



Review



Cite this article: Szekely-Kohn AC, Castellani M, Espino DM, Baronti L, Ahmed Z, Manifold WGK, Douglas M. 2025 Machine learning for refining interpretation of magnetic resonance imaging scans in the management of multiple sclerosis: a narrative review. *R. Soc. Open Sci.* **12:** 241052.
<https://doi.org/10.1098/rsos.241052>

Received: 25 June 2024

Accepted: 17 November 2024

Subject Category:

Computer Science and Artificial Intelligence

Subject Areas:

neuroscience, computational biology, artificial intelligence

Keywords:

artificial intelligence, computational methods, machine learning, magnetic resonance imaging (MRI), multiple sclerosis (MS)

Author for correspondence:

Adam C. Szekely-Kohn

e-mail: axs1090@student.bham.ac.uk

Machine learning for refining interpretation of magnetic resonance imaging scans in the management of multiple sclerosis: a narrative review

Adam C. Szekely-Kohn¹, Marco Castellani¹, Daniel M. Espino¹, Luca Baronti², Zubair Ahmed^{4,3}, William G. K. Manifold⁵ and Michael Douglas^{4,3,6,7}

¹School of Engineering, ²School of Computer Science, and ³Institute of Inflammation and Ageing, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

⁴University Hospitals Birmingham NHS Foundation Trust, Edgbaston, Birmingham B15 2GW, UK

⁵The Royal London Hospital, Barts Health NHS Trust, Whitechapel Road, London E1 1FR, UK

⁶Department of Neurology, Dudley Group NHS Foundation Trust, Russells Hall Hospital, Birmingham DY1 2HQ, UK

⁷School of Life and Health Sciences, Aston University, Birmingham, UK

DOI ACS-K, 0009-0004-1005-603X; DME, 0000-0001-7608-5619;
LB, 0000-0002-2883-706X

Multiple sclerosis (MS) is an autoimmune disease of the brain and spinal cord with both inflammatory and neurodegenerative features. Although advances in imaging techniques, particularly magnetic resonance imaging (MRI), have improved the process of diagnosis, its cause is unknown, a cure remains elusive and the evidence base to guide treatment is lacking. Computational techniques like machine learning (ML) have started to be used to understand MS. Published MS MRI-based computational studies can be divided into five categories: automated diagnosis; differentiation between lesion types and/or MS stages; differential diagnosis; monitoring and predicting disease progression; and synthetic MRI dataset generation. Collectively, these approaches show promise in assisting with MS diagnosis, monitoring of disease activity and prediction of future progression, all potentially contributing to disease management. Analysis quality using ML is highly dependent on the dataset size and variability used for training. Wider public access would mean larger datasets for experimentation, resulting in higher-quality analysis, permitting for more

conclusive research. This narrative review provides an outline of the fundamentals of MS pathology and pathogenesis, diagnostic techniques and data types in computational analysis, as well as collating literature pertaining to the application of computational techniques to MRI towards developing a better understanding of MS.

1. Introduction

A chronic inflammatory condition, multiple sclerosis (MS), is characterized by damage to myelin within the central nervous system (CNS)—demyelination—and the resultant loss of function (see figure 1). Several potential factors have been identified as significant risk factors for MS, including previous infection with Epstein–Barr virus [1,2], geographical latitude [3], month of birth [4] and certain genetic polymorphisms [5,6]. In common with many autoimmune diseases, MS disproportionately affects women, with a female-to-male ratio of 3:1 [7].

Myelin is composed of proteins and lipids and acts to increase the efficiency of signal propagation from a cell body to its dendrites/terminals [8]. Myelin production and maintenance are dependent on oligodendrocytes, a specialized type of glial cell [9]. Several types of inflammatory cells are likely to be involved in the pathogenesis of MS, including CD4⁺ T-lymphocytes, which recognize antigen peptide fragments bound to major histocompatibility complex class II proteins. These antigenic fragments may be derived from a range of autoantigens such as myelin basic proteins, myelin proteolipid protein, myelin oligodendrocyte glycoprotein and myelin-associated glycoprotein [10,11]. T-lymphocyte activation in the peripheral blood or secondary lymphoid tissues results in subsequent crossing of the blood–brain barrier (BBB) and damage within the CNS.

1.1. Multiple sclerosis pathology

At a cellular level, MS is fundamentally a demyelinating disease, stripping axons of their myelin sheath, leading to energy deficiencies and ion imbalances in brain tissue [12], ultimately resulting in axonal destruction [13]. This process is depicted in figure 2.

Myelin-specific T-lymphocytes, in conjunction with tissue-resident cells, generate proinflammatory cytokines (such as interleukin-1, interleukin-6, tumour necrosis factor- α and interferon- γ (IFN- γ)) that have local effects including amplification of the inflammatory cascade and increasing influx of other cell populations from the peripheral blood (e.g. B-lymphocytes) [14]. IFN- γ can also cause damage to the oligodendrocytes, limiting remyelination capability [15]. Alongside T-lymphocytes, B-cells produce antibodies that target cell surface proteins expressed on the myelin sheath and oligodendrocytes [16,17]. Amplification of the inflammatory phase is followed by a period of inflammatory resolution, likely to involve additional cell populations such as regulatory T-lymphocytes [18]. As the inflammatory event resolves, tissue restorative and remodelling processes begin (remyelination) but with super-added areas of tissue scarring (gliosis). Over time, remyelination efficiency decreases and gliosis becomes more prominent [19]. The inflammatory processes lead to foci of inflammatory activity within the CNS [20], which can be visualized on magnetic resonance imaging (MRI) and produce clinical correlates depending on their location. The associated symptomatology can include visual disturbance [21] (from optic neuritis or brainstem lesions), limb weakness [22,23] and sensory disturbance [24,25] (from spinal cord lesions).

1.2. A brief history of multiple sclerosis

The pathology of MS was first formally documented at the end of the nineteenth century with the first key observations on MS through the work of Carl Frommann [26] and Jean-Martin Charcot [27]. Frommann focused on spinal cord lesions and the axonal destruction present within those tissues [26], whereas the work of Charcot centred on brain lesions, in which he correctly postulated that MS was primarily a demyelinating disease resulting in axonal injury [27]. This view was subsequently supported by the work of Joseph Babinski [28]. The subsequent research of Max Fraenkel and Alfons Jakob highlighted the roles of macrophages during demyelination and axonal destruction [29]. Using an animal model of demyelination, Thomas Rivers was the first to propose that MS was primarily an

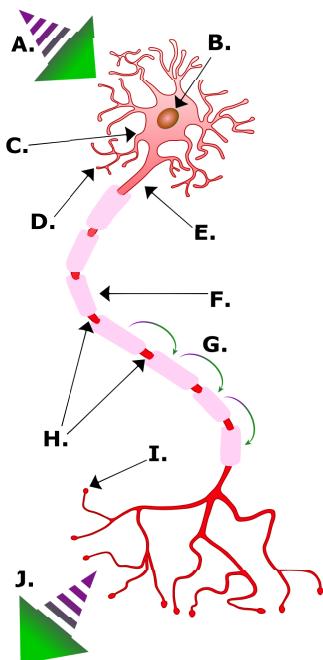


Figure 1. A labelled illustration of a myelinated neural cell. A, nerve signal input; B, cell nucleus; C, soma; D, dendrites; E, axon; F, myelin sheath; G, nerve signal; H, nodes of Ranvier; I, synaptic terminals; J, nerve signal output (original image).

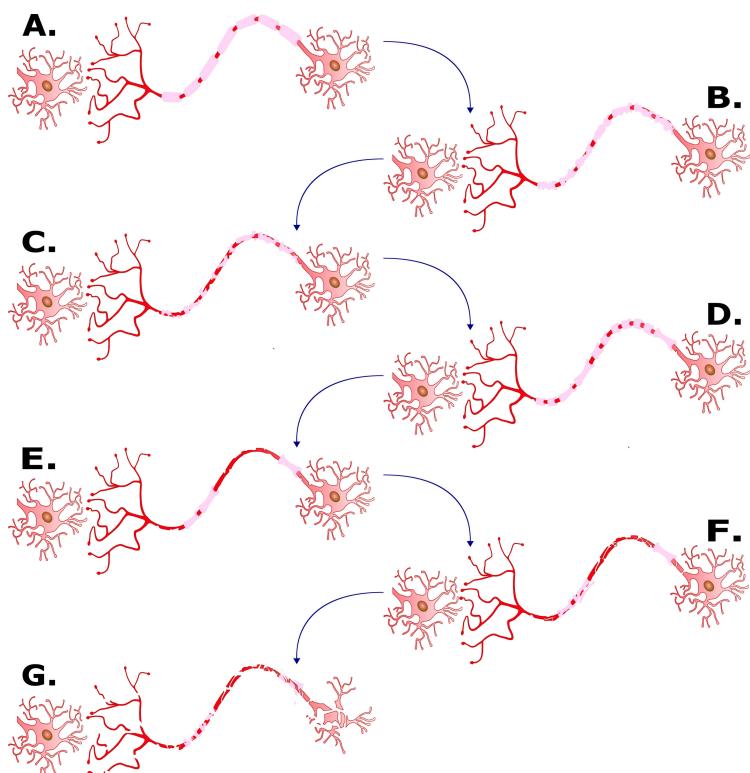


Figure 2. A graphical depiction of the demyelination process. A, healthy neural cell; B, demyelination begins to take place due to malfunctioning immune system; C, myelin is stripped from the axon; D, oligodendrocytes and Schwann cells induce partial remyelination; E, subsequent inflammation event(s) destroys the capacity for remyelination; F, axon begins to degenerate; G, neural cell degenerates and disintegrates (original image).

autoimmune disease and not an infective (particularly viral) condition [30], although evidence for a role for a viral trigger, particularly Epstein–Barr virus, continues to accumulate [1,2,31]. Alongside a primary role for tissue inflammation-mediated demyelination, additional evidence highlights that grey matter damage is also prominent, providing another neurodegenerative element to MS pathology [32].

Imaging techniques, in particular MRI, helped further the understanding of MS. MRI became increasingly important in the diagnosis and monitoring of MS patients from the 1990s onwards, with a steady increase in the range of modalities of MRI used in visualization of various aspects of the disease [33,34]. Alongside advances in MRI, there has been a significant increase in therapies to reduce inflammatory disease activity and consequent disability accumulation (disease-modifying treatments), from earlier studies on interferons, most prominently to a range of anti-CD20 treatments. Although effective, a significant proportion of people with MS continue to experience clinical or radiological disease activity and/or progressive accumulation of disability and none of the presently available treatments convincingly reverse established damage or disability [35,36]. The worldwide prevalence of MS is steadily increasing [37], meaning that the development of new treatments remains a key priority.

2. Multiple sclerosis subtypes and treatments

The following four well-established subtypes of MS exist based on clinical understanding: (i) clinically isolated syndrome (CIS), (ii), relapsing–remitting MS (RRMS), (iii) primary progressive MS (PPMS), and (iv) secondary progressive MS (SPMS) [38]. These groups are primarily defined by their patterns of clinical disease activity (clinical relapses and accumulation of disability), supported by MRI. A further less well-established category exists known as radiologically isolated syndrome, in which patients are diagnosed (through MRI) with pre-symptomatic demyelination, with radiology consistent with MS. The clinical significance and management of this category of patient are the subject of debate and ongoing research [39].

2.1. Clinically isolated syndrome

The CIS diagnosis involves a single clinical event, usually the result of a focal inflammatory demyelinating lesion within CNS with resultant symptoms consistent with MS [40]. By definition, the episode must be 24 h in duration or greater and without evidence for an alternative trigger, such as infection [41]. The inflammatory damage typically manifests itself in a spatially isolated area, namely a single lesion. This is illustrated in figure 3, in which the red line forming a raised peak is representative of a clinical attack that later subsides, the residual long-term effect of which is increased disability (the level of baseline disability does not return to a baseline level following the event).

The clinical presentation of the CIS can include optic neuritis or transverse myelitis, or less specific symptoms such as fatigue or cognitive issues [40]. With respect to treatment, CIS often resolves itself without the need to administer medication; however, corticosteroids such as methylprednisolone are often used to reduce inflammation and speed up patient recovery [43]. It should be noted, however, that steroid treatment does not impact either the likelihood of subsequent relapses or general long-term prognosis, i.e. is not considered disease modifying [44]. CIS can be a single event in the life of a person with no recurrence, but if a person experiences two or more episodes, this would confirm disease chronicity and their diagnosis should be changed to ‘clinically definite’ MS [45]. Even in the absence of clinically evident symptoms or signs, if an individual with CIS is monitored radiologically (via MRI) over time and there is evidence of MRI change (for example, the appearance of new MRI lesions), this again suggests disease chronicity and that a change in diagnosis to MS should be made accordingly.

2.2. Relapsing–remitting multiple sclerosis

The most common form of MS is RRMS, with approximately 85% of all MS cases being diagnosed with this subtype [46,47] and often following an episode previously labelled as CIS (the first clinical event). RRMS is characterized by at least two separate bouts of inflammation in the brain [48], which evolve over 48–72 h, the symptoms of which can last anywhere between two and eight weeks before subsiding. Sometimes an affected patient will fully recover from an episode, although it is more common for recovery to be only partial [49], so that the level of any resulting disability will fluctuate and even accumulate over time. This can be seen in figure 4, in which disability fluctuates with time and clinical attacks are episodic. The greater the number of relapses, the greater the likelihood there is of permanent disability [50]. Disease-modifying therapies (DMTs) aim to reduce inflammation, reduce clinical relapses and reduce the resultant accumulation of disability [51].

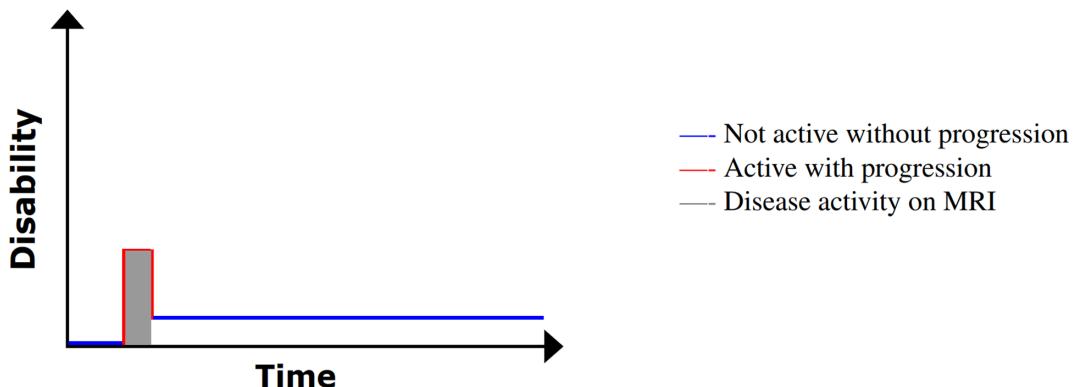


Figure 3. An illustration of disability progression with time for CIS. Image adapted from My-MS.org [42].

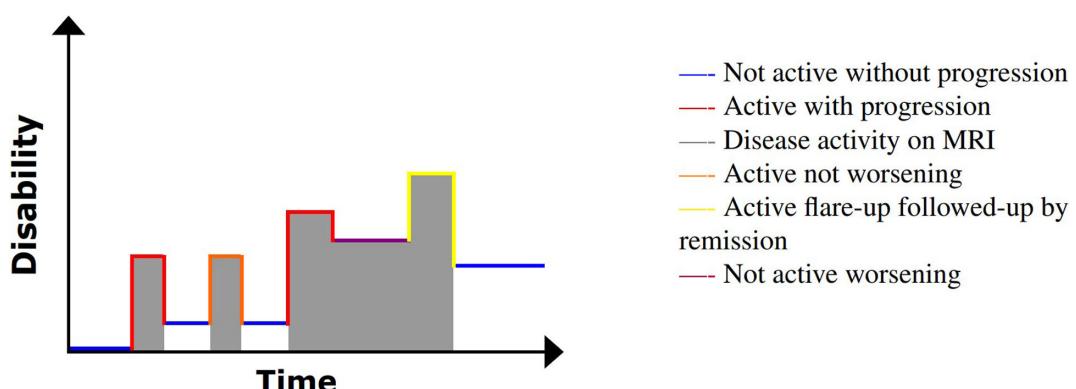


Figure 4. An illustration of disability progression with time for RRMS. Image adapted from My-MS.org [42].

As with CIS, corticosteroids are often prescribed to reduce acute tissue inflammation [52], but steroid-related side effects are significant including insomnia, increased blood pressure, loss of muscle, weight gain, diabetes, cataracts, mood disorders and osteoporosis [44,53]. As a result, steroids are typically administered as a short 'pulse' for 3–5 days.

There are now multiple DMTs that can be used in RRMS treatment [54]. Depending on the specific type, DMT can be administered as oral medication, subcutaneous or intramuscular injection or intravenous infusions. The best-established therapies were preparations of interferon- β with demonstrable effects on clinical and radiological MS activity, acting through ill-defined mechanisms. Interferons are naturally occurring proteins, divided into the following three subtypes: α , β and γ [55]. It is interferon- β that is used to treat MS [54,56,57]. Two varieties exist, $\beta-1a$ and $\beta-1b$, produced in either bacterial or mammalian cell cultures [58].

Over the subsequent years, the development of biological (infused antibodies) agents and small molecules have revolutionized the clinical management of MS. One is natalizumab (branded as Tysabri), a recombinant antibody that binds to the adhesion molecule ($\alpha 4$ -integrin) on white blood cells [59], reducing inflammatory cell transmigration across the BBB, resulting in reduced radiological and clinical disease activity.

2.3. Secondary progressive multiple sclerosis

A further stage in the clinical evolution of MS includes a more prominent role for non-relapse-mediated disease progression—SPMS, in which the relapse–remission cycle is less clinically and radiologically evident [60,61], typically evolving into a more continuous accumulation of disability without clear evidence of recovery from disability. This is demonstrated in figure 5; clinical attacks do occur, but worsening in disability can occur without an inflammatory exacerbation.

The term 'progression' is defined as the continual deterioration of neurological symptoms (or signs) for at least six months [38,61]. The condition is often retrospectively diagnosed [60]. SPMS is more challenging to treat than RRMS as DMTs have less well-defined effects on progression [62].

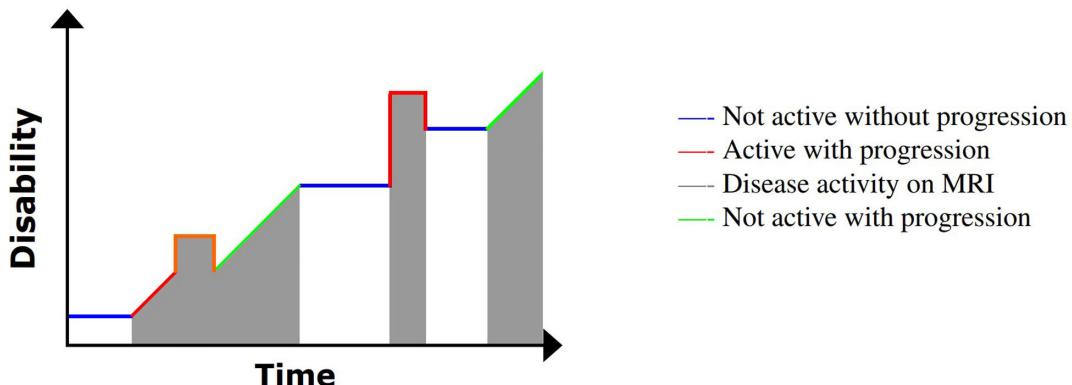


Figure 5. An illustration of disability progression with time for SPMS. Image adapted from My-MS.org [42].

For non-active SPMS, the current standard treatment is primary through physiotherapy to improve strength, balance and mobility and occupational therapy for energy conservation and speech and language therapy to maintain bulbar function [63].

2.4. Primary progressive multiple sclerosis

The final MS subtype, PPMS, is characterized from the stage of diagnosis with continuous accumulation of disability (frequently mobility) and no sustained periods of remission [64]. This continuous disability progression is displayed in figure 6.

Between 10 and 15% of MS patients are diagnosed with PPMS [65]. There are several characteristics that distinguish it from other MS subtypes: lesions are more frequently located in the spinal cord and surrounding white matter, whereas cerebral lesions are more limited as compared to other MS subtypes [66]. Another feature that differentiates PPMS is that while every other subtype of MS disproportionately affects women, the gender distribution of PPMS is approximately equal [64]. In a similar manner to SPMS, DMTs for PPMS are more limited. At the time of writing, the only approved treatment for PPMS is the anti-CD20 antibody treatment ocrelizumab (also licensed for the treatment of RRMS) [35,67]. Although effective, it is, in essence, an anti-inflammatory, and like other DMTs, its impact is reduced in the latter stages of the disease when inflammatory activity is less prominent. Similar to SPMS, the mainstays of treatment are physiotherapy, occupational therapy and speech-language therapy, all of which are aimed at mitigating against the functional decline caused by the disease [63].

3. Regional lesion and atrophy progression in multiple sclerosis

The brain primarily comprises white and grey matter, which are both essential to healthy body functions, with damage to either causing serious negative consequences to a person. These tissue types are affected by MS differently, and understanding the pathogenesis with respect to both tissue types can offer a more complete insight into MS disease progression. A labelled brain slice of an MS patient in the horizontal plane is presented in figure 7, illustrating the distribution of white and grey matter, lesions and important areas of the brain.

3.1. White matter

In general, the damage inflicted on MS patients is most radiologically evident where white matter predominates. White matter is primarily located in the innermost layer of the cortex but also encases the spinal cord. Its constituents include glia such as oligodendrocytes and fibrous astrocytes, and it is light in colour due to its high myelin content. In the early stages of the disease, when relapses are most frequent [68], tissue damage is concentrated in the deeper regions of the brain such as the subarachnoid space and cortex, where white matter is prevalent [69,70]. Focal confluent plaques are common in the white matter of the periventricular region of the brain [13,68]. It has been reported that in these plaques inflammation is more prominent in earlier stages of their formation [68,71]. Their

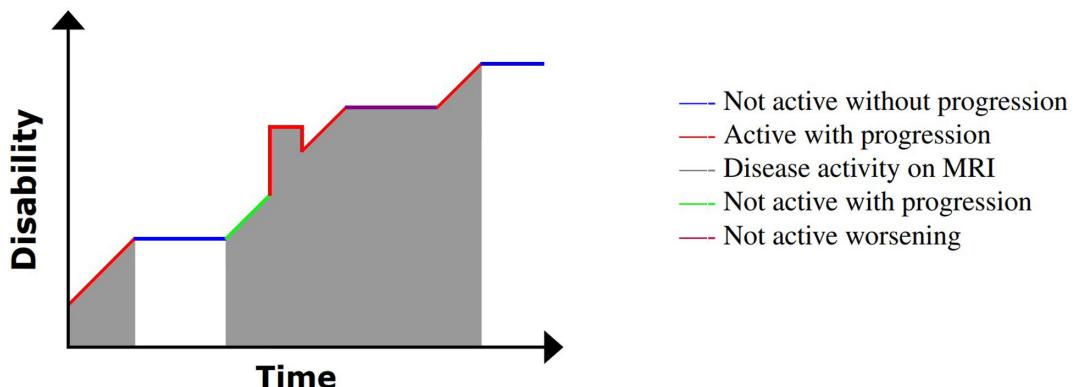


Figure 6. An illustration of disability progression with time for PPMS. Image adapted from My-MS.org [42].

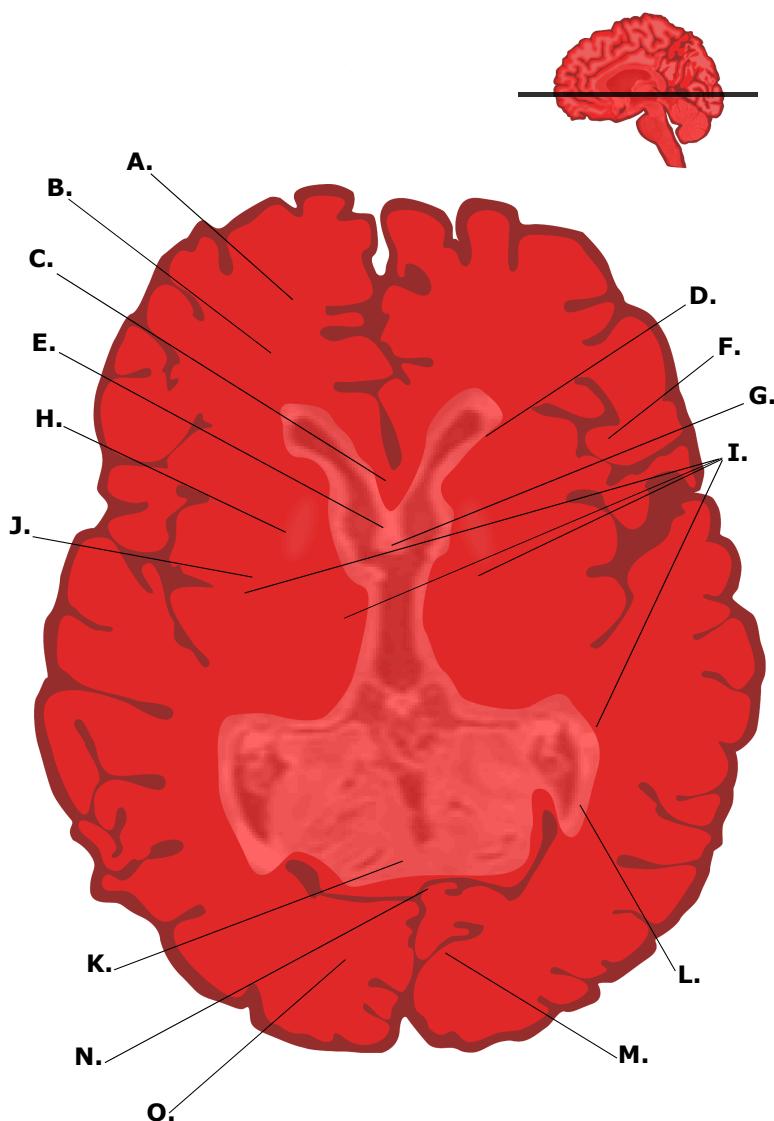


Figure 7. A labelled graphical illustration of a brain slice of an MS patient in the horizontal plane, where pale white spots are indicative of lesion location. A, frontal lobe; B, cerebral white matter; C, corpus callosum genu; D, lateral ventricle, anterior horn; E, septum pellucidum; F, cerebral grey matter; G, third ventricle; H, globus pallidus; I, white matter lesions; J, putamen; K, corpus callosum trunk; L, lateral ventricle, posterior horn; M, longitudinal cerebral fissure; N, posterior forceps; O, occipital lobe (original image).

size and frequency increase with disease progression [72]. Microscopically, there is damage to the glia, the type of cells that help with neuronal protection and maintenance, with degeneration of myelin and loss of oligodendrocytes [68]. Demyelination usually occurs where myelin is in close proximity to activated macrophages or microglia [28]. This process most obviously manifests itself radiologically with the emergence of ‘Dawson’s fingers’ [73], appearing as vividly clear areas of demyelination on MRI scans. As mentioned in previous sections, some repair of myelin can occur during the phase of demyelination, which may correlate with clinical improvement as seen in RRMS. Astrocytic scarring develops during the healing process where any demyelination has occurred, creating a barrier between healthy and damaged or dead tissue and forming lesions [74,75]. Although lesion demarcation is a defence mechanism against uncontrolled CNS inflammation, it similarly impairs post-inflammatory tissue regeneration, likely contributing to tissue scarring and disability accumulation [75].

Active white matter lesions are most common during the early phase of the disease, becoming progressively less common as the disease develops as it transforms into the progressive phase [13]. By contrast, diffuse changes are minimal in normal-appearing white matter (NAWM) but increase as MS reaches a progressive stage [72]. Oxidative injury also occurs in the white matter, even in the early stages of MS [76]. In response to inflammation and the production of oxygen radicals, cellular damage occurs [68], specifically to the neuronal mitochondria, which results in virtual hypoxia: an energy deficiency leading to ionic instability in the damaged cells [13]. The reason for the prominence of this phenomenon as compared to other neuroinflammatory diseases is unknown, but as a consequence, respiratory deficiency in cells correlates with disease duration (and patient age) [68]. The impact of this form of damage becomes more important with disease duration and significantly contributes to oxidative tissue injury. In addition, during the later stages of the disease, the loss of oligodendrocytes becomes increasingly widespread, removing the primary mechanism of repair [77]. This, combined with axon loss, impairs tissue regeneration and repair. In terms of leukocortical lesions, present in both white and grey matter, remyelination tends to be less robust in the white matter than in grey [78,79], possibly due to differences in macroglial populations. Remyelination rates in white matter alone can also be dependent on the affected brain region. Subcortical white matter lesions, for example, are more likely to remyelinate than those situated in the periventricular regions [68,80]. Inflammation in white matter during the progressive phase of MS is similar to age-matched controls [68], suggesting that neurodegenerative processes are dominant in the latter stages of the disease [81]. In a similar manner, damage to the BBB progresses over time; when it becomes sufficiently severe, the level of permeability reaches a nadir such that diffuse injury in the NAWM begins to occur, potentially contributing to brain atrophy [68].

3.2. Grey matter

MS was primarily considered to be a white matter demyelinating disease, but radiological and histological advances over the past 20 years have confirmed that all cerebral tissue types are affected. Grey matter, from disease outset, is significantly impacted [73], but visualization and analysis of the resulting damage have, up until recently, been difficult [82]. As white matter demyelination (readily demonstrable on MRI) does not correlate well with many clinical symptoms (particularly cognitive), research into the drivers and consequences of grey matter injury remains a priority [83]. One key example is the strong correlation between cortical damage (for example, seen as atrophy in MRI) with cognitive deficits [84]. The fundamental pathological difference between white and grey matter lesions is that the former are usually associated with inflammation caused by leukocyte penetration of the BBB, while relatively few leukocytes are found in and around the latter [85]. Axonal and synaptic damage are characteristic of grey matter lesions [86,87]. There is significant variation in cortical demyelination in the early stages of the disease to that of later chronic MS [69]. Large quantities of cortical demyelination are characteristic of MS [88] and tend to be located in the sulci [89] and are associated with localized meningeal inflammation [68,90]. Although present in RRMS, cortical lesions become far more numerous in the progressive stages of MS, correlating with an increase in diffuse microglia reactions in the meninges and perivascular white matter [72]. It has been suggested that sodium channel damage and virtual hypoxia due to mitochondrial damage act as catalysts for grey matter degeneration [91]. It has also been reported that neurons with damaged axons require more oxygen to function, thus further straining oxygen supply and leaving them susceptible to further damage [92].

Grey matter damage is macroscopically characterized by cortical thinning—the shrinking of the cortex, which gradually increases with disease duration [82]—and does not seem to significantly vary with the course of the disease [93]. This phenomenon is largely confined to the frontal and temporal regions of the brain [82,94]. During the CIS and relapsing-remitting phases of relapsing MS, atrophic damage was first observed in the posterior cingulate cortex and precuneus, which was then seen to spread to the middle cingulate cortex, brainstem and thalamus [95]. Progression of atrophy in people with PPMS shares some common features with those suffering from CIS and RRMS but with some differences. The thalamus, cuneus, precuneus and pallidum were the first regions affected, with brainstem and posterior cingulate cortex injury subsequently occurring [96]. While the cerebellum, caudate and putamen appear to atrophy in the early stages of disease progression in the relapsing form, they become atrophic later in PPMS [95]. It is clear, however, that irrespective of the course of the disease, damage to regions of grey matter across the brain is unevenly distributed [97] and that damage does not develop in a stochastic manner [95,98].

3.3. Lesions

Lesions can be categorized by anatomical location, according to their visual appearance on MRI or inflammatory activity [20,99,100]. These classifications help in understanding the nature and implications of the lesions observed. The anatomical locations are as follows: periventricular (relating to lesions surrounding the ventricles), cortical/juxtacortical (lesions in and around the cortex), infratentorial lesions (pertaining to the cerebellum and brainstem) and those affecting the spinal cord and the optic nerves [99,101], with the latter often acting as an early indicator of MS [102]. Cortical lesions can be further divided into the following four categories: type I—leukocortical lesions, affecting cortical grey matter and juxtacortical white matter [103]; type II—perivenous intracortical lesions, affecting exclusively grey matter, which are usually small in size; type III—subpial lesions, which occur in the white matter on the brain surface, protruding inwards; and finally type IV—which occur throughout the grey matter within the cortex [20].

In terms of visual intensity, lesions can be divided into the following two categories: T2-weighted hyperintense lesions and T1-weighted hypointense lesions. T2 lesions appear brighter than surrounding white matter on T2, fluid-attenuated inversion recovery (FLAIR) and proton density (PD) scans (hyperintense). This relationship is inverted on T1 scans, in which lesions are either darker or indistinguishable from surrounding tissue. T1 lesions can be further divided into two subcategories: those that enhance when gadolinium contrast agent is administered (5 min prior to a T1 scan) and those that do not. The former, known as both gadolinium-enhancing lesions and paramagnetic lesions, indicate localized increased permeability (breakdown) of the BBB and the presence of current or recent inflammatory activity [104]. Without the injection of contrast agent, these lesions can be more easily missed by a radiologist or clinician. Other features may include ‘black holes’, focal areas of T1 hypointensity images. These do not enhance after a gadolinium injection and are thought to indicate foci of severe CNS damage relative to other lesion types [105] with permanent axonal loss [104]. Other non-enhancing lesions can occur, and in contrast to black hole lesions, they appear bright on T2 scans but of a similar intensity to surrounding tissue on T1 scans. They are typically chronic but represent less severe damage [106].

Lesions often develop and are capable of both enlarging and shrinking. This behaviour is associated with the level of cerebral inflammation [13]. As such, lesions can be categorized based on the level of inflammatory activity, which is typically related to the age of the lesion [74]. The categories are as follows: acute or active lesions, focal lesions, inactive lesions, chronic lesions, diffuse lesions and atrophic lesions. Newly formed lesions are, by definition, active and indicate both recent inflammation and demyelination [13]. At the cellular level, they contain macrophages and microglia throughout the lesion. Focal lesions can develop from active lesions and appear sharply demarcated, suggesting localized inflammation [74]. These lesions are indicative of more severe damage, including axonal loss, and can ultimately develop into black hole lesions. Inactive lesions manifest after the demyelination process in an area of lesion tissue has been completed. They have a hypocellular centre, with microglia limited to the periphery of the lesion [74]. While chronic lesions can remain latent, inflammatory activity can resume, potentially leading to further demyelination, lesion enlargement and atrophy. Chronic lesions in MS include several types, each with distinct characteristics. Slowly expanding lesions (or smouldering lesions) gradually enlarge due to low-level inflammation and are typically linked to degenerative processes [107]. These lesions often appear

hyperintense on T2-weighted images. Black hole lesions (previously described) may also become chronic, with intermittent inflammation contributing to their progression [106]. Additionally, diffuse lesions (sometimes referred to as diffuse white matter injury) [72] are chronic but do not directly arise from other chronic lesions. Instead, they are associated with general cerebral inflammation, affecting larger regions of white matter. This diffuse injury tends to be exacerbated by bouts of inflammation (relapse and remission) [108]. Unlike active lesions, diffuse lesions do not have distinct borders [109]. Additionally, diffuse lesions can be linked to atrophic damage and overall brain volume loss [110]. Atrophy is a consequence of disease progression and cumulative damage.

4. Clinical disability metrics

There are a number of addition key terms, criteria and definitions used in MS. The McDonald criteria provide a consensus definition of the disease, with a particular focus on clinical and radiological (MRI) diagnosis, with updates over time, based on evolving research. Following diagnosis, quantification of MS-related disability over times is key, most typically using the expanded disability status scale (EDSS) and, to a lesser extent, the multiple sclerosis functional composite (MSFC).

4.1. McDonald criteria

Currently, the most commonly accepted formal diagnostic criteria for MS are the McDonald criteria, which replaced the Poser criteria in 2001 [111]. The pre-requisites for an MS diagnosis should include two clinical attacks, alongside the presence of at least a single inflammatory (active) lesion, MRI should be the imaging method of choice to demonstrate the latter. A number of revisions have subsequently amended, expanded and calibrated earlier definitions [100,101,112].

4.2. Expanded disability status scale

For practical purposes, the clinical definition of MS is based on the EDSS developed by Kurtzke [113]. A numerical scale of 0–10 is used and increases in increments of 0.5 to indicate the level of disability of a patient. This measure is subject to criticism, with relative lack of sensitivity, particularly in the latter stages of disability, where the EDSS is primarily dependent on mobility [114]. The EDSS has advantages of moderate simplicity, reasonable reproducibility and is well established and therefore remains a key element of MS clinical research, particularly clinical trials.

4.3. Multiple sclerosis functional composite

Due to the limitations of EDSS, an alternative metric known as the MSFC was devised [115,116]. It is arguably a more comprehensive measure of disability and attempts to improve on the EDSS. MSFC comprises three subsidiary metrics: a timed 25-foot walk (T25FT), assessing lower limb mobility over limited distance; a 9-hole peg test (9HPT), assessing upper limb mobility using a breadboard test (performed at least twice, once with each hand); and a paced auditory serial addition test (PASAT), a cognitive test assessing mental processing power by getting a patient to add up a series of 60 single digit numbers read out at 2 or 3 s intervals. The results of the T25FT and 9HPT tests are given in units of time, whereas the PASAT score is the number of correct answers given. Results are subsequently converted into generic metrics known as z scores [117]. These z scores are based on the test results of a patient and the average results of a healthy control population. The lower the MSFC score, the greater the disability, negatively correlating with EDSS [118], the scores of which increase with disability.

5. Computational methods applied to multiple sclerosis using magnetic resonance imaging

Fundamental questions surrounding the nature of MS remain unresolved, notably its precipitating cause, and optimal choice of treatment and management following the inflammatory phase. Moreover, other areas such as recognition of early disease and manual lesion identification can often be difficult

and sometimes time-consuming for health professionals. Computational techniques may assist in overcoming these problems, unpicking the ambiguities of MS and improving diagnostic processes. At present, however, the application of computational methods applied to medical imaging within the areas of consideration remains relatively narrow. The general areas in which machine learning (ML), MS and MRI coalesce are in (i) diagnosis of MS and segmentation of lesions—a core aim of research upon which other areas are built, (ii) differentiating MS subtypes, (iii) prediction of disease progression, and (iv) distinguishing between MS and similarly presenting diseases.

5.1. Review methodology

The search strategy used for paper inclusion was fairly simple. The MEDLINE database was accessed through the PubMed search engine, and no date restrictions were implemented in the search process. Keywords used in the search were 'Multiple Sclerosis', 'Machine Learning' and 'MRI', revealing 171 technical studies published both in journals and at conferences and authored in English. Of these, the ones that did not include all aspects of the search were removed from the collection of studies considered. Furthermore, those studies comparing MS to its mimics were discarded from this review as it would require the inclusion of substantially more information unrelated to MS from a biological perspective, rendering this work unjustifiably lengthy. As a consequence, 25 studies were discarded. Currently, no reviews exist on the comparison of MS patient MRI to that of individuals with similarly presenting diseases using ML. There are studies that discuss MS and its mimics in general [119].

Recent systematic reviews around MS have focused on the diagnostic performance of AI in MS [120], ML algorithms used in MS analysis [121] and general AI application to MS [122,123]. There has also been work published on how AI can be applied to MS, but it is not always specific to MRI and lacks important contextual information and comparative details regarding the efficacy of segmentation techniques in published studies [124,125]. Additionally, a recent study paper has evaluated how ML can be applied to evaluate clinical measures in the context of MS progression [126]. However, there is a need to evaluate the application of ML more clearly as regards MS subtypes, which might provide valuable information on the fundamental biological causes of MS, the pathophysiological events occurring during the course of the disease (for example, mix between relapse activity and progression), with implications for clinical outcomes and treatment pathways, potentially highlighting gaps in knowledge. The future development of imaging-based technologies that exploit ML will depend on whether these clinical parameters are specifically addressed in ML datasets. Furthermore, this is the first study to compile all results from publicly available datasets (see appendix B), which enables the objective evaluation of the current state-of-the-art in ML techniques as applied to MRI datasets. This offers a benchmark for those wishing to directly compare the efficacy of any novel model using publicly accessible data.

5.2. Automated diagnosis and segmentation of multiple sclerosis lesions

The most rudimentary problems in the study of MS using computational techniques and MRI are its general diagnosis and the segmentation and quantification of brain tissue of MS patients. In general, the former involves distinguishing healthy brains from those of MS patients, while the latter focuses on the identification and segmentation of lesions or specific brain areas. In a clinical setting, lesion segmentation tends to be performed manually by medical professionals, which is time-consuming, often inexact and inconsistent. Automatic diagnosis and segmentation techniques, if perfected and adopted by hospitals, would free up the time for clinicians, allow for consistency of diagnosis to be applied to all patients (removing the necessity for manual delineation) and enable patients with the most aggressive and malignant disease courses to be rapidly prioritized by clinics. A variety of characteristics have been extracted in investigations assessing MS. These include lesion pattern, volume, shape and quantity.

5.2.1. Objectives

From the inception of this field, pattern recognition has been pivotal to lesion assessment and behind most AI-based techniques that utilize MRI as an input [127], providing a basis from which lesions can be quantified [128]. Manual selection of likely lesion points was sometimes incorporated [129,130],

acting as a basis from which analysis could be performed. The necessity for human input, however, was quickly rendered anachronistic, and end-to-end automatic segmentation techniques became the expected norm.

Automatic segmentation of lesions has been a hugely popular topic and is a wide-ranging field, with a plethora of studies conducted over the last 25 years [131–190]. These studies utilize a myriad of computational techniques and a variety of approaches, offering information about different aspects of MS. Since most readily available MRI modalities display white matter with the greatest clarity, segmentation of white matter lesions proved to be the easiest and thus most popular to investigate initially.

Segmentation of specific areas of the brain has also been a popular area for study in MS. Brain atrophy is a significant consequence of MS, with a reported rate of volume loss of between 0.5 and 1.35% per year [191–193]. Some studies aimed simply to distinguish between white matter, grey matter and cerebrospinal fluid (CSF) [164,182,184]. Other studies concentrated on the quantification of specific brain tissues like grey matter [194], white matter [181] and cortical thickness [195] for the purpose of monitoring cerebral atrophy or used cerebral spatial features [196] to gauge brain tissue degradation and deformation. Assessments of atrophy of specific areas of the brain, notably the thalamus [197] and the cerebellum [147], have been undertaken. General diagnosis of MS has also been explored [198–200]. These investigations are sometimes challenging because study participants tend to be comprising healthy controls and CIS sufferers or patients recently diagnosed with MS. The benefit of such studies is that they assess the possibility of a diagnosis of MS in individuals in whom clinicians have been unable to arrive at a satisfactory conclusion.

5.2.2. Magnetic resonance imaging in lesion segmentation

T2 MRI was initially the standard sequence for identifying MS lesions [127,128,130,143,166]. The notable limitation of using T2 scans exclusively is that lesions appear indistinguishable from CSF [104], which is especially problematic when attempting to identify lesions around gyri and sulci, the latter of which is filled with CSF. Remedial measures implemented have been the inclusion of PD scans as a separate sequence [152,163,187,201] or combining the two sequences [151,202], to enhance lesion visibility. As the field has progressed, it has become standard practice to include multiple MRI sequences for each patient, maximizing the results produced by the methods implemented, with separate T1, T2 and FLAIR images [132,153,168,170,174,176]. A number of other imaging types have been introduced, including functional MRI [203] and diffusion tensor imaging (DTI) [204,205], but their utilization in studies of this nature has thus far been limited.

5.2.3. Algorithms and methods

A wide range of algorithms have been used in diagnosis and segmentation. Variations on fuzzy connectivity—a seed-growing method in which neighbouring pixels of a similar intensity are considered to be connected—were initially a popular method for segmentation in MS studies [130,162,165,166,169,173,177]. Supervised learning techniques, however, like support vector machines (SVM) [145,168,170,181,186,195,203,205,206] and K-nearest neighbours [144,164,178] soon became more popular owing to their lesser computational expense. In 2015, the field of biomedical segmentation underwent a revolution with the publication of the seminal paper ‘U-Net: convolutional networks for biomedical image segmentation’, the architecture of which can be seen in figure 8 [207].

Arguably, its main benefit has been its ability to produce reliable results from relatively little data input. In the International Symposium on Biomedical Imaging (ISBI) cell tracking challenge, it beat the next best method by over 30%. Convolution neural networks (CNNs) became widespread in medical segmentation, including automatic MS lesion delineation. While a number of unrelated CNN architectures have been used, V-net [139], GoogLeNet [133] and other 3D-CNN architectures [208], the vast majority credit the architecture proposed by [207] as their inspiration [132,134,135,146,148,150,152,180,182,185,188,209–212].

5.2.4. Metrics

There are four general feature categories that can be segmented in brain MRI: grey matter, white matter, CSF and lesions. Lesions are the most difficult to segment as they are often small and have a

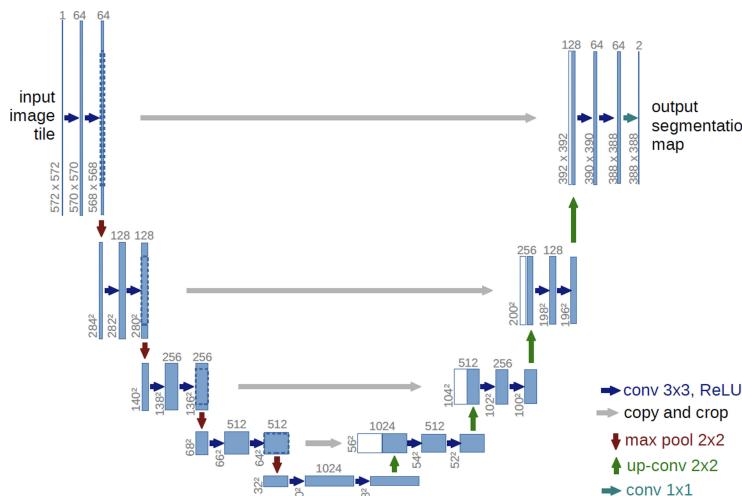


Figure 8. U-Net architecture: blue boxes represent multi-channel feature maps. The amount of channels is indicated by the number above the box. White boxes represent replicated feature maps. Arrows represent specific operations utilized [207].

heterogeneous distribution. Since no single metric encapsulates every pertinent piece of information necessary to fully assess an algorithm, multiple methods for gauging its efficacy for lesion detection and segmentation have been formulated.

The metrics conventionally used are the ones outlined and discussed in detail in appendix A: volume difference, sensitivity, specificity, positive predictive value and Dice score. At least one of these metrics acts as a measure to evaluate nearly all computational techniques involving the use of MRI discussed in subsequent sections. Refer back to this section for further information.

5.2.5. Inferences and conclusions

The foundational work for automatic lesion segmentation was proposed by [127], in which computer-assisted methods were proposed and out-performed manual segmentation by 5%. The initial standard set for methods was to detect 100% of white matter lesions while avoiding false negative detection [128,162]. From early on, a variety of implementations had been reported that improved results, including varying window size in order to alter image scale [162] and the use of multidimensional feature spaces [164]. One of the major hurdles, as previously mentioned, was false lesion detection, not well accounted for by the sensitivity, specificity and accuracy metrics. Methods have been proposed to limit this phenomenon, but this has a tendency to compromise overall sensitivity [168,169]. Smaller lesions tend to be more difficult to detect [152], and it was reported in a study that results for lesions $\leq 20 \mu\text{l}$ in size scored 0.69/0.71/0.67 TPR/FPR/DSC only to improve to a near-perfect 0.99/0.02/0.91 when lesions only $\geq 500 \mu\text{l}$ were studied [182].

Some studies used solely T1 scans expressly for the purpose of segmenting black hole lesions [165,166]. This is a more challenging task due to there being a greater similarity in intensity between these and other surrounding brain tissue. Although some of these lesions are also visible in T2 scans, it was also suggested that there is the greater correlation between clinical disability scores and black holes than those found in T2 images [213], thus making them of greater importance. T2 lesion investigations are not, however, likely to be completely irrelevant as a later study found a linear relationship between grey matter volume and T2 lesion load [194].

In terms of cortical lesion detection, type I and type III lesions were detected with a high degree of accuracy, 83 and 70%, respectively [190]. Type II lesions (with the possible inclusion of type IV lesions), by contrast, proved more elusive (53%). This may be due to their smaller size and that they are predominantly located in the grey matter. Modalities outside of T1, T2 and FLAIR have also been investigated. Diffusion-weighted imaging, in particular DTI, utilizes the anisotropic water diffusion properties of tissue [104] and is known to yield information at an earlier point in time than other more commonly used MRI sequences. As automatic lesion delineation is more challenging when MS is in its prodromal stages [214], DTI could be helpful in earlier diagnosis. Commowick *et al.* [131] were the first to outline a framework by means of which this could be implemented in lesion detection [204]. It was found that DTI was capable of locating points of interest where lesions ultimately go on to manifest

themselves even before their visible formation [215]. A further conclusion reached about imaging types is that with access available to multiple MRI sequences, the inclusion of all possible image types could be beneficial for training in order to maximize the information available to an algorithm. An investigation segmenting enhancing lesions [150] included five multispectral image sets and produced better results than when any single image set was used in isolation.

5.2.6. Limitations

One of the most fundamental limitations in the studies has been data size; some studies are limited to relatively few study participants (five MS [128], two MS and two controls [216], 16 MS [181]). The validity of ML techniques and the success of the outcome are very much dependent on the quantity of available data. With as few as two patients in certain studies, the conclusions that can be drawn are inherently limited. Greater availability of public datasets for all imaging protocols would mitigate against this issue [188]. In the case of MS, this is especially problematic because of the heterogeneity of the condition in terms of the multiple pathological processes and lesion types [163]: a difficulty outlined in previous reviews [217,218] and papers [212]. If only one MS sub-diagnosis is included [196], or if there is a class imbalance [180,199,210], the ability of a method to generalize across all MS types is severely reduced. Furthermore, with a small dataset, it is unlikely that every disease permutation is accounted for which also has the potential to skew analysis in favour of a less commonly occurring subtype (if one is included). Therefore, in order to create a universal software program for clinical use, examples of every lesion variety and disease subtype must be included in the training data. The absence of even a single lesion type is likely to significantly impact the efficacy in terms of detection [175].

Regarding imaging itself, it is best that a variety of imaging protocols are utilized. In an early study, gadolinium-enhancing lesions were hypothesized to appear bright against NAWM in T2 images, but in reality, contrast was sometimes poor [161]. With poor contrast, the training process can be affected, leading to false-negative and false-positive lesion detection [162,167]. In another study, it was reported that chronic lesions were not detectable on FLAIR images owing to low contrast [168], whereas they were perceptible using T2-weighted images in other studies discussed. It has also been reported that smaller lesions [135,170] and limited lesion load [198] make accurate detection and segmentation difficult, especially in situations of low MRI contrast. For some lesion types, such as black holes [165,201], cortical and intercortical lesions [137,190,209], true prevalence is difficult to establish, even for some radiologists, especially when an unsuitable MRI sequence is utilized. All of these limitations illustrate the importance of using multiple imaging protocols to maximize the chance of lesion detection in studies in which a complete assessment of lesion activity is being attempted. It should be noted, however, that the algorithms used in a collection of studies using the publicly available Medical Image Computing and Computer-Assisted Interventions (MICCAI) 2016 dataset were sensitive to scanners and centres in which training data were not collected [131], suggesting that using a variety of scanners might not be of prime importance for engineering a generalizable software.

While datasets have expanded as the field has developed with studies including progressively larger cohorts, validation and ensuring a reliably delineated ground truth arguably remain the greatest limitations in this set of investigations [130,135,143,164,194,219]. Even in published competitions like that of Styner *et al.* [220], the reliability of the ground truth is questionable and has implications for the results of many other studies [141,144,154,157–159]. One publication associated with the most recently available MS MRI lesion dataset (MICCAI 2016 dataset) stated that there was significant discrepancy between the seven experts when defining ground truth for lesion location [131]. It is often difficult to enlist the time of expert radiologists or clinicians to perform ground truth segmentation, and some datasets use only one or two medical experts for this [176]. In an ideal scenario, generating ground truth should be done with the participation of as many experts as possible. Each segmentation performance should then be cross-referenced between themselves and averaged, with any anomalous outlying areas excluded, in order to obtain the most accurate reference image for algorithmic methods to be trained on or tested against.

5.3. Differentiating between multiple sclerosis subtypes and lesion types

Any diagnosis made should be as precise, or accurate, as possible. With MS, however, this is not a straightforward process because of the overlap that exists in the different subtypes. For example, it is

still possible to experience a relapse after a diagnosis of SPMS has been made, and the EDSS scale, upon which conventional diagnosis is dependent, is itself, at least partially, based upon clinical features and self-assessment rather than purely objective measures. Furthermore, the varying presentations of MS lesions can also be categorized, which is useful as they are indicative of different types of damage, for example, enhancing lesions are associated with inflammation, diffusely abnormal white matter is linked to atrophy and black hole lesions are indicative of axonal damage [74,221]. There is always cross-over between MS subtypes and lesion types and lesions can be identified based on visual appearance, inflammatory activity and anatomical location—all of which can also overlap. Computational techniques can make use of objective lesion characteristics, such as volume, shape, visual intensity and location, and have been used to assess how closely lesion presentation reflects subtype diagnosis or lesion type to find a reliable empirical metric.

5.3.1. Objectives

Various permutations of studies have been attempted, using either clinical subtypes or lesion characteristics as reference points. These have included several comparative investigations of MS subtypes. RRMS (chronic MS) comparisons against SPMS (acute MS) were and continue to be a popular area of research as the former tends to evolve into the latter [153,222–226]. Precision in determining the point in time when one subtype transforms into another is critically important for optimizing treatment because of changes in levels of inflammation, damage caused by demyelination and level of patient disability. Inclusion of other subtypes has been used in multi-classification studies, for example, separating the progressive subtypes into SPMS and PPMS [225], and multiple binary studies, comparing controls, CIS, RRMS, SPMS and PPMS patients in a single study [224]. Other standards used for making binary comparisons have included contrasting early-stage and late-stage MS, individuals with high and low lesion load and those with benign and malignant MS [227].

Linking demonstrable neurological deterioration to decline in motor functionality is the basis for the field of neuroscience. Thus, level of cognitive impairment, based on neuropsychological testing, has also been explored; the three classifications used are cognitively preserved, mildly cognitively impaired and cognitively impaired [228]. This study attempted to correlate demyelination and neural atrophy with physical and cognitive disability but excluded conventional MS subtype definitions.

As interest in the application of ML to medicine has grown, clinicians and radiologists have become more active in contributing to the supervised learning studies, leading to their increased recruitment to assist with providing more precise categorization of lesions, for example, in white matter, rim and intracortical lesions [229], and infratentorially versus supratentorially located lesions [230]. Other investigations have been carried out that directly build upon an aforementioned multi-classification study [225], but that use the following feature definitions: persistent black holes, persistent grey holes, acute black holes, acute grey holes, non-black or grey holes and NAWM [231]. Distinguishing between gadolinium-enhancing (paramagnetic) lesions, which become more visible after administering gadolinium contrast, and those that do not (non-enhancing lesions) has also been undertaken using artificial neural networks [222,232]. More recently, distinction between different types of cerebral tissue damage has been examined, diffusely abnormal white matter indicating widespread myelin damage and focal lesions [221], the difference of which may indicate severity of the MS and inform treatment decisions. Arguably, however, the most intriguing approach requires the least clinical contribution. It has become possible to sort MRI datasets based on lesion appearance using unsupervised learning. This was initially carried out using exclusively CIS patients [233], but this was later extended to use all subtypes [234]. It is possible that ultimately avoidance of the use of subjective metrics will prove fruitful and yield greater insight into as-yet unexplored areas of MS.

5.3.2. Magnetic resonance imaging in lesion and subtype identification

The MRI sequences used in these studies can be divided into three types: T1, T2 and FLAIR. T1 was initially the most popular MRI sequence used as data for computational techniques [221,222,227,228], with administration of gadolinium in some instances to enhance image quality and to help distinguish between paramagnetic and inactive lesions [222,232], highlighting active inflammation. Little justification has been given in any of these investigations for the choice of T1 over the other options, but it is known that lesions best demarcated using T1 MRI correlate most strongly with physical disability [235], so this may have been one consideration. Alternative enhancement agents have also

been investigated, notably ultrasmall superparamagnetic iron oxide (UPSIO) [233], with which T1, T2 and FLAIR MRI were explored in some capacity. The UPSIO highlighted a separate collection of lesions to those highlighted with gadolinium. These lesions were indicative of a more aggressive disease course—thereby emphasizing the usefulness of UPSIO. In practice, the choice of MRI is more useful for highlighting inflammation activity and lesion type as opposed to the MS subtype. It was also demonstrated that T2 MRI highlighted a greater number of lesions making it beneficial for finding points from which fresh areas of demyelination could develop, highlighting both diffuse and focal damage [221,233]. Generally, more recent papers tend to use at least two types of MRI: FLAIR and T1 [223,224,230], T1 and T2 [229,231], FLAIR, DTI and T2 [225] as well as all three [153,234]. The greatest benefits of using multiple MRI sequences are that it increases the dataset size as well as also presenting the neural network with essentially new information, thereby making it more effective in improving subtype identification [236].

5.3.3. Metrics

In studies using computational techniques for examining MS subtypes or lesion subcategories, the two simplest metrics for assessment were number of lesions present and the lesion volume. By simply counting pixels or voxels making up an image and then utilizing a method to count objects of interest, both metrics can be calculated. Since lesions can, however, develop in areas of the brain which if damaged do not necessarily lead to disability, they are limited metrics [222]. Consequently, most papers, even from the inception of the field, concentrate on lesion topology [222,223,228,230,233]. Lesion volume [221,223,234] and number of lesions [224], however, have been used in numerous papers in combination with other metrics because it is known that physical damage in specific areas often results in disability, and with this being the case, lesion location has been used as a metric to distinguish one subtype from another [229,231,236]. As technology has improved and MS has been increasingly understood, more focused studies have been undertaken. The markedly more difficult task of investigating MS lesion manifestation in grey matter has also been conducted [227]. One particular study sought to establish which available metrics offered the best results for training an algorithm to distinguish between subtypes. All data were temporally longitudinal and included lesion characteristics and clinical and metabolic features [224]. As the field progresses, more specific data types like nuclear magnetic resonance spectroscopy signals are likely to be increasingly explored [226], the most immediate advantage of which is that they are computationally inexpensive to process relative to those from MRI images.

5.3.4. Algorithms and methods

A number of models have been used to differentiate MS lesion types and subtypes from each other. The first paper published on the topic used a combination of fuzzy logic (fuzzy connectivity) and input from human experts who selected the location of lesions by direct inspection [222]. As the field has developed, human input, post-data pre-processing and model construction have become less of a requirement. A wide variety of supervised learning methods have been utilized to sort data labelled by clinical experts including SVMs [224,226,227], CNNs [153,221,236], MLPs/DNN [231], random forest algorithms [224], gradient boosting [229] and Bayesian neural networks [223]. Unsupervised clustering methods have also proved popular in categorizing lesions into groups with similar characteristics [233,234]. Supervised learning techniques tend to have the capability of being more accurate because they include a ground truth, but there is a danger that in the training process overfitting to a particular dataset can occur, the phenomenon in which an algorithm fits so well to a training set that it is unable to perform well on unseen data. This can be particularly problematic if a single scanner brand is used in generating a dataset or if one subtype is disproportionately represented. The main benefit of unsupervised over supervised learning is that less pre-processing is necessary for the former and so needing less human intervention. In this context, the use of clustering eliminates any conscious or unconscious human-generated bias and established preconceptions by only utilizing the tangible characteristics of the data available. Nevertheless, the importance of diagnoses and patient evaluations should not be minimized as they can assist in improving algorithmic outcomes. In one study, for example, in which a clustering algorithm was applied, it was demonstrated that inclusion of clinical data in addition to the clustering technique improved algorithm accuracy [234].

5.3.5. Inferences and conclusions

Lesion volume was found to be statistically significantly larger in RRMS than in SPMS [222], and an inverse relationship was found to exist between both disease duration and lesion quantity in terms of both number and volume. There was, however, no demonstrable correlation found between lesion volume and EDSS [222,225] due to lesions manifesting themselves in silent regions of the brain, i.e. those not responsible for bodily functions, suggesting a need for a more nuanced approach. Supervised learning methods were able to effectively differentiate between MS data, based on a variety of criteria with considerable success, including early-stage MS and late-stage MS (85%), low and high lesion load (83%) and benign and malignant MS disease courses (77%) [227]. In terms of differentiating between conventionally accepted subtypes, identification rates ranged between 71% and approximately 96% [153,224,226]. It was found that metabolic data were a more reliable means than MRI for discerning the differences between RRMS and PPMS, whereas MRI was better for differentiating between RRMS and SPMS [153,224]. Some of the best results were achieved using a pre-processing technique known as class activation mapping (95.42%), but it should be emphasized that overfitting could have contributed to these results as relatively few patients participated, and only a single scanner was used in the study [153].

Other data types have also proved to be effective in determining MS subtype, including isotropic fraction (reflecting cellularity) and fibre fraction (representing apparent axonal density) [225]. Using a hybrid DNN–DBSI methodology, based on lesion appearance, a prediction accuracy of 93.4% was achieved [231]. Additionally, a link was found between MRI appearance and level of cognitive impairment based on a measure known as centrality [228], illustrating the efficacy of MRI for tracking MS development. Studies on anatomical location of lesions have also been conducted [229,236], and it was found to be easier to identify infratentorial lesions (87%) than the supratentorial ones (62%) [236]. In both papers that explored the use of unsupervised methods, three lesion categories were found that had no correlation to established subtypes [233,234], consistent with a paper previously discussed in the MS progression section [237]. When CIS was explored in isolation, two of the lesion categories found contained a greater proportion of hypointense lesions, indicative of greater disease activity and lesion load. The *p*-values of each group indicated that the reclassification was statistically significant (0.0001, 0.0021 and 0.0192) [233]. In a study in which all MS subtypes were explored, they were characterized in greater detail as cortex-led, NAWM-led and lesion-led [234]. The ‘lesion-led phenotype’ had the highest probability of developing 24 week confirmed disability progression (CDP), and it was also the only one that showed any significant response to treatment. The results showed promise for assisting with the medical stratification of patients for the purposes of optimizing treatment based on predicted disease progression. The concordance metric produced for predicting CDP (after a 24 week period), however, was not entirely favourable even with the inclusion of clinical information (0.63) [234]. Diffuse lesions have been found to be particularly difficult to detect, with a Dice coefficient of only 0.49 achieved compared to 0.81 for focal lesions in the same study [221]. This is likely a consequence of the subtle appearance of this type of damage on MRI scans.

5.3.6. Validation

Validation methods can be divided into two categories: human expert and computational. Some of the earlier studies employed clinicians to validate study results [222,223], but this process is both time-consuming and susceptible to human error. Computational methods have therefore been preferentially utilized to improve the quality of validation and to reduce the need for human input. Leave-one-out (or multiple data points) cross-validation methods became particularly popular for achieving this [224,227,229,234], in which an algorithm is trained using a dataset with one or more random pieces of data left out for testing. The process is then repeated over a defined number of epochs and with a different portion of the dataset being used for testing in each epoch. Other methods simply define the training, testing and/or validation datasets [153,231,236]. As in the case of the leave-one-out method, multiple epochs can be applied to improve the accuracy of an algorithm in the classification process but with the caution that this could lead to overfitting, in which case it may perform poorly on unseen testing data. Furthermore, if too many epochs are defined, the process can become prohibitively computationally expensive.

5.3.7. Limitations

Early papers required the need for a human operator [222]. Most of these studies have their specific limitations, such as the need for a human operator in the earlier ones, but there are also universal problems that span almost the entire range. Of these, one of the most significant has been dataset size with only between 19 and 100 patients in most instances [153,222–227,229,231,233]. Only a few publications exceeded this number, notably 332 [228], 1800 [236] and 9390 [234] MS participants. The more data that are available, the more effectively an algorithm can be trained. The drawback of using fewer patients is that potentially it generates an insufficient quantity of data to create a method that effectively differentiates MS subtypes, especially if the data are demographically homogeneous, such as when they are taken from a similar age group, geographic location or use only one or a limited number of scanners. Quality of scanner used can also be limiting since lower frequency MRI reduces the detectable amount of feature detail [224]. It is also important to stress that although the use of high-quality images improves accuracy, there is little practical sense in mismatching the use of high-resolution equipment for training with images from lesser quality scanners. For this reason, it was necessary to exclude some data, in particularly from more up-to-date and sophisticated scanners, which generated sensitive metrics in order to avoid skewing data and clustering images based on quality rather than on visible features [234]. Discussion has also arisen over whether MS could be in fact a group of diseases with similar characteristics and manifestations rather than a single disease entity with divergent courses [238]. Research into understanding individual subtypes with their respective courses may also help shed light on answering this auxiliary question. A gold standard for identifying certain lesion types does not exist. For example, in the case of diffuse lesions—a newer subject of interest—there is no agreed-upon clinical standard for identifying diffusely abnormal white matter [221].

5.4. Predicting multiple sclerosis disease progression

Monitoring MS and its progress is of fundamental importance to guarantee that patients receive both optimum management and assessment of the efficacy of their treatment. As outlined, regional atrophy evolves over time, and the areas of the brain affected vary between patients. For example, the ventricular system is particularly affected with evolution of RRMS, whereas in progressive forms of disease the cortex is pre-disposed to this [193,239]. These appearances therefore govern the management that would need to be implemented subsequently.

5.4.1. Objectives

Elliot *et al.* were the first to implement a technique using ML methods involving longitudinal data on MS patients to predict the likelihoods of new lesions appearing, enlargement of any existing lesions and lesion resolution [240]. Other studies exploring MS progression, based on lesion development and appearance, have used a variety of criteria including image intensity, deformation fields [241,242], lesion volume [243–245], lesion enlargement [236], lesion age [246] and total lesion load [244]. Exploration of the rate of deterioration of MS symptoms comparing algorithmic performance against that of clinicians has also been a central theme of studies [247,248]. Due to the availability of T2 MRI and because MS is considered to be primarily a demyelinating disease, attempts to correlate the progression of white matter demyelination to physical or cognitive disability have proven to be a common recurrent theme in studies [237,249–252].

Monitoring cerebral atrophy and alteration in brain structure have also been used as objectives in research to monitor disease progression. One of the earliest studies in this area aimed to correlate morphological changes in deformation fields with patterns in the spatial distribution [242], which are a direct consequence of morphological changes caused by atrophy and lesion development. The benefit of this approach lies in the fact that the further an MS brain structure diverges from the norm, the greater the prospect of disability. Understanding how changes in the brain affect disability is vitally important for improving understanding of progression in MS. Models have been developed to attempt to correlate changes in the entire brain with progression of physical disability [247,252–256], but other models created have concentrated on specific areas of the brain, for example, thalamus [257], corpus callosum [258] and cerebellum [259]. MS is a phasic condition in which either CIS develops into RRMS or PPMS and RRMS develop into SPMS. General investigations on MS progression have

been conducted involving all subtypes [253], but RRMS has been the subtype most widely studied in relation to MS progression [232,237,240,246,251,259], most likely because it is the most common form of MS. Progression from CIS to RRMS has also been a particular area of focus [241,249,252,254,260–263], as not all patients diagnosed with CIS ultimately progress to RRMS. Having the ability to gauge the likelihood of this happening may ultimately yield aetiological understanding, and if fully comprehended, lead to clinical protocol that might minimize the chance of progression. These arguments can be made similarly for investigating RRMS to SPMS transition [245]. Research has also been conducted to identifying the most significant features required for diagnosis and, in so doing, reducing the need for data collection in future studies [254]. One technique, recursive feature elimination analysis, was implemented to explore the most indicative characteristics of subtype transition [262]. Experiments have also been undertaken to assess the impact of certain medications on MS progression. Both interferon users, relative to a control group [263], and ocrelizumab users, relative to interferon users [232], were analysed to assess their efficacy in curtailing the advancement of MS.

5.4.2. Metrics

The most common way in which analysis is conducted in predicting MS progress is through use of two or more sets of serial scans. Normally, a group of patients, who fit certain criteria (e.g. time since diagnosis and MS subtype), are inducted into a study for whom various data types including MRI scans and clinical test scores are collected to establish a health baseline for each study participant. After a designated period of time, each patient is re-assessed at least once to gauge disease progression. These re-assessments can be carried out at either regular or irregular intervals and can incorporate further MRI scans or disability evaluations. In most cases, the health of each patient will have deteriorated illustrated by the metrics collected. Lesions may be seen to have visibly grown or diminished in size and changes in brain morphology may have also evolved on serial scanning. Changes in numerical data can be tracked from longitudinal MRI datasets such as lesion load, total number of lesions and lesion pattern. Lesion load has been factored into the analysis of many papers tracking MS progression [236,237,240,241,243,245,248,252,254–257,260–263]. This has enabled straightforward comparisons to be made at different points in time. MRI texture analysis has also been a popular method. This assesses the heterogeneity of an image through the use of statistical models [249,251], ultimately extracting information such as tissue abnormalities that may precede lesion manifestation associated with a specific condition. Brain atrophy, to which disability can be reliably correlated, has commonly been used as a metric [255,256,258,259,263], but this is difficult to assess because of the ‘reserve capacity’ of the human brain [264,265]. Much of the research available utilizes more than one predictor, in order to determine which of them is most prognostic of disability progression [255,262]. Level of disability can be gauged through a variety of clinical assessment tools of which the most commonly used in the investigations outlined is undoubtedly the EDSS [232,237,243,245–249,251–256,258,261–263]. Other metrics do exist and have been employed both instead of and in addition to EDSS: these include the MSFC [242,259], McDonald criteria [237,243,253,254,260] and symbol digit modalities test (SDMT) [247,252,258]. These alternatives allow other aspects of MS disability, especially cognitive impairment and upper body mobility, to be considered, which EDSS does not assess sufficiently [114].

5.4.3. Duration of observation

Timespan of observation is a key consideration of temporal MS studies. MS is a disease that often lasts for decades. With the improvement of medical care, life expectancy of people with the disease has increased, and it is not uncommon to die from unrelated causes. In a 60-year Norwegian study, it was shown that MS patients have only a 7.1 year reduction in life expectancy over the general population [266]. As MS is often a near-enough lifelong condition, it is clearly favourable to monitor a patient as frequently as possible. On the whole, yearly clinical and/or radiological re-evaluation appears to be an optimal interval. This ensures that a study is cost-effective, with a single MRI brain scan costing approximately £200 (NHS) and \$980 in the UK and USA, respectively. Investigation of specific subtypes may require more specific temporal parameters. For instance, CIS needs to be assessed as quickly as possible after diagnosis as a second relapse indicates a diagnosis of RRMS, and this can happen quickly. Several studies have shown that the average time between CIS onset and assessment was under two months [261,262]. Most studies based wholly on MRI include at least a single year

of observation [236,237,240,241,250,251], but usually more [232,243,246–248,253,255,256,258,260,263], in which an initial scan and at least one other subsequent scan are collected per patient. Some studies, like the aforementioned ones focusing on CIS, used solely clinical assessment in follow-up appointments as opposed to MRI. A number of more general studies refrain from performing any subsequent MRI scans and base outcomes solely on clinical diagnoses made by medical practitioners. These include the MSFC [242], McDonald criteria [254], EDSS [245,257], fall status and atrophy damage [259] and SDMT [252]. Although this has practical benefits, the absence of follow-up scans limits the scope of the study by omitting important information that allows for the visualization of progression in terms of demyelination and atrophy.

5.4.4. Algorithms and methods

A variety of algorithms have been used to study MS progression. SVMs have been particularly popular, especially when using data in addition to baseline MRI scans in supervised classification problems [237,245,248,249,251,261–263]. Similarly, random forest algorithms [245,246,252,260] and multi-layer perceptrons have been used for supervised classification problems [256]. Regression models have also been incorporated into studies [241,245,250], two of which incorporate SuBLIME [267], which involves probability assessment of voxels making up a lesion after implementation of pre-processing. CNNs have proved to be highly useful in visualizing the progression of MS on MRI scans, especially in terms of demyelination and/or brain atrophy [242,247,253–255], generating images as outputs. Again, U-Net is often cited as the main influence in architectural design [232,236,242,243,257,258].

5.4.5. Inferences and conclusions

Initial studies demonstrated that it was possible to make predictions using MRI scans. With the benefit of amplitude modulation–frequency modulation to extract features, it was possible to determine the likelihood of an MS patient having an EDSS of greater than two with 86% precision [249]. Other early studies focused primarily on documenting methodology rather than on generating significant results and illustrated either single or multiple methods for making predictions using computational techniques based on MRI scanning [240,242,250]. Some of these were conducted around binary classification, comparing tissue types at different points in time. Classification results in this study exceeded 85% for all but one of the seven comparisons [251]. As regards predicting conversion from CIS to clinically definite MS, results varied between 70.4 and 92.9% [254,260–263]. Accuracy in results, based on an average taken between sensitivity and specificity, declined over time with results dropping in one study from 71.4% at the 1-year mark to 68% at the 3-year mark. Sensitivity decreased while specificity increased over this period [261]. Addition of combinations of various features including MRI lesion geometry, EDSS, SDMT, sex, demographic and clinical features improved prediction accuracy [247,261,263]. SVMs have been one of the most effective ML techniques for assessing MS disease progression. An *R*-value of 0.8 was generated predicting EDSS using multiple features [245]. Under certain circumstances, matching EDSS to disease subtype was inaccurate due to RRMS and SPMS patients, and SPMS and PPMS patients both having the potential to produce similar EDSS [253].

In another study, which made use of multiple variables as well as implementing class imbalance (weighting towards sensitivity), an accuracy of 86% was achieved, with T2 lesion volume ranked as one of the best indicators of MS progression alongside age [248]. The significance of T2 lesion load was consistent with findings from other studies [256]. Other informative metrics included lesion load in the main white matter tracts in the cerebellum [237], cerebellar volume [259], lesion patterns [252], cerebellar lesion patterns (88% chance of worsening) [255] and T1-unenhancing baseline lesions [232]. Area under the curve, based upon true positive rate (TPR) and true negative rate, scored between 0 (no measure of separability) and 1 (excellent measure of separability), is another metric that has been used to predict disease progression. It was found that a combination of features, including lesion patterns, proved to be effective in predicting worsening cognitive and functional status in MS patients with scores between 0.79 and 0.90 [252,256,268]. A key conclusion was that radiological features were more accurate classifiers in predicting disease progression than clinical ones [245,246,256]. Predictions were also made using Dice, for segmentation for which results ranged from 0.56 to 0.908 [232,241,243], with contrast-enhancing lesions achieving the best results [243]. Demonstration and quantification of cerebral atrophy, using MRI, has also been used as tool for exploring disease progression in MS. The

feature can be used to calculate disability levels and to predict future disability and cognitive decline ($p = 0.025$) [257]. Corpus callosal atrophy was similarly shown to be able to predict disability based on EDSS up to 10 years into the future, with $p < 0.001$ produced [258].

5.4.6. Limitations

The limitation of insufficient dataset size also applies to researching MS progression. Some studies have included as few as 20 patients [250]. The issue is compounded by MS having the potential to evolve in different ways. This unpredictability means that within a small sample, a certain subtype(s) may become either under- or over-represented skewing the results. This could especially be a problem in any generic study about MS. For example, a study conducted on transition of CIS to RRMS contained markedly few patients who did not progress to having a second relapse [263]. Furthermore, some papers treat MS as a homogeneous condition [253,254], despite there being numerous universally accepted pathogeneses, and that a number of recent computational studies have divided MS into three pathways based upon lesion pattern [233,234,237]. In studies concentrating on MS progression, study participants sometimes miss examinations, leaving gaps in datasets. For example, one study of 83 patients that used 455 scans was illustrative of this (as 455 is not equally divisible by 83) [253] and was common across relevant studies [240,261]. Data collection is often also taken at irregular intervals or not immediately after an incident of interest takes place. An issue unique to MS progression and deterioration is that some papers only include one MRI volume per patient, normally performed at the initial assessment, subsequently preferring to use clinical assessment metrics in isolation as a measure of disability [242,245,252,254,257–259], meaning that there is no way in which physical tissue damage can be correlated with disability. Additionally, the duration of monitoring is crucial. All papers cited described carrying out at least one follow-up assessment within a year of initial contact. Only five of over 30 of the papers discussed in this section involved patient data collection of at least 5 years [243,246,248,253,256]. MS is a disease that can debilitate an individual for decades, so it could be argued that tracking progression for only a single year is insufficient for gleaning the maximum amount of useful information.

Many studies do not include healthy control groups, which would be helpful to ensure that results are unique to MS [259]. This would help with overcoming some problems as one algorithm designed to estimate EDSS had trouble with predicting a score of zero [255]. EDSS is often used in studies either as a metric for disease progression or as an additional piece of data alongside MRI scanning to establish rate of deterioration. Nevertheless, EDSS does not take account of cognitive and visual decline, or upper limb disability [245,248,255,256], which also correlate to lesion development and brain atrophy. The MSFC, comprising mobility, cognitive and auditory tests, is arguably a more complete metric for gauging different aspects of disability [242,259].

Larger studies have tended to utilize data from multiple centres, either nationally or internationally. This has resulted in the use of a variety of scanners providing images of non-uniform resolution and size. In order to ensure uniform image characteristics when using different scanners, images produced by higher resolution scanners need to be reduced in quality, erasing useful information for neural networks to learn from in their training process [258]. The reduction in quality is especially important if an unsupervised ML technique is used to categorize scans, as it is feasible that it would use resolution as a criterion for sorting. Alternatively, it is possible to use images taken by the same scanner, a strategy which has been implemented in some investigations [232,241,243]. This does, however, limit its applicability to clinical and industrial work if access to the scanner that has been used to train the software is not available. Establishing ground truth for lesion location can also be problematic. For example, manual T2 lesion segmentation was reported as being only 62% [219], demonstrating the ongoing challenge of automatic lesion delineation. Moreover, the number of clinicians involved in manual segmentation upon which the ground truth is based was as few as two in some instances [232,247]. A further limitation in these studies is that there is not always record of whether a study participant has been medicated, making it difficult to establish the impact of treatment on data used in an experiment [195]. Treatment itself may influence the way in which MS manifests itself in individuals, as in lesion pattern, shape and development. Only two studies were found that specifically investigated the effect of treatment [232,263]. Without taking this variable into account, it is difficult to evaluate to what extent MS progression is governed by treatment and how much is part of the normal disease process. This reservation is further illustrated by the fact that the use of gadolinium as a contrast agent is thought to possibly underestimate lesion age [246].

6. Future developments

Larger datasets are essential to improve the predictive capacity of any algorithm [241,249,261]. Datasets have gradually increased in size over time, with pooling of national and international data and greater public availability. Some of the more recent studies utilize data from over 1000 patients [234], but many tend to not involve more than a single MRI per patient. Data augmentation methods and production of synthetic data both have the potential to improve dataset size [254]. It would be a great technological advance to create a commercially viable product that incorporates data from multiple scanners to make a software package as universally applicable as possible [241,246,257]. Ultimately, one of the primary aims of research in this area is to produce functional software to aid clinicians in assessing and diagnosing patients by speeding up these processes and increasing diagnostic precision.

The inclusion of further features in analysis is likely to be beneficial, especially to determine which features are most influential in predicting disease progression [256,260,261]. Some of the features proposed include grey matter atrophy, intrathecal synthesis of oligoclonal bands and genomic data [261,262], as well as consideration of smaller subregions of tissue within the brain in isolation, such as dividing the cerebellum into subregions to better assess the tissue damage–disability relationship [259]. In relation to predicting progression, extending the timeframe for monitoring patients [237] and increasing the number of data points within the assessment period [253] have been proposed to obtain a fuller picture of MS pathogenesis. There is a strong argument for using MSFC as a disability metric as opposed to EDSS, as it covers different aspects of disability more comprehensively [237,242]. Encompassing multiple MRI protocols has been recommended because, as previously mentioned, certain lesions are more easily discernible, depending on the relaxation time used to generate the MRI image [232,252,256].

7. Conclusions

This review gives a comprehensive assessment and overview of the application of ML techniques to MRI scans performed on MS patients. Various aspects of MS have been described and discussed, including pathology, history, subtypes, white and grey matter damage, lesion development, clinical disability metrics and publicly available MRI datasets. The principal sections of the review explore three ways in which ML techniques have been applied to MS: automated diagnosis and segmentation, differentiating between subtypes and predicting disease progression. It is apparent that the vast majority of techniques used at the time of publication utilized some variation on U-Net, a type of CNN, as a basis for initial lesion segmentation. It is also clear that unsupervised learning techniques are being explored to limit human bias, a problem that has hindered studies in establishing a ground truth for data. Greater data accessibility would assist in augmenting both the volume and rate of quality research and would allow for increased heterogeneity in datasets for both MS subtypes and imaging quality. The likelihood is that a reliable and generalizable computational framework for MS assessment will increase with greater public access to larger quantities of medical data. Achieving this objective would allow for MS subtypes to be differentiated more easily from one another and for the prediction of disease progression to be estimated with a higher degree of confidence. Such a breakthrough would mean that individuals with the most aggressive forms of MS could be prioritized for earlier and more intensive and focused treatment.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. This article has no additional data.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. A.C.S.-K.: conceptualization, data curation, investigation, methodology, resources, software, visualization, writing—original draft, writing—review and editing; M.C.: funding acquisition, project administration, supervision, writing—review and editing; D.M.E.: conceptualization, funding acquisition, project administration, supervision, writing—review and editing; L.B.: supervision, writing—review and editing; Z.A.: methodology, project administration, supervision, writing—review and editing; W.G.K.M.: writing—review and editing; M.D.: supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interests. We declare we have no competing interests.

Funding. A.C.S.-K. was funded by the Engineering and Physical Sciences Research Council (grant number: EP/T517926/1).

Acknowledgements. The authors would like to thank Professor Thomas Montenegro-Johnson and Dr Gabor Szekely (MD) for their insightful discussion and advice on structure and MS, respectively, and for their comments on the manuscripts.

Appendices

The appendices outline the few publicly available datasets of MS brain MRI scans paired with ground truth segmentation, as well as published segmentation results for various studies that have participated in the challenge. Appendix A provides context to how lesion segmentation is assessed, both generally and in each of the datasets published. A variety of metrics used to describe and explain lesion segmentation results are provided to assist the reader in understanding the tabulated data in appendix B.

Appendix A. Evaluation of lesion segmentation

Predictably, the most obvious validation method in determining whether an algorithm accurately counts the number of lesions present on an MRI image is by actually counting. This is known as total lesion load (TLL). This is complemented by the volume difference (VD) metric, which yields the fractional difference between ground truth lesion volume (LV_{gt}) and segmentation lesion volume (LV_{seg}):

$$VD = \frac{|LV_{gt} - LV_{seg}|}{LV_{gt}}. \quad (A1)$$

TLL and VD are both limited performance measures because they do not indicate whether the location of lesions is accurate, nor do they yield any information about lesion size or shape. The most frequently occurring set of metrics used in the literature to address these limiting factors are those relating to the overlap between ground truth data (the point of reference data considered to be real) and the output of the method. They are based on four variables, true positive (TP)—volume of correctly predicted lesion presence; true negative (TN)—volume of correctly predicted lesion absence; false positive (FP)—volume of incorrectly predicted lesion presence; and false negative (FN)—volume of incorrectly predicted lesion absence. The metrics, all of which yield values between zero and one, include *Dice similarity* or *similarity coefficient* (DSC) [269]:

$$DSC = \frac{2 \times TP}{FP + FN + 2 \times TP}, \quad (A2)$$

which is perhaps the standard for assessing segmentation performance and is essentially a measure of similarity between ground truth image and the output produced by a method across all images. For context, a score of 0.7 is considered to be an excellent result in automatic lesion segmentation [270,271] and 0.8 is considered to be near perfect [219]. *Sensitivity*, also known as TPR, is given by

$$TPR = \frac{TP}{TP + FN} \quad (A3)$$

and describes the fraction of correctly detected lesions. *Specificity* is denoted by

$$TNR = \frac{TN}{FP + TN} \quad (A4)$$

and indicates the fraction of healthy tissue correctly detected. The positive predictive value (PPV) is described below by

$$PPV = \frac{TP}{TP + FP} \quad (A5)$$

and reports the percentage of correctly identified lesions over all lesions detected. *Accuracy* (*Acc*) is given by

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (A6)$$

and denotes the fraction of correctly identified tissue. False-positive rate can be calculated through using

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (\text{A7})$$

and informs on the percentage of incorrectly detected lesion volume present in all healthy tissue. In addition, in published MS lesion segmentation challenges [131,220,272,273], overall scores achieved by participating algorithms are often calculated based on a weighted amalgamation of all the metrics taken, yielding a comprehensive assessment of a method. Overall scores are the basis for the rankings seen in the tables in subsequent sections, score formulae vary from challenge to challenge. However, metrics can be misleading in cases in which many more lesions are detected than what are present in the ground truth due to small lesion size relative to the quantity of surrounding tissue. This can result in high scores in situations in which algorithms have performed badly [217]. Even when the same metric is employed to assess techniques, it is fundamentally necessary to keep other variables, including image properties and the corresponding masks/labels, the same; otherwise, it can become problematic when comparing the results recorded across papers if the data used are not identical [219,274]. Dice scores for lesion detection have varied between 0.53 and 0.86 [132–151,201,208–211,219,275,276], but different datasets are used within these papers, some of which are publicly available, while others are not. Certain studies focus on detection of all lesions, others solely on white matter lesions and others on cortical lesions. It should be further stressed that the quality of images and masks is crucial, which may explain the large variation in results across segmentation studies using different data repositories and, therefore, the limitations when applied to certain datasets.

In the 2015 ISBI competition [272], DSCs exceeding 0.7 even 7 years after publication have not been achieved across all published entrants [137,139,140,152,153,260,277–280]. Nevertheless, other works have been published, which have yielded significantly better results: 0.86 [211] and 0.78 [150] in which private datasets with restricted access were used. It should be noted that in the best performance, lesions ≤ 0.2 ml in volume were excluded from the study and that an automatic segmentation prior (MRIAP) was utilized in generating the ground truth before manual experts adjusted them [211]. Ideally, no lesions should be ignored, and all segmentation for the ground truth should be manually performed for validation purposes. It is possible that methods used have improved; it is also, however, plausible that the improvement in scores are a consequence of either larger datasets, better image quality, more homogeneous datasets (i.e. all data produced from the same scanner type) or more precise image segmentation for image labels/masks by radiologists. This is supported by information documented in [153], in which improved results using the same methods were demonstrated with a second dataset for the method described along with a number of others, 0.697 [280], 0.723 [260] and 0.755 [153]. It is therefore imperative that when assessing methods to compare results produced on identical data to ensure results are as a consequence of the properties of the method proposed as opposed to those of the dataset (such as manual segmentation of ground truth or image resolution). This point is most starkly illustrated when examining the works found that ostensibly perform well using the 2016 dataset compared to previous studies published using the same data, achieving Dice scores of 0.824 [132], 0.866 [281] and 0.76 [282]. In all three cases, only the 15 patients from the training set are used in the study for both training and testing. The two-dimensional images comprising the 15 three-dimensional scans are redistributed so that 80% of them are used for training and 20% are used for testing. This breaks the rubric of the challenge, where the challenge states that 15 patients should be used for training and a separate dataset provided comprising 38 patients should be used for testing. It should also be noted that the exclusion of the testing set provided means that data from one of the four centres collected in this study are entirely excluded. As a consequence, this skews the results to make the method used appear better than it actually is and means that the methods are not tested as extensively as they might have been had data from all the scanners collected been used.

Appendix B. Publicly available magnetic resonance imaging datasets

A variety of MRI sequences exist, each with its own benefits in visualizing tissue types. The most commonly used MRI sequences are T1-weighted, T2-weighted and FLAIR. The two variables that distinguish these from each other are repetition time (TR) and time echo (TE). The former describes the time between externally emitted radio frequency (RF) pulses and the latter is the time between the emission of an RF signal and the reception of an echo. Signal times are shortest in T1 MRI and

Table 1. Results using 2008 MICCAI MS lesion segmentation dataset. VD = volume difference; TPR = true positive rate; FPR = false positive rate.

method	UNC rater			CHB rater			overall	
	VD	TPR	FPR	VD	TPR	FPR	score	rank
[140]	62.5	55.5	46.8	40.8	68.7	46.0	87.12	1
[287]	46.9	43.9	32.3	113.4	53.5	24.2	86.93	2
[141]	46.3	47.0	43.5	51.3	52.7	42.0	86.11	3
[154]	37.8	42.0	44.1	53.4	51.8	45.1	84.46	4
[147]	64.7	71.8	69.1	118.6	84.0	69.7	84.34	5
[288]	58.1	59.0	64.7	96.8	71.3	62.8	84.16	6
[289]	33.1	63.8	69.7	59.3	68.0	68.6	84.25	7
[152]	63.5	47.1	52.7	52.0	56.0	49.8	84.07	8
[290]	56.9	37.7	34.6	113.7	42.9	30.5	83.92	9
[155]	88.6	55.5	62.6	45.1	68.1	63.2	83.25	12
[156]	65.2	44.9	43.2	158.9	55.4	40.5	82.34	14
[157]	49.0	39.7	60.8	73.0	47.9	52.7	82.16	15
[158]	45.4	51.2	76.7	52.4	59.0	71.5	82.07	16
[159]	57.9	49.1	76.3	86.4	58.2	70.6	80.00	22
[201]	63.7	49.0	74.9	85.2	55.8	70.5	79.90	24
[144]	96.0	48.4	83.6	46.4	60.1	81.2	78.19	30
[291]	183.7	20.0	18.0	387.5	24.6	18.2	77.96	31

longest in FLAIR. Differences in appearance between tissues are entirely dependent on water content that dictates the resonant frequency of the tissue. The closer to its resonant frequency a tissue is, the brighter it appears. Based on the research, five publicly available datasets associated with real patients were found. While some early published synthetic data do exist, e.g. BrainWeb [283–286], they will only be briefly mentioned here since their usefulness has been arguably superseded by the resources detailed below.

B.1. 2008 MICCAI multiple sclerosis lesion segmentation challenge

The MICCAI anonymized dataset comprises 45 3D MRI scans collected by both the University of North Carolina (UNC), Chapel Hill and Boston Children's Hospital (BCH). 3T Siemens scanners were used in both instances [220]. Only two experts manually segmented the images from UNC, and just one expert segmented those provided by BCH. Multiple MRI modalities were used for each patient: T1, T2, FLAIR, DTI-derived fractional anisotropy and mean diffusivity images. Resolution was $0.5 \times 0.5 \times 0.5 \text{ mm}^3$ [220]. The training portion of the dataset was composed of 20 patients, all of which underwent a 3D scan, 10 from each respective scanning centre, whereas 25 cases make up the testing set, 15 from CHB and 10 from UNC. The dataset can be accessed through the *NeuroImaging Tools and Resources Collaboratory* website: <https://www.nitrc.org/projects/msseg> (last accessed: 19 February 2023). An up-to-date leaderboard can be found at the following address: http://www.ia.unc.edu/MSseg/results_table.php (last accessed: 19 February 2023). Only results associated with a publication were included in table 1. Best results of each participant were recorded (two of the three best results were produced by [140]).

B.2. 2012 Institute of Neurology and Genetics Nicosia, Cyprus, multiple sclerosis dataset

The dataset is longitudinal with each of 38 patients being scanned twice with an interval of between 6 and 12 months. At the time of the initial scan, each of the patients had to be diagnosed with CIS. The gender distribution was approximately even, with 17 men and 21 women. A T2 scanning

Table 2. Results using 2015 ISBI longitudinal MS lesion segmentation dataset. DSC = Dice score; PPV = positive predictive value; TPR = true positive rate; LFPR = lesion false positive rate; LTPR = lesion true positive rate; VD = volume difference.

method	DSC	PPV	TPR	LFPR	LTPR	VD	score	rank
[297] ^a	0.669	0.886	0.575	0.125	0.538	0.396	93.358	1
[153]	0.639	0.914	0.530	0.122	0.533	0.436	93.317	2
[297]	0.643	0.908	0.533	0.124	0.520	0.428	93.210	3
[151]	0.655	0.886	0.555	0.140	0.544	0.399	93.238	4
[243]	0.681	0.855	0.603	0.157	0.538	0.366	93.030	5
[280]	0.687	0.849	0.615	0.168	0.554	0.359	93.087	6
[298]	0.643	0.887	0.546	0.132	0.480	0.421	92.730	7
[299]	0.643	0.879	0.534	0.105	0.441	0.399	92.637	8
[300]	0.662	0.838	0.590	0.151	0.491	0.372	92.554	9
[277]	0.584	0.921	0.456	0.414	0.087	0.497	92.486	10
[137]	0.611	0.899	0.490	0.410	0.139	0.454	92.118	11
[160]	0.630	0.845	0.537	0.487	0.201	0.405	92.076	12
[139] ^a	0.639	0.813	0.567	0.120	0.366	0.356	91.633	13
[138]	0.631	0.787	0.555	0.367	0.153	0.338	91.331	14
[301]	0.633	0.830	0.548	0.242	0.347	0.412	90.591	15
[302]	0.605	0.775	0.535	0.266	0.367	0.365	90.283	16
[139]	0.627	0.789	0.555	0.498	0.568	0.352	90.070	17
[303]	0.568	0.611	0.570	0.474	0.353	0.343	87.72	18
[304]	0.501	0.549	0.523	0.577	0.429	0.571	86.92	19

^aSubmitted results that improved on those associated with publications.

sequence was used with the repetition time (TR) = 4408 ms and echo time (TE) = 100 ms, a slice thickness of 5 mm, a field of view of 230 mm and pixel resolution of 2.226 mm⁻¹. In addition, 10 healthy candidates were included as controls. All lesions thought to be clinically significant were delineated by a neurologist and validated by a radiologist [292–295]. The dataset can be located and freely downloaded from the *Department of Computer Science, University of Cyprus* (Nicosia) website: <http://ehealthlab.cs.ucy.ac.cy/index.php/32-software/218-datasets> (last accessed: 19 February 2023).

B.3. 2015 Longitudinal multiple sclerosis lesion segmentation challenge

The dataset was published in association with the ISBI [271,296]. The data were collected from University College Hospital, London. The data repository was separated into training and testing data of 5 and 14 subjects, respectively. The dataset was longitudinal. Four of the five participants in the training set were scanned at four points in time, and the remaining patient was scanned at five [272]. The test data were subdivided into set A, comprising ten subjects, and set B, comprising four subjects. In set A, eight of the subjects were scanned four times and two were scanned five times. In set B, three patients were scanned four times, and one patient was scanned five times. Four different scanning modalities were used for each patient: T1-weighted (TR = 10.3 ms, TE = 6 ms, res = 0.82 × 0.82 × 1.17 mm³), double spin-echo PD-weighted (TR = 4177 ms, TE = 12.31 ms, res = 0.82 × 0.82 × 2.2 mm³), T2-weighted (TR = 4177 ms, TE = 80 ms, res = 0.82 × 0.82 × 2.2 mm³) and T2-weighted FLAIR (inversion time (TI) = 835 ms, TE = 68 ms, res = 0.82 × 0.82 × 2.2 mm³) all using a 3T Philips system [272,296]. Two raters were used in the manual delineation of all images, one of whom had 4 years of experience and the other had 10 [272]. The ratio of subjects is skewed both in favour of women (15:4) and relapsing–remitting patients (16:3); notably, no male primary progressive patients are present in the dataset. The dataset can be downloaded from the *smart-stats tools website on the longitudinal MS lesion segmentation challenge* webpage: <https://smart-stats-tools.org/lesion-challenge-2015>.

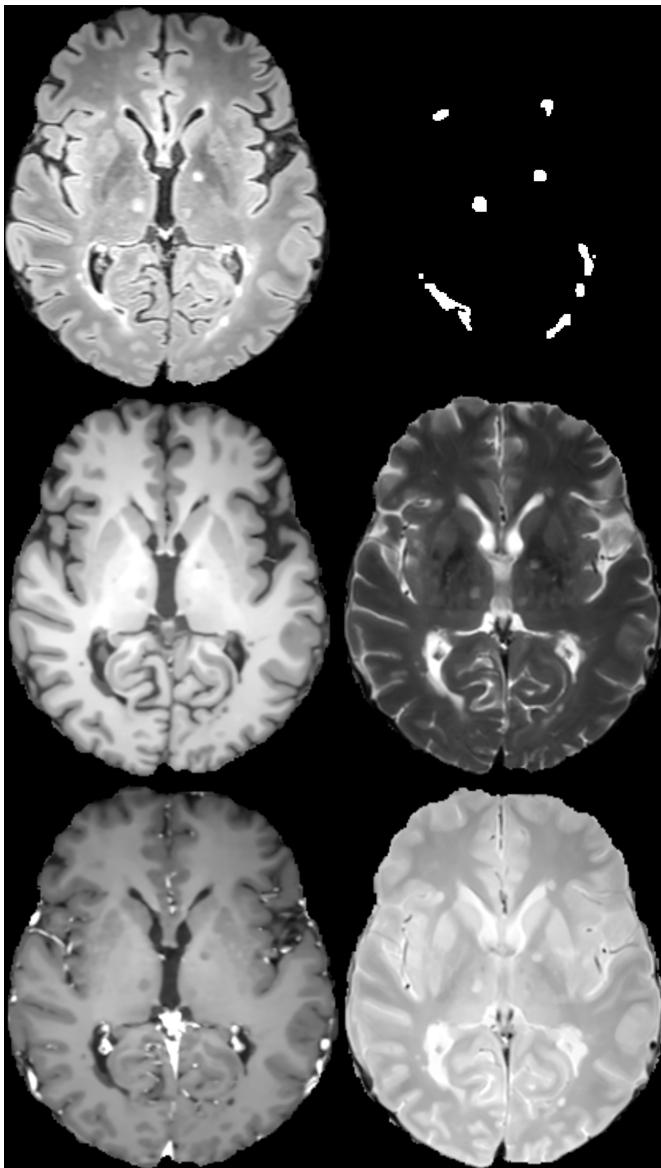


Figure 9. Example MRI scans in the horizontal plane taken from the MICCAI 2016 MS dataset of the same patient using different imaging protocols [204]. From left to right: top row, 1. FLAIR and 2. Consensus position of lesions based upon the manual delineation of seven radiologists; middle row, 3. T1 and 4. T2; bottom row, 5. Gadolinium enhanced and 6. Proton density.

Table 2 illustrates a selection of published results found using the dataset. The leaderboard can be found at the following address: <https://smart-stats-tools.org/lesion-challenge>. Only results associated with a publication were included in table 2.

B.4. 2016 MICCAI multiple sclerosis lesion segmentation challenge

This is the second dataset published by MICCAI, comprising 53 individuals with MS [131,305]. The data were collected from various hospitals across France using different MRI scanners produced by different manufacturers: centre 01, University Hospital of Rennes (Siemens Verio 3T); centre 03 University Hospital of Bordeaux (General Electrics Discovery 3T); centre 07, University Hospital of Lyon (Siemens Aura 1.5T) and centre 08, University Hospital of Lyon (Philips Ingenia 1.5T). Centres 01, 07 and 08 contributed five patients each to the training data—giving a total of 15 patients—and 10 patients each to the testing data; centre 03 also provided an additional eight patients, giving a total of 38 patients in the testing dataset [305]. Sample scans of the various imaging protocols can be seen in figure 9. Although the dataset comprises 3D scans, the images are shown in the horizontal plane as

Table 3. Results using 2016 MICCAI MS lesion segmentation dataset. DSC = Dice; TPR = true positive rate; PPV = positive predictive value.

method	DSC	TPR	PPV
[306]	0.566	0.456	0.287
[307]	0.571	0.358	0.526
[308]	0.60	0.53	0.80
[309]	0.502	X	X
[310]	0.651	0.689	X
[311] ^a	0.63	0.57	X
[311] ^b	0.60	0.69	X
[281] ^c	0.866	0.856	X
[132] ^c	0.824	0.762	X
[280] ^c	0.76	0.65	X

^aLeMan-PV method.^bPV-CNN method.^cStudies that use dataset but do not adhere to the set.**Table 4.** Results using 2021 MICCAI MS lesion segmentation dataset. DSC = Dice; TPR = true positive rate.

method	DSC	TPR	accuracy
[312]	0.557	X	X
[313]	0.75	X	X
[314]	0.728	0.666	X
[315]	0.437	X	0.65
[316]	0.62	0.58	X
[317]	0.54	0.58	X
[318]	0.36	0.75	X
[319]	0.743	X	X
[320]	0.564	X	X
[321]	0.508	X	X
[322]	0.451	0.718	X

opposed to sagittal or coronal planes, since lesions appear most conspicuous from this angle and are thus most informative for the purpose of this review.

The gender ratio was 38:15 in favour of women across all centres. This dataset includes a wide range of MRI types. FLAIR, T1, T2, PD and Gd-enhanced MRI scans are included. TE, TR and resolutions vary from both centre to centre and scan type to scan type; see the reference paper for further details [305]. Seven experts (junior radiologists) delineated the images based on both the T2 and FLAIR images; consensus segmentation masks were then generated using iterative software, using the combined delineation performance of each of the experts [305]. Lesions smaller than 3 mm³ were removed as they were considered to be an unreliable basis for true lesions [100]. The dataset is downloadable after setting up an account on the *Shanoir platform* using the following webpage: <https://shanoir.irisa.fr/shanoir-ng/challenge-request>.

No leaderboard of results is publicly accessible; therefore, there may be discrepancies across the different studies so that direct comparison may or may not be appropriate. Not all participants of the challenge published their results explicitly, although they were plotted on a graph by the event organizers, giving an indication of the most effective entrees [131]. Results found are displayed in table 3.

B.5. 2021 MICCAI multiple sclerosis lesion segmentation challenge

The 2021 MICCAI dataset was the largest to ever be made publicly available. It comprises 100 MS patients, 40 for training and 60 for testing and the FLAIR MRI sequence was used in its creation. The dataset is longitudinal, with two scans associated with each patient, and the follow-up scan collected between 1 and 3 years after the initial one [273]. A total of 15 different MRI scanners were used in the data collection process: three General Electric, six Philips and six Siemens. Lesion segmentation of ground truth for each MRI scan was performed by four neuroradiologists, then a consensus was formed under the guidance and supervision of an expert neuroradiologist. Like the previous dataset, it is downloadable after setting up an account on the *Shanoir platform* using the same webpage: <https://shanoir.irisa.fr/shanoir-ng/challenge-request>.

No leaderboard of results is publicly accessible; therefore, there may be discrepancies across the different studies so that direct comparison may or may not be appropriate. Furthermore, at the time of writing no analysis has been published with respect to the results of the challenge. Published results found associated with the dataset can be seen in table 4.

References

- Ascherio A, Munger KL, Lennette ET, Spiegelman D, Hernán MA, Olek MJ, Hankinson SE, Hunter DJ. 2001 Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study. *JAMA* **286**, 3083–3088. (doi:[10.1001/jama.286.24.3083](https://doi.org/10.1001/jama.286.24.3083))
- Ascherio A, Munger KL. 2010 Epstein–Barr virus infection and multiple sclerosis: a review. *J. Neuroimmune Pharmacol.* **5**, 271–277. (doi:[10.1007/s11481-010-9201-3](https://doi.org/10.1007/s11481-010-9201-3))
- Kurtzke JF. 1997 Geography in multiple sclerosis. *J. Neurol.* **215**, 1–26. (doi:[10.1191/135245801669625359](https://doi.org/10.1191/135245801669625359))
- Dobson R, Giovannoni G, Ramagopalan S. 2013 The month of birth effect in multiple sclerosis: systematic review, meta-analysis and effect of latitude. *J. Neurol. Neurosurg. Psychiatry* **84**, 427–432. (doi:[10.1136/jnnp-2012-303934](https://doi.org/10.1136/jnnp-2012-303934))
- Kurtzke JF, Beebe GW, Norman JE. 1979 Epidemiology of multiple sclerosis in U.S. veterans: 1. Race, sex, and geographic distribution. *Neurology* **29**, 1228–1235. (doi:[10.1212/wnl.29.9_part_1.1228](https://doi.org/10.1212/wnl.29.9_part_1.1228))
- Barrie W *et al.* 2024 Elevated genetic risk for multiple sclerosis emerged in steppe pastoralist populations. *Nature* **625**, 321–328. (doi:[10.1038/s41586-023-06618-z](https://doi.org/10.1038/s41586-023-06618-z))
- Leray E, Moreau T, Fromont A, Edan G. 2016 Epidemiology of multiple sclerosis. *Rev. Neurol.* **172**, 3–13. (doi:[10.1016/j.neurol.2015.10.006](https://doi.org/10.1016/j.neurol.2015.10.006))
- Poitelon Y, Kopec AM, Belin S. 2020 Myelin fat facts: an overview of lipids and fatty acid metabolism. *Cells* **9**, 812. (doi:[10.3390/cells9040812](https://doi.org/10.3390/cells9040812))
- Kuhn S, Gritti L, Crooks D, Dombrowski Y. 2019 Oligodendrocytes in development, myelin generation and beyond. *Cells* **8**, 1424. (doi:[10.3390/cells8111424](https://doi.org/10.3390/cells8111424))
- Steinman L. 1999 Assessment of animal models for MS and demyelinating disease in the design of rational therapy. *Neuron* **24**, 511–514. (doi:[10.1016/s0896-6273\(00\)81107-1](https://doi.org/10.1016/s0896-6273(00)81107-1))
- Friese MA, Fugger L. 2005 Autoreactive CD8+ T cells in multiple sclerosis: a new target for therapy? *Brain* **128**, 1747–1763. (doi:[10.1093/brain/awh578](https://doi.org/10.1093/brain/awh578))
- Lee JY, Taghian K, Petratos S. 2014 Axonal degeneration in multiple sclerosis: can we predict and prevent permanent disability? *Acta Neuropathol. Commun.* **2**, 97. (doi:[10.1186/s40478-014-0097-7](https://doi.org/10.1186/s40478-014-0097-7))
- Lassmann H. 2018 Multiple sclerosis pathology. *Cold Spring Harb. Perspect. Med.* **8**, a028936. (doi:[10.1101/cshperspect.a028936](https://doi.org/10.1101/cshperspect.a028936))
- Navikas V, Link H. 1996 Review: cytokines and the pathogenesis of multiple sclerosis. *J. Neurosci. Res.* **45**, 322–333. (doi:[10.1002/\(sici\)1097-4547\(19960815\)45:43.0.co;2-b](https://doi.org/10.1002/(sici)1097-4547(19960815)45:43.0.co;2-b))
- Schmitz T, Chew LJ. 2008 Cytokines and myelination in the central nervous system. *Sci. World J.* **8**, 1119–1147. (doi:[10.1100/tsw.2008.140](https://doi.org/10.1100/tsw.2008.140))
- Li R *et al.* 2015 Proinflammatory GM-CSF-producing B cells in multiple sclerosis and B cell depletion therapy. *Sci. Transl. Med.* **7**, 310ra166. (doi:[10.1126/scitranslmed.aab4176](https://doi.org/10.1126/scitranslmed.aab4176))
- van Langelaar J, Rijvers L, Smolders J, van Luijn MM. 2020 B and T cells driving multiple sclerosis: identity, mechanisms and potential triggers. *Front. Immunol.* **11**, 760. (doi:[10.3389/fimmu.2020.00760](https://doi.org/10.3389/fimmu.2020.00760))
- Zozulya AL, Wiendl H. 2008 The role of regulatory T cells in multiple sclerosis. *Nat. Clin. Pract. Neurol.* **4**, 384–398. (doi:[10.1038/ncpneuro0832](https://doi.org/10.1038/ncpneuro0832))
- Lassmann H, Brück W, Lucchinetti C, Rodriguez M. 1997 Remyelination in multiple sclerosis. *Mult. Scler. J.* **3**, 133–136. (doi:[10.1177/13524589700300213](https://doi.org/10.1177/13524589700300213))
- Filippi M *et al.* 2019 Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. *Brain* **142**, 1858–1875. (doi:[10.1093/brain/awz144](https://doi.org/10.1093/brain/awz144))
- Ebers GC. 1985 Optic neuritis and multiple sclerosis. *Arch. Neurol.* **42**, 702–704. (doi:[10.1001/archneur.1985.04060070096025](https://doi.org/10.1001/archneur.1985.04060070096025))
- Barnes MP, Kent RM, Semlyen JK, McMullen KM. 2003 Spasticity in multiple sclerosis. *Neurorehabilit. Neural Repair* **17**, 66–70. (doi:[10.1177/0888439002250449](https://doi.org/10.1177/0888439002250449))
- Howes SFH. 1927 Psychoses with multiple sclerosis. *Boston Med. Surg. J.* **196**, 310–315. (doi:[10.1056/nejm192702241960804](https://doi.org/10.1056/nejm192702241960804))

24. Costello F. 2016 Vision disturbances in multiple sclerosis. *Semin. Neurol.* **36**, 185–195. (doi:10.1055/s-0036-1579692)
25. Beiske AG, Pedersen ED, Czukoj B, Myhr K-M. 2004 Pain and sensory complaints in multiple sclerosis. *Eur. J. Neurol.* **11**, 479–482. (doi:10.1111/j.1468-1331.2004.00815.x)
26. Frommann C. 1878 *Untersuchungen über die Gewebsveränderungen bei der Multiplen Sklerose des Gehirns und Rückenmarks*. Jena, Germany: Gustav Fischer Verlag.
27. Charcot JM. 1879 *Lectures on the diseases of the nervous system, delivered at la salpêtrière*, vol. 2187, 2nd edn. Henry C. Lea. (doi:10.1037/12839-000)
28. Babinski J. 1885 Recherches sur l'anatomie pathologique de la sclérose en plaque et étude comparative des diverses variétés de la sclérose de la moelle. *Arch. Physiol.* **5**, 186–207.
29. Fraenkel M, Jakob A. 1913 Zur Pathologie der multiplen Sklerose mit besonderer Berücksichtigung der akuten Formen. *Z. Für Die Gesamte Neural. Und Psychiatr.* **14**, 565–603. (doi:10.1007/bf02867902)
30. Rivers TM, Schwentker FF. 1935 Encephalomyelitis accompanied by myelin destruction experimentally produced in monkeys. *J. Exp. Med.* **61**, 689–702. (doi:10.1084/jem.61.5.689)
31. Bjornevik K *et al.* 2022 Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* **375**, 296–301. (doi:10.1126/science.abj8222)
32. Brownell B, Hughes JT. 1962 The distribution of plaques in the cerebrum in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **25**, 315. (doi:10.1191/135245801669625359)
33. Gize RW, Mishkin FS. 1970 Brain scans in multiple sclerosis. *Radiology* **97**, 297–300. (doi:10.1191/135245801669625359)
34. Kirshner HS, Tsai SI, Runge VM, Price AC. 1985 Magnetic resonance imaging and other techniques in the diagnosis of multiple sclerosis. *Arch. Neurol.* **9**, 859–863. (doi:10.1191/135245801669625359)
35. Hauser SL, Cree BA. 2020 Treatment of multiple sclerosis: a review. *Am. J. Med.* **133**, 1380–1390. (doi:10.1191/135245801669625359)
36. McGinley MP, Goldschmidt CH, Rae-Grant AD. 2021 Diagnosis and treatment of multiple sclerosis: a review. *JAMA* **325**, 765–779. (doi:10.1191/135245801669625359)
37. Walton C, King R, Rechtman L, Kaye W, Leray E, Marrie RA, Robertson N. 2020 Rising prevalence of multiple sclerosis worldwide: insight from the Atlas of MS. *Mult. Scler. J.* **26**, 1816–1821. (doi:10.1191/135245801669625359)
38. Lublin FD *et al.* 1996 Defining the clinical course of multiple sclerosis: result of an international survey. *Neurology (Eponicon)* **46**, 907–911. (doi:10.1191/135245801669625359)
39. Lebrun-Frenay C, Kantarci O, Siva A, Azevedo CJ, Makhani N, Pelletier D, Okuda DT. 2023 Radiologically isolated syndrome. *Lancet Neurol.* **22**, 1075–1086. (doi:10.1016/s1474-4422(23)00281-8)
40. Miller DH, Chard DT, Ciccarelli O. 2012 Clinically isolated syndromes. *Lancet Neurol.* **11**, 157–169. (doi:10.1016/s1474-4422(11)70274-5)
41. Miller D *et al.* 2008 Differential diagnosis of suspected multiple sclerosis: a consensus approach. *Mult. Scler. J.* **14**, 1157–1174. (doi:10.1177/135245808096878)
42. My-MS.org. 2013 *Types of multiple sclerosis*. See https://my-ms.org/ms_types.htm.
43. Marcus JF, Waubant EL. 2013 Updates on clinically isolated syndrome and diagnostic criteria for multiple sclerosis. *Neurohospitalist* **3**, 65–80. (doi:10.1177/1941874412457183)
44. Ciccone A, Beretta S, Brusaferri F, Galea I, Protti A, Spreafico C. 2008 Corticosteroids for the long-term treatment in multiple sclerosis. *Cochrane Database Syst. Rev.* CD006264. (doi:10.1002/14651858.cd006264.pub2)
45. Miller D, Barkhof F, Montalban X, Thompson A, Filippi M. 2005 Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. *Lancet Neurol.* **4**, 281–288. (doi:10.1016/s1474-4422(05)70071-5)
46. Goldenberg MM. 2012 Multiple sclerosis review. *Pharm. Ther.* **37**, 175. (doi:10.1191/135245801669625359)
47. McKay KA, Kwan V, Duggan T, Tremlett H. 2015 Risk factors associated with the onset of relapsing-remitting and primary progressive multiple sclerosis: a systematic review. *BioMed Res. Int.* **2015**, 817238. (doi:10.1155/2015/817238)
48. Poser CM *et al.* 1983 New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann. Neurol.* **13**, 227–231. (doi:10.1002/ana.410130302)
49. Ghasemi N, Razavi S, Nikzad E. 2017 Multiple sclerosis: pathogenesis, symptoms, diagnoses and cell-based therapy. *Cell J.* **19**, 1. (doi:10.22074/cellj.2016.4867)
50. Filippi M *et al.* 2020 Identifying progression in multiple sclerosis: new perspectives. *Ann. Neurol.* **88**, 438–452. (doi:10.1002/ana.25808)
51. Gohil K. 2015 Multiple sclerosis: progress, but no cure. *Pharm. Therapeut.* **40**, 604–605. (doi:10.1191/135245801669625359)
52. Pozzilli C, Marinelli F, Romano S, Bagnato F. 2004 Corticosteroids treatment. *J. Neurol. Sci.* **223**, 47–51. (doi:10.1016/j.jns.2004.04.019)
53. Lester RS, Knowles SR, Shear NH. 1998 The risks of systemic corticosteroid use. *Dermatol. Clin.* **16**, 277–288. (doi:10.1016/s0733-8635(05)70010-3)
54. Goodin DS. 2008 Disease-modifying therapy in multiple sclerosis: update and clinical implication. *Neurology* **71**, S8–S13. (doi:10.1212/WNL.0b013e31818f3d8b)
55. De Andrea M, Ravera R, Gioia D, Gariglio M, Landolfo S. 2002 The interferon system: an overview. *Eur. J. Paediatr. Neurol.* **6**, A41–A46. (doi:10.1053/ejpn.2002.0573)
56. Ebers GC. 1998 Randomised double-blind placebo-controlled study of interferon β-1a in relapsing/remitting multiple sclerosis. *Lancet* **352**, 1498–1504. (doi:10.1016/S0140-6736(98)03334-0)

57. Jakimovski D, Kolb C, Ramanathan M, Zivadinov R, Weinstock-Guttman B. 2018 Interferon β for multiple sclerosis. *Cold Spring Harb. Perspect. Med.* **8**, a032003. (doi:10.1101/cshperspect.a032003)
58. Zettl UK, Hecker M, Aktas O, Wagner T, Rommer PS. 2018 Interferon β -1a and β -1b for patients with multiple sclerosis: updates to current knowledge. *Expert Rev. Clin. Immunol.* **14**, 137–153. (doi:10.1080/1744666x.2018.1426462)
59. Ransohoff RM. 2007 Natalizumab for multiple sclerosis. *N. Engl. J. Med.* **356**, 2622–2629. (doi:10.1056/nejmct071462)
60. Lublin FD *et al.* 2014 Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* **83**, 278–286. (doi:10.1191/135245801669625359)
61. Lorscheider J *et al.* 2016 Defining secondary progressive multiple sclerosis. *Brain* **139**, 2395–2405. (doi:10.1093/brain/aww173)
62. Lorscheider J *et al.* 2017 Anti-inflammatory disease-modifying treatment and short-term disability progression in SPMS. *Neurology* **89**, 1050–1059. (doi:10.1212/wnl.0000000000004330)
63. Campbell E, Coulter EH, Mattison PG, Miller L, McFadyen A, Paul L. 2016 Physiotherapy rehabilitation for people with progressive multiple sclerosis: a systematic review. *Arch. Phys. Med. Rehabil.* **97**, 141–151. (doi:10.1016/j.apmr.2015.07.022)
64. Thompson AJ, Polman CH, Miller DH, McDonald WI, Brochet B, Filippi M, Montalban X, De Sá J. 1997 Primary progressive multiple sclerosis. *Brain* **120**, 1085–1096. (doi:10.1093/brain/120.6.1085)
65. Miller DH, Leary SM. 2007 Primary-progressive multiple sclerosis. *Lancet Neurol.* **6**, 903–912. (doi:10.1016/S1474-4422(07)70243-0)
66. Thompson AJ, Kermode AG, Wicks D, MacManus DG, Kendall BE, Kingsley DPE, McDonald WI. 1991 Major differences in the dynamics of primary and secondary progressive multiple sclerosis. *Ann. Neurol.* **29**, 53–62. (doi:10.1002/ana.410290111)
67. Montalban X *et al.* 2017 Ocrelizumab versus placebo in primary progressive multiple sclerosis. *N. Engl. J. Med.* **376**, 209–220. (doi:10.1056/nejmoa1606468)
68. Lassmann H. 2014 Mechanisms of white matter damage in multiple sclerosis. *Glia* **62**, 1816–1830. (doi:10.1002/glia.22597)
69. Lucchinetti CF *et al.* 2011 Inflammatory cortical demyelination in early multiple sclerosis. *N. Engl. J. Med.* **365**, 2188–2197. (doi:10.1056/nejmoa1100648)
70. Popescu BFG, Pirko I, Lucchinetti CF. 2013 Pathology of multiple sclerosis: where do we stand? *Continuum Lifelong Learn. Neurol.* **19**, 901–921. (doi:10.1212/01.CON.0000433291.23091.65)
71. Barnett MH, Prineas JW. 2004 Relapsing and remitting multiple sclerosis: pathology of the newly forming lesion. *Ann. Neurol.* **55**, 458–468. (doi:10.1002/ana.20016)
72. Kutzelnigg A, Lucchinetti CF, Stadelmann C, Brück W, Rauschka H, Bergmann M, Schmidbauer M, Parisi JE, Lassmann H. 2005 Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain* **128**, 2705–2712. (doi:10.1093/brain/awh641)
73. Dawson JW. 1916 The histology of disseminated sclerosis. *Edinb. Med. J.* **17**, 229–241.
74. Kuhlmann T, Ludwin S, Prat A, Antel J, Brück W, Lassmann H. 2017 An updated histological classification system for multiple sclerosis lesions. *Acta Neuropathol.* **133**, 13–24. (doi:10.1007/s00401-016-1653-y)
75. Ponath G, Park C, Pitt D. 2018 The role of astrocytes in multiple sclerosis. *Front. Immunol.* **9**, 217. (doi:10.3389/fimmu.2018.00217)
76. Fischer MT *et al.* 2012 NADPH oxidase expression in active multiple sclerosis lesions in relation to oxidative tissue damage and mitochondrial injury. *Brain* **135**, 886–899. (doi:10.1093/brain/aws012)
77. Kuhlmann T, Miron V, Cuo Q, Wegner C, Antel J, Brück W. 2008 Differentiation block of oligodendroglial progenitor cells as a cause for remyelination failure in chronic multiple sclerosis. *Brain* **131**, 1749–1758. (doi:10.1093/brain/awn096)
78. Chang A *et al.* 2012 Cortical remyelination: a new target for repair therapies in multiple sclerosis. *Ann. Neurol.* **72**, 918–926. (doi:10.1002/ana.23693)
79. Werkman IL, Lentferink DH, Baron W. 2021 Macroglial diversity: white and grey areas and relevance to remyelination. *Cell. Mol. Life Sci.* **78**, 143–171. (doi:10.1007/s00018-020-03586-9)
80. Patrikios P *et al.* 2006 Remyelination is extensive in a subset of multiple sclerosis patients. *Brain* **129**, 3165–3172. (doi:10.1093/brain/awl217)
81. Frischer JM, Bramow S, Dal-Bianco A, Lucchinetti CF, Rauschka H, Schmidbauer M, Laursen H, Sorensen PS, Lassmann H. 2009 The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain* **132**, 1175–1189. (doi:10.1093/brain/awp070)
82. Geurts JJG, Barkhof F. 2008 Grey matter pathology in multiple sclerosis. *Lancet Neurol.* **7**, 841–851. (doi:10.1016/S1474-4422(08)70191-1)
83. Barkhof F. 2002 The clinico-radiological paradox in multiple sclerosis revisited. *Curr. Opin. Neurol.* **15**, 239–245. (doi:10.1097/00019052-200206000-00003)
84. Amato MP *et al.* 2004 Neocortical volume decrease in relapsing-remitting MS patients with mild cognitive impairment. *Neurology* **63**, 89–93. (doi:10.1212/01.wnl.0000129544.79539.d5)
85. Prins M, Schul E, Geurts J, van der Valk P, Drukker B, van Dam A. 2015 Pathological differences between white and grey matter multiple sclerosis lesions. *Ann. N. Y. Acad. Sci.* **1351**, 99–113. (doi:10.1111/nyas.12841)
86. Peterson JW, Bö L, Mörk S, Chang A, Trapp BD. 2001 Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Ann. Neurol.* **50**, 389–400. (doi:10.1002/ana.1123)
87. Klaver R, De Vries HE, Schenk GJ, Geurts JJG. 2013 Grey matter damage in multiple sclerosis: a pathology perspective. *Prion* **7**, 66–75. (doi:10.4161/pri.23499)
88. Fischer MT *et al.* 2013 Disease-specific molecular events in cortical multiple sclerosis lesions. *Brain* **136**, 1799–1815. (doi:10.1093/brain/awt110)
89. Howell OW *et al.* 2011 Meningeal inflammation is widespread and linked to cortical pathology in multiple sclerosis. *Brain* **134**, 2755–2771. (doi:10.1093/brain/awr182)

90. Maglizzi R, Howell O, Vora A, Serafini B, Nicholas R, Puopolo M, Reynolds R, Aloisi F. 2007 Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. *Brain* **130**, 1089–1104. (doi:10.1093/brain/awm038)
91. Stys PK. 2004 Axonal degeneration in multiple sclerosis: is it time for neuroprotective strategies? *Ann. Neurol.* **55**, 601–603. (doi:10.1002/ana.20082)
92. Trapp BD, Stys PK. 2009 Virtual hypoxia and chronic necrosis of demyelinated axons in multiple sclerosis. *Lancet Neurol.* **8**, 280–291. (doi:10.1016/s1474-4422(09)70043-2)
93. Tsagkas C *et al.* 2020 Longitudinal patterns of cortical thinning in multiple sclerosis. *Hum. Brain Mapp.* **41**, 2198–2215. (doi:10.1002/hbm.24940)
94. Sailer M *et al.* 2003 Focal thinning of the cerebral cortex in multiple sclerosis. *Brain* **126**, 1734–1744. (doi:10.1191/135245801669625359)
95. Eshaghi A *et al.* 2009 Progression of regional grey matter atrophy in multiple sclerosis. *Brain* **141**, 1665–1677. (doi:10.1191/135245801669625359)
96. Eshaghi A, Bodini B, Ridgway GR, García-Lorenzo D, Tozer DJ, Sahraian MA, Thompson AJ, Ciccarelli O. 2014 Temporal and spatial evolution of grey matter atrophy in primary progressive multiple sclerosis. *NeuroImage* **86**, 257–264. (doi:10.1016/j.neuroimage.2013.09.059)
97. Haider L *et al.* 2016 The topography of demyelination and neurodegeneration in the multiple sclerosis brain. *Brain* **139**, 807–815. (doi:10.1093/brain/awv398)
98. Rocca MA, Mesaros S, Pagani E, Sormani MP, Comi G, Filippi M. 2010 Thalamic damage and long-term progression of disability in multiple sclerosis. *Radiology* **257**, 463–469. (doi:10.1148/radiol.10100326)
99. Filippi M *et al.* 2016 MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol.* **15**, 292–303. (doi:10.1016/s1474-4422(15)00393-2)
100. Thompson AJ *et al.* 2018 Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173. (doi:10.1016/s1474-4422(17)30470-2)
101. Polman CH *et al.* 2011 Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* **69**, 292–302. (doi:10.1002/ana.22366)
102. Kale N. 2016 Optic neuritis as an early sign of multiple sclerosis. *Eye Brain* **8**, 195–202. (doi:10.2147/eb.s54131)
103. Sethi V, Yousry TA, Muhlert N, Ron M, Golay X, Wheeler-Kingshott C, Miller DH, Chard DT. 2012 Improved detection of cortical MS lesions with phase-sensitive inversion recovery MRI. *J. Neurol. Neurosurg. Psychiatry* **83**, 877–882. (doi:10.1136/jnnp-2012-303023)
104. Trip SA. 2005 Imaging in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **76**, iii11–iii18. (doi:10.1136/jnnp.2005.073213)
105. Simon JH, Jacobs L, Simonian N. 2004 T1-hypointense lesions (T1 black holes) in mild-to-moderate disability relapsing multiple sclerosis. In *Early indicators, early treatments, neuroprotection in multiple sclerosis*, pp. 135–139. Milan, Italy: Springer. (doi:10.1007/978-88-470-2117-4_14)
106. Soon D, Tozer D, Altmann D, Tofts P, Miller D. 2007 Quantification of subtle blood-brain barrier disruption in non-enhancing lesions in multiple sclerosis: a study of disease and lesion subtypes. *Mult. Scler. J.* **13**, 884–894. (doi:10.1177/1352458507076970)
107. Giovannoni G *et al.* 2022 Smouldering multiple sclerosis: the ‘real MS’. *Ther. Adv. Neurol. Disord.* **15**, 17562864211066751. (doi:10.1177/17562864211066751)
108. Andelova M *et al.* 2022 Brainstem lesions are associated with diffuse spinal cord involvement in early multiple sclerosis. *BMC Neurol.* **22**, 270. (doi:10.1186/s12883-022-02778-z)
109. Lycklama G, Thompson A, Filippi M, Miller D, Polman C, Fazekas F, Barkhof F. 2003 Spinal-cord MRI in multiple sclerosis. *Lancet Neurol.* **2**, 555–562. (doi:10.1016/s1474-4422(03)00504-0)
110. Ge Y, Gonen O, Inglese M, Babb JS, Markowitz CE, Grossman RI. 2004 Neuronal cell injury precedes brain atrophy in multiple sclerosis. *Neurology* **62**, 624–627. (doi:10.1212/wnl.62.4.624)
111. McDonald WI *et al.* 2001 Recommended diagnostic criteria for multiple sclerosis: guidelines from International Panel on the Diagnosis of Multiple Sclerosis. *Ann. Neurol. Off. J. Am. Neurol. Assoc. Child Neurol. Soc.* **50**, 121–127. (doi:10.1191/135245801669625359)
112. Polman CH *et al.* 2005 Diagnostic criteria for multiple sclerosis: 2005 revisions to the ‘McDonald criteria’. *Ann. Neurol.* **58**, 840–846. (doi:10.1002/ana.20703)
113. Kurtzke JF. 1983 Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* **33**, 1444–1452. (doi:10.1212/wnl.33.11.1444)
114. Meyer-Moock S, Feng YS, Maeurer M, Dippel FW, Kohlmann T. 2014 Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurol.* **14**, 58. (doi:10.1186/1471-2377-14-58)
115. Fischer J, Rudick R, Cutter G, Reingold S. 1999 The multiple sclerosis functional composite measure (MSFC): an integrated approach to MS clinical outcome assessment. *Mult. Scler. J.* **5**, 244–250. (doi:10.1177/135245859900500409)
116. Rudick R, Cutter G, Reingold S. 2002 The multiple sclerosis functional composite: a new clinical outcome measure for multiple sclerosis trials. *Mult. Scler. J.* **8**, 359–365. (doi:10.1191/135245802ms845oa)
117. Tiftikçioğlu BI. 2018 Multiple sclerosis functional composite (MSFC): scoring instructions. *Noro. Psikiyat. Ars.* **55**, S46–S48. (doi:10.29399/npa.23330)
118. Cohen JA, Cutter GR, Fischer JS, Goodman AD, Heidenreich FR, Jak AJ, Kniker JE. 2001 Use of the multiple sclerosis functional composite as an outcome measure in a phase 3 clinical trial. *Arch. Neurol.* **58**, 961–967. (doi:10.1001/archneur.58.6.961)
119. Aliaga ES, Barkhof F. 2014 MRI mimics of multiple sclerosis. In *Handbook of clinical neurology, multiple sclerosis and related disorders*, vol. 122, pp. 291–316. Amsterdam, The Netherlands: Elsevier. (doi:10.1016/b978-0-444-52001-2.00012-1)

120. Nabizadeh F, Masrouri S, Ramezannezhad E, Ghaderi A, Sharifi AM, Soraneh S, Naser Moghadasi A. 2022 Artificial intelligence in the diagnosis of multiple sclerosis: a systematic review. *Mult. Scler. Relat. Disord.* **59**, 103673. (doi:[10.1016/j.msard.2022.103673](https://doi.org/10.1016/j.msard.2022.103673))
121. Vázquez-Marrufo M, Sarrias-Arrabal E, García-Torres M, Martín-Clemente R, Izquierdo G. 2023 A systematic review of the application of machine-learning algorithms in multiple sclerosis. *Neurología* **38**, 577–590. (doi:[10.1016/j.nrleng.2020.10.013](https://doi.org/10.1016/j.nrleng.2020.10.013))
122. Naji Y, Mahdaoui M, Klevor R, Kissani N. 2023 Artificial intelligence and multiple sclerosis: up-to-date review. *Cureus* **15**, e45412. (doi:[10.7759/cureus.45412](https://doi.org/10.7759/cureus.45412))
123. Amin M, Martínez-Heras E, Ontaneda D, Prados Carrasco F. 2024 Artificial intelligence and multiple sclerosis. *Curr. Neurol. Neurosci. Rep.* **24**, 233–243. (doi:[10.1007/s11910-024-01354-x](https://doi.org/10.1007/s11910-024-01354-x))
124. Hartmann M, Fenton N, Dobson R. 2021 Current review and next steps for artificial intelligence in multiple sclerosis risk research. *Comput. Biol. Med.* **132**, 104337. (doi:[10.1016/j.combiomed.2021.104337](https://doi.org/10.1016/j.combiomed.2021.104337))
125. Jasperse B, Barkhof F. 2023 Machine learning in multiple sclerosis. In *Machine learning for brain disorders neuromethods*, pp. 899–919. New York, NY: Springer. (doi:[10.1007/978-1-0716-3195-9_28](https://doi.org/10.1007/978-1-0716-3195-9_28))
126. Yousef H, Malagurski Tortei B, Castiglione F. 2024 Predicting multiple sclerosis disease progression and outcomes with machine learning and MRI-based biomarkers: a review. *J. Neurol.* **271**, 6543–6572. (doi:[10.1007/s00415-024-12651-3](https://doi.org/10.1007/s00415-024-12651-3))
127. Raff U, Newman FD. 1990 Lesion detection in radiologic images using an autoassociative paradigm: preliminary results. *Med. Phys.* **17**, 926–928. (doi:[10.1111/1.596449](https://doi.org/10.1111/1.596449))
128. Raff U, Vargas PF, Rojas GM, Scherzinger AL, Simon JH. 1997 Quantitation of T2 lesion load in multiple sclerosis with magnetic resonance imaging: a pilot study of a probabilistic neural network approach. *Acad. Radiol.* **4**, 431–437. (doi:[10.1016/s1076-6332\(97\)80051-7](https://doi.org/10.1016/s1076-6332(97)80051-7))
129. Filippi M, Horsfield MA, Tofts PS, Barkhof F, Thompson AJ, Miller DH. 1995 Quantitative assessment of MRI lesion load in monitoring the evolution of multiple sclerosis. *Brain* **118**, 1601–1612. (doi:[10.1093/brain/118.6.1601](https://doi.org/10.1093/brain/118.6.1601))
130. Udupa JK, Wei L, Samarasekera S, Miki Y, van Buchem MA, Grossman RI. 1997 Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Trans. Med. Imaging* **16**, 598–609. (doi:[10.1109/42.640750](https://doi.org/10.1109/42.640750))
131. Commowick O *et al.* 2018 Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Sci. Rep.* **8**, 13650. (doi:[10.1038/s41598-018-31911-7](https://doi.org/10.1038/s41598-018-31911-7))
132. Aljamaa A, Nikravan Shalmani A, Bayat P. 2021 Multiple sclerosis lesion segmentation from brain MRI using U-Net based on wavelet pooling. *Int. J. Comput. Assist. Radiol. Surg.* **16**, 1459–1467. (doi:[10.1007/s11548-021-02327-y](https://doi.org/10.1007/s11548-021-02327-y))
133. Ansari SU, Javed K, Qaisar SM, Jillani R, Haider U. 2021 Multiple sclerosis lesion segmentation in brain MRI using inception modules embedded in a convolutional neural network. *J. Healthc. Eng.* **2021**, 1155. (doi:[10.1155/2021/4138137](https://doi.org/10.1155/2021/4138137))
134. Billot B, Cerri S, Leemput KV, Dalca AV, Iglesias JE. Joint segmentation of multiple sclerosis lesions and brain anatomy in MRI scans of any contrast and resolution with CNNs. *Proc. IEEE Int. Symp. Biomed. Imaging* **2021**, 1971–1974. (doi:[10.1109/isbim2021.9434127](https://doi.org/10.1109/isbim2021.9434127))
135. Rosa FL, Abdulkadir A, Fartaria MJ, Rahmazadeh R, Galbusera R, Barakovic M, Granziera C, Cuadra MB. 2020 Multiple sclerosis cortical and WM lesion segmentation at 3 T MRI: a deep learning method based on FLAIR and MP2RAGE. *Neuroimage Clin.* **27**, 102335. (doi:[10.1016/j.nicl.2020.102335](https://doi.org/10.1016/j.nicl.2020.102335))
136. Gessert N, Krüger J, Opfer R, Manogaran P, Kitzler HH, Schippling S, Schlaefke A. 2020 Multiple sclerosis lesion activity segmentation with attention-guided two-path CNNs. *Comput. Med. Imaging Graph.* **84**, 101772. (doi:[10.1016/j.compmedimag.2020.101772](https://doi.org/10.1016/j.compmedimag.2020.101772))
137. Aslani S, Dayan M, Storelli L, Filippi M, Murino V, Rocca MA, Sona D. 2019 Multi-branch convolutional neural network for multiple sclerosis lesion segmentation. *NeuroImage* **196**, 1–15. (doi:[10.1016/j.neuroimage.2019.03.068](https://doi.org/10.1016/j.neuroimage.2019.03.068))
138. Valverde S *et al.* 2019 One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks. *Neuroimage Clin.* **21**, 101638. (doi:[10.1016/j.nicl.2018.101638](https://doi.org/10.1016/j.nicl.2018.101638))
139. Birenbaum A, Greenspan H. 2017 Multi-view longitudinal CNN for multiple sclerosis lesion segmentation. *Eng. Appl. Artif. Intell.* **65**, 111–118. (doi:[10.1016/j.engappai.2017.06.006](https://doi.org/10.1016/j.engappai.2017.06.006))
140. Valverde S *et al.* 2017 Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. *NeuroImage* **155**, 159–168. (doi:[10.1016/j.neuroimage.2017.04.034](https://doi.org/10.1016/j.neuroimage.2017.04.034))
141. Guizard N, Coupé P, Fonov VS, Manjón JV, Arnold DL. 2015 Rotation-invariant multi-contrast non-local means for MS lesion segmentation. *NeuroImage Clin.* **8**, 376–389. (doi:[10.1016/j.nicl.2015.05.001](https://doi.org/10.1016/j.nicl.2015.05.001))
142. Cerasa A *et al.* 2012 A cellular neural network methodology for the automated segmentation of multiple sclerosis lesions. *J. Neurosci. Methods* **203**, 193–199. (doi:[10.1016/j.jneumeth.2011.08.047](https://doi.org/10.1016/j.jneumeth.2011.08.047))
143. García-Lorenzo D, Lecoeur J, Arnold DL, Collins DL, Barillot C. 2009 Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts. In *Medical image computing and computer-assisted intervention*, pp. 584–591. Berlin, Germany: Springer. (doi:[10.1007/978-3-642-04271-3_71](https://doi.org/10.1007/978-3-642-04271-3_71))
144. Anbeek P, Vincken KL, Viergever MA. 2008 Automated MS-lesion segmentation by k-nearest neighbor classification. *MIDAS J.* (doi:[10.54294/ksvxf6](https://doi.org/10.54294/ksvxf6))
145. Steenwijk MD, Pouwels PJW, Daams M, Caan MWA, Richard E, Barkhof F, Vrenken H. 2013 Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (kNN-Ttps). *Neuroimage Clin.* **3**, 462–469. (doi:[10.1016/j.nicl.2013.10.003](https://doi.org/10.1016/j.nicl.2013.10.003))
146. Salem M, Valverde S, Cabezas M, Pareto D, Oliver A, Salvi J, Rovira À, Lladó X. 2020 A fully convolutional neural network for new T2-w lesion detection in multiple sclerosis. *NeuroImage* **25**, 102149. (doi:[10.1016/j.nicl.2019.102149](https://doi.org/10.1016/j.nicl.2019.102149))

147. Ghribi O, Sellami L, Slima MB, Mhiri C, Dammak M, Hamida AB. 2018 Multiple sclerosis exploration based on automatic MRI modalities segmentation approach with advanced volumetric evaluations for essential feature extraction. *Biomed. Signal Process. Control* **40**, 473–487. (doi:10.1016/j.bspc.2017.07.008)
148. Gaj S, Ontaneda D, Nakamura K. 2021 Automatic segmentation of gadolinium-enhancing lesions in multiple sclerosis using deep learning from clinical MRI. *PLoS ONE* **16**, e0255939. (doi:10.1371/journal.pone.0255939)
149. Hindsholm AM, Cramer SP, Simonsen HJ, Frederiksen JL, Andersen F, Højgaard L, Ladefoged CN, Lindberg U. 2022 Assessment of artificial intelligence automatic multiple sclerosis lesion delineation tool for clinical use. *Clin. Neuroradiol.* **32**, 643–653. (doi:10.1007/s00062-021-01089-z)
150. Coronado I, Gabr RE, Narayana PA. 2021 Deep learning segmentation of gadolinium-enhancing lesions in multiple sclerosis. *Mult. Scler. J.* **27**, 519–527. (doi:10.1177/1352458520921364)
151. Chen Z, Wang X, Huang J, Lu J, Zheng J. 2022 Deep attention and graphical neural network for multiple sclerosis lesion segmentation from MR imaging sequences. *IEEE J. Biomed. Health Informatics* **26**, 1196–1207. (doi:10.1109/jbhi.2021.3109119)
152. Brosch T, Tang LYW, Yoo Y, Li DKB, Traboulsee A, Tam R. 2016 Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation. *IEEE Trans. Med. Imaging* **35**, 1229–1239. (doi:10.1109/tmi.2016.2528821)
153. Zhang H, Zhang J, Li C, Sweeney EM, Spincemaille P, Nguyen TD, Gauthier SA, Wang Y, Marcille M. 2021 ALL-Net: anatomical information lesion-wise loss function integrated into neural network for multiple sclerosis lesion segmentation. *Neuroimage Clin.* **32**, 102854. (doi:10.1016/j.nicl.2021.102854)
154. Tomas-Fernandez X, Warfield SK. 2015 A model of population and subject (MOPS) intensities with application to multiple sclerosis lesion segmentation. *IEEE Trans. Med. Imaging* **34**, 1349–1361. (doi:10.1109/tmi.2015.2393853)
155. Essa E, Aldesouky D, Hussein SE, Rashad MZ. 2020 Neuro-fuzzy patch-wise R-CNN for multiple sclerosis segmentation. *Med. Biol. Eng. Comput.* **58**, 2161–2175. (doi:10.1007/s11517-020-02225-6)
156. Roura E, Oliver A, Cabezas M, Valverde S, Pareto D, Vilanova JC, Ramió-Torrentà L, Rovira À, Lladó X. 2015 A toolbox for multiple sclerosis lesion segmentation. *Neuroradiology* **57**, 1031–1043. (doi:10.1007/s00234-015-1552-2)
157. Bricq S, Collet C, Armspach JP. 2008 MS lesion segmentation based on hidden Markov chains. *Med. Image Anal.* **12**, 639–652. (doi:10.1016/j.media.2008.03.001)
158. Geremia E, Clatz O, Menze BH, Konukoglu E, Criminisi A, Ayache N. 2011 Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images. *NeuroImage* **57**, 378–390. (doi:10.1016/j.neuroimage.2011.03.080)
159. Souplet JC, Lebrun C, Ayache N, Malandain G. 2008 An automatic segmentation of T2-FLAIR multiple sclerosis lesions. *MIDAS J.* (doi:10.54294/6eyg0w)
160. Andermatt S, Pezold S, Cattin PC. 2018 Automated segmentation of multiple sclerosis lesions using multi-dimensional gated recurrent units. In *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries*, pp. 31–42. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-75238-9_3)
161. Goldberg-Zimring D, Achiron A, Miron S, Faibel M, Azhari H. 1998 Automated detection and characterization of multiple sclerosis lesions in brain MR images. *Magn. Reson. Imaging* **16**, 311–318. (doi:10.1016/s0730-725x(97)00300-7)
162. Boudraa AO, Dehak SM, Zhu YM, Pachai C, Bao YG, Grimaud J. 2000 Automated segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy clustering. *Comput. Biol. Med.* **30**, 23–40. (doi:10.1016/s0010-4825(99)00019-0)
163. Ait-Ali LS, Prima S, Hellier P, Carsin B, Edan G, Barillot C. 2005 STREM: a robust multidimensional parametric method to segment MS lesions in MRI. In *Medical image computing and computer-assisted intervention*, pp. 409–416. Berlin, Germany: Springer. (doi:10.1007/11566465_51)
164. He R, Sajja BR, Narayana PA. 2005 Implementation of high-dimensional feature map for segmentation of MR images. *Ann. Biomed. Eng.* **33**, 1439–1448. (doi:10.1007/s10439-005-5888-3)
165. Datta S, Sajja BR, He R, Wolinsky JS, Gupta RK, Narayana PA. 2006 Segmentation and quantification of black holes in multiple sclerosis. *NeuroImage* **29**, 467–474. (doi:10.1016/j.neuroimage.2005.07.042)
166. Horsfield MA et al. 2007 Incorporating domain knowledge into the fuzzy connectedness framework: application to brain lesion volume estimation in multiple sclerosis. *IEEE Trans. Med. Imaging* **26**, 1670–1680. (doi:10.1109/tmi.2007.901431)
167. Kawa J, Pietka E. 2007 Kernelized fuzzy c-means method in fast segmentation of demyelination plaques in multiple sclerosis. In *29th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007*. (doi:10.1109/EMBS.2007.4353620)
168. Yamamoto D, Arimura H, Kakeda S, Magome T, Yamashita Y, Toyofuku F, Ohki M, Higashida Y, Korogi Y. 2010 Computer-aided detection of multiple sclerosis lesions in brain magnetic resonance images: false positive reduction scheme consisted of rule-based, level set method, and support vector machine. *Comput. Med. Imaging Graph.* **34**, 404–413. (doi:10.1016/j.compmedimag.2010.02.001)
169. Aymerich FX, Sobrevilla P, Montseny E, Rovira A. 2011 Fuzzy approach toward reducing false positives in the detection of small multiple sclerosis lesions in magnetic resonance images. In *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011*, pp. 5694–5697. (doi:10.1109/EMBS.2011.6091378)
170. Kuwazuru J, Arimura H, Kakeda S, Yamamoto D, Magome T, Yamashita Y, Ohki M, Toyofuku F, Korogi Y. 2012 Automated detection of multiple sclerosis candidate regions in MR images: false-positive removal with use of an ANN-controlled level-set method. *Radiol. Phys. Technol.* **5**, 105–113. (doi:10.1007/s12194-011-0141-2)
171. Shepherd T, Prince SJD, Alexander DC. 2012 Interactive lesion segmentation with shape priors from offline and online learning. *IEEE Trans. Med. Imaging* **31**, 1698–1712. (doi:10.1109/TMI.2012.2196285)

172. Weiss N, Rueckert D, Rao A. 2013 Multiple sclerosis lesion segmentation using dictionary learning and sparse coding. In *Medical image computing and computer-assisted intervention*, pp. 735–742. Berlin, Germany: Springer. (doi:10.1007/978-3-642-40811-3_92)
173. Bijar A, Khayati R, Peñalver Benavent A. 2013 Increasing the contrast of the brain MR FLAIR images using fuzzy membership functions and structural similarity indices in order to segment MS lesions. *PLoS ONE* **8**, e65469. (doi:10.1371/journal.pone.0065469)
174. Sweeney EM, Vogelstein JT, Cuzzocreo JL, Calabresi PA, Reich DS, Crainiceanu CM, Shinohara RT. 2014 A comparison of supervised machine learning algorithms and feature vectors for MS lesion segmentation using multimodal structural MRI. *PLoS ONE* **9**, e95753. (doi:10.1371/journal.pone.0095753)
175. Cabezas M, Oliver A, Valverde S, Beltran B, Freixenet J, Vilanova JC, Ramió-Torrentà L, Rovira À, Lladó X. 2014 BOOST: a supervised approach for multiple sclerosis lesion segmentation. *J. Neurosci. Methods* **237**, 108–117. (doi:10.1016/j.jneumeth.2014.08.024)
176. Gao J, Li C, Feng C, Xie M, Yin Y, Davatzikos C. 2014 Non-locally regularized segmentation of multiple sclerosis lesion from multi-channel MRI data. *Magn. Reson. Imaging* **32**, 1058–1066. (doi:10.1016/j.mri.2014.03.006)
177. Deshpande H, Maurel P, Barillot C. 2015 Classification of multiple sclerosis lesions using adaptive dictionary learning. *Comput. Med. Imaging Graph.* **46**, 2–10. (doi:10.1016/j.compmedimag.2015.05.003)
178. Fartaria MJ *et al.* 2016 Automated detection of white matter and cortical lesions in early stages of multiple sclerosis. *J. Magn. Reson. Imaging* **43**, 1445–1454. (doi:10.1002/jmri.25095)
179. Knight J, Taylor GW, Khademi A. 2018 Voxel-wise logistic regression and leave-one-source-out cross validation for white matter hyperintensity segmentation. *Magn. Reson. Imaging* **54**, 119–136. (doi:10.1016/j.mri.2018.06.009)
180. Narayana PA, Coronado I, Sujit SJ, Wolinsky JS, Lublin FD, Gabr RE. 2020 Deep learning for predicting enhancing lesions in multiple sclerosis from noncontrast MRI. *Radiology* **294**, 398–404. (doi:10.1148/radiol.2019191061)
181. Al-iedani O, Ribbons K, Gholizadeh N, Lechner-Scott J, Quadrelli S, Lea R, Andronesi O, Ramadan S. 2020 Spiral MRSI and tissue segmentation of normal-appearing white matter and white matter lesions in relapsing remitting multiple sclerosis patients. *Magn. Reson. Imaging* **74**, 21–30. (doi:10.1016/j.mri.2020.09.001)
182. Narayana PA, Coronado I, Sujit SJ, Sun X, Wolinsky JS, Gabr RE. 2020 Are multi-contrast magnetic resonance images necessary for segmenting multiple sclerosis brains? A large cohort study based on deep learning. *Magn. Reson. Imaging* **65**, 8–14. (doi:10.1016/j.mri.2019.10.003)
183. Nair T, Precup D, Arnold DL, Arbel T. 2020 Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation. *Med. Image Anal.* **59**, 101557. (doi:10.1016/j.media.2019.101557)
184. Gordon S, Kodner B, Goldfryd T, Sidorov M, Goldberger J, Raviv TR. 2021 An atlas of classifiers—a machine learning paradigm for brain MRI segmentation. *Med. Biol. Eng. Comput.* **59**, 1833–1849. (doi:10.1007/s11517-021-02414-x)
185. Rakic M, Vercruyssen S, Van Eynghoven S, de la Rosa E, Jain S, Van Huffel S, Maes F, Smeets D, Sima DM. 2021 icobrain ms 5.1: combining unsupervised and supervised approaches for improving the detection of multiple sclerosis lesions. *NeuroImage* **31**, 102707. (doi:10.1016/j.nicl.2021.102707)
186. Peng Y *et al.* 2021 Prediction of unenhanced lesion evolution in multiple sclerosis using radiomics-based models: a machine learning approach. *Mult. Scler. Relat. Disord.* **53**, 102989. (doi:10.1016/j.msard.2021.102989)
187. Iswisi AFA, Karan O, Rahebi J. 2021 Diagnosis of multiple sclerosis disease in brain magnetic resonance imaging based on the Harris Hawks optimization algorithm. *BioMed Res. Int.* **2021**, 9804042. (doi:10.1155/2021/3248834)
188. McKinley R *et al.* 2021 Simultaneous lesion and brain segmentation in multiple sclerosis using deep neural networks. *Sci. Rep.* **11**, 1087. (doi:10.1038/s41598-020-79925-4)
189. Krishna Priya R, Chacko S. 2021 Improved particle swarm optimized deep convolutional neural network with super-pixel clustering for multiple sclerosis lesion segmentation in brain MRI imaging. *Int. J. Numer. Methods Biomed. Eng.* **37**, e3506. (doi:10.1002/cnm.3506)
190. Rosa FL, Beck ES, Maranzano J *et al.* 2022 Multiple sclerosis cortical lesion detection with deep learning at ultra-high-field MRI. *NMR Biomed.* **35**, e4730. (doi:10.1002/nbm.4730)
191. Giovannoni G *et al.* 2016 Brain health: time matters in multiple sclerosis. *Mult. Scler. Relat. Disord.* **9**, S5–S48. (doi:10.1016/j.msard.2016.07.003)
192. Lycklama à Nijeholt GJ. 2005 Reduction of brain volume in MS. MRI and pathology findings. *J. Neurol. Sci.* **233**, 199–202. (doi:10.1016/j.jns.2005.03.016)
193. Andrávizou A *et al.* 2019 Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options. *Autoimmun. Highlights* **10**, 7. (doi:10.1186/s13317-019-0117-5)
194. Nakamura K, Fisher E. 2009 Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *NeuroImage* **44**, 769–776. (doi:10.1016/j.neuroimage.2008.09.059)
195. Zhong J, Chen DQ, Nantes JC, Holmes SA, Hodaie M, Koski L. 2017 Combined structural and functional patterns discriminating upper limb motor disability in multiple sclerosis using multivariate approaches. *Brain Imaging Behav.* **11**, 754–768. (doi:10.1007/s11682-016-9551-4)
196. Yoo Y *et al.* 2018 Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. *NeuroImage Clin.* **17**, 169–178. (doi:10.1016/j.nic.2017.10.015)
197. Majdi MS, Keerthivasan MB, Rutt BK, Zahr NM, Rodriguez JJ, Saranathan M. 2020 Automated thalamic nuclei segmentation using multi-planar cascaded convolutional neural networks. *Magn. Reson. Imaging* **73**, 45–54. (doi:10.1016/j.mri.2020.08.005)
198. Saccà V *et al.* 2019 Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state fMRI connectivity data. *Brain Imaging Behav.* **13**, 1103–1114. (doi:10.1007/s11682-018-9926-9)

199. Mostardeiro TR, Panda A, Campeau NG, Witte RJ, Larson NB, Sui Y, Lu A, McGee KP. 2021 Whole brain 3D MR fingerprinting in multiple sclerosis: a pilot study. *BMC Med. Imaging* **21**, 88. (doi:10.1186/s12880-021-00620-5)
200. Buyukturkoglu K et al. 2022 Machine learning to investigate superficial white matter integrity in early multiple sclerosis. *J. Neuroimaging* **32**, 36–47. (doi:10.1111/jon.12934)
201. Shiee N, Bazin PL, Ozturk A, Reich DS, Calabresi PA, Pham DL. 2010 A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *NeuroImage* **49**, 1524–1535. (doi:10.1016/j.neuroimage.2009.09.005)
202. Bouman PM, Strijbis VI, Jonkman LE, Hulst HE, Geurts JJ, Steenwijk MD. 2022 Artificial double inversion recovery images for (juxta)cortical lesion visualization in multiple sclerosis. *Mult. Scler.* **28**, 541–549. (doi:10.1177/13524585211029860)
203. Azarmi F, Miri Ashtiani SN, Shalbaf A, Behnam H, Daliri MR. 2019 Granger causality analysis in combination with directed network measures for classification of MS patients and healthy controls using task-related fMRI. *Comput. Biol. Med.* **115**, 103495. (doi:10.1016/j.combiomed.2019.103495)
204. Commowick O, Fillard P, Clatz O, Warfield SK. 2008 Detection of DTI white matter abnormalities in multiple sclerosis patients. In *Medical image computing and computer-assisted intervention*, pp. 975–982. Berlin, Germany: Springer. (doi:10.1007/978-3-540-85988-8_116)
205. Zurita M, Montalba C, Labb   T, Cruz JP, Tejos C, Ciampi E, C  rcamo C, Sitaram R, Uribe S. 2018 Characterization of relapsing-remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion MRI data. *NeuroImage Clin.* **20**, 724–730. (doi:10.1016/j.nicl.2018.09.002)
206. Hackmack K, Paul F, Weygandt M, Allefeld C, Haynes JD. 2012 Multi-scale classification of disease using structural MRI and wavelet transform. *NeuroImage* **62**, 48–58. (doi:10.1016/j.neuroimage.2012.05.022)
207. Ronneberger O, Fischer P, Brox T. 2015 U-Net: convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention*, pp. 234–241. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-24574-4_28)
208. Eitel F et al. 2019 Uncovering convolutional neural network decisions for diagnosing multiple sclerosis on conventional MRI using layer-wise relevance propagation. *NeuroImage Clin.* **24**, 102003. (doi:10.1016/j.nicl.2019.102003)
209. Gabr RE et al. 2020 Brain and lesion segmentation in multiple sclerosis using fully convolutional neural networks: a large-scale study. *Mult. Scler. J.* **26**, 1217–1226. (doi:10.1177/1352458519856843)
210. Gros C et al. 2019 Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. *NeuroImage* **184**, 901–915. (doi:10.1016/j.neuroimage.2018.09.081)
211. Zhang Y, Duan Y, Wang X, Zhuo Z, Haller S, Barkhof F, Liu Y. 2022 A deep learning algorithm for white matter hyperintensity lesion detection and segmentation. *Neuroradiology* **64**, 727–734. (doi:10.1007/s00234-021-02820-w)
212. Narayana PA, Coronado I, Sujit SJ, Wolinsky JS, Lublin FD, Gabr RE. 2020 Deep-learning-based neural tissue segmentation of MRI in multiple sclerosis: effect of training set size. *J. Magn. Reson. Imaging* **51**, 1487–1496. (doi:10.1002/jmri.26959)
213. Truyen L, van Waesberghe J, van Walderveen MAA, van Oosten BW, Polman CH, Hommes OR, Ader HJA, Barkhof F. 1996 Accumulation of hypointense lesions ('black holes') on T spin-echo MRI correlates with disease progression in multiple sclerosis. *Neurology* **47**, 1469–1476. (doi:10.1212/wnl.47.6.1469)
214. Siar H, Tesnhehlab M. 2019 Diagnosing and classification tumors and MS simultaneous of magnetic resonance images using convolution neural network. *2019 7th Iranian Joint Congress on Fuzzy and Intelligent Systems, Bojnord, Iran, 29–31 January 2019*. (doi:10.1109/cfis.2019.8692148)
215. Drake-P  rez M, Boto J, Fitsiori A, Lovblad K, Vargas MI. 2018 Clinical applications of diffusion weighted imaging in neuroradiology. *Insights Into Imaging* **9**, 535–547. (doi:10.1007/s13244-018-0624-3)
216. Nedjati-Gilani GL, Schneider T, Hall MG, Cawley N, Hill I, Ciccarelli O, Drobniak I, Wheeler-Kingshott CAMG, Alexander DC. 2017 Machine learning based compartment models with permeability for white matter microstructure imaging. *NeuroImage* **150**, 119–135. (doi:10.1016/j.neuroimage.2017.02.013)
217. Garc  a-Lorenzo D, Francis S, Narayanan S, Arnold DL, Collins DL. 2013 Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Med. Image Anal.* **17**, 1–18. (doi:10.1016/j.media.2012.09.004)
218. Danelakis A, Theoharis T, Verganelakis DA. 2018 Survey of automated multiple sclerosis lesion segmentation techniques on magnetic resonance imaging. *Comput. Med. Imaging Graph.* **70**, 83–100. (doi:10.1016/j.compmedimag.2018.10.002)
219. Schmidt P et al. 2012 An automated tool for detection of FLAIR-hyperintense white-matter lesions in multiple sclerosis. *NeuroImage* **59**, 3774–3783. (doi:10.1016/j.neuroimage.2011.11.032)
220. Styner M, Lee J, Chin B, Chin MS, Commowick O, Tran HH, Markovic-Plese S, Jewells V, Warfield S. 2008 3D segmentation in the clinic: a grand challenge II: MS lesion segmentation. *MIDAS J.* (doi:10.54294/lmkqvm)
221. Musall BC et al. 2024 Detection of diffusely abnormal white matter in multiple sclerosis on multiparametric brain MRI using semi-supervised deep learning. *Sci. Rep.* **14**, 17157. (doi:10.1038/s41598-024-67722-2)
222. Miki Y et al. 1997 Computer-assisted quantitation of enhancing lesions in multiple sclerosis: correlation with clinical classification. *AJNR Am. J. Neuroradiol* **18**, 705–710.
223. Khayati R, Vafadust M, Towhidkhah F, Nabavi SM. 2008 A novel method for automatic determination of different stages of multiple sclerosis lesions in brain MR FLAIR images. *Comput. Med. Imaging Graph.* **32**, 124–133. (doi:10.1016/j.compmedimag.2007.10.003)
224. Ion-M  rgineanu A, Kocevar G, Stamile C, Sima DM, Durand-Dubief F, Van Huffel S, Sapppay-Marinier D. 2017 Machine learning approach for classifying multiple sclerosis courses by combining clinical data with lesion loads and magnetic resonance metabolic features. *Front. Neurosci* **11**, 398. (doi:10.3389/fnins.2017.00398)

225. Shirani A, Sun P, Trinkaus K, Perantie DC, George A, Naismith RT, Schmidt RE, Song S, Cross AH. 2019 Diffusion basis spectrum imaging for identifying pathologies in MS subtypes. *Ann. Clin. Transl. Neurol.* **6**, 2323–2327. (doi:10.1002/acn3.50903)
226. Ekşi Z, Çakiroğlu M, Öz C, Aralasmak A, Karadeli HH, Özcan ME. 2020 Differentiation of relapsing-remitting and secondary progressive multiple sclerosis: a magnetic resonance spectroscopy study based on machine learning. *Arq. De Neuro Psiquiatr.* **78**, 789–796. (doi:10.1590/0004-282x20200094)
227. Bendfeldt K *et al.* 2012 Multivariate pattern classification of gray matter pathology in multiple sclerosis. *NeuroImage* **60**, 400–408. (doi:10.1016/j.neuroimage.2011.12.070)
228. Eijlers AJC, Meijer KA, Wassenaar TM, Steenwijk MD, Uitdehaag BMJ, Barkhof F, Wink AM, Geurts JJJG, Schoonheim MM. 2017 Increased default-mode network centrality in cognitively impaired multiple sclerosis patients. *Neurology* **88**, 952–960. (doi:10.1212/wnl.0000000000003689)
229. Conti A, Treaba CA, Mehndiratta A, Barletta VT, Mainero C, Toschi N. 2021 An interpretable machine learning model to explain the interplay between brain lesions and cortical atrophy in multiple sclerosis. In *2021 43rd Annual Int. Conf. of the IEEE Engineering in Medicine & Biology Society, Mexico City, Mexico, 1–5 November 2021*, pp. 3757–3760. (doi:10.1109/EMBC46164.2021.9629526)
230. Krüger J, Ostwaldt AC, Spies L, Geisler B, Schlaefer A, Kitzler HH, Schippling S, Opfer R. 2022 Infratentorial lesions in multiple sclerosis patients: intra- and inter-rater variability in comparison to a fully automated segmentation using 3D convolutional neural networks. *Eur. Radiol.* **32**, 2798–2809. (doi:10.1007/s00330-021-08329-3)
231. Ye Z *et al.* 2020 Deep learning with diffusion basis spectrum imaging for classification of multiple sclerosis lesions. *Ann. Clin. Transl. Neurol.* **7**, 695–706. (doi:10.1002/acn3.51037)
232. Krishnan AP, Song Z, Clayton D, Gaetano L, Jia X, de Crespigny A, Bengtsson T, Carano RAD. 2022 Joint MRI T1 unenhancing and contrast-enhancing multiple sclerosis lesion segmentation with deep learning in OPERA trials. *Radiology* **302**, 662–673. (doi:10.1148/radiol.211528)
233. Crimi A *et al.* 2014 Predictive value of imaging markers at multiple sclerosis disease onset based on gadolinium- and USPIO-enhanced MRI and machine learning. *PLoS ONE* **9**, e93024. (doi:10.1371/journal.pone.0093024)
234. Eshagh A *et al.* 2021 Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nat. Commun.* **12**, 2078. (doi:10.1038/s41467-021-22265-2)
235. van Walderveen MA, Kamphorst W, Scheltens P, van Waesberghe JH, Ravid R, Valk J, Polman CH, Barkhof F. 1998 Histopathologic correlate of hypointense lesions on T1-weighted spin-echo MRI in multiple sclerosis. *Neurology* **50**, 1282–1288. (doi:10.1212/wnl.50.5.1282)
236. Krüger J, Opfer R, Gessert N, Ostwaldt AC, Manogaran P, Kitzler HH, Schlaefer A, Schippling S. 2020 Fully automated longitudinal segmentation of new or enlarged multiple sclerosis lesions using 3D convolutional neural networks. *NeuroImage Clin.* **28**, 102445. (doi:10.1016/j.nicl.2020.102445)
237. Muthuraman M *et al.* 2020 Covarying patterns of white matter lesions and cortical atrophy predict progression in early MS. *Neuro. Neuroimmunol. Neuroinflammation* **7**, 681. (doi:10.1212/nxi.0000000000000681)
238. Rolak LA. 2003 Multiple sclerosis: it's not the disease you thought it was. *Clin. Med. Res.* **1**, 57–60. (doi:10.3121/cmr.1.1.57)
239. Pagani E, Rocca MA, Gallo A, Rovaris M, Martinelli V, Comi G, Filippi M. 2005 Regional brain atrophy evolves differently in patients with multiple sclerosis according to clinical phenotype. *AJNR Am. J. Neuroradiol.* **26**, 341–346.
240. Elliott C, Francis SJ, Arnold DL, Collins DL, Arbel T. 2010 Bayesian classification of multiple sclerosis lesions in longitudinal MRI using subtraction images. In *Medical image computing and computer-assisted intervention*, pp. 290–297. Berlin, Germany: Springer. (doi:10.1007/978-3-642-15745-5_36)
241. Salem M, Cabezas M, Valverde S, Pareto D, Oliver A, Salvi J, Rovira À, Lladó X. 2018 A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *NeuroImage* **17**, 607–615. (doi:10.1016/j.nicl.2017.11.015)
242. Brosch T, Yoo Y, Li DKB, Traboulsee A, Tam R. 2014 Modeling the variability in brain morphology and lesion distribution in multiple sclerosis by deep learning. In *Medical image computing and computer-assisted intervention*, pp. 462–469. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-10470-6_58)
243. Brugnara G *et al.* 2020 Automated volumetric assessment with artificial neural networks might enable a more accurate assessment of disease burden in patients with multiple sclerosis. *Eur. Radiol.* **30**, 2356–2364. (doi:10.1007/s00330-019-06593-y)
244. Karaca Y, Zhang YD, Cattani C, Ayan U. 2017 The differential diagnosis of multiple sclerosis using convex combination of infinite kernels. *CNS Neural. Disord. Drug Targets* **16**, 36–43. (doi:10.2174/1871527315666161024142439)
245. Pontillo G *et al.* 2021 A combined radiomics and machine learning approach to overcome the clinicoradiologic paradox in multiple sclerosis. *AJNR Am. J. Neuroradiol.* **42**, 1927–1933. (doi:10.3174/ajnr.A7274)
246. Sweeney EM, Nguyen TD, Kuceyeski A, Ryan SM, Zhang S, Zexter L, Wang Y, Gauthier SA. 2021 Estimation of multiple sclerosis lesion age on magnetic resonance imaging. *NeuroImage* **225**, 117451. (doi:10.1016/j.neuroimage.2020.117451)
247. Storelli L *et al.* 2022 A deep learning approach to predicting disease progression in multiple sclerosis using magnetic resonance imaging. *Investig. Radiol.* **57**, 423–432. (doi:10.1097/rli.0000000000000854)
248. Zhao Y, Healy BC, Rotstein D, Guttmann CRG, Bakshi R, Weiner HL, Brodley CE, Chitnis T. 2017 Exploration of machine learning techniques in predicting multiple sclerosis disease course. *PLoS ONE* **12**, e0174866. (doi:10.1371/journal.pone.0174866)
249. Loizou CP, Murray V, Pattichis MS, Seimenis I, Pantziaris M, Pattichis CS. 2011 Multiscale amplitude-modulation frequency-modulation (AM-FM) texture analysis of multiple sclerosis in brain MRI images. *IEEE Trans. Inf. Technol. Biomed.* **15**, 119–129. (doi:10.1109/titb.2010.2091279)
250. Lesjak Ž, Pernuš F, Likar B, Špiclin Ž. 2016 Validation of white-matter lesion change detection methods on a novel publicly available MRI image database. *Neuroinformatics* **14**, 403–420. (doi:10.1007/s12021-016-9301-1)

251. Loizou CP, Pantzaris M, Pattichis CS. 2020 Normal appearing brain white matter changes in relapsing multiple sclerosis: texture image and classification analysis in serial MRI scans. *Magn. Reson. Imaging* **73**, 192–202. (doi:10.1016/j.mri.2020.08.022)
252. Buyukturkoglu K *et al.* 2021 Classifying multiple sclerosis patients on the basis of SDMT performance using machine learning. *Mult. Scler. J.* **27**, 107–116. (doi:10.1177/1352458520958362)
253. Marzullo A, Kocevar G, Stamile C, Calimeri F, Terracina G, Durand-Dubief F, Sappay-Marinier D. 2019 Prediction of multiple sclerosis patient disability from structural connectivity using convolutional neural networks. In *2019 41st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Berlin, Germany*, 23–27 July 2019, pp. 2087–2090. (doi:10.1109/embc.2019.8856845)
254. Yoo Y, Tang LYW, Li DKB, Metz L, Kolind S, Traboulsee AL, Tam RC. 2019 Deep learning of brain lesion patterns and user-defined clinical and MRI features for predicting conversion to multiple sclerosis from clinically isolated syndrome. *Comput. Methods Biomed. Eng.* **7**, 250–259. (doi:10.1080/21681163.2017.1356750)
255. Roca P *et al.* 2020 Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI. *Diagn. Interv. Imaging* **101**, 795–802. (doi:10.1016/j.diii.2020.05.009)
256. Tommasin S *et al.* 2021 Machine learning classifier to identify clinical and radiological features relevant to disability progression in multiple sclerosis. *J. Neural.* **268**, 4834–4845. (doi:10.1007/s00415-021-10605-7)
257. Dwyer M *et al.* 2021 DeepGRAI (Deep Gray Rating via Artificial Intelligence): fst, feasible, and clinically relevant thalamic atrophy measurement on clinical quality T2-FLAIR MRI in multiple sclerosis. *Neuroimage* **30**, 102652. (doi:10.1016/j.nicl.2021.102652)
258. Platten M, Brusini I, Andersson O, Ouellette R, Piehl F, Wang C, Granberg T. 2021 Deep learning corpus callosum segmentation as a neurodegenerative marker in multiple sclerosis. *J. Neuroimaging* **31**, 493–500. (doi:10.1111/jon.12838)
259. Eskut N, Koc AM, Köskderelioglu A *et al.* 2023 Is there a relationship between fall status, cognition and cerebellar lobule volume in patients with multiple sclerosis? *Appl. Neuropsychol.* **30**, 2021. (doi:10.1080/23279095.2021.1962881)
260. Zhang H, Alberts E, Pongratz V, Mühlau M, Zimmer C, Wiestler B, Eichinger P. 2019 Predicting conversion from clinically isolated syndrome to multiple sclerosis: an imaging-based machine learning approach. *Neuroimage* **21**, 101593. (doi:10.1016/j.nicl.2018.11.003)
261. Wottschel V, Alexander DC, Kwok PP, Chard DT, Stromillo ML, De Stefano N, Thompson AJ, Miller DH, Ciccarelli O. 2015 Predicting outcome in clinically isolated syndrome using machine learning. *NeuroImage* **7**, 281–287. (doi:10.1016/j.nicl.2014.11.021)
262. Wottschel V *et al.* 2019 SVM recursive feature elimination analyses of structural brain MRI predicts near-term relapses in patients with clinically isolated syndromes suggestive of multiple sclerosis. *Neuroimage Clin.* **24**, 102011. (doi:10.1016/j.nicl.2019.102011)
263. Bendfeldt K *et al.* 2019 MRI-based prediction of conversion from clinically isolated syndrome to clinically definite multiple sclerosis using SVM and lesion geometry. *Brain Imaging Behav.* **13**, 1361–1374. (doi:10.1007/s11682-018-9942-9)
264. Rocca MA, Colombo B, Falini A, Ghezzi A, Martinelli V, Scotti G, Comi G, Filippi M. 2005 Cortical adaptation in patients with MS: a cross-sectional functional MRI study of disease phenotypes. *Lancet Neurol.* **4**, 618–626. (doi:10.1016/s1474-4422(05)70171-x)
265. Sumowski JF, Leavitt VM. 2013 Cognitive reserve in multiple sclerosis. *Mult. Scler. J.* **7**, 1122–1127. (doi:10.1191/135245801669625359)
266. Lunde HMB, Assmus J, Myhr KM, Bø L, Grytnes N. 2017 Survival and cause of death in multiple sclerosis: a 60-year longitudinal population study. *J. Neurol. Neurosurg. Psychiatry* **88**, 621–625. (doi:10.1136/jnnp-2016-315238)
267. Sweeney EM, Shinohara RT, Shea CD, Reich DS, Crainiceanu CM. 2013 Automatic lesion incidence estimation and detection in multiple sclerosis using multisquence longitudinal MRI. *AJNR Am. J. Neuroradiol.* **34**, 68–73. (doi:10.3174/ajnr.a3172)
268. Pinto MF, Oliveira H, Batista S, Cruz L, Pinto M, Correia I, Martins P, Teixeira C. 2020 Prediction of disease progression and outcomes in multiple sclerosis with machine learning. *Sci. Rep.* **10**, 21038. (doi:10.1038/s41598-020-78212-6)
269. Dice LR. 1945 Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302. (doi:10.2307/1932409)
270. Bartko JJ. 1991 Measurement and reliability: statistical thinking considerations. *Schizophr. Bull.* **17**, 483–489. (doi:10.1093/schbul/17.3.483)
271. Anbeek P, Vincken KL, van Osch MJP, Bisschops RHC, van der Grond J. 2004 Probabilistic segmentation of white matter lesions in MR imaging. *Neuroimage* **21**, 1037–1044. (doi:10.1016/j.neuroimage.2003.10.012)
272. Carass A *et al.* 2017 Longitudinal multiple sclerosis lesion segmentation: resource and challenge. *Neuroimage* **148**, 77–102. (doi:10.1016/j.neuroimage.2016.12.064)
273. Commowick O, Cervenansky F, Cotton F, Dojat M. 2021 MSSEG-2 challenge proceedings: multiple sclerosis new lesions segmentation challenge using a data management and processing infrastructure. In *24th Int. Conf. on Medical Image Computing and Computer Assisted Intervention, Strasbourg, France*. <https://inria.hal.science/hal-03358968>.
274. Uebersax JS. 1987 Diversity of decision-making models and the measurement of interrater agreement. *Psychol. Bull.* **101**, 140–146. (doi:10.1037//0033-2909.101.1.140)
275. Hermann I, Golla AK, Martínez-Heras E, Schmidt R, Solana E, Llufrui S, Gass A, Schad LR, Zöllner FG. 2021 Lesion probability mapping in MS patients using a regression network on MR fingerprinting. *BMC Med. Imaging* **21**, 107. (doi:10.1186/s12880-021-00636-x)
276. Rachmadi MF *et al.* 2020 Limited one-time sampling irregularity map (LOTS-IM) for automatic unsupervised assessment of white matter hyperintensities and multiple sclerosis lesions in structural brain magnetic resonance images. *Comput. Med. Imaging Graph.* **79**, 101685. (doi:10.1016/j.compmedimag.2019.101685)
277. Hashemi SR, Mohseni Salehi SS, Erdogmus D, Prabhu SP, Warfield SK, Gholipour A. 2019 Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: application to multiple sclerosis lesion detection. *IEEE Access* **7**, 2018. (doi:10.1109/access.2018.2886371)
278. Feng Y, Pan H, Meyer C, Feng X. 2019 A self-adaptive network for multiple sclerosis lesion segmentation from multi-contrast MRI with various imaging protocols. In *IEEE 16th Int. Symp. on Biomedical Imaging, Venice, Italy*, 8–11 April 2019, pp. 472–475. (doi:10.1109/isbi.2019.8759522)

279. Roy S, Butman JA, Reich DS. 2018 Multiple sclerosis lesion segmentation from brain MRI via fully convolutional neural networks. <http://arxiv.org/abs/1803.09172>
280. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. 2021 nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211. (doi:10.1038/s41592-020-01008-z)
281. Ghosal P, Prasad PKC, Nandi D. 2019 A light weighted deep learning framework for multiple sclerosis lesion segmentation. In *2019 5th Int. Conf. on Image Information Processing, Shimla, India, 15–17 November 2019*, pp. 526–531. (doi:10.1109/ICIIP47207.2019.8985674)
282. Kumar A et al. 2019 A dense U-Net architecture for multiple sclerosis lesion segmentation. In *TENCON 2019: 2019 IEEE Region 10 Conf., Kochi, India, 17–20 October 2019*, pp. 662–667. (doi:10.1109/TENCON.2019.8929615)
283. Kwan RKS, Evans AC, Pike GB. 1996 An extensible MRI simulator for post-processing evaluation. In *Visualization in biomedical computing*, pp. 135–140. Berlin, Germany: Springer. (doi:10.1007/bfb0046947)
284. Cocosco CA, Kollokian V, Kwan RKS, Evans AC. 1997 Brainweb: online interface to a 3D MRI simulated brain database. *Neuroimage* **5**, S425.
285. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, Evans AC. 1998 Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imaging* **17**, 463–468. (doi:10.1109/42.712135)
286. Kwan RKS, Evans AC, Pike GB. 1999 MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans. Med. Imaging* **18**, 1085–1097. (doi:10.1109/42.816072)
287. Jesson A, Wolinsky T. 2015 Hierarchical MRF and random forest segmentation of MS lesions and healthy tissues in brain MRI. In *Proc. 2015 Longitudinal Multiple Sclerosis Lesion Segmentation Challenge*, vol. 7, pp. 33–41. (doi:10.1191/135245801667520627)
288. Jerman T, Galimzianova A, Pernuš F, Likar B, Špiclin Ž. 2015 Combining unsupervised and supervised methods for lesion segmentation. In *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries*, pp. 45–56. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-30858-6_5)
289. Manjón JV, Coupé P, Raniga P, Xia Y, Desmond P, Fripp J, Salvado O. 2018 MRI white matter lesion segmentation using an ensemble of neural networks and overcomplete patch-based voting. *Comput. Med. Imaging Graph.* **69**, 43–51. (doi:10.1016/j.compmedimag.2018.05.001)
290. Strumia M, Schmidt FR, Anastasopoulos C, Granziera C, Krueger G, Brox T. 2016 White matter MS-lesion segmentation using a geometric brain model. *IEEE Trans. Med. Imaging* **35**, 1636–1646. (doi:10.1109/tmi.2016.2522178)
291. Sudre CH, Cardoso MJ, Bouvy WH, Biessels GJ, Barnes J, Ourselin S. 2015 Bayesian model selection for pathological neuroimaging data applied to white matter lesion segmentation. *IEEE Trans. Med. Imaging* **34**, 2079–2102. (doi:10.1109/tmi.2015.2419072)
292. Loizou CP, Pantziaris M, Seimenis I, Pattichis CS. 2009 Brain MR image normalization in texture analysis of multiple sclerosis. In *2009 9th Int. Conf. on Information Technology and Applications in Biomedicine, Larnaka, Cyprus, 4–7 November 2009*, pp. 1–5. (doi:10.1109/ITAB.2009.5394331)
293. Loizou CP, Murray V, Pattichis MS, Seimenis I, Pantziaris M, Pattichis CS. 2010 Multiscale amplitude-modulation frequency-modulation (AM–FM) texture analysis of multiple sclerosis in brain MRI images. *IEEE Trans. Inf. Technol. Biomed.* **15**, 119–129. (doi:10.1109/titb.2010.2091279)
294. Loizou CP, Kyriacou EC, Seimenis I, Pantziaris M, Petroudi S, Karaolis M, Pattichis CS. 2013 Brain white matter lesion classification in multiple sclerosis subjects for the prognosis of future disability. *Intell. Decis. Technol.* **7**, 3–10. (doi:10.3233/IDT-120147)
295. Loizou CP, Petroudi S, Seimenis I, Pantziaris M, Pattichis CS. 2015 Quantitative texture analysis of brain white matter lesions derived from T2-weighted MR images in MS patients with clinically isolated syndrome. *J. Neuroradiol.* **42**, 99–114. (doi:10.1016/j.neurad.2014.05.006)
296. Carass A et al. 2017 Longitudinal multiple sclerosis lesion segmentation data resource. *Data Brief* **12**, 346–350. (doi:10.1016/j.dib.2017.04.004)
297. Zhang H, Valcarcel AM, Bakshi R, Chu R, Bagnato F, Shinohara RT, Hett K, Oguz I. 2019 Multiple sclerosis lesion segmentation with tiramisu and 2.5D stacked slices. In *Medical image computing and computer assisted intervention*, pp. 338–346. Cham, Switzerland: Springer. (doi:10.1007/978-3-030-32248-9_38)
298. Zhang H, Zhang J, Wang R, Zhang Q, Gauthier SA, Spincemaille P, Nguyen TD, Wang Y. 2021 Geometric loss for deep multiple sclerosis lesion segmentation. In *2021 IEEE 18th Int. Symp. on Biomedical Imaging, Nice, France, 13–16 April 2021*, pp. 24–28. (doi:10.1109/ISBI48211.2021.9434085)
299. Hou B, Kang G, Xu X, Hu C. 2019 Cross attention densely connected networks for multiple sclerosis lesion segmentation. In *2019 IEEE Int. Conf. on Bioinformatics and Biomedicine, San Diego, CA, USA, 18–21 November 2019*, pp. 2356–2361. (doi:10.1109/BIBM47256.2019.8983149)
300. Ma T, Zhang H, Ong H, Vora A, Nguyen TD, Gupta A, Wang Y, Sabuncu MR. 2021 Ensembling low precision models for binary biomedical image segmentation. In *2021 IEEE Winter Conf. on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021*, pp. 325–334. (doi:10.1109/WACV48630.2021.00037)
301. Fenneteau A, Bourdon P, Helbert D, Fernandez-Maloigne C, Habas C, Guillemin R. 2021 Investigating efficient CNN architecture for multiple sclerosis lesion segmentation. *J. Med. Imaging* **8**. (doi:10.1117/1.jmi.8.1.014504)
302. Maier O, Handels H. 2015 MS lesion segmentation in MRI with random forests. In *Proc. 2015 Multiple Sclerosis Lesion Segmentation Challenge (MSSEG-2015), Brooklyn, New York, USA*, pp. 1–2.
303. Valcarcel AM, Linn KA, Vandekar SN, Satterthwaite TD, Muschelli J, Calabresi PA, Pham DL, Martin ML, Shinohara RT. 2018 MiMoSA: an automated method for intermodal segmentation analysis of multiple sclerosis brain lesions. *J. Neuroimaging* **28**, 389–398. (doi:10.1111/jon.12506)
304. Ghafoorian M et al. 2017 Deep multi-scale location-aware 3D convolutional neural networks for automated detection of lacunes of presumed vascular origin. *Neuroimage Clin.* **14**, 391–399. (doi:10.1016/j.nicl.2017.01.033)
305. Commissick O et al. 2021 Multiple sclerosis lesions segmentation from multiple experts: the MICCAI 2016 challenge dataset. *Neuroimage* **244**, 118589. (doi:10.1016/j.neuroimage.2021.118589)

306. Beaumont J, Commowick O, Barillot C. 2016 Automatic multiple sclerosis lesion segmentation from intensity-normalized multi-channel MRI. In *Proc. 1st MICCAI Challenge on Multiple Sclerosis Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 9–16. https://inserm.hal.science/inserm-01397806v1/file/MSSEG_Challenge_Proceedings.pdf.
307. Beaumont J, Commowick O, Barillot C. 2016 Multiple sclerosis lesion segmentation using an automated multimodal graph cut. In *Proc. 1st MICCAI Challenge on Multiple Sclerosis Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 1–8. https://inserm.hal.science/inserm-01397806v1/file/MSSEG_Challenge_Proceedings.pdf.
308. Knight J, Khademi A. 2016 MS lesion segmentation using FLAIR MRI only. In *Proc. 1st MICCAI Challenge on Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure*, pp. 21–28. https://inserm.hal.science/inserm-01397806v1/file/MSSEG_Challenge_Proceedings.pdf.
309. Mahbod A, Wang C, Smedby O. 2016 Automatic multiple sclerosis lesion segmentation using hybrid artificial neural networks. In *MSSEG Challenge Proc.: Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure*, pp. 29–36. https://inserm.hal.science/inserm-01397806v1/file/MSSEG_Challenge_Proceedings.pdf.
310. Vera-Olmos F, Melero H, Malpica N. 2016 Random forest for multiple sclerosis lesion segmentation. In *Proc. 1st MICCAI Challenge on Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure*, pp. 81–86. https://inserm.hal.science/inserm-01397806v1/file/MSSEG_Challenge_Proceedings.pdf.
311. La Rosa F, Fartaria MJ, Kober T, Richiardi J, Granziera C, Thiran JP, Cuadra MB. 2019 Shallow vs deep learning architectures for white matter lesion segmentation in the early stages of multiple sclerosis. In *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries*, pp. 142–151. Cham, Switzerland: Springer. (doi:10.1007/978-3-030-11723-8_14)
312. Prados F, Kanber B. 2021 Detecting new multiple sclerosis lesions using a mixed approach. In *MSSEG2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 1–4. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
313. Kang J, Jung S, Yim J. 2021 Segmentation of new multiple sclerosis lesions in longitudinal MRI analysis using a multi-stage 3D patch-wise deep learning algorithm. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 9–12. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
314. Gibicar A, Mitha S, Khademi A. 2021 Segmentation of new multiple sclerosis lesions using an ensemble of SC U-Nets with multi channel patch-based inputs. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 17–20. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
315. Löhr T, Paetzold JC, Sekobouyiina A, Shit S, Ezhev I, Wiestler B, H.Menze B. 2021 MSSEG-2 new MS lesions detection and segmentation challenge using a data management and processing infrastructure. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure*, pp. 21–24. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
316. Ashtari P, Barile B, Van Huffel S, Sapppay-Marinier D. 2021 Longitudinal multiple sclerosis lesion segmentation using pre-activation u-net. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 45–52. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
317. Efird C, Miller D, Cobzas D. 2021 A Unet pipeline for segmentation of new MS lesions. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure*, pp. 53–56. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
318. La Rosa F, Thiran JP, Bach M. 2021 A subtraction image-based method to detect new appearing multiple sclerosis lesions on single-contrast FLAIR MRI. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 65–68. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
319. Basaran BD, Matthews P, Bai W. 2021 Cascaded networks for new MS lesion detection. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure*, pp. 77–80. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
320. Siddiquee MMR, Myronenko A. 2021 Robust 3D MRI segmentation of multiple sclerosis lesions. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 81–84. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
321. McKinley R, Wagner F, Wiest R. 2021 Detection of lesion change in multiple sclerosis using a cascade of 3D-to-2D networks. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 97–100. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.
322. Andresen J, Uzunova H, Ehrhardt J. 2021 New multiple sclerosis lesion detection with convolutional neural registration networks. In *MSSEG-2 Challenge Proc.: Multiple Sclerosis New Lesions Segmentation Challenge using a Data Management and Processing Infrastructure*, pp. 111–114. https://inria.hal.science/hal-03358968/file/MSSEG2_Challenge_Proceedings.pdf.