Journal of NeuroEngineering and Rehabilitation

REVIEW

Open Access

# Integrating big data and artificial intelligence to predict progression in multiple sclerosis: challenges and the path forward

Hamza Khan[1,2,3], Sofie Aerts[1,4], Ilse Vermeulen[1,2], Henry C. Woodruff[3,5], Philippe Lambin[3,5†] and Liesbet M. Peeters[1,2*†]

**Abstract**

Multiple sclerosis (MS) remains a complex and costly neurological condition characterised by progressive disability, making early detection and accurate prognosis of disease progression imperative. While artificial intelligence (AI) combined with big data promises transformative advances in personalised MS care, integration of multimodal, real-world datasets, including clinical records, magnetic resonance imaging (MRI), and digital biomarkers, remains limited. This perspective paper identifies a critical gap between technical innovation and clinical implementation, driven by methodological constraints, evolving regulatory frameworks, and ethical concerns related to bias, privacy, and equity. We explore this gap through three interconnected lenses: the underuse of integrated real-world data, the barriers posed by regulation and ethics, and emerging solutions. Promising strategies such as federated learning, regulatory initiatives like DARWIN-EU and the European Health Data Space, and patient-led frameworks including PROMS and CLAIMS, offer structured pathways forward. Additionally, we highlight the growing relevance of foundation models for interpreting complex MS data and supporting clinical decision-making. We advocate for harmonised data infrastructures, patient-centred design, explainable AI, and real-world validation as core pillars for future implementation. By aligning technical, regulatory, and ethical domains, stakeholders can unlock the full potential of AI to enhance prognosis, personalise care, and improve outcomes for people with MS.

**Keywords** Multiple sclerosis, Artificial intelligence, Big data, Disability progression, Real-world data, Radiomics, Multimodal integration, Ethics, EU regulations

†Philippe Lambin and Liesbet M. Peeters contributed equally to this article.

*Correspondence:
Liesbet M. Peeters
liesbet.peeters@uhasselt.be
[1]Biomedical Research Institute (BIOMED), University MS Center, Hasselt University, Agoralaan Building C, 3590 Diepenbeek, Belgium
[2]Data Science Institute (DSI), Hasselt University, Agoralaan Building D, 3590 Diepenbeek, Belgium
[3]The D-Lab, Department of Precision Medicine, GROW Research Institute for Oncology and Reproduction, Maastricht University, Maastricht, Netherlands
[4]UHasselt, Rehabilitation Research Centre (REVAL), Faculty of Rehabilitation Sciences, Wetenschapspark 7, 3590 Diepenbeek, Belgium
[5]Department of Radiology and Nuclear Medicine, GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, Netherlands

## Introduction

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system with a complex pathogenesis involving inflammatory and neurodegenerative processes. The global prevalence of MS has increased from 2.1 million in 2008 to 2.9 million in 2023 [1]. The total yearly cost of MS management, estimated at €15.5 billion in Europe alone, is largely driven by progressive disability, related productivity losses and informal care costs [2]. Conventionally, MS is classified into 'relapsing-remitting' and 'progressive' MS. The latter is further divided into primary and secondary progressive MS, based on the presence or absence of a preceding relapsing-remitting phase. 'Progressive MS' is characterised by a gradual, continuous worsening of neurological disability over time, usually without relapses. However, recent evidence suggests that the clinical course of MS is better viewed as a continuum, with different pathophysiological processes affecting individuals at different times and to differing extents [3].

Early detection of MS progression is imperative because progressive MS has limited treatment options [4]. Moreover, accurate prognostic information might improve the quality of life of people with MS (PwMS) [5]. Identifying and quantifying progression in MS remains a challenge due to differing patient demographics, heterogeneity in lesion location and disease mechanisms [3]. Diagnosis of progression is currently based on worsening of the physical examination and persistent symptoms [6]. While many biomarkers have been proposed to detect progression, they often lack sufficient sensitivity or specificity at the individual patient level, limiting their clinical usefulness. As a result, they are rarely adopted in daily practice. Consequently, progression is often diagnosed retrospectively with an average delay of 2–3 years [7].

Big data and artificial intelligence (AI) have the potential to play a role in overcoming the complexity of MS progression [8]. AI can be defined as processes that allow machines or computers to exhibit intelligent behaviour. AI is typically classified into different types, such as machine learning (ML) and deep learning (DL). ML gives computers the ability to learn without being explicitly programmed and refers to computational algorithms for gathering and making sense of evidence derived from large volumes of data, thereby allowing or facilitating human judgment and decision-making. DL consists of ML algorithms with a brain-like algorithmic structure called artificial neural networks [9]. To date, AI has already shown promise in improving diagnostic and prognostic performances, workflow, and cost-effectiveness in various fields such as oncology [10], Parkinson's disease [11], depression [12], and epilepsy [13]. Given the increased use of AI tools assisting in medical care, the European Union (EU) has broadly classified such tools as high-risk software-as-a-medical-device (SaMD) under the EU MDR 2017/745 [14]. To clarify which frameworks govern AI when used on medical data, Table 1 contrasts the main European regulations/guidelines, indicating whether they are mandatory in clinical research, in routine care, or both.

While there have been advancements in AI [15], its use in MS clinical settings to predict disability progression remains limited, representing a significant gap [16]. This limited adoption stems largely from the lack of

**Table 1** Key regulatory frameworks for AI applied to medical data

| Framework | Jurisdiction/Issuer | Relevance to research | Relevance to clinical routine | Core relevance for AI on datasets | Typical validation/oversight requirements |
|---|---|---|---|---|---|
| General Data Protection Regulation GDPR (2016/679) [70] | European Union | Yes | Yes | Lawful processing of personal health data; data minimisation; cross-border transfer rules | Data-protection impact assessment, informed consent, pseudonymisation/anonymisation |
| EU Medical Device Regulation MDR (2017/745) [14] | European Union | +/- | Yes | Classifies diagnostic/prognostic AI software as Software-as-a-Medical-Device (SaMD) | CE-marking (European Conformity marking), clinical evaluation, post-market surveillance, quality-management system |
| EU AI Act (2024) [17] | European Union | Exempt for AI systems and models developed and used only for the purpose of the scientific research and developments | Yes | All "high-risk" clinical AI must meet horizontal lifecycle rules | Risk-management system, data/algorithm transparency, human oversight, conformity assessment |
| European Health Data Space (EHDS) [62] | European Union | Yes | Yes | Legal & technical infrastructure for secure cross-border health-data exchange | Common data models, mandatory metadata, secondary-use permits |

+/- MDR obligations apply to research prototypes only if they are used to support clinical decision-making in patient care; pure exploratory research is usually exempt but must still follow GDPR and ethics rules

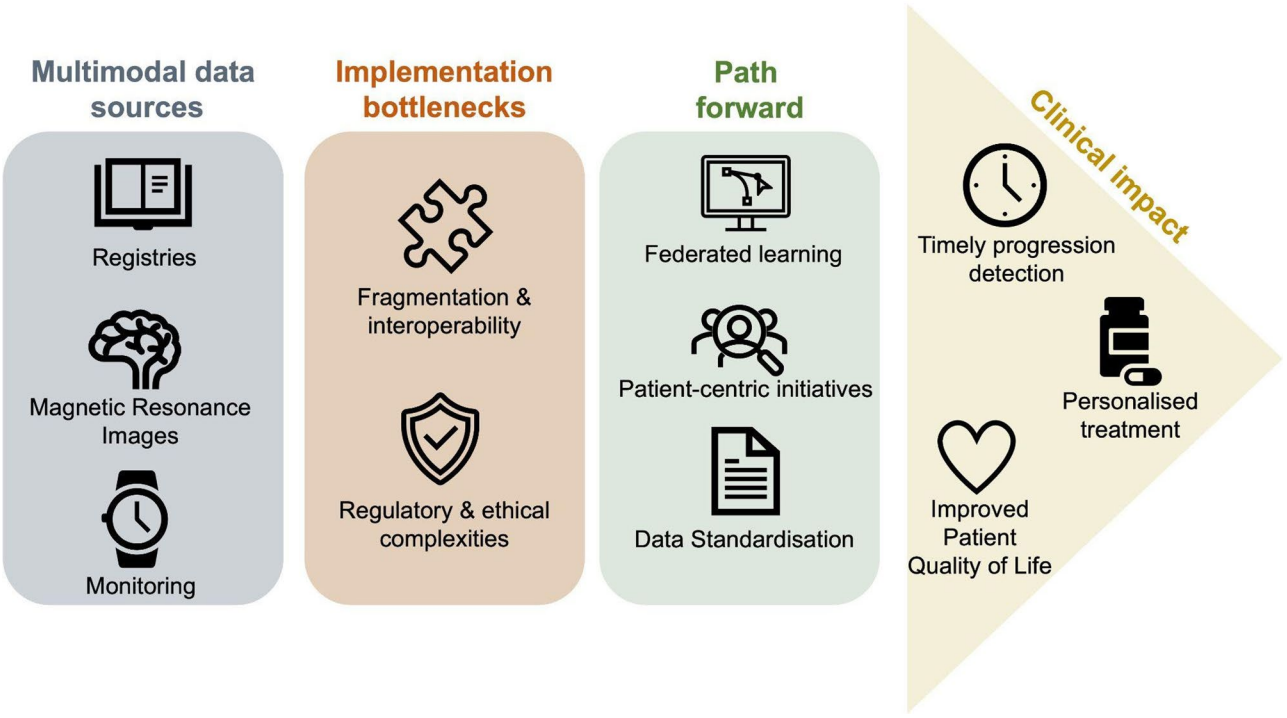


**Fig. 1** A Path Forward for Artificial Intelligence in Multiple Sclerosis Progression Prediction. The figure illustrates three interlinked challenges and enablers: the current gap in multimodal data integration, persistent regulatory barriers, and actionable solutions such as federated learning, inclusive design, and ethical oversight to support real-world clinical implementation for predicting disability progression in people with Multiple Sclerosis

availability of multimodal, diverse datasets and longitudinal validations, which are essential for clinically reliable AI algorithms. Furthermore, the 2024 EU AI Act imposes rigorous conformity assessments for SaMD, requiring evidence of data quality, risk mitigation, and robust post-market surveillance [17, 18]. In this perspective paper, we examine this gap and propose a structured way forward. Specifically, we argue that, although AI applications in MS are advancing, their potential to predict disability progression remains underutilised due to barriers in multimodal data integration and data harmonisation, leading to insufficient clinical implementation. Figure 1 summarises our perspective, highlighting three interconnected areas that will structure our argument: first, we identify the key gaps in how real-world data and AI are currently used in MS prognostication; second, we explore why these innovations have yet to be implemented at scale, focusing on regulatory, ethical, and methodological hurdles; and third, we propose a practical path forward, outlining strategies to bridge these barriers and enable meaningful clinical adoption.

## The current gap: integrating real-world data and AI for MS prognosis

Despite rapid developments in AI applications across MS, the integration of high-dimensional, real-world data (RWD) streams remains scarce. Data from clinical records, MRI scans, laboratory tests, paraclinical tests (e.g., evoked potentials), and digital monitoring are typically analysed in isolation, limiting the full potential of AI to reveal MS progression patterns. Nevertheless, the role of RWD in MS research can not be understated.

The term 'Real-World Data' is used with various definitions across the literature. In this paper, RWD is pragmatically defined as any data gathered within the context of standard care, distinct from experimental data obtained from randomised clinical trials (RCT). RWD is increasingly used to address clinical questions related to MS disease diagnosis, prognosis, and treatment [19]. Some large-scale initiatives have begun to demonstrate the potential of well-integrated RWD. Examples that stand out include the COVID-19 in MS Global Data Sharing Initiative (GDSI) [20] and the Big MS Data Network [21]. The latter, initiated in 2014, comprises data from over 350,000 PwMS across multiple European countries and has significantly advanced our understanding of disease progression and treatment effectiveness [22]. Similarly, the GDSI also illustrates a successful international data collaboration that has the potential to effectively utilise large-scale, multicentric data [20]. However, such efforts are resource-intensive and often emerge reactively during crises. To enable routine clinical AI, foundational investments and efforts in data standardisation and integration are required.

MS registries represent one of the most immediate opportunities to build this foundation. Existing for over 70 years, they have recently proliferated, with over 500,000 PwMS globally contributing data globally to varying registries [23], which can contribute to the development of prognostic models in MS [24]. There is extensive literature on prognostic MS models, summarised in several recent reviews [25, 26]. However, none have been widely adopted in clinical practice. Before introducing solutions, it is important to understand why current AI-based prognostic models in MS struggle. Despite the scale of data collection, fragmentation persists. Registries often vary in structure, granularity, and data standards, making integration across systems difficult [18]. Moreover, these registries typically collect clinical outcomes, while imaging, digital monitoring, or paraclinical test results are stored elsewhere. As a result, AI models trained on such data are constrained in scope and generalisability.

Another reason why most prognostic studies are limited is a small sample size, typically fewer than 200 subjects, restricted data sources, and limited accuracy, rarely surpassing 90% area under the receiver operating characteristic curve (AUROC) [27]. Methodological challenges such as poor calibration, selection bias, limited use of high-dimensional data, and insufficient external validation remain barriers to clinical implementation [25, 26]. Moreover, these weaknesses also reflect the limitations of relying on narrow or poorly harmonised data sources. In this paper, we mainly explore magnetic resonance imaging (MRI), which is the clinical cornerstone of MS diagnosis and monitoring [28], and patient-reported outcome measures in MS (PROMS), given their potential to incorporate patient perspectives. AI models that take into account patient-reported outcomes (PROs) and clinician-assessed outcomes (CAOs) have shown promise in one proof-of-concept study, where an AI algorithm using PROs and CAOs data correctly classified current MS disease course with ~86% accuracy and predicted progression to a secondary phase with ~82% accuracy, supporting the idea that PROs and CAOs can meaningfully aid prognostication and thus might be useful in a multimodal setting [29].

While multimodal data integration is ideal, many ML and DL applications continue to rely heavily on single modalities, particularly MRI. One of the reasons why MRI remains the central modality of neuroradiology research in MS is due to its ability to capture spatial and temporal dissemination of lesions [28]. Considerable work using MRI has been done in the field of MS diagnosis, predicting disease course and disability progression, and disease-modifying therapy (DMT) response [30–32]. Yet, traditional MRI approaches face critical limitations, encapsulated in the "clinico-radiological paradox," where clinical symptoms do not always align with visible MRI lesions [33, 34].

Recent studies suggest that quantitative MRI features, such as radiomics, combined with AI algorithms, have the capability to overcome some of these limitations by detecting subtle structural changes and microstructural damages not visible to conventional radiological assessments [35, 36]. Radiomics is a computational technique for extracting large amounts of quantifiable features from medical images [37]. The features extracted from the images comprise features related to size, shape, intensity, texture etc., that can be correlated with biological factors or clinical outcomes and help identify potential fast responders to DMTs [38]. A recent study conducted in 2024 identified a radiomic "warning sign" on brain MRI. By analysing normal-appearing white matter (NAWM) on a prior scan, their ML model (XGBoost) could distinguish regions where new MS lesions would later form . The model showed high internal performance with an accuracy of 0.91 and reasonable external validity achieving 0.74 accuracy on an independent dataset. The most important radiomic features reflected tissue heterogeneity and suggested that radiomics features from NAWM could serve as an imaging biomarker for MS progression [39].

Similarly, a review conducted by Yousef et al. (2024) found that a multimodal approach with quantitative MRI features, in most cases, improves prognostic accuracy compared to approaches without MRI [26]. Moreover, DL models that incorporate longitudinal MRI or multimodal health data have further improved prediction performance. For example, a study conducted in 2024 used innovative DL models with incorporated longitudinal MRI sequences to enhance long-term predictions. A convolutional neural network with long short-term memory (CNN-LSTM) using sequential spinal cord MRI achieved an AUROC of approximately 0.74 over a six-year prediction horizon, while a Vision Transformer-based model notably outperformed this, reaching an AUROC of around 0.84 [40]. Additionally, another study integrated longitudinal electronic health records (EHR), MRI features, and clinical notes using a deep neural network, and reported up to a 19% higher AUROC compared to models using any single modality [41].

All these studies point to a shift to explore options that go beyond conventional lesion-based assessments, and consequently, it has indeed sparked interest among the scientific community to explore radiomics-based AI algorithms to predict MS disease progression [42–45]. However, despite promising results [35, 36, 46, 47], the use of radiomics currently faces challenges in its integration into MS clinical practice. For instance, the lack of harmonisation of MRI protocols and multi-centre studies poses significant challenges that must be addressed

in order to validate its clinical utility [48], which further led to a non-recommendation in 2021 MAGNIMS-CMSC-NAIMS guidelines [49]. Moreover, even state-of-the-art ML models may struggle with the inherent unpredictability of MS in unseen cases. For example, one multi-cohort study found that while ML could classify subjects with existing high disability with good accuracy (area under the curve (AUC) ~ 0.83), it failed to significantly predict which subjects would worsen over 2–5 years [50]. The study further points towards the need for continued research in refining these MRI-based models and addressing issues such as limited longitudinal data, cohort bias, and overfitting, with an eye towards improving robustness and clinical utility [50].

Ultimately, while MRI offers valuable snapshots of disease burden and holds significant promise when combined with AI, it remains inherently limited by its scanner variability and resource demands. Additionally, complementary approaches are required to capture the everyday health status of PwMS and enable real-time progression monitoring. One avenue could be the use of smartphone applications and wearables, which are capable of recording continuous functional and mental assessments; such tools hold promise given recent evidence that patient-reported outcomes, particularly fatigue and depression scores, can independently predict disability progression in relapsing MS, even in the absence of relapses [51].

Smartphones and wearable monitoring devices are increasingly available, opening possibilities for individual monitoring in daily life. A recent online survey [52] of healthcare professionals working in the care of people with epilepsy, MS and depression showed that the majority of healthcare professionals are positive about implementing remote monitoring in patient care plans. Most PwMS have no experience using mobile health applications for their MS, with only a minority being current users of this technology [53]. Although compliance with remote monitoring remains a challenge [54], continuous, passive data collection can capture disease-related changes earlier and more consistently than episodic clinical assessments [55, 56].

Wearables and smartphones are also equipped with multiple sensor types (e.g., accelerometer, touchscreen, microphone), and therefore provide the opportunity to extract digital biomarkers for cognitive function, dexterity, gait, and balance through both active tests and passive monitoring. Recent scientific publications have showcased the use of AI and digital tools for remote monitoring of people with MS, highlighting the potential importance of these devices and softwares in MS care [57–60].

Indeed, digital health tools have recently shown their value in capturing the clinical manifestation of MS in PwMS. Multiple medical device smartphone apps for the telemonitoring of PwMS have already been developed - e.g., Floodlight [58] and iCompanion [53]. However, rigorous clinical validation and strategic oversight remain limited. Furthermore, most digital monitoring tools are being developed independently of clinical and imaging datasets. This fragmentation impairs efforts to build integrated prediction models that reflect the full disease burden, incorporating the perspective of PwMS. Therefore to bridge the gap, a truly multimodal approach is needed linking real-world clinical data, MRI, and digital monitoring tools.

**Regulatory and ethical barriers beyond data fragmentation**

While we emphasise the importance of big data and AI in MS, it is important to know why it is not merely a technical problem. In our opinion, it reflects deeper challenges at the intersection of regulation and ethics. In this section, we examine two key components that contribute to the persistence of the gap: evolving regulatory frameworks and ethical concerns around bias.

In Europe, a comprehensive regulatory shift is underway with the proposed European AI Act, which is in force as of 2025. This legislation classifies all AI systems used in healthcare as "high risk," mandating strict requirements for quality, transparency, and oversight [61]. AI tools for MS must demonstrate conformity and robust safety evidence, ensuring AI-driven decision support is reliable and unbiased before use [61]. Similarly, updates to the EU Medical Device Regulation (MDR) [14] also impact AI software and require post-market surveillance and risk management.

Complementing these regulatory changes, major EU-level initiatives, such as the European Health Data Space (EHDS) and DARWIN-EU (Data Analysis and Real World Interrogation Network), have emerged to facilitate the integration and harmonisation of health data. EHDS aims to standardise data collection, improve interoperability, and enable secure cross-border data exchange across Europe [62]. Specifically, the DARWIN-EU initiative by the European Medicines Agency (EMA) aims to create a sustainable platform to access and analyse healthcare data across Europe. DARWIN-EU focuses on using RWD to assess the safety, effectiveness, and post-marketing surveillance of medicines. It promotes methodological rigour, transparency, and harmonisation in RWD analytics, providing practical guidance to ensure high-quality evidence generation [62–64].

Even though these initiatives represent major progress and aim to protect subjects and ensure trust in AI, they also create additional complexity that slows down the adoption of AI models in clinical MS care. Researchers who are often at the forefront of developing these tools lack the regulatory expertise, technical infrastructure, or financial resources to meet such stringent certification

demands. For example, demonstrating real-world safety and performance across heterogeneous populations often requires multi-site validation studies, long-term monitoring, and formal quality management systems that are rarely feasible outside large industry-sponsored trials [65].

Furthermore, while the EHDS and DARWIN-EU initiatives aim to promote data interoperability and reuse, their implementation is still at an early stage. Many clinical centres are not yet integrated into these infrastructures, and cross-border sharing remains limited by governance and consent constraints. Thus, the intended benefits of these frameworks, while promising, are not yet accessible to most AI developers or clinical users.

To sum it up, although regulation is essential for responsible AI, it currently operates as a barrier for many early-stage innovations, especially in the context of MS where multimodal data and patient variability introduce additional complexity. Without clearer operational guidance and aligned incentives [65], the regulatory burden risks reinforcing the gap between research and real-world implementation.

As AI becomes embedded in MS care, careful attention must be paid to ethics, bias, and equity to ensure AI-driven healthcare benefits all subjects fairly. Bias primarily arises from underrepresented or unrepresentative training data. In an AI pipeline, the training dataset is the subset of data on which the algorithm actually "learns" its parameters, whereas the validation dataset is kept separate and used only to measure how well the trained model generalises to unseen cases. A bias in the training data would translate into a biased model, leading to performance disparities across diverse populations [66].

Additional ethical barriers include limited explainability, unclear accountability, privacy risks and unequal access to AI augmented health care [67–69]. The FUTURE-AI guidelines attempt to address this, recommending early identification of bias, diversified dataset inclusion, and transparent evaluation of model fairness. They recommend techniques such as data augmentation, re-sampling, or algorithmic adjustments that can help reduce bias, but they must be transparently reported and validated [61].

However, implementing these safeguards again introduces practical challenges. AI-driven MS care often relies on pooling large amounts of personal health data. Strong safeguards are needed to protect this sensitive information. European regulators have closely aligned AI requirements with the General Data Protection Regulation (GDPR) [70] to ensure data protection by design, requiring informed consent and transparency around how patient data is used, especially in continuous monitoring scenarios. In practice, GDPR compliance can delay multi-site data integration, particularly when local interpretations differ or when institutional review boards lack AI-specific guidance. This slows down model development and restricts access to diverse, representative training datasets.

An often overlooked aspect is the risk of AI-induced disparities in access to care. Advanced AI tools and digital platforms might not be equally available to all clinics or subjects. This can lead to inequitable access to AI-guided MS care. Furthermore, equity concerns also demand that AI systems be validated in underrepresented subgroups yet such datasets are often unavailable, or require costly oversampling and annotation. Policies promoting AI integration into public healthcare systems, subsidisation of validated digital tools, and explicit incorporation of social determinants of health into AI algorithms are essential strategies to ensure equity [61, 71].

Taken together, ethical, privacy, and equity concerns are necessary safeguards, but they also raise the threshold for real-world implementation. Without institutional support, funding, or dedicated guidance, these concerns may unintentionally stall or prevent AI deployment in clinical MS care.

### How we can move forward: bridging barriers with practical strategies

In recent years, several emerging frameworks and initiatives have pointed the way forward. These approaches, ranging from technical architectures to regulatory sandboxes and patient-led design, demonstrate practical methods for implementation while addressing the ethical, methodological, and operational complexities outlined previously.

The era of big data has enabled the training of more generalisable AI models, despite the limited availability of RWD in MS [16]. One example is a 2023 international study using the MSBase registry (15,240 PwMS from 146 MS centres) to predict 2-year disability progression . Using only routinely collected clinical variables (EDSS scores, relapses, treatments, etc.), machine-learning models achieved an external validation AUC of ~0.71 for predicting confirmed progression. An external validation AUC of 0.71 means that, for 71% of randomly selected pairs consisting of one progressive and one non-progressive patient, the model assigns a higher risk score to the progressive patient (counting ties as half correct). Given the arguably modest discriminative capability, the authors of the study concluded that these models are ready for clinical impact evaluation [72]. This real-world evidence signals that AI tools could be ready to assist clinicians in prognosis using data available in everyday practice.

However, sharing patient-level data across borders remains a major obstacle to AI development. To further scale data-driven discovery, federated learning

approaches have gained traction. Federated models refer to AI models training on distributed datasets across multiple hospitals or countries without pooling patient-level data. Thus, each site keeps its data locally and only model updates or learned parameters are shared, thereby addressing privacy and data-sharing barriers. Pirmani et al. (2023) describe a 3-layer federated analysis pipeline for global MS research, enabling multi-site collaboration via shared model updates instead of raw data [73]. Such frameworks have already demonstrated the ability to train lesion segmentation algorithms across different centres while preserving data confidentiality [74]. By incorporating data from diverse populations and scanners, federated learning can produce more robust AI models and reduce biases caused by single-centre training. However, federated learning, while solving the challenge of centralised data sharing, does not inherently resolve the deeper issue of data interoperability. Without standardised data schemas, ontologies, and harmonised semantic structures, AI models trained via federated methods risk learning institution-specific biases or missing key clinical nuances. Data standardisation initiatives like Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), and international MS-specific minimum datasets offer promising blueprints for semantic alignment across sites and should be considered as essential complements to any federated AI pipeline [75–78]. These methods offer a technical path forward that is inherently aligned with regulatory and ethical demands. In addition to RWD, RCT datasets can also serve as valuable complementary resources. While RCTs provide high-quality evidence under controlled conditions, combining them with RWD could enhance AI-based predictions of MS progression by merging the internal validity of trial data with the broader population relevance of real-world evidence. Although the primary focus of this paper remains on RWD, future implementation strategies could explore such hybrid approaches to strengthen both model robustness and generalisability.

To address the disconnect between innovation and regulation, regulatory sandboxes, which are essentially safe spaces for supervised testing of AI tools, can serve as vital bridges between research and clinical practice. These frameworks provide a controlled environment where AI developers, clinicians, and regulators can co-design and test high-risk AI tools before formal deployment. The European AI Act explicitly encourages such frameworks as part of its high-risk classification process [79]. By providing provisional pathways and early feedback, sandboxes reduce the risk of costly redesigns after full development.

Building on their aforementioned contributions, EHDS and DARWIN-EU provide blueprints for methodological rigour, with the former aiming to standardise data structures and secure cross-border sharing, and the latter focusing on post-marketing safety, effectiveness, and RWD analysis for healthcare interventions. Moreover, DARWIN-EU supports a federated infrastructure, enabling the reuse of healthcare data from across the EU for better evidence-generation [62–64]. These projects not only provide guidance, but they model what sustainable, GDPR-compliant, multicentre AI research could look like. Their value lies not only in guidance, but in building the infrastructure and legal framework needed for AI to move safely from research to clinic.

It is also important to note that AI systems for MS must not only be technically performant but also clinically relevant and patient-centred. Professional and scientific organisations have stepped up with practical guidance. An international consortium of experts published the FUTURE-AI guidelines in 2025, articulating good practices for trustworthy AI in healthcare [61].

Moreover, patient-centric initiatives such as PROMS, which is a global multi-stakeholder collaboration aiming to address these gaps by mapping and standardising patient-reported outcomes, can provide structured, meaningful data on the patient's lived experiences. Similarly, another initiative called IMPROVE, which is short for "Framework to IMPROVE the Integration of Patient-Generated Health Data to Facilitate Value-Based Healthcare", is an EU project on patient-generated health data, which also aims to use PROs, patient experience measures and other patient-generated health data [80]. These initiatives allow AI to prioritise what matters most: patient outcomes. Engaging PwMS in the co-design of algorithms, feature selection, and usability testing ensures greater trust and adoption of AI solutions [81].

Bridging the final mile from prototype to clinical reality, initiatives like the Clinical Impact through AI-assisted MS Care (CLAIMS) represent efforts to bridge the gap between research prototypes and clinical workflows. CLAIMS integrates regulatory, technical, and clinical stakeholders to identify and validate AI tools that are ready for impact assessment and deployment in MS clinics [82]. These types of programmes illustrate how a coordinated, cross-disciplinary approach can expedite the safe and effective integration of AI into routine care.

Together, these frameworks and initiatives demonstrate that the technical feasibility of AI for MS progression is no longer in question. What is needed now is a structured, multi-layered implementation with sustained investment in ethical oversight, patient engagement, data harmonisation, and real-world evaluation.

## Conclusion - towards clinical implementation

To meaningfully translate the promise of Big Data and AI into clinical value for PwMS, a shift is required from fragmented, siloed efforts to coordinated, patient-driven, and

ethically sound integration strategies. This paper highlights three interlinked imperatives: acknowledging the persistent gap in multimodal integration; understanding why this gap persists; and identifying practical, collaborative pathways to overcome it.

Standardisation and interoperability remain foundational hurdles for integrating AI into clinical practice. Initiatives such as EHDS and DARWIN-EU facilitate cross-border data sharing, but harmonising RWD collection methods across institutions is equally critical. For example, hospitals EHR systems can adopt a consistent structure by incorporating "Fast Health Interoperability Resources" (FHIR) [83] and/or OMOP common data models, which can standardise MS data collection across sites and lead to robust interoperability. Projects such as PROMS and CLAIMS advocate for effective strategies leading to the creation of common data frameworks, resulting in consistent, reliable, and generalisable AI models [81, 82]. However, despite their potential, practical challenges such as inconsistent coding practices, lack of semantic alignment, and vendor lock-in can hinder the broader goal of achieving true interoperability. Therefore, a continued international collaboration in harmonising clinical, imaging, and digital data will be instrumental in advancing AI towards widespread clinical adoption.

AI tools must be developed through active patient involvement, ensuring they meaningfully address the priorities of PwMS. Engaging patients and advocacy groups in defining relevant outcomes, dataset priorities, and design decisions enhance the relevance and acceptance of AI systems. The global PROMS initiative exemplifies such patient-led collaboration by embedding patient-reported outcomes into digital health technologies [81, 84]. Furthermore, explainable AI (XAI) principles foster transparency, enabling clinicians to make informed decisions rather than relying solely on opaque "black-box" predictions [61, 85]. Clear regulatory frameworks, such as FUTURE-AI, further promote the clinical integration of trustworthy and understandable AI models.

Regular and transparent evaluation of AI tools within real-world clinical settings is crucial. Structured monitoring and post-market surveillance, aligned with regulatory guidelines, will ensure that AI remains effective, robust, and safe in routine care [61]. Establishing continuous evaluation as a standard practice can promptly identify biases or unintended consequences, safeguarding patient care and maintaining clinical trust in AI solutions.

Emerging opportunities and unexplored avenues present additional pathways to strengthen AI integration in MS care. For example, combining multimodal data, including advanced MRI and computed tomography, digital biomarkers from wearables, and EHR, could enhance predictive accuracy and personalisation but remains underutilised. Similarly, adaptive AI systems capable of dynamically updating predictions based on continuous, real-time patient data hold significant promise, albeit their practical realisation remains limited. Foundation models (FMs), including vision-language models (VLMs), represent another promising direction. Recent studies demonstrate their ability to interpret MRI scans without task-specific training, achieving accuracy levels comparable to specialised models, though clinical-grade performance currently requires fine-tuning [86, 87]. Transformer-based imaging models, such as Video Vision Transformers (ViViT), have effectively predicted long-term disability progression from longitudinal MRI sequences, highlighting their potential to capture subtle temporal disease dynamics [40]. However, critical challenges such as explainability, data diversity, regulatory approval, and embedded biases must be addressed to enable practical clinical adoption of these powerful tools [61].

Furthermore, a need for a shift from episodic, clinic-centred follow-ups to continuous, home-based disease management is an avenue that should be explored. In this model, PROs, wearable-derived metrics for gait and cognition, and cloud-hosted imaging data are seamlessly integrated into AI-powered dashboards accessible to both clinicians and PwMS. This vision of a "digital hospital-at-home," discussed in recent literature [88], holds the potential to transform long-term MS care. However, patient adoption will be critical, as continuous monitoring may be perceived as intrusive without clear consent and transparent governance. Real-world implementation will depend on enabling factors such as national reimbursement frameworks, equitable broadband access, and sustained political investment in digital health infrastructure. Without these enablers, high-income healthcare systems are likely to adopt such innovations first, exacerbating the existing "digital divide" in MS care. To avoid new international disparities, we advocate for publicly funded open-source toolkits, multilateral financing, and global minimum-dataset standards that ensure AI solutions remain affordable and technically interoperable across diverse economic settings.

While various complementary approaches to AI implementation in MS exist, prioritising data harmonisation, patient-centred explainability, continuous real-world validation, and exploration of emerging innovative avenues provide a structured and ethically sound path forward. Addressing these critical themes proactively will enable AI to contribute meaningfully to MS management, delivering clinically relevant improvements that enhance the quality of life for PwMS.

**Abbreviations**

| | |
|---|---|
| AI | Artificial intelligence |
| AUROC | Area under the receiver operating characteristic curve |

| | |
|---|---|
| AUC | Area under the curve (same metric as AUROC, sometimes written without "ROC") |
| B-data / Big Data | (Capitalised as concept; not expanded because it is a common phrase) |
| CDM | Common Data Model |
| CLAIMS | Clinical L-impact through AI-assisted Multiple-sclerosiS Care initiative |
| CNN-LSTM | Convolutional neural network – long short-term memory architecture |
| DL | Deep learning |
| DMT | Disease-modifying therapy |
| EDSS | Expanded Disability Status Scale |
| EMA | European Medicines Agency |
| EHDS | European Health Data Space |
| EU | European Union |
| EU MDR | European Union Medical Device Regulation |
| FHIR | Fast Health Interoperability Resources |
| FMs | Foundation models |
| FUTURE-AI | Framework for Trustworthy and Responsible AI in Healthcare (consortium name) |
| GDPR | General Data Protection Regulation |
| GDSI | COVID-19 in MS Global Data-Sharing Initiative |
| HR / LR | High-resolution / Low-resolution (MRI acquisition) |
| IMPROVE | Framework to IMPROVE the Integration of Patient-Generated Health Data to Facilitate Value-Based Healthcare |
| LSTM | Long short-term memory (neural-network layer) |
| ML | Machine learning |
| MRI | Magnetic resonance imaging |
| MS | Multiple sclerosis |
| NAWM | Normal-appearing white matter |
| OMOP | Observational Medical Outcomes Partnership (data model) |
| PRO | Patient-Reported Outcome Measures |
| PROMS (initiative) | Patient-Reported Outcome Measures in MS (multi-stakeholder consortium) |
| PwMS | People with multiple sclerosis |
| RWD | Real-world data |
| RCT | Randomised Controlled Trials |
| ViViT | Video Vision Transformer |
| VLMs | Vision-language models |
| XAI | Explainable artificial intelligence |

**Acknowledgements**
The authors acknowledge the assistance of ChatGPT4-o, an AI language model developed by OpenAI, for its support in structuring and refining the content of this paper.

**Author contributions**
H.K. conceived the perspective, drafted the manuscript, created the figure, and coordinated revisions.S.A. co-conceived the manuscript focus, contributed critical revisions, and refined the visualisation.I.V. carried out literature verification and contributed to manuscript revisions.H.W. provided methodological guidance and substantive editorial feedback.P.L. supervised the project and provided strategic oversight and critical review.L.P. co-conceived the concept, supervised the work, and contributed substantial revisions.All authors reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

**Funding**
Hamza Khan and Sofie Aerts are supported by the Special Research Fund of Hasselt University ( BOF19DOCMA10, BOF22DOC18, respectively).
This research received funding from the Flemish Government under the Onderzoeksprogramma Artificiele Intelligentie (AI) Vlaanderen program, Stichting Multiple Sclerosis Research (19-1040 MS). The funding bodies had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

**Data availability**
No datasets were generated or analysed during the current study.

## Declarations

**Ethics approval and consent to participate**
Not applicable.

**Consent for publication**
Not applicable.

**Competing interests**
The authors declare no competing interests.

Received: 26 April 2025 / Accepted: 8 September 2025
Published online: 29 September 2025

**References**

1. Number of people with MS. | Atlas of MS [Internet]. [cited 2025 Mar 22]. Available from: https://www.atlasofms.org/map/global/epidemiology/number-of-people-with-ms
2. Kobelt G, Thompson A, Berg J, Gannedahl M, Eriksson J, MSCOI Study Group. New insights into the burden and costs of multiple sclerosis in Europe. Mult Scler Houndmills Basingstoke Engl. 2017;23:1123–36.
3. Kuhlmann T, Moccia M, Coetzee T, Cohen JA, Correale J, Graves J, et al. Multiple sclerosis progression: time for a new mechanism-driven framework. Lancet Neurol. 2023;22:78–88.
4. Inojosa H, Proschmann U, Akgün K, Ziemssen T. Should we use clinical tools to identify disease progression?? Front Neurol. 2021;11:628542.
5. Dennison L, Brown M, Kirby S, Galea I. Do people with multiple sclerosis want to know their prognosis? A UK nationwide study. PLoS ONE. 2018;13:e0193407.
6. Voigt I, Inojosa H, Wenk J, Akgün K, Ziemssen T. Building a monitoring matrix for the management of multiple sclerosis. Autoimmun Rev. 2023;22:103358.
7. Katz Sand I, Krieger S, Farrell C, Miller AE. Diagnostic uncertainty during the transition to secondary progressive multiple sclerosis. Mult Scler Houndmills Basingstoke Engl. 2014;20:1654–7.
8. Naji Y, Mahdaoui M, Klevor R, Kissani N. Artificial intelligence and multiple sclerosis: Up-to-Date review. Cureus. 2023;15:e45412.
9. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349:255–60.
10. Morin O, Vallières M, Braunstein S, Ginart JB, Upadhaya T, Woodruff HC, et al. An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pan-cancer prognostication. Nat Cancer. 2021;2:709–22.
11. Chen OY, Lipsmeier F, Phan H, Prince J, Taylor KI, Gossens C, et al. Building a Machine-Learning framework to remotely assess parkinson's disease using smartphones. IEEE Trans Biomed Eng. 2020;67:3491–500.
12. Matcham F, Leightley D, Siddi S, Lamers F, White KM, Annas P, et al. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): recruitment, retention, and data availability in a longitudinal remote measurement study. BMC Psychiatry. 2022;22:136.
13. Jirsa V, Wang H, Triebkorn P, Hashemi M, Jha J, Gonzalez-Martinez J, et al. Personalised virtual brain models in epilepsy. Lancet Neurol. 2023;22:443–54.
14. Gilbert S. The EU passes the AI act and its implications for digital medicine are unclear. NPJ Digit Med. 2024;7:135.
15. Reeve K, On BI, Havla J, Burns J, Gosteli-Peter MA, Alabsawi A, et al. Prognostic models for predicting clinical disease progression, worsening and activity in people with multiple sclerosis. Cochrane Database Syst Rev. 2023;9:CD013606.
16. Amin M, Martínez-Heras E, Ontaneda D, Prados Carrasco F. Artificial intelligence and multiple sclerosis. Curr Neurol Neurosci Rep. 2024;24:233–43.
17. Aboy M, Minssen T, Vayena E. Navigating the EU AI act: implications for regulated digital medical products. Npj Digit Med. 2024;7:237.
18. Trojano M, Kalincik T, Iaffaldano P, Amato MP. Interrogating large multiple sclerosis registries and databases: what information can be gained? Curr Opin Neurol. 2022;35:271–7.
19. Cohen JA, Trojano M, Mowry EM, Uitdehaag BM, Reingold SC, Marrie RA. Leveraging real-world data to investigate multiple sclerosis disease behavior, prognosis, and treatment. Mult Scler Houndmills Basingstoke Engl. 2020;26:23–37.

20. Peeters LM, Parciak T, Walton C, Geys L, Moreau Y, De Brouwer E, et al. COVID–19 in people with multiple sclerosis: A global data sharing initiative. Mult Scler Houndmills Basingstoke Engl. 2020;26:1157–62.
21. Hillert J, Magyari M, Soelberg Sørensen P, Butzkueven H, Van Der Welt A, Vukusic S, et al. Treatment switching and discontinuation over 20 years in the big multiple sclerosis data network. Front Neurol. 2021;12:647811.
22. Big Multiple Sclerosis Data Network. – The home of MS Real World Evidence! [Internet]. [cited 2025 Mar 15]. Available from: https://bigmsdata.org/
23. Geys L, Parciak T, Pirmani A, McBurney R, Schmidt H, Malbaša T, et al. Int J MS Care. 2021;23:261–8. The Multiple Sclerosis Data Alliance Catalogue: Enabling Web-Based Discovery of Metadata from Real-World Multiple Sclerosis Data Sources.
24. Trojano M, Tintore M, Montalban X, Hillert J, Kalincik T, Iaffaldano P, et al. Treatment decisions in multiple sclerosis - insights from real-world observational studies. Nat Rev Neurol. 2017;13:105–18.
25. Hartmann M, Fenton N, Dobson R. Current review and next steps for artificial intelligence in multiple sclerosis risk research. Comput Biol Med. 2021;132:104337.
26. Yousef H, Malagurski Tortei B, Castiglione F. Predicting multiple sclerosis disease progression and outcomes with machine learning and MRI-based biomarkers: a review. J Neurol. 2024;271:6543–72.
27. Hossain MZ, Daskalaki E, Brüstle A, Desborough J, Lueck CJ, Suominen H. The role of machine learning in developing non-magnetic resonance imaging based biomarkers for multiple sclerosis: a systematic review. BMC Med Inf Decis Mak. 2022;22:242.
28. Montalban X, Gold R, Thompson AJ, Otero-Romero S, Amato MP, Chandraratna D, et al. ECTRIMS/EAN guideline on the Pharmacological treatment of people with multiple sclerosis. Eur J Neurol. 2018;25:215–37.
29. Brichetto G, Monti Bragadin M, Fiorini S, Battaglia MA, Konrad G, Ponzio M, et al. The hidden information in patient-reported outcomes and clinician-assessed outcomes: multiple sclerosis as a proof of concept of a machine learning approach. Neurol Sci. 2020;41:459–62.
30. Aslam N, Khan IU, Bashamakh A, Alghool FA, Aboulnour M, Alsuwayan NM, et al. Multiple sclerosis diagnosis using machine learning and deep learning: challenges and opportunities. Sensors. 2022;22:7856.
31. Moazami F, Lefevre-Utile A, Papaloukas C, Soumelis V. Machine learning approaches in study of multiple sclerosis disease through magnetic resonance images. Front Immunol. 2021;12:700582.
32. Ponce de Leon-Sanchez ER, Dominguez-Ramirez OA, Herrera-Navarro AM, Rodriguez-Resendiz J, Paredes-Orta C. Mendiola-Santibañez JD. A deep learning approach for predicting multiple sclerosis. Micromachines. 2023;14:749.
33. Uitdehaag BMJ. Disability outcome measures in phase III clinical trials in multiple sclerosis. CNS Drugs. 2018;32:543–58.
34. Fouad K, Popovich PG, Kopp MA, Schwab JM. The neuroanatomical–functional paradox in spinal cord injury. Nat Rev Neurol. 2021;17:53–62.
35. Khan H, Woodruff HC, Giraldo DL, Werthen-Brabants L, Mali SA, Amirrajab S et al. Leveraging Hand-Crafted Radiomics on Multicenter FLAIR MRI for Predicting Disability Progression in People with Multiple Sclerosis [Internet]. medRxiv; 2025 [cited 2025 Mar 15]. p. 2025.01.23.25320971. Available from: https://www.medrxiv.org/content/https://doi.org/10.1101/2025.01.23.25320971v1
36. Lavrova E, Lommers E, Woodruff HC, Chatterjee A, Maquet P, Salmon E, et al. Exploratory radiomic analysis of conventional vs. Quantitative brain MRI: toward automatic diagnosis of early multiple sclerosis. Front Neurosci. 2021;15:679941.
37. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the Bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14:749–62.
38. Rogers W, Thulasi Seetha S, Refaee TAG, Lieverse RIY, Granzier RWY, Ibrahim A, et al. Radiomics: from qualitative to quantitative imaging. Br J Radiol. 2020;93:20190948.
39. Kelly BS, Mathur P, McGuinness G, Dillon H, Lee EH, Yeom KW et al. A Radiomic Warning-Sign of Progression on Brain MRI in Individuals with MS. Am J Neuroradiol [Internet]. 2024 [cited 2025 Feb 14]; Available from: https://www.ajnr.org/content/early/2024/01/11/ajnr.A8104
40. Mayfield JD, Murtagh R, Ciotti J, Robertson D, Naqa IE. Time-Dependent deep learning prediction of multiple sclerosis disability. J Imaging Inf Med. 2024;37:3231–49.
41. Zhang K, Lincoln JA, Jiang X, Bernstam EV, Shams S. Predicting multiple sclerosis severity with multimodal deep neural networks. BMC Med Inf Decis Mak. 2023;23:255.
42. Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, et al. Harmonization of multi-site diffusion tensor imaging data. NeuroImage. 2017;161:149–70.
43. La Rosa F, Abdulkadir A, Fartaria MJ, Rahmanzadeh R, Lu P-J, Galbusera R, et al. Multiple sclerosis cortical and WM lesion segmentation at 3T MRI: a deep learning method based on FLAIR and MP2RAGE. NeuroImage Clin. 2020;27:102335.
44. Eshaghi A, Young AL, Wijeratne PA, Prados F, Arnold DL, Narayanan S, et al. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. Nat Commun. 2021;12:2078.
45. Barquero G, La Rosa F, Kebiri H, Lu P-J, Rahmanzadeh R, Weigel M, et al. RimNet: A deep 3D multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis. NeuroImage Clin. 2020;28:102412.
46. Pontillo G, Tommasin S, Cuocolo R, Petracca M, Petsas N, Ugga L, et al. A combined radiomics and machine learning approach to overcome the clinicoradiologic paradox in multiple sclerosis. AJNR Am J Neuroradiol. 2021;42:1927–33.
47. Caba B, Gafson A, Lombard A, Liu D, Cafaro A, Battistella E, et al. Machine Learning-Based prediction of new multiple sclerosis lesion formation using radiomic features from Pre-Lesion normal appearing white matter (S26.009). Neurology. 2022;98:1687.
48. Gafson A, Craner MJ, Matthews PM. Personalised medicine for multiple sclerosis care. Mult Scler Houndmills Basingstoke Engl. 2017;23:362–9.
49. Wattjes MP, Ciccarelli O, Reich DS, Banwell B, de Stefano N, Enzinger C, et al. 2021 MAGNIMS–CMSC–NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. Lancet Neurol. 2021;20:653–70.
50. Noteboom S, Seiler M, Chien C, Rane RP, Barkhof F, Strijbis EMM, et al. Evaluation of machine learning-based classification of clinical impairment and prediction of clinical worsening in multiple sclerosis. J Neurol. 2024;271:5577–89.
51. Gil-Perotin S, Bernad L, Reddam S, Ferrer-Pardo C, Navarro-Quevedo S, Solís-Tarazona L. Patient's perspective in clinical practice to assess and predict disability in multiple sclerosis. Sci Rep. 2022;12:18238.
52. Andrews JA, Craven MP, Lang AR, Guo B, Morriss R, Hollis C, et al. Making remote measurement technology work in multiple sclerosis, epilepsy and depression: survey of healthcare professionals. BMC Med Inf Decis Mak. 2022;22:125.
53. Griffin N, Kehoe M. A questionnaire study to explore the views of people with multiple sclerosis of using smartphone technology for health care purposes. Disabil Rehabil. 2018;40:1434–42.
54. Canali S, Schiaffonati V, Aliverti A. Challenges and recommendations for wearable devices in digital health: data quality, interoperability, health equity, fairness. PLOS Digit Health. 2022;1:e0000104.
55. Dillenseger A, Weidemann ML, Trentzsch K, Inojosa H, Haase R, Schriefer D, et al. Digital biomarkers in multiple sclerosis. Brain Sci. 2021;11:1519.
56. Masanneck L, Gieseler P, Gordon WJ, Meuth SG, Stern AD. Evidence from clinicaltrials.gov on the growth of digital health technologies in neurology trials. NPJ Digit Med. 2023;6:23.
57. Wendt K, Trentzsch K, Haase R, Weidemann ML, Weidemann R, Aßmann U, et al. Transparent quality optimization for machine Learning-Based regression in neurology. J Pers Med. 2022;12:908.
58. Creagh AP, Dondelinger F, Lipsmeier F, Lindemann M, De Vos M. Longitudinal trend monitoring of multiple sclerosis ambulation using smartphones. IEEE Open J Eng Med Biol. 2022;3:202–10.
59. Schumann P, Scholz M, Trentzsch K, Jochim T, Śliwiński G, Malberg H, et al. Detection of fall risk in multiple sclerosis by gait Analysis-An innovative approach using feature selection ensemble and machine learning algorithms. Brain Sci. 2022;12:1477.
60. Trentzsch K, Melzer B, Stölzer-Hutsch H, Haase R, Bartscht P, Meyer P, et al. Automated analysis of the Two-Minute walk test in clinical practice using accelerometer data. Brain Sci. 2021;11:1507.
61. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ. 2025;388:e081554.
62. European Health Data Space Regulation (EHDS) - European Commission [Internet]. 2025 [cited 2025 Mar 15]. Available from: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en
63. Peeters LM. The arisal of data spaces: why I am excited and worried. Front Immunol [Internet]. 2024 [cited 2025 Jan 17];15. Available from: https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2024.1461361/full.

64. Agency EM. Data Analysis and Real World Interrogation Network (DARWIN EU) | European Medicines Agency (EMA) [Internet]. 2021 [cited 2025 Mar 15]. Available from: https://www.ema.europa.eu/en/about-us/how-we-work/big-data/real-world-evidence/data-analysis-real-world-interrogation-network-darwin-eu
65. van Kolfschooten H, van Oirschot J. The EU artificial intelligence act (2024): implications for healthcare. Health Policy. 2024;149:105152.
66. Rouzrokh P, Khosravi B, Faghani S, Moassefi M, Vera Garcia DV, Singh Y, et al. Mitigating bias in radiology machine learning: 1. Data handling. Radiol Artif Intell. 2022;4:e210290.
67. Jacobs BM, Schalk L, Tregaskis-Daniels E, Tank P, Hoque S, Peter M, et al. The relationship between ethnicity and multiple sclerosis characteristics in the united kingdom: A UK MS register study. Mult Scler Houndmills Basingstoke Engl. 2024;30:1544–55.
68. Cartwright J, Kipp K, Ng AV. Innovations in multiple sclerosis care: the impact of artificial intelligence via machine learning on clinical research and Decision-Making. Int J MS Care. 2023;25:233–41.
69. Cerezo-Martínez P, Nicolás-Sánchez A, Castro-Toledo FJ. Analyzing the European institutional response to ethical and regulatory challenges of artificial intelligence in addressing discriminatory bias. Front Artif Intell [Internet]. 2024 [cited 2025 Jul 26];7. Available from: https://www.frontiersin.org/journals/artificial-intelligence/articles/https://doi.org/10.3389/frai.2024.1393259/full
70. Regulation–2016/679 - EN. - gdpr - EUR-Lex [Internet]. [cited 2025 Mar 15]. Available from: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng
71. McLennan S, Fiske A, Tigard D, Müller R, Haddadin S, Buyx A. Embedded ethics: a proposal for integrating ethics into the development of medical AI. BMC Med Ethics. 2022;23:6.
72. Brouwer ED, Becker T, Werthen-Brabants L, Dewulf P, Iliadis D, Dekeyser C, et al. Machine-learning-based prediction of disability progression in multiple sclerosis: an observational, international, multi-center study. PLOS Digit Health. 2024;3:e0000533.
73. Pirmani A, Brouwer ED, Geys L, Parciak T, Moreau Y, Peeters LM. JMIR Med Inf. 2023;11:e48030. The Journey of Data Within a Global Data Sharing Initiative: A Federated 3-Layer Data Analysis Pipeline to Scale Up Multiple Sclerosis Research.
74. Bai L, Wang D, Wang H, Barnett M, Cabezas M, Cai W, et al. Improving multiple sclerosis lesion segmentation across clinical sites: A federated learning approach with noise-resilient training. Artif Intell Med. 2024;152:102872.
75. De Moor G, Sundgren M, Kalra D, Schmidt A, Dugas M, Claerhout B, et al. Using electronic health records for clinical research: the case of the EHR4CR project. J Biomed Inf. 2015;53:162–73.
76. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. Stud Health Technol Inf. 2015;216:574–8.
77. Mandl KD, Gottlieb D, Mandel JC. Integration of AI in healthcare requires an interoperable digital data ecosystem. Nat Med. 2024;30:631–4.
78. Szarfman A, Levine JG, Tonning JM, Weichold F, Bloom JC, Soreth JM, et al. Recommendations for achieving interoperable and shareable medical data in the USA. Commun Med. 2022;2:86.
79. Schmidt J, Schutte NM, Buttigieg S, Novillo-Ortiz D, Sutherland E, Anderson M, et al. Mapping the regulatory landscape for artificial intelligence in health within the European union. NPJ Digit Med. 2024;7:229.
80. Folkvord F, Carlson JI, Ottaviano M, Carvajal D, Gonzalez LH, van de Schoot R, et al. Using patient-generated health data more efficient and effectively to facilitate the implementation of value-based healthcare in the EU – Innovation report. Comput Struct Biotechnol J. 2024;24:672–8.
81. Zaratin P, Vermersch P, Amato MP, Brichetto G, Coetzee T, Cutter G, et al. The agenda of the global patient reported outcomes for multiple sclerosis (PROMS) initiative: progresses and open questions. Mult Scler Relat Disord. 2022;61:103757.
82. Praet J, Anderhalten L, Comi G, Horakova D, Ziemssen T, Vermersch P, et al. A future of AI-driven personalized care for people with multiple sclerosis. Front Immunol. 2024;15:1446748.
83. Index -. FHIR v5.0.0 [Internet]. [cited 2025 Jul 26]. Available from: https://www.hl7.org/fhir/?
84. Helme A, Kalra D, Brichetto G, Peryer G, Vermersch P, Weiland H et al. Artificial intelligence and science of patient input: a perspective from people with multiple sclerosis. Front Immunol [Internet]. 2025 [cited 2025 Mar 15];16. Available from: https://www.frontiersin.org/journals/immunology/articles/https://doi.org/10.3389/fimmu.2025.1487709/full
85. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3:e745–50.
86. Falet J-PR, Nobile S, Szpindel A, Barile B, Kumar A, Durso-Finley J, et al. The role of AI for MRI-analysis in multiple sclerosis—A brief overview. Front Artif Intell. 2025;8:1478068.
87. Kelly BS, Duignan S, Mathur P, Dillon H, Lee EH, Yeom KW, et al. Can ChatGPT4-vision identify radiologic progression of multiple sclerosis on brain MRI? Eur Radiol Exp. 2025;9:9.
88. Pandit JA, Pawelek JB, Leff B, Topol EJ. The hospital at home in the USA: current status and future prospects. Npj Digit Med. 2024;7:48.


## Publisher's note