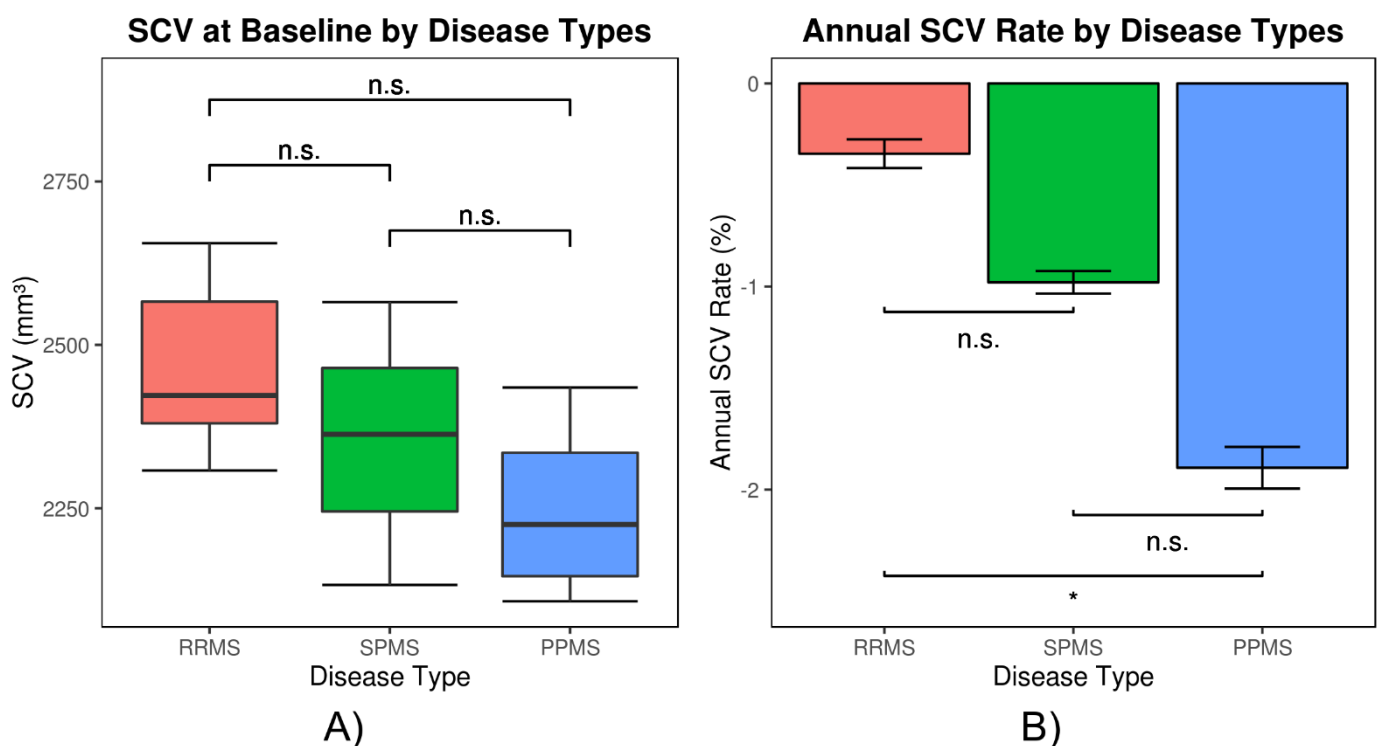


Figure 23. SCV display comparing RRMS (red), SPMS (green) and PPMS (blue) patients. A) At baseline, SCV was not significantly different between groups but B) the annual SCV rate was significantly higher for the PPMS when compared to the RRMS ($p < 0.05$). RRMS and SPMS as well as SPMS and PPMS did not differ significantly in terms of annual SCV rate.

Between-group differences in MRI measures

In the LMER analyses PPMS had lower SCV on average, as well as a faster SCV loss over time, when compared with RRMS patients (Figure 24). PPMS had also by trend faster SCV loss over time compared to the SPMS patients. Furthermore, male sex was significantly correlated with a higher SCV average.

No significant differences were demonstrated between groups in terms of TBV and T2LV. The statistical models are exhibited in detail in Table 9.

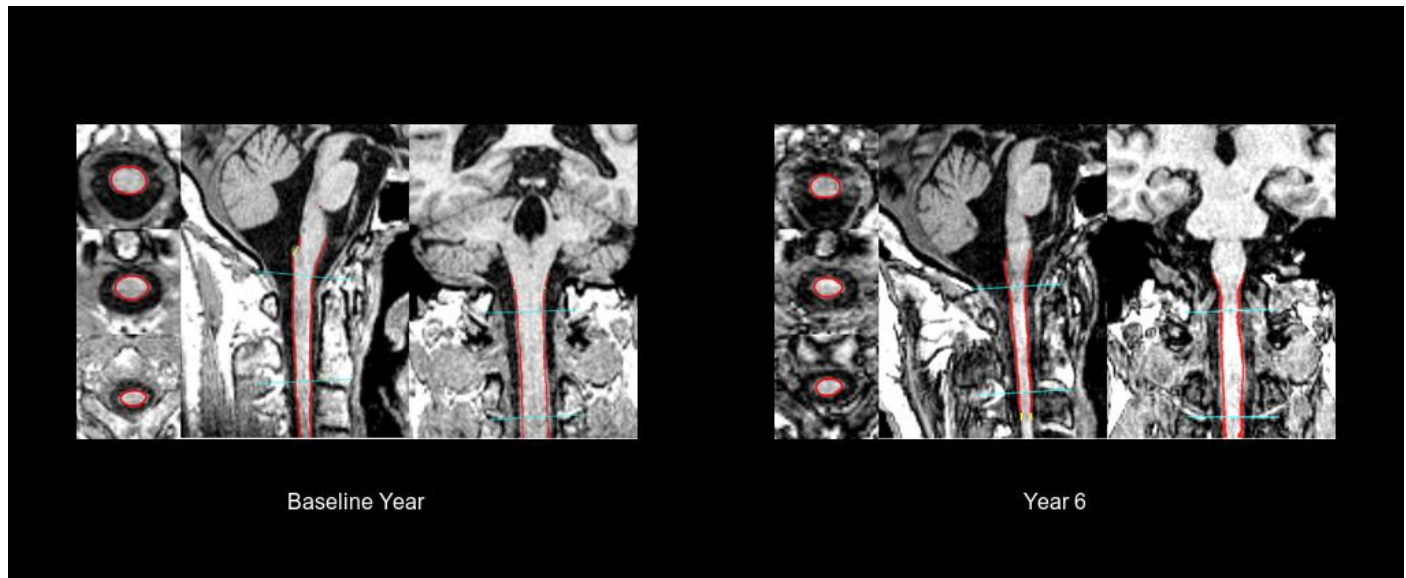


Figure 24. *Segmented MR imaging of one representative PPMS patient at baseline and 6 years after. Notice the significant diffuse spinal cord volume loss in all three planes of the patient's images.*

Associations between SCV, TBV and T2LV across disease types

Separate LMER analyses were performed to specifically investigate the extent of association between SCV, TBV, T2LV in each MS disease type. In PPMS, only age was associated with SCV loss over time. Male sex had a trend to a higher SCV average. Neither TBV nor T2LV were significantly related to SCV. In RRMS and SPMS, only TBV was associated with the average SCV, but no factors were associated with SCV loss over time. T2LV did not have a significant effect on SCV. Analyses details are exhibited in Table 10.

Table 10. Comparison of SCV, TBV and T2LV between disease types.

Variable		SCV (mm ³)		TBV (cm ³)		T2LV (mm ³)	
		Average	Loss over time	Average	Loss over time	Average	Increase over time
Age		n.s.	n.s.	B=-3.59, p<0.01	B=-0.19, p<0.05	n.s.	n.s.
Sex (Males)		B=200, p<0.05	n.s.	B=-55.2, p<0.001	n.s.	n.s.	n.s.
Disease Duration		n.s.	n.s.	B=-3.61, p<0.05	n.s. (p=0.056)	B=564, p<0.001	n.s.
Disease Type	PPMS vs RRMS	B=-258, p<0.05	B=-29.1, p<0.01	n.s.	n.s.	n.s.	n.s.
	PPMS vs SPMS	n.s.	n.s. (B=-21.5, p=0.066)				
	RRMS vs SPMS	n.s.	n.s.				
		Final model : R ² _m =21%, R ² _c =99%		Final model : R ² _m =36%, R ² _c =98%		Final model : R ² _m =25%, R ² _c =100%	

SCV= Spinal Cord Volume; SD=Standard Deviation; TBV= Total Brain Volume; T2LV= Brain T2 Lesion Volume; PPMS=Primary Progressive MS; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS; n.s.=not significant for any comparisons between PPMS, RRMS and SPMS; B=regression coefficient; R²_m= marginal R-squared; R²_c= conditional R-squared

Analysis was performed with linear mixed effect models with a random intercept and slope.

Table 11. Associations between MRI measures across disease types

Variable	PPMS		RRMS		SPMS	
	Average SCV	SCV loss over time	Average SCV	SCV loss over time	Average SCV	SCV loss over time
Age	n.s.	B=3.66, p<0.05	n.s.	n.s.	n.s.	n.s.
Sex	n.s. (p=0.088)	n.s.	n.s.	n.s.	n.s.	n.s.
Disease Duration	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
TBV (mm ³)	n.s.	n.s.	B=6 x 10 ⁻⁴ , p<0.05	n.s.	B=6 x 10 ⁻⁴ , p<0.05	n.s.
T2LV (mm ³)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Final model : R ² _m =5%, R ² _c =99%		Final model : R ² _m =4%, R ² _c =98%		Final model : R ² _m =2%, R ² _c =99%	

SCV= Spinal Cord Volume; SD=Standard Deviation; TBV= Total Brain Volume; T2LV= Brain T2 Lesion Volume; PPMS=Primary Progressive MS; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS; n.s.=not significant for any comparisons between PPMS, RRMS and SPMS; B=regression coefficient; R²_m= marginal R-squared; R²_c= conditional R-squared

Analysis was performed with linear mixed effect models with a random intercept and slope.

SCV changes and disability

In the LMER analyses, TBV ($B=-9 \times 10^{-7}$, $p<0.05$) and SCV ($B=-2 \times 10^{-4}$, $p<0.05$) were significantly correlated with the average EDSS over 6 years, but not with the EDSS increase over time. Age ($B=0.004$, $p<0.05$), disease duration ($B=0.01$, $p<0.01$), and disease type (PPMS vs RRMS: $B=0.46$, $p<0.001$; PPMS vs SPMS: not significant; RRMS vs SPMS: $B=0.42$, $p<0.001$) were also associated with the average EDSS, but not with the EDSS increase over time. Gender and T2LV did not contribute to the explanation of the EDSS. The final model accounted for 93%, while fixed effects alone accounted for 51% of the EDSS variance.

Regarding T25fwt, TBV and T2LV did not significantly correlate to the T25fwt results while SCV had a trend to correlate with the average T25fwt speed ($p=0.077$), but not with the deterioration in T25fwt over time. Disease duration ($B=-0.004$, $p<0.01$) and disease type (PPMS vs RRMS: $B=-0.06$, $p<0.05$; PPMS vs SPMS: not significant; RRMS vs SPMS: $B=-0.06$, $p<0.001$) had a significant effect on the average T25fwt of the whole cohort of MS patients. No factors were associated with a faster deterioration in T25fwt over time, although sex had a trend toward significance ($p=0.066$). The final model accounted for 94%, while fixed effects alone accounted for 34% of the T25fwt variance.

Separate LMER analyses were performed to specifically investigate the extent of association between SCV, TBV, T2LV to clinical outcomes in each MS disease type. Numerical details are displayed in Table 14.

In PPMS, SCV correlated with EDSS increase over time and had a trend toward significance with the average EDSS over 6 years. TBV and T2LV were not significantly associated with EDSS. Disease duration also correlated significantly to EDSS increase over time. TBV and SCV correlated with average T25fwt, while TBV and T2LV, but not SCV, were associated with the T25fwt worsening over time.

In RRMS, while age, SCV and TBV were associated with the average EDSS, but not with the EDSS increase over time, T2LV did not have a significant effect on the EDSS. Further, in regard of the T25fwt only SCV and T2LV were correlated with the T25fwt worsening over time, but not with the average T25fwt. Age and disease duration were associated with the average T25fwt.

In SPMS, SCV, TBV and T2LV were not associated with the EDSS. Only male sex was significantly associated with the average EDSS. No factors were significantly correlated with the T25fwt, although there was a trend for a correlation between the SCV and the T25fwt worsening over time.

Table 12. Associations between MRI measures and clinical outcomes across disease types.

Clinical Outcome	Variable	PPMS		RRMS		SPMS	
		Outcome's Average	Outcome's Change Over Time	Outcome's Average	Outcome's Change Over Time	Outcome's Average	Outcome's Change Over Time
EDSS	Age	n.s.	B=3.66, p<0.05	B=0.01, p<0.01	n.s.	n.s.	n.s. (p=0.054)
	Sex (Males)	n.s. (p=0.088)	n.s.	n.s.	n.s.	B=0.23, p<0.05	n.s.
	Disease Duration	n.s.	B=-0.003, p<0.05	n.s.	n.s.	n.s. (p=0.088)	n.s.
	TBV (mm ³)	n.s.	n.s.	B=-10 ⁻⁶ , p<0.05	n.s.	n.s.	n.s.
	SCV (mm ³)	n.s. (p=0.078)	B=-8x10 ⁻⁵ , p<0.01	B=-5 x 10 ⁻⁴ , p<0.05	n.s.	n.s.	n.s.
	T2LV (mm ³)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
		Final model : R ² _m =13%, R ² _c =89%		Final model : R ² _m =36%, R ² _c =89%		Final model : R ² _m =21%, R ² _c =93%	
T25fwt	Age	n.s.	n.s.	B=-10 ⁻³ , p<0.05	n.s.	n.s.	n.s.
	Sex (Males)	n.s.	n.s.	B=-0.003, p<0.05	n.s.	n.s.	n.s.
	Disease Duration	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	TBV (mm ³)	B=7x10 ⁻⁷ , p<0.01	B=-2x10 ⁻⁸	n.s.	n.s.	n.s.	n.s.
	SCV (mm ³)	B=9x10 ⁻⁵ , p<0.05	n.s.	n.s.	B=10 ⁻⁵ , p<0.05	n.s.	n.s. (p=0.055)
	T2LV (mm ³)	n.s.	B=8x10 ⁻⁷ , p<0.05	n.s.	B=-5 x 10 ⁻⁷ , p<0.05	n.s.	n.s.
		Final model : R ² _m =63%, R ² _c =94%		Final model : R ² _m =41%, R ² _c =85%		Final model : R ² _m =8%, R ² _c =96%	

EDSS=Expanded Disability Status Scale; T25fwt=Timed 25-foot walk test; SCV= Spinal Cord Volume; SD=Standard Deviation; TBV= Total Brain Volume; T2LV= Brain T2 Lesion Volume; PPMS=Primary Progressive MS; RRMS=Relapsing Remitting MS; SPMS=Secondary Progressive MS; n.s.=not significant for any comparisons between PPMS, RRMS, SPMS; B=regression coefficient; R²_m= marginal R-squared; R²_c= conditional R-squared

Analysis was performed with linear mixed effect models with a random intercept and slope.

Discussion

The present study investigated the upper cervical cord volume in a group of primary progressive MS patients in comparison to age-, sex- and disease duration-matched relapse-onset patients over a follow-up period of 6 years. Our work demonstrated a marked and accelerated loss of SCV in PPMS compared to relapse-onset MS, pointing out a notable characteristic difference of this patient group. PPMS patients clearly showed a preferential loss of SC over brain volume and interestingly, SCV was the only measure correlating with the disability progression over time as measured by the EDSS in this patient group. These study results give important insight on the role of SC atrophy in PPMS. The long follow-up period with annual measurements and the well-matched disease groups in our study are especially valuable in expanding to the existing literature of partially controversial results regarding the dominant role of SC atrophy as a distinctive feature between MS disease types^{44,45,97,140,143}.

On average, PPMS patients were found to have significantly reduced SC volumes compared to RRMS, as suggested by previous studies focusing on SC cross-sectional area^{44,45,140}. Moreover PPMS SCV declined significantly faster over the years compared to RRMS as well as by trend when compared with SPMS. To our knowledge, this is the first longer-term follow-up study demonstrating such a striking between-group difference of SCV loss over time^{44,136}. Furthermore, both average TBV or T2LV and their decline over time did not differ between groups and thus potentially underlining the significance of the observed increased SCV loss in PPMS.

Interestingly, while in PPMS SCV and TBV seemed to be widely independent both in terms of average volume and volume loss over time (with preferential loss in SC), atrophy in relapse-onset patients seems to be in part a global process with a coherent impact on both the brain and the SC. This is supported by Ingle and colleagues finding no correlation between brain and spinal cord cross-sectional atrophy in a group of PPMS patients¹³⁶.

The latter findings lead to an ongoing discussion in the field hypothesizing different underlying pathomechanisms in PPMS compared to relapse-onset MS and supporting the PPMS may represent a separate clinical entity¹⁴¹. In PPMS neuronal injury has been reported to predominantly affect motor related brain structures, while in SPMS and RRMS a more widespread pattern was shown¹⁴⁵. Even from a clinical perspective, PPMS patients rather show motor-related than visual or sensory disturbances^{140,146}. Based on that, it is possible that there is a selective kind of neuronal damage on the level of the cord, explaining early neurodegeneration and diffuse pathology of the SC from the very onset of PPMS and continuous throughout the course of the disease^{5,44,45,95,118,133–143}. This pathology is then being depicted as spinal cord volume loss in MR images. Furthermore, a secondary axonal and neuronal loss in the SC through Wallerian degeneration of the projections of motor control areas is possible. This focal volume loss in the brain may be undetectable in measurements such as TBV, but highly evident in the SC, where the corticospinal tract and motor neurons take up a significant percentage of the cross-section. This would also be in line with our results providing independent information of brain versus spinal MRI atrophy measurements both cross-sectionally and

longitudinally. Last, obviously also a combination of both pathomechanisms, leading to cumulative volume loss in the cord, is also possible.

Independent of the disease type SCV contributed significantly in the explanation of clinical disability, even after the introduction of TBV into the model, underlining the cords significance and prognostic capabilities discussed in the literature ^{5,44,45,77,134}. When analysing each individual disease type separately, only PPMS patients showed a strong relation to disease progression. At the same time TBV and T2LV were not relevant contributors to the explanation of the EDSS and disease progression. This confirms previous results from cross-sectional studies and one longitudinal study of PPMS patients ^{95,136–139,143}, as well as qualitative MRI studies showing the importance of SC pathology beyond brain metrics in the explanation of PPMS disability ^{118,142,147}. In contrast to PPMS, in RRMS TBV and T2LV seemed to play a bigger role in the explanation of clinical outcomes. However, even in RRMS SCV was a strong predictor of physical disability, measured both as EDSS and T25fw. Concerning SPMS patients in our study no MRI factors had a significant effect on clinical outcomes. This might be explained by the relatively small number of SPMS patients, in combination with the limitation of clinical scores catching underlying clinical progression, especially in such highly disabled patients as the SPMS.

The development of SC related physical disability in our group of PPMS patients in parallel to preferential SCV loss as compared to other MS-subtypes, emphasizes the inclusion of SCV measurements not only as a marker in clinical trials for emerging therapeutic agents, but also in the clinical routine in order to optimise individualised medicine in this patient group. SCV metrics may be the answer to the unmet need for reliable disease markers in PPMS. In a recent successful clinical trial investigating the therapeutic effect of Ocrelizumab compared to placebo in PPMS ¹⁴⁸, a difference of 17.5% reduction in brain atrophy was evident in the treatment group. Using this effect as a meaningful atrophy reduction, our study demonstrated rather achievable numbers of patients per arm needed (94 patients for 1-year, 59 patients for 2-year follow up time; analysis not shown here) to reach such therapeutic effects in the SC, even with a fairly low size effect at a considerably high power (95%) and 5% significance. Cawley et al. ¹⁴³ reported a minimum sample size of 57 per arm aiming to detect a 50% reduction in percentage change in SC cross-sectional area in PPMS at 80% power and 5% significance in a 1-year neuroprotective clinical trial, whereas based on our results a minimum sample size of 7 per arm is needed to demonstrate the same effect. Although this finding may reflect a higher sensitivity of our method, the interpretation of these results requires caution because of the small sample size in both studies. However, with several agents currently being under investigation in phase II and III clinical trials in PPMS ¹⁴⁹, our work suggests that SCV loss should be included as valuable outcome, being able to display even a small but clinically relevant therapeutic effect requiring a relatively low sample size. Utilizing our pipeline, this should be feasible by measuring SCV within brain scans, which are already included in most clinical trials.

Few limitations have to be considered when interpreting our results. First, the small sample size of each individual MS subgroup and especially of PPMS requires caution when evaluating the findings, since it may

have affected our ability to reveal correlations between the MRI variables and clinical data. The propensity score matching allowed us to analyse the data of the small number of PPMS patients in comparison to the more frequently occurring relapsing-remitting and secondary progressive disease types. We believe that we could reduce imbalance between disease subgroups produced by time- and disease-related phenomena in CNS structures by using this matching strategy. Further, the strong correlation of SC volume and clinical disability despite the small number of PPMS patients as well as the long follow-up time of our cohort makes us believe that the data may nevertheless be representative for this disease type.

Furthermore, the current retrospective study aimed to analyse follow-up data of a selected group of patients with MS. A loss of follow-up data over time cannot be excluded resulting in a potential bias related to the incomplete follow-up, although the proportion of MS patients with different disease types seems stable over 6 years. Moreover, there was no available sample of healthy controls recruited in this study, making a distinction between MS-related SC atrophy and any SCV decrease due to normal aging impossible. However, according to previous studies age-dependent SCV loss in the normal population seems to be significantly lower, suggesting that the observed loss of SCV over time could be interpreted as a disease related effect¹⁰³. Unfortunately, to our knowledge there is no longitudinal study investigating the effect of age on SCV loss in healthy controls. Despite the advantages of e.g. shorter acquisition times and reduction of scanning costs, using brain scans for SCV evaluation may be hampered by geometrical distortion and/or reduced signal intensity at the edge of the field of view potentially influencing segmentation quality of the cervical cord¹⁵⁰. To avoid such interference we evaluated the SCV well beyond the edge of the FOV which should grant robust measures. However, studies actually comparing brain versus dedicated spinal cord images and segmentation outcome should be performed to address this issue. Another disadvantage of brain scans is the insufficient grey/white-matter segregation within the cord making SC grey matter atrophy estimations impossible. Finally, the MR-protocol of our study did not include a T2w sequence covering the area of volume measurement over all time points analysed here and therefore the effect of SC lesion volume was not taken into account.

Acknowledgements

We are very grateful to all participants and staff involved in the GeneMSA cohort study, in particular Alain Thoeni.

Potential conflicts of interest statement

C. Tsagkas, S. Pezold, Y. Naegelin, M. Amann, C. Stippich, P. Cattin, O. Bieri: Nothing to report.

S. Magon: Travel support from Biogen.

L. Gaetano: Novartis advisory board.

J. Wuerfel: CEO of MIAC AG, Basel, Switzerland; speaker honoraria (Bayer, Biogen, Novartis, Teva); advisory boards and research grants (Biogen, Novartis); supported by the German Ministry of Science (BMBF/KKNMS) and German Ministry of Economy (BMW).

T. Sprenger: The current and previous employers of TS have received compensation for his serving on scientific advisory boards or speaking fees from Novartis, ATI, Electrocore, Sanofi Genzyme, Actelion, Jansen, Teva, Mitsubishi Pharma Europe and Biogen Idec.

L. Kappos: Author's institution (University Hospital Basel) has received in last 3 years and used exclusively for research support: steering committee, advisory board, and consultancy fees (Actelion, Addex, Bayer HealthCare, Biogen Idec, Biotica, Genzyme, Lilly, Merck, Mitsubishi, Novartis, Ono Pharma, Pfizer, Receptos, Sanofi, Santhera, Siemens, Teva, UCB, XenoPort); speaker fees (Bayer HealthCare, Biogen Idec, Merck, Novartis, Sanofi, Teva); support of educational activities (Bayer HealthCare, Biogen, CSL Behring, Genzyme, Merck, Novartis, Sanofi, Teva); royalties (Neurostatus Systems GmbH); grants (Bayer HealthCare, Biogen Idec, European Union, Merck, Novartis, Roche Research Foundation, Swiss MS Society, Swiss National Research Foundation).

K. Parmar holds a grant of the Baasch-Medicus foundation. Her institution (University Hospital Basel) received speakers honoraria from Novartis and ExceMED and travel support by Novartis Switzerland.

4. Discussion and Outlook

4.1. Spinal Cord Volume Quantification

The SC is an area of great importance due to its critical role for locomotion and MRI of the SC is a valuable part of the diagnostic work-up in neurological disorders. MRI-based techniques to study SC abnormalities are not as well developed as for the brain. Currently, mainly MRI signal intensity changes within the SC are evaluated in the clinical management of disorders of the central nervous system in order to identify focal intramedullary pathology in terms of a qualitative assessment. However, cross-sectional or longitudinal quantitative measurements of SC volume deliver additional valuable information. In the past, SC atrophy was usually determined by assessing the cross-sectional area of the cervical cord, usually at the C2/C3 level, which has been shown to correlate with clinical measures, although reproducibility was limited and dependent on data quality as well as repositioning^{27,40,41}. Moreover, most of the proposed approaches focused on cervical and mid thoracic levels of the SC^{27-34,36-39,42-46}. Thus, disorders affecting predominantly the lumbar level of the SC, such as hereditary spastic paraplegia cannot be studied with the available quantification methods. In disorders affecting the entire SC it is currently poorly understood, if all SC segments are equally involved. Furthermore, segmentation of the SC into GM and WM with quantification of these two SC compartments in an automated fashion has only recently become possible⁴⁷⁻⁵⁰. This task, however, remains challenging as state-of-the-art MR sequences so far only achieve an in-slice resolution of around 0.5 mm (considering the very small dimensions of the SC GM) while maintaining a good signal-to-noise ratio (SNR) and an acceptable acquisition time. This resolution is barely enough to visualize the SC's butterfly-shaped GM structure. Taking into account, that mean SC volume changes in disorders such as MS range between 0.5-2.2% per year, it becomes apparent that new accurate and reproducible method for the development of such a biomarker. A fast and reliable SC GM quantification methods would allow to study this important structure in larger patient populations as well as to reveal the extent to which the SC GM can predict future outcomes for chronic diseases such as MS.

This thesis presents clinical validation for the quantification of the lumbar SC using the *Cordial* software as an important application addition beyond the cervical SC. This is the first study validating the accuracy and precision of volumetric measurements of the lumbar SC calculated in a semi-automatic fashion and thus delivering a tool for fast and reliable lumbar SC volumetry in clinical settings. One of the most important advantages of the proposed method is that it relies on natural landmarks on the SC itself and in this way bypasses reproducibility obstacles related to the physiologic motion of the SC, which is greater in the lumbar region. The method demonstrated highly reproducible results within the lumbar region and therefore a high potential for application of this metric in longitudinal studies or clinical trials. Despite the fact that many neurological disorders affecting the lumbar SC were till recently considered incurable, the first therapeutic strategies against disorders such as spinal muscular atrophy showed impressive efficacy¹⁵¹. This, however,

creates the need of accurate spinal cord markers to monitor the therapeutic effect over long periods of time. Thus, our work may provide a reproducible surrogate of neuronal injury (or integrity) in this part of the CNS.

Further, the application of Cordial to a large set of long-term brain MRI follow up data of a heterogeneous group of MS patients was performed for the first time. It proved to be feasible within a reasonable time frame and gave important insight into the understanding of the disease. Notably, it also showed that the quantification of SC volume is applicable in a clinical setting using routine brain –and not SC- MRI sequences, acquired on an “older generation” MRI scanner (1.5T). This emphasizes the broad applicability of the proposed marker with minimum requirements or additional financial burden for MS-patients and physicians as well as clinical trials for new therapeutic agents. It also bypasses some financial and time restrictions resulting from requirements of an extra MRI of the SC, which hampered such studies for decades.

Moreover, this thesis proposes a pipeline for the quantification of cervical SC GM and WM. This is one of the first approaches to visualize and quantify SC compartments in a feasible MR-acquisition time without any user-interface interaction. Our work utilized a novel MR-sequence (averaged magnetization inversion recovery acquisitions) with submillimeter in plane resolution ($0.67 \times 0.67\text{mm}^2$), which offers a surprisingly high contrast between GM and WM as well as between WM and CSF, together with a newly proposed variational segmentation algorithm with a shape prior modified for 3D data with a slice similarity prior^{21,152}. This pipeline was able to demonstrate a highly accurate and reproducible SC GM and WM segmentation. This approach will allow studying the relation of different SC volume compartments to SC and brain lesions, determine the degree and relevance of white versus grey matter loss in the SC, depict key areas of spinal cord involvement and to relate spinal cord changes to other advanced (brain and spinal cord) MRI measures longitudinally in MS and other neurological disorders. Since image processing time of our fully automatic method is negligible it seems promising not only for the application in large datasets in research settings but also in clinical routine, providing SC GM and WM measurements effortlessly.

4.2. Spinal Cord Volume Loss in Multiple Sclerosis

MS is an inflammatory-demyelinating disease of the central nervous system and continues to be a major cause of disability among young individuals leading to major health and socioeconomic disadvantages. Neurodegeneration seems to be the key driver for the accrual of physical and neuropsychological disability and atrophy is one of the hallmarks of neurodegeneration in MS. SC atrophy in MS –a measure of neurodegeneration- has previously been reported mainly in cross-sectional studies^{44,45,91-93}. However, previous approaches have been hampered by the relatively low-resolution and contrast of the acquired MR images, long measurement times, artifacts as well as the low reproducibility of the segmentation techniques³. This has led to a lack in longitudinal data concerning SC atrophy in MS and consequently a gap in the clinicians' understanding of the disease. Moreover, the potential of SC atrophy as a disease marker over long periods of time was unknown.

The work presented in this thesis demonstrates a detailed investigation of SC atrophy across all MS subgroups over a fairly long follow-up time using brain MR-images. We examined the upper cervical SC volume, which anatomically contains all efferent and afferent fibers from and to the brain as well as the SC GM, and offers the best possible reproducibility of volumetric SC measurements. This is the first longer-term study of SC volume loss in a rather large dataset of MS patients including all disease phenotypes as well as multiple clinical and MRI metrics beyond SC volume. We were able to investigate in depth the clinical significance of SC volume loss and explore temporal features of the SC volume over the course of six years and in association with clinical and demographic characteristics of MS patients.

In contrast to former studies suggesting a faster atrophy in the SC than in the brain, our study showed comparable SC atrophy rates in RRMS and SPMS with the reported brain atrophy rate in MS, which ranges between 0.5-1.0% per year. SC volume was shown to be particularly accelerated in the group of PPMS patients compared to relapse onset MS, while being the only MR metric associated with disability progression. In contrast brain volume loss in PPMS was similar to relapse onset MS. This shows that the SC is a predilection spot in PPMS and is a strong marker of clinical worsening. In contrast to previous studies, SPMS and RRMS patients did not differ in terms of SC atrophy rate, emphasizing on the similarities rather than the differences between those two disease subtypes. Interestingly, SC volume loss over time was not correlated with brain volumetric metrics, which suggests that atrophy progression is an independent process in the two CNS regions. In RRMS, SC volume loss was also partly related to the number of clinical relapses, which suggest a weak relationship with acute focal inflammatory events.

SC proved to be a strong predictor of physical disability and disease progression, indicating that it may be a suitable marker for monitoring disease activity and severity in all disease types and especially in progressive MS (both SPMS and PPMS). Despite the fact that RRMS and SPMS had similar SC atrophy rates, in SPMS clinical outcomes correlated much stronger compared to RRMS. This implies that the key driver for clinical progression is not the same in the two groups. Indeed, recent studies point to an early thalamic involvement

in MS, which correlates highly with clinical outcomes¹⁵³. It seems that -for unknown reasons- SC pathology is of greater importance in later stages of the disease. SC volume also proved to be the only MRI metric to strongly explain the clinical progression over time as opposed to brain atrophy and lesion measures for all MS patients, which are currently commonly used in clinical trials. These results open a new chapter in the field of MRI-research in MS showing the importance of SC evaluation in these patients underlining not the need to include assessment of SC MRI atrophy metrics in the management of MS patients in the clinic. It also provides new insights about the mechanisms driving physical disability in different MS groups indicating that SC should be included in future clinical trials, especially in progressive MS. Apart from the issues answered, this work gave rise to many questions, namely concerning the pathophysiologic mechanisms leading to the observed between-group differences both in terms of SC atrophy rates and correlation strength between SC volume loss and clinical outcomes.

In view of the results, SC volume metrics may be the answer to the unmet need for reliable disease markers in MS, especially in progressive MS. In a recent successful clinical trial investigating the therapeutic effect of Ocrelizumab compared to placebo in PPMS, a difference of 17.5% reduction in brain atrophy was evident in the treatment group¹⁴⁸. Using this effect as a meaningful atrophy reduction, our study demonstrated rather achievable numbers of patients per arm needed (94 patients for 1-year, 59 patients for 2-year follow up time; analysis not shown here) to reach such therapeutic effects in the SC, even with a fairly low size effect at a considerably high power (95%) and 5% significance. Cawley et al.¹⁴³ reported a minimum sample size of 57 per arm aiming to detect a 50% reduction in percentage change in SC cross-sectional area in PPMS at 80% power and 5% significance in a 1-year neuroprotective clinical trial, whereas based on our results a minimum sample size of 7 per arm is needed to demonstrate the same effect. However, with several agents currently being under investigation in phase II and III clinical trials in PPMS, our work suggests that SCV loss should be included as valuable outcome, being able to display even a small but clinically relevant therapeutic effect requiring a relatively low sample size. Utilizing our pipeline, this should be feasible by measuring SCV within brain scans, which are already included in most clinical trials.

In addition to the – proof of principle – application of these techniques in MS, the same imaging and post-processing strategies should be feasible also in other disorders affecting the SC including traumatic SC injuries, and degenerative SC disorders provided that similar quality 3D MRI data is available.

4.3. Outlook and Future Research

4.3.1. Lumbar Spinal Cord Segmentation

Spinal Muscular Atrophy (SMA) is a neuromuscular genetic disease^{154,155} characterized by a progressive loss of anterior horn motor neurons in the spinal cord and subsequent system-wide muscle atrophy followed by progressive weakness and disability due to mobility impairment, respiratory, gastrointestinal, and functional complications¹⁵⁶. SMA is one of the most devastating neurological diseases in childhood and is the number one cause of death related to genetic dysfunction in children.

Recent advances in understanding SMA etiopathogenesis prompted emergence of promising therapeutic strategies resulting in a number of successful clinical trials being performed worldwide^{151,157}. Further, longitudinal progression is typically slow and difficult to detect especially in milder forms of SMA. Therefore, in the view of recent developments in this field new meaning to patient monitoring was given in this rare disease with a wide phenotypic spectrum, creating the need for reliable biomarkers capable of reliably assessing disease progression and potential treatment effects.

Translating our results from “Manuscript 1”, we intend to implement our method in a small cohort of 19 untreated spinal muscular atrophy patients and matched healthy volunteers described in a previous study¹⁵⁸, in order to study lumbar SC atrophy compared to healthy controls in vivo. This cohort was observed for 54 weeks, so that we aim to observe the SC volume changes in this time period and provide a measure of SC atrophy progression for future assessments. We also intend to investigate the correlation of those measurements with clinical outcomes and clinical progression over time.

4.3.2. Spinal Cord Grey Matter Segmentation

The SC GM was investigated in two recent previous cross-sectional studies using manual delineations of the butterfly-shaped GM. Schlaeger and colleagues observed evidence of GM atrophy of the cervical and thoracic SC compared to healthy controls^{95,97}. These studies also demonstrated marked GM atrophy in progressive MS subtypes compared to RRMS^{95,97}. More importantly, SC GM atrophy was found to be the most significant correlate of clinical disability over brain metrics such as GM and WM^{95,97}. However, more studies are needed to validate these results. Furthermore, the application and validation of automatic computer-based methods for SC GM segmentation is highly needed before further implementation in clinical routine. Finally, the changes of SC GM areas and/or volumes over time has never been investigated so far.

Translating our results from “Manuscript 2”, our goal is to study a new cohort of MS patients cross-sectionally using the developed pipeline. At present more than 80 patients have been scanned at baseline and follow up evaluations (after one year) and first data analysis are still ongoing. We aim to study the suitability of the

developed methods to determine longitudinal changes of the SC compartments in MS and provide brand new insights with regard to this promising biomarker cross-sectionally and longitudinally. We also plan to study the relative contribution of SC metrics (GM and WM, lesions) to physical disability in relation to brain metrics (T2 lesion load, normalized brain volume, volume of GM versus WM). In view of recent studies identifying a MS patient group with absent brain WM demyelinating plaques and pronounced lesion load in the brain GM and the SC¹⁵⁹, we intend to search for patients of this possibly newly discovered disease subtype in our cohort.

4.3.3. Future Volumetric Studies in Multiple Sclerosis

As a continuation of the presented work using a large preexisting longitudinal dataset, we intend to also analyze other CNS structures such as the brain deep GM, the cerebellum and cortical thickness in this cohort. With regard to the SC, we plan to study atrophy progression in an even larger cohort of over 1000 patients observed longitudinally with the purpose of further examining differences between groups MS patients (e.g. RRMS and SPMS, early versus late stages of MS etc.).

The standardized definitions of clinical courses of MS (RRMS, PPMS, and SPMS) were proposed in 1996 and have been used in clinical routine ever since ¹⁶⁰. However, these definitions are purely descriptive and do not provide information about the underlying pathophysiology of the disease. Thus, a classification based on specific pathomechanisms is still a much wanted unmet need and would be crucial for the effective selection of disease-modifying treatments ¹⁶¹. Thus, we intend to investigate the potential of MRI atrophy measurements (including the SC) to objectively classify MS patients with the intention to contribute to individualized medical decision-making in this disorder in the future.

5. Contributions by the PhD student

During my PhD, I had the opportunity to contribute to the planning and conduct of several research projects and finally was also able to explore my own research ideas. Through the courses I followed in the University of Basel and learning from my co-workers and supervisors, I had the opportunity to acquire a number of skills that will be essential for future research work.

Performance of the image post-processing within the aforementioned projects helped me to expand my knowledge of SC anatomy and pathology and I became an experienced reader of conventional and non-conventional SC MR-imaging. I further deepened my statistical skills within advanced courses of the University of Basel including the application of the R statistical software. This enabled me to conduct all statistical analyses of the aforementioned projects with the support of our research group. It will further allow me to implement my own ideas at the level of data mining and analysis in the future. Drafts and revisions of all manuscripts as well as the ethics proposal for the ongoing clinical study were written by myself with the support and help of my co-workers and supervisors. I presented our work at several international conferences as oral and poster presentations (European Committee for Teaching and Research in Multiple Sclerosis, annual *a la recherche sur la Sclérose en Plaques*) and at the end of my PhD was invited to give a lecture on SC MRI in MS at an international congress (26th Annual Meeting of the European Charcot Foundation). The ongoing projects on SC GM changes as well as its preliminary work (Manuscript 2) allowed me to contribute to all aspects of the study design, recruitment and clinical and MRI examination of participants. I further supervised two master students within our ongoing projects.

The work and my gained experience within my PhD thesis lead to several active collaborations on projects including SC atrophy as an outcome measure (SC atrophy in progressive MS receiving Biotin by Dr. N. Collongues – manuscript under preparation; SC volume in SMA vs healthy controls by Prof. Dr. med. D. Fischer – under analyses). Further I was involved in the SC segmentation within the scope of three investigator-initiated studies conducted at MIAC AG (TEMSSO trial, and two investigator-initiated research cohorts).

Finally, the close collaboration within our research group gave me the opportunity to broaden my horizon to other related MRI research on MS patients. I was part of a recently published work focusing on the damage of the lateral geniculate nucleus in MS (in press in *Neurology*), for which I segmented the optic nerves and chiasm and participated in the draft together with all other co-authors. I also conducted all statistical analysis and helped in writing and drafting a currently unpublished work (currently under revision) regarding atrophy of deep GM in MS, whereas I am currently preparing multiple other longitudinal volumetric studies of the same cohort focusing on other CNS structures and new aspects of this disease as a principal investigator.

6. Conclusion and closing remarks

The work presented in this thesis consists of two core features focusing on SC volume quantification: a methodological and a clinical application.

Regarding method development, in “Manuscript 1” we engaged with the validation of the first method reliably and accurately quantifying the lumbar SC volume in a semi-automatic fashion. This method requires minimum user-software interaction and is therefore suitable for further studies examining the lumbar SC with cross-sectional and longitudinal design even in large cohorts. “Manuscript 2” provides a new pipeline for quantification of SC and GM of the cervical cord. This method implemented the newly developed AMIRA MRI sequence and a novel fully automatic method providing accurate and reproducible measurements of the SC compartments with essentially no user-software interaction. Therefore, this method is suitable for the examination of the SC in disorders affecting the SC such as MS even in studies with large-scale and longitudinal setting. Both proposed methods could be implemented in clinical routine in order to provide information to the clinician that are not detectable or quantifiable with the naked eye.

With regard to clinical application of SC volume quantification, this work delivered new exciting insights about SC volume loss in MS patients of all disease subtypes. “Manuscript 3” was the first long-term longitudinal study systematically examining SC volume loss in a large cohort of relapse-onset MS. Our study, observed a mean annual SC volume loss of about 0.4% in RRMS and 0.6% in SPMS, but –in contrast to previous cross-sectional and short-term longitudinal studies- no statistically significant between-group difference was observed. On the other side, “Manuscript 4” focused on a small PPMS population in our cohort and found an accelerated SC volume loss (mean annual atrophy rate about 2%), in contrast to whole brain volume changes, which were similar between PPMS and relapse-onset MS. This points to a preferential SC volume loss in PPMS. Our work verified, that SC volume is a reliable biomarker going hand in hand with clinical changes and being able to predict future outcomes, especially in progressive MS. Finally, we were able to show that SC volume would be able to show clinically meaningful therapeutic effect in clinical trial with a rather small sample size. Our studies were able to shed light to important new aspects of MS and showed that accurate and reliable SC volume measurements would have a major impact on disease- and treatment-monitoring. We presented SC volume as a new promising biomarker, which can be implemented in the clinical setting in the management of MS.

In conclusion, we are confident that we provided the research community with the necessary tools for reliable and time-efficient atrophy quantification in the spinal cord. We were also able to systematically follow SC volume loss over time and contributed to establishing this biomarker in MS. Our research in MS and SC volume quantification in general may ultimately have an impact on understanding and monitoring of other neurological disorders and improve patient management in the foreseeable future.

7. References

1. Stroman, P. W. & Bosma, R. L. Spinal Cord Imaging. in *Neurological Aspects of Spinal Cord Injury* 237–257 (Springer, Cham, 2017). doi:10.1007/978-3-319-46293-6_10
2. Wheeler-Kingshott, C. A. *et al.* The current state-of-the-art of spinal cord imaging: Applications. *NeuroImage* **84**, 1082–1093 (2014).
3. Stroman, P. W. *et al.* The current state-of-the-art of spinal cord imaging: Methods. *NeuroImage* **84**, 1070–1081 (2014).
4. Nair, G., Absinta, M. & Reich, D. S. Optimized T1-MPRAGE Sequence for Better Visualization of Spinal Cord Multiple Sclerosis Lesions at 3T. *Am. J. Neuroradiol.* **34**, 2215–2222 (2013).
5. Kearney, H., Miller, D. H. & Ciccarelli, O. Spinal cord MRI in multiple sclerosis—diagnostic, prognostic and clinical value. *Nat. Rev. Neurol.* **11**, 327–338 (2015).
6. Donaldson Henry H. & Davis David J. A description of charts showing the areas of the cross sections of the human spinal cord at the level of each spinal nerve. *J. Comp. Neurol.* **13**, 19–40 (2004).
7. Lassek Arthur M. & Rasmussen Grant L. A quantitative study of the newborn and adult spinal cords of man. *J. Comp. Neurol.* **69**, 371–379 (2004).
8. Elliott H. Chandler. Cross-sectional diameters and areas of the human spinal cord. *Anat. Rec.* **93**, 287–293 (2005).
9. Cadotte, D. W. *et al.* Characterizing the Location of Spinal and Vertebral Levels in the Human Cervical Spinal Cord. *Am. J. Neuroradiol.* **36**, 803–810 (2015).
10. Ko, H.-Y., Park, J. H., Shin, Y. B. & Baek, S. Y. Gross quantitative measurements of spinal cord segments in human. *Spinal Cord* **42**, 35–40 (2004).
11. Sherman, J. L., Nassaux, P. Y. & Citrin, C. M. Measurements of the normal cervical spinal cord on MR imaging. *AJNR Am. J. Neuroradiol.* **11**, 369–372 (1990).
12. Kameyama, T., Hashizume, Y., Ando, T. & Takahashi, A. Morphometry of the normal cadaveric cervical spinal cord. *Spine* **19**, 2077–2081 (1994).
13. Frostell, A., Hakim, R., Thelin, E. P., Mattsson, P. & Svensson, M. A Review of the Segmental Diameter of the Healthy Human Spinal Cord. *Front. Neurol.* **7**, (2016).

14. Bhadelia, R. A., Bogdan, A. R., Kaplan, R. F. & Wolpert, S. M. Cerebrospinal fluid pulsation amplitude and its quantitative relationship to cerebral blood flow pulsations: a phase-contrast MR flow imaging study. *Neuroradiology* **39**, 258–264 (1997).
15. Figley C.R. & Stroman P.W. Investigation of human cervical and upper thoracic spinal cord motion: Implications for imaging spinal cord structure and function. *Magn. Reson. Med.* **58**, 185–189 (2007).
16. Piché, M. *et al.* Characterization of cardiac-related noise in fMRI of the cervical spinal cord. *Magn. Reson. Imaging* **27**, 300–310 (2009).
17. Yiannakas, M. C. *et al.* Feasibility of grey matter and white matter segmentation of the upper cervical cord in vivo: A pilot study with application to magnetisation transfer measurements. *NeuroImage* **63**, 1054–1059 (2012).
18. Held, P., Dorenbeck, U., Seitz, J., Fründ, R. & Albrich, H. MRI of the abnormal cervical spinal cord using 2D spoiled gradient echo multiecho sequence (MEDIC) with magnetization transfer saturation pulse. A T2* weighted feasibility study. *J. Neuroradiol. J. Neuroradiol.* **30**, 83–90 (2003).
19. Papinutto, N. *et al.* 2D phase-sensitive inversion recovery imaging to measure in vivo spinal cord gray and white matter areas in clinically feasible acquisition times. *J. Magn. Reson. Imaging* n/a-n/a (2014).
doi:10.1002/jmri.24819
20. Weigel Matthias and Bieri Oliver. A Simple and Fast Approach for Spinal Cord Imaging at 3T with High In-Plane Resolution and Good Contrast. in *In Proceedings of the 24th Annual Meeting of ISMRM, Singapore* 4408 (2016).
21. Weigel, M. & Bieri, O. Spinal cord imaging using averaged magnetization inversion recovery acquisitions. *Magn. Reson. Med.* **79**, 1870–1881 (2018).
22. Pratt, W. K. *Digital Image Processing: PIKS Scientific Inside.* (Wiley, 2007).
23. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 3431–3440 (2015).
doi:10.1109/CVPR.2015.7298965
24. Kass, M., Witkin, A. & Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1**, 321–331 (1988).

25. Yuan, J., Bae, E. & Tai, X. C. A study on continuous max-flow and min-cut approaches. in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2217–2224 (2010).
doi:10.1109/CVPR.2010.5539903
26. Pezold, S. *et al.* Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures. in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016* 362–370 (Springer, Cham, 2016).
doi:10.1007/978-3-319-46726-9_42
27. Losseff, N. A. *et al.* Spinal cord atrophy and disability in multiple sclerosis. *Brain* **119**, 701–708 (1996).
28. Leener, B. D., Taso, M., Cohen-Adad, J. & Callot, V. Segmentation of the human spinal cord. *Magn. Reson. Mater. Phys. Biol. Med.* **29**, 125–153 (2016).
29. Coulon, O. *et al.* Quantification of spinal cord atrophy from magnetic resonance images via a B-spline active surface model. *Magn. Reson. Med.* **47**, 1176–1185 (2002).
30. Hickman, S. J., Hadjiprocopis, A., Coulon, O., Miller, D. H. & Barker, G. J. Cervical spinal cord MTR histogram analysis in multiple sclerosis using a 3D acquisition and a B-spline active surface segmentation technique. *Magn. Reson. Imaging* **22**, 891–895 (2004).
31. Horsfield, M. A. *et al.* Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: Application in multiple sclerosis. *NeuroImage* **50**, 446–455 (2010).
32. Mukherjee, D. P. *et al.* Automatic Segmentation of Spinal Cord MRI Using Symmetric Boundary Tracing. *IEEE Trans. Inf. Technol. Biomed. Publ. IEEE Eng. Med. Biol. Soc.* **14**, 1275–1278 (2010).
33. Van Uitert, R., Bitter, I. & Butman, J. A. Semi-automatic spinal cord segmentation and quantification. *Int. Congr. Ser.* **1281**, 224–229 (2005).
34. Carbonell-Caballero, J. *et al.* Accurate quantification methods to evaluate cervical cord atrophy in multiple sclerosis patients. *Magn. Reson. Mater. Phys. Biol. Med.* **19**, 237–246 (2006).
35. Koh, J., Kim, T., Chaudhary, V. & Dhillon, G. Automatic segmentation of the spinal cord and the dural sac in lumbar MR images using gradient vector flow field. *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* **2010**, 3117–3120 (2010).

36. Chen, M. *et al.* Automatic magnetic resonance spinal cord segmentation with topology constraints for variable fields of view. *NeuroImage* **83**, 1051–1062 (2013).
37. Fonov, V. S. *et al.* Framework for integrated MRI average of the spinal cord white and gray matter: The MNI–Poly–AMU template. *NeuroImage* **102, Part 2**, 817–827 (2014).
38. De Leener, B., Cohen-Adad, J. & Kadoury, S. Automatic Segmentation of the Spinal Cord and Spinal Canal Coupled With Vertebral Labeling. *IEEE Trans. Med. Imaging* **34**, 1705–1718 (2015).
39. Yiannakas, M. C. *et al.* Fully automated segmentation of the cervical cord from T1-weighted MRI using PropSeg: Application to multiple sclerosis. *NeuroImage Clin.* **10**, 71–77 (2016).
40. Kidd D *et al.* Spinal cord MRI using multi-array coils and fast spin echo. II. Findings in multiple sclerosis. *Neurology* **43**, 2632–7 (1993).
41. Lin, X., Tench, C. R., Evangelou, N., Jaspan, T. & Constantinescu, C. S. Measurement of spinal cord atrophy in multiple sclerosis. *J. Neuroimaging Off. J. Am. Soc. Neuroimaging* **14**, 20S-26S (2004).
42. Stevenson, V. L. *et al.* Spinal cord atrophy and disability in MS: a longitudinal study. *Neurology* **51**, 234–238 (1998).
43. Valsasina, P., Rocca, M. A., Horsfield, M. A., Copetti, M. & Filippi, M. A longitudinal MRI study of cervical cord atrophy in multiple sclerosis. *J. Neurol.* **262**, 1622–1628 (2015).
44. Lukas, C. *et al.* Cervical spinal cord volume loss is related to clinical disability progression in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **86**, 410–418 (2015).
45. Lukas, C. *et al.* Relevance of Spinal Cord Abnormalities to Clinical Disability in Multiple Sclerosis: MR Imaging Findings in a Large Cohort of Patients. *Radiology* **269**, 542–552 (2013).
46. Amann, M. *et al.* Reliable volumetry of the cervical spinal cord in MS patient follow-up data with cord image analyzer (Cordial). *J. Neurol.* **263**, 1364–1374 (2016).
47. Prados, F. *et al.* Fully automated grey and white matter spinal cord segmentation. *Sci. Rep.* **6**, 36151 (2016).
48. Datta, E. *et al.* Gray matter segmentation of the spinal cord with active contours in MR images. *NeuroImage* **147**, 788–799 (2017).

49. Dupont, S. M. *et al.* Fully-integrated framework for the segmentation and registration of the spinal cord white and gray matter. *NeuroImage* **150**, 358–372 (2017).
50. De Leener, B. *et al.* SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. *NeuroImage* **145**, 24–43 (2017).
51. Ropper, A. H., Samuels, M. A. & Klein, J. *Adams and Victor's Principles of Neurology 10th Edition*. (McGraw Hill Professional, 2014).
52. Reich, D. S., Lucchinetti, C. F. & Calabresi, P. A. Multiple Sclerosis. *N. Engl. J. Med.* **378**, 169–180 (2018).
53. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol.* **16**, 877–897 (2017).
54. Thompson, A. J., Baranzini, S. E., Geurts, J., Hemmer, B. & Ciccarelli, O. Multiple sclerosis. *The Lancet* **391**, 1622–1636 (2018).
55. Kobelt, G. *et al.* New insights into the burden and costs of multiple sclerosis in Europe. *Mult. Scler. J.* **23**, 1123–1136 (2017).
56. Orton, S.-M. *et al.* Sex ratio of multiple sclerosis in Canada: a longitudinal study. *Lancet Neurol.* **5**, 932–936 (2006).
57. Ascherio, A. & Munger, K. L. Epidemiology of Multiple Sclerosis: From Risk Factors to Prevention—An Update. *Semin. Neurol.* **36**, 103–114 (2016).
58. Weinshenker, B. G. *et al.* THE NATURAL HISTORY OF MULTIPLE SCLEROSIS: A GEOGRAPHICALLY BASED STUDY3. MULTIVARIATE ANALYSIS OF PREDICTIVE FACTORS AND MODELS OF OUTCOME. *Brain* **114**, 1045–1056 (1991).
59. Lassmann, H. Mechanisms of white matter damage in multiple sclerosis. *Glia* **62**, 1816–1830 (2014).
60. Lucchinetti Claudia *et al.* Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Ann. Neurol.* **47**, 707–717 (2001).
61. Schirmer, L., Srivastava, R. & Hemmer, B. To look for a needle in a haystack: the search for autoantibodies in multiple sclerosis. *Mult. Scler. J.* **20**, 271–279 (2014).

62. Hohlfeld, R., Dornmair, K., Meinl, E. & Wekerle, H. The search for the target antigens of multiple sclerosis, part 1: autoreactive CD4+ T lymphocytes as pathogenic effectors and therapeutic targets. *Lancet Neurol.* **15**, 198–209 (2016).
63. Mishra, M. K. & Yong, V. W. Myeloid cells — targets of medication in multiple sclerosis. *Nat. Rev. Neurol.* **12**, 539–551 (2016).
64. Prinz, M., Priller, J., Sisodia, S. S. & Ransohoff, R. M. Heterogeneity of CNS myeloid cells and their roles in neurodegeneration. *Nat. Neurosci.* **14**, 1227–1235 (2011).
65. van der Valk, P. & Amor, S. Preactive lesions in multiple sclerosis. *Curr. Opin. Neurol.* **22**, 207–213 (2009).
66. Ludwin, S. K., Rao, V. T., Moore, C. S. & Antel, J. P. Astrocytes in multiple sclerosis. *Mult. Scler. Houndmills Basingstoke Engl.* **22**, 1114–1124 (2016).
67. Absinta, M. *et al.* Direct MRI detection of impending plaque development in multiple sclerosis. *Neurol. Neuroimmunol. Neuroinflammation* **2**, e145 (2015).
68. Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 (2018).
69. Metz Imke *et al.* Pathologic heterogeneity persists in early active multiple sclerosis lesions. *Ann. Neurol.* **75**, 728–738 (2014).
70. Peterson, J. W., Bö, L., Mörk, S., Chang, A. & Trapp, B. D. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Ann. Neurol.* **50**, 389–400 (2001).
71. Bø, L., Vedeler, C. A., Nyland, H. I., Trapp, B. D. & Mørk, S. J. Subpial demyelination in the cerebral cortex of multiple sclerosis patients. *J. Neuropathol. Exp. Neurol.* **62**, 723–732 (2003).
72. Kutzelnigg, A. *et al.* Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain* **128**, 2705–2712 (2005).
73. Lucchinetti, C. F. *et al.* Inflammatory Cortical Demyelination in Early Multiple Sclerosis. *N. Engl. J. Med.* **365**, 2188–2197 (2011).

74. Serafini, B., Rosicarelli, B., Magliozzi, R., Stigliano, E. & Aloisi, F. Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. *Brain Pathol. Zurich Switz.* **14**, 164–174 (2004).
75. Trapp, B. D. *et al.* Axonal transection in the lesions of multiple sclerosis. *N. Engl. J. Med.* **338**, 278–285 (1998).
76. Mahad, D. H., Trapp, B. D. & Lassmann, H. Pathological mechanisms in progressive multiple sclerosis. *Lancet Neurol.* **14**, 183–193 (2015).
77. Weier, K. *et al.* Biplanar MRI for the assessment of the spinal cord in multiple sclerosis. *Mult. Scler. J.* **18**, 1560–1569 (2012).
78. Bot, J. C. J. *et al.* Spinal cord abnormalities in recently diagnosed MS patients: Added value of spinal MRI examination. *Neurology* **62**, 226–233 (2004).
79. Sombekke, M. H. *et al.* Spinal cord lesions in patients with clinically isolated syndrome: A powerful tool in diagnosis and prognosis. *Neurology* **80**, 69–75 (2013).
80. Brownlee, W. J. *et al.* Association of asymptomatic spinal cord lesions and atrophy with disability 5 years after a clinically isolated syndrome. *Mult. Scler. J.* **23**, 665–674 (2017).
81. Petrova, N., Carassiti, D., Altmann, D. R., Baker, D. & Schmierer, K. Axonal loss in the multiple sclerosis spinal cord revisited: Axonal Loss in the MS Spinal Cord Revisited. *Brain Pathol.* **28**, 334–348 (2018).
82. Lycklama à Nijeholt, G. J. *et al.* Sagittal MR of multiple sclerosis in the spinal cord: fast versus conventional spin-echo imaging. *AJNR Am. J. Neuroradiol.* **19**, 355–360 (1998).
83. Laroche, C., Uphaus, T., Prat, A. & Zipp, F. Secondary Progression in Multiple Sclerosis: Neuronal Exhaustion or Distinct Pathology? *Trends Neurosci.* (2016). doi:10.1016/j.tins.2016.02.001
84. Bjartmar, C., Wujek, J. R. & Trapp, B. D. Axonal loss in the pathology of MS: consequences for understanding the progressive phase of the disease. *J. Neurol. Sci.* **206**, 165–171 (2003).
85. Irvine, K. A. & Blakemore, W. F. Remyelination protects axons from demyelination-associated axon degeneration. *Brain* **131**, 1464–1477 (2008).
86. Tomassini, V. *et al.* Neuroplasticity and functional recovery in multiple sclerosis. *Nat. Rev. Neurol.* **8**, 635–646 (2012).

87. Weiss, S., Mori, F., Rossi, S. & Centonze, D. Disability in multiple sclerosis: When synaptic long-term potentiation fails. *Neurosci. Biobehav. Rev.* **43**, 88–99 (2014).
88. Evangelou, N., DeLuca, G. C., Owens, T. & Esiri, M. M. Pathological study of spinal cord atrophy in multiple sclerosis suggests limited role of local lesions. *Brain* **128**, 29–34 (2005).
89. Carassiti, D. *et al.* Neuronal loss, demyelination and volume change in the multiple sclerosis neocortex. *Neuropathol. Appl. Neurobiol.* **44**, 377–390 (2018).
90. Ganter, Prince & Esiri. Spinal cord axonal loss in multiple sclerosis: a post-mortem study. *Neuropathol. Appl. Neurobiol.* **25**, 459–467 (1999).
91. Biberacher, V. *et al.* Atrophy and structural variability of the upper cervical cord in early multiple sclerosis. *Mult. Scler. J.* **21**, 875–884 (2015).
92. Daams, M. *et al.* Mean upper cervical cord area (MUCCA) measurement in long-standing multiple sclerosis: Relation to brain findings and clinical disability. *Mult. Scler. J.* **20**, 1860–1865 (2014).
93. Kearney, H. *et al.* Improved MRI quantification of spinal cord atrophy in multiple sclerosis. *J. Magn. Reson. Imaging* **39**, 617–623 (2014).
94. Bernitsas, E. *et al.* Spinal cord atrophy in multiple sclerosis and relationship with disability across clinical phenotypes. *Mult. Scler. Relat. Disord.* **4**, 47–51 (2015).
95. Schlaeger, R. *et al.* Spinal cord gray matter atrophy correlates with multiple sclerosis disability. *Ann. Neurol.* **76**, 568–580 (2014).
96. Valsasina, P. *et al.* Regional Cervical Cord Atrophy and Disability in Multiple Sclerosis: A Voxel-based Analysis. *Radiology* **266**, 853–861 (2013).
97. Schlaeger, R. *et al.* Association Between Thoracic Spinal Cord Gray Matter Atrophy and Disability in Multiple Sclerosis. *JAMA Neurol.* **72**, 897 (2015).
98. Han, J. *et al.* NMR imaging of the spine. *Am. J. Roentgenol.* **141**, 1137–1145 (1983).
99. Modic, M. T. *et al.* Nuclear magnetic resonance imaging of the spine. *Radiology* **148**, 757–762 (1983).
100. Norman, D. *et al.* Magnetic resonance imaging of the spinal cord and canal: potentials and limitations. *Am. J. Roentgenol.* **141**, 1147–1152 (1983).

101. Pezold, S. *et al.* A semi-automatic method for the quantification of spinal cord atrophy. in *Computational methods and clinical applications for spine imaging : proceedings of the Workshop held at the 16th International Conference on Medical Image Computing and Computer Assisted Intervention, September 22-26, 2013, Nagoya, Japan* S. 143-155 (Springer, 2014).
102. Pezold, S. *et al.* Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features. in *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging* (eds. Yao, J., Glocker, B., Klinder, T. & Li, S.) 107–118 (Springer International Publishing, 2015).
doi:10.1007/978-3-319-14148-0_10
103. Papinutto, N. *et al.* Age, Gender and Normalization Covariates for Spinal Cord Gray Matter and Total Cross-Sectional Areas at Cervical and Thoracic Levels: A 2D Phase Sensitive Inversion Recovery Imaging Study. *PLoS ONE* **10**, e0118576 (2015).
104. Leener, B. D., Cohen-Adad, J. & Kadoury, S. Automatic Segmentation of the Spinal Cord and Spinal Canal Coupled With Vertebral Labeling. *IEEE Trans. Med. Imaging* **34**, 1705–1718 (2015).
105. Yiannakas, M. C., Kakar, P., Hoy, L. R., Miller, D. H. & Wheeler-Kingshott, C. A. M. The Use of the Lumbosacral Enlargement as an Intrinsic Imaging Biomarker: Feasibility of Grey Matter and White Matter Cross-Sectional Area Measurements Using MRI at 3T. *PLOS ONE* **9**, e105544 (2014).
106. Ge, Y. *et al.* Age-Related Total Gray Matter and White Matter Changes in Normal Adult Brain. Part I: Volumetric MR Imaging Analysis. *Am. J. Neuroradiol.* **23**, 1327–1333 (2002).
107. Terribilli, D. *et al.* Age-related gray matter volume changes in the brain during non-elderly adulthood. *Neurobiol. Aging* **32**, 354–368 (2011).
108. Ginsberg, L. Disorders of the spinal cord and roots. *Pract. Neurol.* **11**, 259–267 (2011).
109. Cook, S. A. A. SPINAL DISEASE: NEOPLASTIC, DEGENERATIVE, AND INFECTIVE SPINAL CORD DISEASES AND SPINAL CORD COMPRESSION. *Clinical Gate* (2015).
110. Paquin, M.-É. *et al.* Spinal Cord Gray Matter Atrophy in Amyotrophic Lateral Sclerosis. *Am. J. Neuroradiol.* **39**, 184–192 (2018).

111. Bach Cuadra, M., Duay, V. & Thiran, J.-Ph. Atlas-based Segmentation. in *Handbook of Biomedical Imaging: Methodologies and Clinical Research* (eds. Paragios, N., Duncan, J. & Ayache, N.) 221–244 (Springer US, 2015). doi:10.1007/978-0-387-09749-7_12
112. Wachinger, C. & Golland, P. Atlas-Based Under-Segmentation. *Med. Image Comput. Comput.-Assist. Interv. MICCAI Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* **17**, 315–322 (2014).
113. Horváth, A. *et al.* Variational Segmentation of the White and Gray Matter in the Spinal Cord Using a Shape Prior. in *Computational Methods and Clinical Applications for Spine Imaging: 4th International Workshop and Challenge, CSI 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 17, 2016, Revised Selected Papers* (eds. Yao, J. *et al.*) 26–37 (Springer International Publishing, 2016). doi:10.1007/978-3-319-55050-3_3
114. Lüthi, M., Albrecht, T. & Vetter, T. Probabilistic Modeling and Visualization of the Flexibility in Morphable Models. in *Mathematics of Surfaces XIII* 251–264 (Springer, Berlin, Heidelberg, 2009). doi:10.1007/978-3-642-03596-8_14
115. Asman, A. J., Smith, S. A., Reich, D. S. & Landman, B. A. Robust GM/WM Segmentation of the Spinal Cord with Iterative Non-Local Statistical Fusion. *Lect. Notes Comput. Sci.* 759–767 (2013).
116. Prados, F. *et al.* Spinal cord grey matter segmentation challenge. *NeuroImage* **152**, 312–329 (2017).
117. Perone, C. S., Calabrese, E. & Cohen-Adad, J. Spinal cord gray matter segmentation using deep dilated convolutions. *Sci. Rep.* **8**, 5966 (2018).
118. Abdel-Aziz, K. *et al.* Evidence for early neurodegeneration in the cervical cord of patients with primary progressive multiple sclerosis. *Brain J. Neurol.* **138**, 1568–1582 (2015).
119. De Stefano, N. *et al.* Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology* **74**, 1868–1876 (2010).
120. Zackowski, K. M. *et al.* Sensorimotor dysfunction in multiple sclerosis and column-specific magnetization transfer-imaging abnormalities in the spinal cord. *Brain* **132**, 1200–1209 (2009).
121. Rocca, M. A. *et al.* A multicenter assessment of cervical cord atrophy among MS clinical phenotypes. *Neurology* **76**, 2096–2102 (2011).

122. Kearney, H. *et al.* Spinal cord grey matter abnormalities are associated with secondary progression and physical disability in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **86**, 608–614 (2015).
123. Kearney, H. *et al.* Magnetic resonance imaging correlates of physical disability in relapse onset multiple sclerosis of long disease duration. *Mult. Scler. Houndmills Basingstoke Engl.* **20**, 72–80 (2014).
124. Gass, A. *et al.* MRI monitoring of pathological changes in the spinal cord in patients with multiple sclerosis. *Lancet Neurol.* **14**, 443–454 (2015).
125. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. *Ann. Neurol.* **50**, 121–127 (2001).
126. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, Supplement 1, S208–S219 (2004).
127. Magon, S. *et al.* Label-fusion-segmentation and deformation-based shape analysis of deep gray matter in multiple sclerosis: The impact of thalamic subnuclei on disability. *Hum. Brain Mapp.* **35**, 4193–4203 (2014).
128. Bergamaschi, R. Prognostic Factors in Multiple Sclerosis. in (ed. Neurobiology, B.-I. R. of) **79**, 423–447 (Academic Press, 2007).
129. Kappos, L. *et al.* The 11-year long-term follow-up study from the randomized BENEFIT CIS trial. *Neurology* 10.1212/WNL.0000000000003078 (2016). doi:10.1212/WNL.0000000000003078
130. Reddy, H. *et al.* Evidence for adaptive functional changes in the cerebral cortex with axonal injury from multiple sclerosis. *Brain* **123**, 2314–2320 (2000).
131. Inglese, M. MRI measures of neuroprotection and repair in multiple sclerosis. *J. Neurol. Sci.* **311**, S16–S23 (2011).
132. Favaretto, A., Lazzarotto, A., Margoni, M., Poggiali, D. & Gallo, P. Effects of disease modifying therapies on brain and grey matter atrophy in relapsing remitting multiple sclerosis. *Mult. Scler. Demyelinating Disord.* **3**, 1 (2018).
133. Thompson, A. J. *et al.* Primary progressive multiple sclerosis. *Brain* **120**, 1085–1096 (1997).
134. Nijeholt, G. J. *et al.* Brain and spinal cord abnormalities in multiple sclerosis. Correlation between MRI parameters, clinical subtypes and symptoms. *Brain* **121**, 687–697 (1998).

135. Kantarci, O. H. *et al.* Primary Progressive Multiple Sclerosis Evolving From Radiologically Isolated Syndrome. *Ann. Neurol.* **79**, 288–294 (2016).
136. Ingle, G. T., Stevenson, V. L., Miller, D. H. & Thompson, A. J. Primary progressive multiple sclerosis: a 5-year clinical and MR study. *Brain* **126**, 2528–2536 (2003).
137. Galego, O., Gouveia, A., Batista, S., Moura, C. & Machado, E. Brain atrophy and physical disability in primary progressive multiple sclerosis: A volumetric study. *Neuroradiol. J.* **28**, 354–358 (2015).
138. Ruggieri S, Petracca M, Miller A & et al. Association of deep gray matter damage with cortical and spinal cord degeneration in primary progressive multiple sclerosis. *JAMA Neurol.* **72**, 1466–1474 (2015).
139. Aymerich, F. X. *et al.* Cervical Cord Atrophy and Long-Term Disease Progression in Patients with Primary-Progressive Multiple Sclerosis. *Am. J. Neuroradiol.* (2018). doi:10.3174/ajnr.A5495
140. Bieniek, M. *et al.* Cord atrophy separates early primary progressive and relapsing remitting multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **77**, 1036–1039 (2006).
141. Antel, J., Antel, S., Caramanos, Z., Arnold, D. L. & Kuhlmann, T. Primary progressive multiple sclerosis: part of the MS disease spectrum or separate disease entity? *Acta Neuropathol. (Berl.)* **123**, 627–638 (2012).
142. Agosta, F. *et al.* Quantification of cervical cord pathology in primary progressive MS using diffusion tensor MRI. *Neurology* **64**, 631–635 (2005).
143. Cawley, N. *et al.* Spinal cord atrophy as a primary outcome measure in phase II trials of progressive multiple sclerosis. *Mult. Scler.* **23**, (2017).
144. Zeydan, B. *et al.* Cervical spinal cord atrophy. *Neurol. - Neuroimmunol. Neuroinflammation* **5**, (2018).
145. Maarouf, A. *et al.* Topography of brain sodium accumulation in progressive multiple sclerosis. *Magn. Reson. Mater. Phys. Biol. Med.* **27**, 53–62 (2013).
146. Stevenson, V. L. *et al.* Primary and transitional progressive MS: A clinical and MRI cross-sectional study. *Neurology* **52**, 839–839 (1999).
147. Narayana, P. A. *et al.* Multicentre proton magnetic resonance spectroscopy imaging of primary progressive multiple sclerosis. *Mult. Scler.* **10**, S73–S78 (2004).

148. Montalban, X. *et al.* Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N. Engl. J. Med.* **376**, 209–220 (2017).
149. Ontaneda, D., Fox, R. J. & Chataway, J. Clinical trials in progressive multiple sclerosis: lessons learned and future perspectives. *Lancet Neurol.* **14**, 208–223 (2015).
150. Liu, Z. *et al.* Cervical cord area measurement using volumetric brain magnetic resonance imaging in multiple sclerosis. *Mult. Scler. Relat. Disord.* **4**, 52–57 (2015).
151. Mendell, J. R. *et al.* Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017).
152. Horváth, A. *et al.* Variational Segmentation of the White and Gray Matter in the Spinal Cord Using a Shape Prior. in 26–37 (Springer, Cham, 2016). doi:10.1007/978-3-319-55050-3_3
153. Eshaghi, A. *et al.* Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann. Neurol.* **83**, 210–222 (2018).
154. Lunn, M. R. & Wang, C. H. Spinal muscular atrophy. *Lancet Lond. Engl.* **371**, 2120–2133 (2008).
155. Mercuri, E., Bertini, E. & Iannaccone, S. T. Childhood spinal muscular atrophy: controversies and challenges. *Lancet Neurol.* **11**, 443–452 (2012).
156. Shababi, M., Lorson, C. L. & Rudnik-Schöneborn, S. S. Spinal muscular atrophy: a motor neuron disorder or a multi-organ disease? *J. Anat.* **224**, 15–28 (2014).
157. Therapy for Spinal Muscular Atrophy. *N. Engl. J. Med.* **378**, 487–488 (2018).
158. Bonati, U. *et al.* Longitudinal characterization of biomarkers for spinal muscular atrophy. *Ann. Clin. Transl. Neurol.* **4**, 292–304 (2017).
159. Trapp, B. D. *et al.* Cortical neuronal densities and cerebral white matter demyelination in multiple sclerosis: a retrospective study. *Lancet Neurol.* **17**, 870–884 (2018).
160. Lublin, F. D. & Reingold, S. C. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology* **46**, 907–911 (1996).

161. Coetzee, T. & Thompson, A. J. Unified understanding of MS course is required for drug development. *Nat. Rev. Neurol.* **14**, 191–192 (2018).

8. Curriculum Vitae

Nationality: Greek

DOB: 26.07.1990

Education and Professional Career

April 2016 – Present

Neurology residency in the Neurological Policlinic, Department of Neurology, University Hospital Basel, Switzerland

May 2015 – Present

PhD in clinical research at University of Basel, Department of Neurology, University Hospital Basel, Switzerland

Title: „Spinal Cord Volume Quantification and Clinical Application in Multiple Sclerosis”

December 2014

Advanced Life Support Provider Certificate, European Resuscitation Council

July 2014 – Jan. 2014

Amyotrophic Lateral Sclerosis research, Neuroimmunological Laboratory, Eginition University Hospital, Athens, Greece

Sept. 2008 – Aug. 2014

Medical Studies at the Faculty of Medicine, National and Kapodistrian University of Athens

Final grade: “Very Good” (7.52/10)

Sept. 2005 – July 2008

1st Lykeion Agiou Dimitriou, Greek High School

State Panhellenic Examinations final grade: “Distinction” (19.25/20)

Scholarships and Awards

January to December 2019	Competitive 1-year grant of the special program „Nachwuchsförderung Klinische Forschung“, University of Basel Project “Structural MRI biomarkers in multiple sclerosis”
15-17 November 2018	Travel Grant for 26th Annual Meeting, European Charcot Foundation
2 February 2018	Travel Grant for 13th annual ARSEP MRI workshop
25-28 October 2017	Travel Grant Award for the 7th joint ACTRIMS-ECTRIMS Congress
14-17 September 2016	Travel Grant Award for the 32nd ECTRIMS Congress
March 2012	Full 3-year Scholarship “Mavrokordatou” through exams in Pathophysiology, Anatomic Pathology and Essay
September 2007	All - Greek top 400 high school students

Publications

- Athina Papadopoulou, Laura Gaetano, Armanda Pfister, Anna Altermatt, Charidimos Tsagkas, Felix Morency, Alexander U. Brand, Martin Hardmeier, Mallar M. Chakravarty, Maxime Descoteaux, Ludwig Kappos, Till Sprenger, and Stefano Magon, Damage of the lateral geniculate nucleus in MS: assessing the missing node of the visual pathway (in press in Neurology)
- Horvath A*, Tsagkas C*, Andermatt S, Pezold S, Parmar K, Cattin P. Spinal Cord Gray Matter-White Matter Segmentation on Magnetic Resonance AMIRA Images with MD-GRU. MICCAI 2018 CSI workshop. arXiv:180802408 [cs]. 2018;<http://arxiv.org/abs/1808.02408>. Accessed August 22, 2018. * equally contributing authors
- Barro, C., Benkert, P., Disanto, G., Tsagkas, C., Amann, M., Naegelin, Y., Leppert, D., Gobbi, C., Granziera, C., Yaldizli, O., Michalak, Z., Wuerfel, J., Kappos, L., Parmar, K., Kuhle, J. Serum neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in multiple sclerosis – Brain. 2018 May 30. doi: 10.1093/brain/awy154.
- Tsagkas, C., Magon, S., Gaetano, L., Pezold, S., Naegelin, Y., Amann, M., Stippich, C., Cattin, P., Wuerfel, J., Bieri, O., Sprenger, T., Kappos, L., Parmar, K., 2018. Preferential Spinal Cord Volume Loss in Primary Progressive Multiple Sclerosis – Mult Scler. 2018 May 1:1352458518775006. doi: 10.1177/1352458518775006
- Tsagkas, C., Magon, S., Gaetano, L., Pezold, S., Naegelin, Y., Amann, M., Stippich, C., Cattin, P., Wuerfel, J., Bieri, O., Sprenger, T., Kappos, L., Parmar, K., 2018. Spinal Cord Volume Loss: A Marker of Disease Progression in Multiple Sclerosis – Neurology. 2018 Jul 24;91(4):e349-e358. doi: 10.1212/WNL.0000000000005853
- Tsagkas, C., Altermatt, A., Bonati, U., Pezold, S., Reinhard, J., Amann, M., Cattin, P., Wuerfel, J., Fischer, D., Parmar, K., Fischmann, A., 2018. Reliable and fast volumetry of the lumbar spinal cord using cord image analyser (Cordial). Eur Radiol. 2018 Apr 30. doi: 10.1007/s00330-018-5431-1
- Pezold, S., Horváth, A., Fundana, K., Tsagkas, C., Andělová, M., Weier, K., Amann, M., Cattin, P.C., 2016. Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures, in: Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016. Presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, pp. 362–370. https://doi.org/10.1007/978-3-319-46726-9_42

Posters and Presentations

- “Spinal Cord MRI”, 26th Annual Meeting of the European Charcot Foundation, 15-17 November 2018, Baveno, Italy – Invited oral presentation
- “Cortical thinning in multiple sclerosis: a 6-year longitudinal study”, 32nd European Committee for Teaching and Research in Multiple Sclerosis, 10-12 October 2018, – Poster presentation
- “Reliability Of Automatic Spinal Cord Grey Matter Segmentation Using Averaged Magnetization Inversion Recovery Acquisitions”, 13th annual a la recherche sur la Sclérose en Plaques (ARSEP) MRI workshop, February 2nd 2018, Paris – Oral presentation
- “Preferential Spinal Cord Volume Loss in Primary Progressive Multiple Sclerosis”, 7th joint American Committee for Teaching and Research in Multiple Sclerosis European Committee for Teaching and Research in Multiple Sclerosis, 25-28 October 2017, Paris – Poster presentation
- “Preferential Spinal Cord Volume Loss in Primary Progressive Multiple Sclerosis”, Department Biomedical Engineering (DBE) Research Day 2017, 23 August 2017, University Hospital Basel – Poster presentation
- „CNS MRI-markers in Multiple Sclerosis“, Seminar “Bench to Bedside in Neuroscience”, 4 April 2017, University Basel – Oral presentation
- “Preferential Spinal Cord Volume Loss in Primary Progressive Multiple Sclerosis”, Clinical Research Day 2017, 19 January 2017, University Hospital Basel – Poster presentation
- “Spinal Cord Atrophy-a new biomarker?”, Research Retreat 2016, 1 December 2016, University Hospital Basel – Oral presentation
- “Spinal Cord Volume Loss Predicts Disease Activity and Disability over Time in Multiple Sclerosis”, 32nd European Committee for Teaching and Research in Multiple Sclerosis, London, 14-17 September 2016 – Oral presentation
- “Spinal Cord Volume Loss Predicts Disease Activity and Disability over Time in Multiple Sclerosis”, Department Biomedical Engineering (DBE) Research Day, 7 September 2016, University Hospital Basel – Poster presentation
- “Spinal Cord Volume Loss in Multiple Sclerosis Patients: A 7-Year Longitudinal Study”, 31st European Academy of Neurology, 28-31 May 2016, Copenhagen – Oral presentation
- “Spinal Cord Volume Loss in Multiple Sclerosis Patients: A 7-Year Longitudinal Study”, Clinical Research Day 2016, 28 January 2016, University Hospital Basel– Poster presentation

PhD Courses

- Programmieren lernen mit R (2 ECTS)
- Data-processing pipelines in Life Science (2 ECTS)
- From bench to bedside (1 ECTS)
- Clinical Neurosciences (3 ECTS)
- Neuropsychology und behavioral neurology (4 ECTS)
- Self-Branding and Self-Promotion (1 ECTS)
- Neurological Diseases (2 ECTS)
- Principles of Medical Imaging (3 ECTS)
- Designing Your PowerPoint Presentation (0 ECTS)

Teaching

- Seminar Bench to Bedside in Neuroscience, spring semester 2017, CNS MRI-markers in Multiple Sclerosis
- Supervision of two Master medical students
- Autumn semester 2018, Arzt-Patienten-Unterricht (AP), Neurologie USB

Other Certifications

- Introduction to MR-Safety, 17 June 2015, Klinik für Radiologie und Nuklearmedizin, University Hospital Basel
- Good Clinical Practice: Research Ethics Evaluation; Introduction to Research Ethics; Informed Consent; Good Clinical Practice (GCP), 27 May 2015
- Advanced Life Support Provider Certificate, European Resuscitation Council, December 2014

Languages

- Greek – excellent (native speaker)
- English – proficient (both written and spoken)
- German – proficient (both written and spoken)