BMJ Open

Multimodal diagnostics in multiple sclerosis: predicting disability and conversion from relapsing-remitting to secondary progressive disease course – protocol for systematic review and meta-analysis

Yauhen Statsenko ,[1] Darya Smetanina ,[1] Teresa Arora,[2] Linda Östlundh,[3] Tetiana Habuza,[4] Gillian Lylian Simiyu,[1] Sarah Meribout,[5] Tatsiana Talako,[1] Fransina Christelle King,[6] Iryna Makhnevych,[1] Juri Gelovani,[7,8] Karuna M Das,[1] Klaus Neidl-Van Gorkom,[1] Taleb M Almansoori,[9] Fatmah Al Zahmi,[10,11] Miklós Szolics,[9] Fatma Ismail,[12] Milos Ljubisavljevic[6]

To cite: Statsenko Y, Smetanina D, Arora T, *et al*. Multimodal diagnostics in multiple sclerosis: predicting disability and conversion from relapsing-remitting to secondary progressive disease course – protocol for systematic review and meta-analysis. *BMJ Open* 2023;**0**:e068608. doi:10.1136/bmjopen-2022-068608

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-068608).

Received 26 September 2022
Accepted 03 May 2023

For numbered affiliations see end of article.

Correspondence to
Dr Yauhen Statsenko;
e.a.statsenko@uaeu.ac.ae

ABSTRACT

Background The number of patients diagnosed with multiple sclerosis (MS) has increased significantly over the last decade. Identifying the transition from relapsing-remitting to secondary progressive MS presents a challenge. Since the number of available methods to examine patients with MS is limited, both the diagnostics and the prognostication of disease progression would benefit from the multimodal approach combining the evidence obtained from disparate radiologic modalities, neurophysiological evaluation, cognitive assessment and molecular diagnostics. This systematic review will analyse the advantages of multimodal studies in predicting the risk of conversion to secondary progressive MS.

Methods and analysis Peer-reviewed publications available in Web of Science, Medline/PubMed, Scopus, Embase and CINAHL databases will be used. We will consider in vivo studies reporting the predictive value of diagnostic methods. Selected publications will be processed through Covidence software for automatic deduplication and blind screening. Two reviewers will use a predefined template to extract the data from the eligible studies. We will analyse (1) sensitivity, specificity and accuracy, area under the curve, positive predictive value and negative predictive value in classification models predicting the risk of secondary progression and (2) accuracy of the regression models forecasting disability scores expressed as the ratio of mean absolute error to the range of values. Then, we will create ranking charts representing the performance of algorithms predicting disability level and MS progression. Finally, we will compare the predictive power of radiological and radiomical correlates of clinical disability and cognitive impairment in patients with MS.

Ethics and dissemination The study does not require ethical approval because we will analyse publicly available literature. The project results will be published in a peer-review journal and presented at scientific conferences.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The protocol is prepared according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for systematic review and is registered with the international database, PROSPERO database.
⇒ The systematic review compares distinct diagnostic modalities, and their settings for predicting clinical disability and the conversion from relapsing-remitting to secondary progressive multiple sclerosis. This helps identify the most suitable tool for confirming the disease stage and monitoring its progression.
⇒ A notable limitation of this systematic review is the uneven distribution of published studies regarding each diagnostic method to be used in the analysis.

PROSPERO registration number CRD42022354179.

INTRODUCTION

The number of patients diagnosed with multiple sclerosis (MS) increased from 2.3 to 2.8 million worldwide between 2013 and 2020.[1] These global statistics are published every 5 years in the Atlas of MS, serving as the official database for the Multiple Sclerosis International Federation. The latest report included 115 countries covering 87% of the world's population. Notably, data are missing for most African and several Central and Southeast Asian countries and do not include the total population of countries where MS

clinics fail to report the total number of cases.[2 3] The atlas data are limited for paediatric patients.[2]

The disease rates differ markedly among the countries. In the early 2000s, the highest incidence of MS was reported in North America and Northern Europe, and it was lower in Central Africa and Asia.[4 5] Hypothetically, the occurrence of MS rises in relation to increasing distance from the equator. However, this tendency is not supported by the similarities in MS incidence in South and North Europe (137–187 vs 167 cases per 100 000 individuals in Italy and Iceland, respectively).[6] Worldwide, the highest incidence was reported in the Italian region Sardinia and the Canadian province Saskatchewan (330 vs 314).[4 7] Before 2000s, the Persian Gulf countries were considered a low-risk zone for MS. However, recent studies reported a rise in the number of MS cases in these countries with an average of 31–55 individuals per 100 000 people.

Challenges in predicting disability and risk of conversion from relapsing-remitting to secondary progressive disease course

Predicting MS progression has always been an issue of attention for research scientists and clinical praticioners. For instance, the conversion from relapsing-remitting MS (RRMS) to a secondary progressive MS form (SPMS) remains challenging to predict. On average, nearly 80% of patients with RRMS develop SPMS within 20 years from the onset. In 50% of patients, the transition to SPMS occurs within 10 years after the first episode.

Over the past decade, several studies raised concerns about identifying the factors that account for the RRMS-to-SPMS transition. However, no uniform clinical, imaging, pathological or immunological criterion that reliably marks or predicts such a transition was described.[8] The SPMS diagnosis is based on retrospective analysis of the apparent increase in physical disability over the previous 6–12 months,[9] and the timing of conversion is an essential predictor of physical and cognitive dysfunction.[8 10] The exact time point of RMS-to-SPMS conversion can be missed due to the lack of clear diagnostic threshold criteria. Reports on the forecast of RRMS-to-SPMS conversion provide limited information on the predictive value of diagnostic findings received with MRI, molecular imaging and neurophysiological tests. For example, MS progression is known to correlate with the subarachnoid space enlargement due to parenchymal loss.[11 12] Other radiologic predictors for disease disability and conversion to SPMS refer to the number of cortical lesions, atrophied lesion volume, smouldering plaques (slowly expanding lesions) and spinal cord lesions.[12–15] Still, the prognostication of disease course is challenging.

Disparities in the previous results of studies make risk assessment of clinical and cognitive disabilities difficult. Studying clinical disability, authors reported conflicting findings about its correlation with radiological markers.[16] Some papers reported a clinical-radiological paradox which is a mismatch between clinical and radiological measures.[17 18] Contrarily, a recent article showed a strong association between the volume reduction in brain structures and the Expanded Disability Status Scale (EDSS).[19] Some studies on cognitive disability revealed that the clinical type of MS does not necessarily correlate with the cognitive dysfunction level. For example, Ntoskou *et al*[20] reported that patients with RRMS and SPMS showed similar results in cognitive tests on verbal learning, semantic fluency and processing speed. Individual variance in cognitive reserve may contribute to this phenomenon.

Diagnostic value of radiological, functional and genetic findings

MRI is a method of choice to support the clinical diagnosis of MS.[21] Other imaging and neurophysiological modalities can potentially assist in disease detection, differentiation and progress assessment. MRI is one of the components proposed in McDonald criteria for diagnosing MS. However, the application of MRI varies among different forms of the disease. Commonly, MRI is used to identify patients with the clinically isolated syndrome suggestive of the RRMS onset and patients with insidious neurological progression suspected for the primary progressive MS. The confirmation of the MS type is based on the T2 lesions count, the lesion distribution and dissemination in time or space.[22] However, other neurological diseases may also manifest with such lesions.

Numerous approaches were tested to decrease the number of false or misdiagnosed MS cases.

1. Studies evaluated the diagnostic value of MRI modalities. Patients with RRMS have acute demyelinating plaques and vasogenic oedema that can be identified with postcontrast T1-weighted (T1w) and diffusion-weighted imaging (DWI). Although DWI-MRI findings are consistent with the T1GAD-MRI sequence, Yousefi *et al* found contrast-enhanced imaging to be superior to DWI. The latter had 66.99% sensitivity (Sn) and 99.76% specificity (Sp) in detecting acute MS lesions from the total number of plaques in patients with active relapses when T1GAD was used as a standard.[23] Another critical identifier of MS is the paramagnetic rim at the edge of non-gadolinium-enhancing lesions, which are characteristic of an aggressive disease form. Three-dimensional echo-planar imaging detects the paramagnetic rim more accurately than T1w brain imaging with routine settings.[24]
2. The combined analysis of imaging modalities was used for an advanced MS lesion detection. For instance, Cetin *et al* compared the classification accuracy of different modalities in segmenting brain tissues with and without MS lesions in the same dataset. The study showed that the criteria based on a combination of T1w, T2 fluid-attenuated inversion recovery (FLAIR), and conventional T2 sequences identified MS lesions with the sensitivity of 90% reflecting the portion of only successfully classified MS lesions among all 'perceived' MS lesions and specificity of 65%.[25] In a combined analysis of FLAIR and T2 sequences, patients with MS (pwMS) were distinguished from those with small vessel disease

with 96%–100% sensitivity and over 80% specificity.[26] Joint evaluation of FLAIR and FLAIR* images modestly improved diagnostic accuracy for MS. In a study with healthy adults and patients with other neurological pathologies serving as controls, the detection of MS cases improved when the images were considered together (0.93 vs 0.98 area under the curve (AUC) averaged across different raters).[27]

3. Bioengineers developed a radiomics signature of MS from diffusion tensor imaging. It depicts nerve bundles and can differentiate patients with MS from healthy controls with 87% sensitivity and 91.7% specificity.[28] Radial diffusivity increases in response to demyelination, and axial diffusivity decreases with axonal damage.[29] Advanced diagnostics may result from postprocessing (image segmentation) and analysis of radiomics.
4. Another way to improve diagnostic accuracy is the modification of existing MRI protocols and scanners. Seven Tesla MRI scanners are superior in detecting chronic inflammation compared with the machines with a three Tesla magnification.[30]
5. The development of multiparametric quantitative (q) MRI enables radiologists to detect microstructural changes in tissue composition. Subtle or diffuse tissue desintegration due to gliosis, demyelination, axonal loss and infiltration of immune cells may occur then the conventional MRI appear normal. The qMRI could considerably improve follow-up studies of patients with MS by assessing tissue remodelling over time.[31]

Molecular imaging

Positron emission tomography (PET) detects neuroinflammation and successfully distinguishes between RRMS and SPMS.[32] PET is also helpful in differentiating MS lesions from gliomas.[33] However, the most commonly used radiotracer—fluorodeoxyglucose—is not efficient in brain PET studies since the glucose uptake is too high. Reasonably, researchers are looking for other markers of neuroinflammation, for example, translocator protein 18 (TSPO), cannabinoid and adenosine receptors, astrogliosis and sphingosine 1-phosphate receptors.[34]

Electroencephalography (EEG) has a potential to diagnose MS at an early onset since non-invasive EEG is used to evaluate the structural and functional connectivity. Hence, it can indicate disconnection among brain regions caused by the demyelination in MS. EEG detects an increase in slow frequencies and decrease in the alpha band in 40%–79% of patients with MS.[35] EEG with photic stimulation can distinguish patients with MS from healthy controls with 80% accuracy.[36] The data on the application of EEG for diagnosing MS are lacking, but the method holds promise as an adjuvant modality when assessing pwMS.

Evoked potentials are also used in MS diagnostics. Studies on motor evoked potential indicate that patients with MS show a prolonged latency, increased central motor conduction time and reduced signal amplitudes. The increase in central motor conduction time is more common than the prolongation of the silent period, yet all the reported findings reflect the clinical disability level.[37 38] The multifocal visual evoked potential (mfVEP) studies can also assess abnormalities that patients with MS exhibit in their visual field, for example, diminished intensity delayed nerve conduction velocity and wave cancellation.[39] A study compared the detection of optic neuritis in patients with MS with mfVEP, Humphrey visual field and optic coherence tomography (OCT). The optic neuritis history was determined by clinical signs and symptoms. Patients with MS without optic neuritis served as a control group. The research publication reported 89% sensitivity for detecting the damage to the optic nerve in MS cases with mfVEP, which is considerably higher than the sensitivity of OCT (62%) and Humphrey visual field assessment (72%).[40] The vestibular evoked myogenic potential studies detect brain stem dysfunction typical of MS. In a study with a cross-sectional design, the method discriminated between healthy controls and pwMS with the sensitivity reaching 70%.[41] Notably, results in vestibular-evoked myogenic potentials do not correlate with the defects detected with VEP.[41]

Molecular biology and genetic tests

Clinical diagnostics of MS can be complemented by analysing blood serum and cerebrospinal fluid because molecular markers are highly sensitive to neuroinflammation.[42] MicroRNAs (miRNAs) of serum exosomes are significantly dysregulated in patients with MS.[43] The deficiency in exosomal expression of specific miRNAs correlates with radiological and clinical signs of the acute phase of RRMS,[42] while other miRNAs demonstrate an increased expression in the primary progressive form of the disease.[44] The concentration of myeloid microvesicles in the cerebrospinal fluid also rises in patients with MS. The number of microvesicles reflects the number of enhancing lesions and predicts disability in RRMS and SPMS patients.[45] The intrathecal synthesis of oligoclonal IgG is considered to be the immunological hallmark of MS: oligoclonal IgG bands are associated with increased levels of disease activity and disability.[46] Worsening of the patient's condition is also associated with higher levels of neurofilament light (Nfl) in blood serum or plasma.[47] Nfl is a marker of neuronal injury in many neurodegenerative pathologies.[48] The elevated concentration of Nfl is commonly observed in patients with pronounced cognitive dysfunction.[49]

Diagnostics of cognitive impairment is also relevant to patients with MS as it is detected in 30%–60% of cases. It is a highly debatable question how to test the impairment with the lengthy batteries of neuropsychological tests: brief repeatable battery of neuropsychological tests and the minimal assessment of cognitive function in MS.[50] In order to cover nearly all cognitive domains, these batteries consist of 7–14 tests.[51] Such a comprehensive assessment seems to be abundant since MS affects mostly two domains: information processing speed and episodic

memory.[52] Slowed articulation rate is a reliable (91% sensitivity) discriminator between patients with MS with and without a decline in information processing speed as measured with the Symbol Digit Modalities Test and Paced Auditory Serial Addition Test-3.[53] The test results correlate with the articulation rate which is a marker of cognitive impairment.[53]

Multimodal examination seems to hold new promise to enhance diagnostic precision in medicine. A new diagnostic software confirms MS and other neurological diseases based on demographic and clinical features.[54] A clinical decision support system was shown to distinguish patients with RRMS from those with nine other pathologies (meningitis, cerebral palsy, migraine, cluster headache, stroke, epilepsy, Parkinson, Huntington and Alzheimer's disease). The system performance reached 99% accuracy and 100% sensitivity when clinical and non-clinical data were used as predictors. The clinical predictors included MS symptoms and signs (the number and the duration of clinical attacks), MRI data (lesion type, location, quantity), laboratory findings (the number of oligoclonal bands, the IgG index) and VEP measurements. The non-clinical predictors were age, gender, previous neurological symptoms, family medical history and a viral infection such as HIV.[55]

Prognosic potential of diagnostic data

Prediction of disease progression received special attention in the past decades. The approaches to forecast the disease course are as follows. First, the burden of cortical lesions may correlate with disease severity. Quantifying the severity of damage in lesions can help physicians to distinguish RRMS from SPMS. When fractional anisotropy is measured, diffusion tensor MRI discriminates between MS types with 85% sensitivity and 65% specificity. When the mean diffusivity is calculated, the performance drops to 62% sensitivity and 75% specificity.[56] Second, molecular imaging quantifies microglial activation which increases as the disease progresses.[57] Binding 11C-PK11195 tracer with TSPO is commonly used to detect microglial activation in the cortical grey matter of patients with MS.[58] A major advantage of TSPO-PET is the identification of diffuse inflammation around lesions[58] and the reflection of clinical disability.[32 59] TSPO-radioligand uptake or the distribution volume ratio of TSPO-PET is used in combination with other clinical and radiological variables to predict disease progression. However, models fed with these data have an insufficient sensitivity (52.9%–55%) and specificity of 95% for predicting progression in the entire MS cohort.[59 60] Third, NfL and the glial fibrillar acidic protein (GFAP) are candidates for MS-associated pathologies. The levels of these biomarkers in CSF correlate positively with the increase in neurological disability. NfL and GFAP categorise SPMS and RRMS patients with 54%–57% sensitivity and 84%–89% specificity.[61] Forth, neurophysiological biomarkers can discriminate clinical subtypes of MS. For example, abnormalities in somatosensory temporal discrimination threshold (STDT) and short intracortical inhibition (SICI) reflect neurodegenerative processes which play an important role in SPMS pathophysiology. Compared with SICI, STDT has a lower sensitivity (94.4% vs 58.8%) and higher specificity (67.9% vs 54.3%) in differentiating the MS subtypes.[62] The preliminary literature analysis showed a low classification accuracy of the discussed methods. To obtain conclusive evidence on the applicability of the tools for predicting MS conversion and progression, we aim to perform a systematic review and meta-analysis.

OBJECTIVES

We aim to analyse the advantages of the multimodal approach in predicting MS progression, specifically, in the RRMS-to-SPMS conversion. The objectives of this project will be as follows:

- Explore which settings of diagnostic methods correlate with the accuracy of MS identification. These settings may include the strength of the magnetic field, parameters of MRI scanning sequences (eg, T1, FLAIR, susceptibility-weighted imaging (SWI)), the type of MRI contrast and PET tracers, the injection time, the number of EEG electrodes and the miRNAs expression profiles.
- Rank the diagnostic methods for MS identification and progression prediction by sensitivity and specificity.
- Find the most reliable predictors for MS progression and disability level.
- Compare the predictive power of radiological findings and radiomics data as indicators of clinical disability and cognitive impairment in patients with MS.

METHODS AND ANALYSIS

To prepare the protocol, we followed the checklist of the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) checklist. The PRISMA-P checklist is available in online supplemental material file 1 .

Study design and data source

A comprehensive systematic review and meta-analysis will cover the literature on MS, its subtypes differentiation and monitoring of the disease progression. To perform the literature search, we will use five databases: Web of Science, Medline/PubMed, Scopus, Excerpta Medica Database Guide and CINAHL. We will extract papers written in English and published from January 1990 to December 2022. The keywords and medical subject headings will be as follows: MS, relapsing-remitting, secondary-progressive, progression, sensitivity and specificity, area under the receiver-operating characteristic (ROC) curve, mean absolute error (MAE). We will also include each type of diagnostic method into the search strings. The detailed search strategy is presented in online supplemental file 1. Our preliminary search indicated a greater number of studies devoted to MRI diagnostics rather than other methods. For this reason, the review will have

Table 1 Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria	
	For literature	For participants
General criteria		
1. Original peer-reviewed studies written in English and published from January 1990 to December 2022 2. In vivo studies 3. Small study cohort (8–500 patients with RRMS and/or SPMS) 4. Studies with a longitudinal and cross-sectional design 5. Female and male participants of any age 6. Individuals free from primary mental disorders, head injuries, on-MS related central nervous system pathologies	1. Grey literature 2. Editorial letters, reviews and protocol papers 3. Case studies 4. Studies that did not report sensitivity and specificity 5. Surgical interventional studies 6. Exposure of the participants to any factor that can potentially affect results 7. Nationwide studies and cohorts with over 500 patients with MS	1. Mental and psychological disorders (F00–F99 in ICD-10) 2. Cerebrovascular diseases (I60–I69) 3. Organic pathologies of the central nervous system (eg, brain and meninges tumours—C71, D32-33) 4. Head injuries (S00–S09)
Subobjectives 1–2		
7. Disease progression, cognitive impairment in MS	Same criteria as listed above	Same criteria as listed above
Subobjectives 3–4		
8. Scores on the expanded disability status scale disability status scale or MS severity score or age-related MS severity 9. Score in Mini-Mental State Examination or Brief Repeatable Neuropsychological Battery or Symbol Digit Modalities Test or the Minimal Assessment of Cognitive Function in MS	Same criteria as listed above	Same criteria as listed above

MS, multiple sclerosis; RRMS, relapsing-remitting MS; SPMS, secondary progressive MS.

a disparity in the number of analysed papers for each proposed method.

Eligibility criteria

The review will include in vivo MS studies with a cross-sectional and longitudinal design. We will consider the provided treatment for the meta-analysis and include interventional studies covering the disease progression. This study will only include papers published in peer-reviewed journals, and no grey literature will be considered. We will exclude protocol papers, editorial letters, reviews and case studies. The selected studies must report the sensitivity, specificity and accuracy of diagnostic modalities. Moreover, we will not consider papers which only reported accuracy. Furthermore, we will target scientific publications which reported data on men and women of any age, including paediatrics. Participants should be free from primary mental disorders, head injuries and central nervous system pathologies other than MS. Since we focus on the accuracy of predicting RRMS-to-SPMS conversion, the papers for review should compare the patients whose disease form progressed into the confirmed SPMS with those who sustained the relapsing-remitting disease course.

The literature inclusion and exclusion criteria are listed in table 1

Study records

Selection process

Two reviewers will conduct an initial search and screen the articles independently. They will review the titles and abstracts of the studies and select only those meeting the inclusion criteria. Then, they will review the full text to confirm the eligibility of the study. The selected papers will be uploaded to Covidence for automatic deduplication and blinded screening. Reproducible search strings for all databases will be appended to the review. The researchers will record the selection process and results according to the 2020 PRISMA statement. Furthermore, they will depict the selection process and outcomes in a PRISMA flow chart.

Data extraction

The research team will create an online form containing specified measures and study characteristics, which will be analysed in the review. This will include disease form, diagnostic modality, sample size and studied biomarkers. The necessary measures will include the performance metrics of predictive algorithms listed in subsection 3.1. From eligible papers, we will also extract data on EDSS, MSSS and ARMSS, and cognitive examinations mentioned in table 1. These scores will be used for the correlation of a disability level with radiological findings in pwMS. We will pay particular attention to the acquisition setting of

medical images (eg, the strength of the magnetic field) and the comparison groups/golden diagnostic standards in the reviewed studies. The extracted information will be grouped by such settings to allow adequate data analysis.

Quality assessment of individual studies

For the risk assessment, we will resort to the quality assessment tool for observational cohort and cross-sectional studies.[63] Two reviewers will use the assessment criteria and identify studies with the lowest risk of bias. In case of disputes, a third reviewer will decide if a study should be included. The research team will assess the risk of bias with the following criteria: sample size, gender of participants, diagnostic method, the strength of the magnetic field of the scanner (1.5, 3 Tesla or above), MRI scanning sequences (eg, T1w, FLAIR, SWI), type of studies (primary diagnostics or follow-up). To avoid selection bias between the papers, we will consider the studies conducted on relatively small cohorts (8–500 patients with RRMS and/or SPMS).

We will analyse (1) sensitivity, specificity and accuracy, AUC, positive predictive value, and negative predictive value in classification models predicting the risk of secondary progression and (2) accuracy of the regression models forecasting disability scores expressed as the ratio of MAE to the range of values. The publication bias will be assessed with a funnel plot created by plotting each estimate against the sample size as it was done in meta-analytical studies on diagnostic accuracy by Gong *et al* and Qu *et al*.[64 65] The plots will be constructed with *metafor* package for meta-analysis in R.[66] The package has a programme implementation of the 'trim and fill' method which allows us to calculate the number of studies needed for constructing a symmetric funnel plot.[67] Researchers can use disparate thresholds of disability scales to confirm the MS progression. To overcome this reporting bias and construct an appropriate summary ROC curve, we will use the Steinhausen random effect model.[68] The model allows to determine optimal cut-offs in the meta-analysis of the diagnostic and prognostic test accuracy. The model is implemented in diagmeta R package.[69] To do the calculation, we will collect true positive, false positive, true negative, false negative from eligible articles.[70] If these parameters are not reported, we will request the details from the corresponding author of the particular paper.

Data analysis and synthesis

Our data analysis will follow the objectives of the study (see figure 1 for the study pipeline).

We will review the diagnostic and prognostic power of methods for detecting MS and describe radiological correlates of the disease severity (the first study objective). Hence, we will focus on MRI, PET, electrophysiological methods, cognitive assessment and molecular lab tests. We will look into genetic and epigenetic markers. For the comparison, we will review the metrics of success specified in 'Quality assessment of individual studies' subsection. These variables will be examined directly with the methods described below.

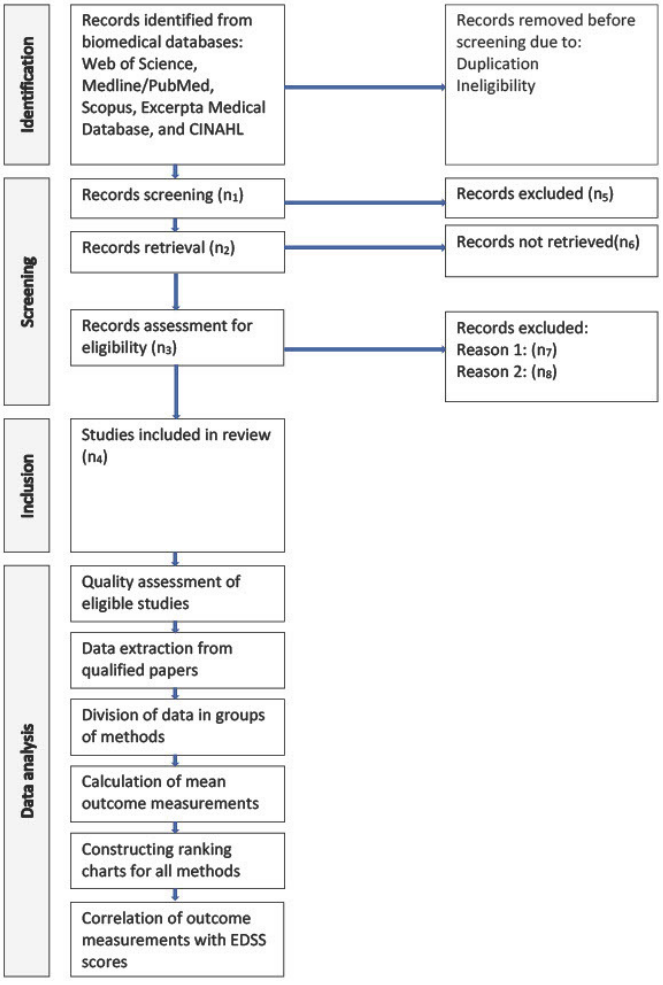

Figure 1 Study pipeline. EDSS, Expanded Disability Status Scale.

Once the data are extracted into a predefined workbook, we will group them by the diagnostic method. To generalise the results beyond the included studies, we will use the random-effects model while conducting the meta-analysis. For the analysis, we expect to receive enough studies (over 5) per each diagnostic method. We will evaluate the normality assumption of all the collected findings (Sn, Sp, ACC, AUC, MAE, MAE/range) with Shapiro-Wilk test.[71] Commonly, the results of diagnostic accuracy studies are distributed non-normally. If this is the case, we will use the bivariate generalised linear mixed model function from metafor R package to avoid the unnecessary normality assumption within studies.[72] The model will be also employed to calculate the true positives and true negatives.

The simultaneous consideration of a set of statistical inferences can lead to the multiple testing problem. To resolve the problem, we will apply multiple comparison corrections. For example, we will use the Bonferroni correction that is the best-known solution for making statistical tests more stringent.[73] We will divide the critical

p value (0.05) by the number of pairwise comparisons being made on the dataset. The modified p value will be used to assess the statistical power of the study. In our analysis, we will consider two outcome measures ('RRMS-to-SPMS conversion' and 'disability progression') and use n variables reflecting results in diagnostic tests: two outcome measures will be measured against *n* hypothesised predictors. A Bonferroni adjusted significance level of alpha will be calculated to account for increased possibility of false-positive results. We will limit the number of tested hypothesis to a maximum of 10, otherwise the risk of false-negative results will increase.

Multiple testing results in the between-study heterogeneity which will be assessed with the Higgins-Thompson I^2 test.[74] We expect years lived with MS and EDSS score to be the sources of heterogeneity in the meta-analysis. The calculation of I^2 statistics will be done with dmetar R package.[75] I^2 values of 75% and above signal a high level of variability among the results of individual studies. If this is the case, we will resort to a narrative systematic review instead of the meta-analysis. To avoid the heterogeneity due to setting variance of diagnostic tests, we will conduct the subgroup analysis with metafor package in R.[76]

If the distribution of variables is normal and I^2 value is below 75%, we will model the sensitivity and specificity values with the bivariate linear mixed model implemented in meta.dt R package.[77] Calculation of the pooled performance metric will be made to assess summary performance metrics for each diagnostic method. We will also construct a hierarchical summary ROC curve for the prediction of the disease progression. The analysis will be conducted with the mada R package which is a common tool for meta-analyses of the diagnostic/prognostic power.[78]

To rank the diagnostic methods for MS identification and progression prediction (the second study objective), we will create forest plots and summary ROC space presenting performance of the algorithms trained on various diagnostic findings separately and in combination.[79] In this subobjective, the performance metrics remain the same as in the previous one. The difference in performance will be confirmed by a significance level ≤0.05. We will adopt standard approaches to compare the distribution of outcome measurements among different diagnostic procedures used to predict the MS progression. Mada package will be used to accomplish these tasks.[78]

In objectives 3 and 4, we will study associations between distinct markers of MS progression and the disability level (eg, scores on EDSS, MS Severity Score and Age-Related MS Severity). This part of the analysis will be carried out on papers reporting the aforementioned scores for studied cohorts.

Review status

The review started in October 2022 and it will be completed in March 2024.

Potential amendments

We predefined the inclusion and exclusion criteria and conducted a preliminary search to avoid possible amendments. However, any necessary changes during the review preparation will be reported by updating the online registered PROSPERO protocol.

Patients and public involvement

The study does not involve patients or members of the public.

ETHICS AND DISSEMINATION

The systematic review does not require an ethical approval. The study findings will be published in a peer-reviewed journal and presented as a poster or presentation at scientific conferences.

Author affiliations
[1]Radiology, United Arab Emirates University College of Medicine and Health Sciences, Al Ain, Abu Dhabi Emirate, UAE
[2]College of Natural & Health Sciences, Zayed University, Dubai, UAE
[3]National Medical Library, United Arab Emirates University College of Medicine and Health Sciences, Al Ain, Abu Dhabi, UAE
[4]Department of Computer Science, UAE University College of Information Technology, Al Ain, Abu Dhabi, UAE
[5]University of Constantine 3, Constantine, Algeria
[6]Physiology, UAE University College of Medicine and Health Sciences, Al Ain, Abu Dhabi, UAE
[7]Biomedical Engineering Department, Wayne State University College of Engineering, Detroit, Michigan, USA
[8]Mahidol University Faculty of Medicine Siriraj Hospital, Bangkok, Thailand
[9]United Arab Emirates University, Al Ain, UAE
[10]Neurology, Mediclinic Parkview Hospital, Dubai, UAE
[11]Mohammed Bin Rashid University Of Medicine and Health Sciences, Dubai, UAE
[12]Pediatrics, United Arab Emirates University College of Medicine and Health Sciences, Al Ain, Abu Dhabi, UAE

Twitter Yauhen Statsenko @StatsenkoE

Contributors The authors' contributions are as follows: YS, JG, ML, FI, MS, KMD, KN-VG, TA and FAZ specified the research questions and the study design. LÖ, YS and TMA prepared the protocol draft. LÖ will perform the literature search. DS, GLS, SM, TT, FCK and IM will perform the abstract, title and full-text screening. Statistics expert TH will perform the data analysis and synthesis.

Funding This work was supported by Aspire grant AARE19-060 (PI: ML), UAEU StartUp grant 12M045 (PI: YS).

Competing interests The authors declare that the research will be conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Yauhen Statsenko http://orcid.org/0000-0002-7713-3333
Darya Smetanina http://orcid.org/0000-0002-4462-8823

REFERENCES

1 Walton C, King R, Rechtman L, *et al*. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Mult Scler* 2020;26:1816–21.
2 Federation MI. *Atlas of ms 3rd edition*. 2020.
3 Benjak T, Štefančić V, Draušnik Ž, *et al*. Prevalence of multiple sclerosis in Croatia: Data from national and non-governmental organization registries. *Croat Med J* 2018;59:65–70.
4 Al-Sakran LH, Marrie RA, Blackburn DF, *et al*. Establishing the incidence and prevalence of multiple sclerosis in Saskatchewan. *Can J Neurol Sci* 2018;45:295–303.
5 Hosseinzadeh A, Baneshi MR, Sedighi B, *et al*. Incidence of multiple sclerosis in Iran: A nationwide, population-based study. *Public Health* 2019;175:138–44.
6 Pirttisalo A-L, Soilu-Hänninen M, Sipilä JOT. Multiple sclerosis epidemiology in Finland: Regional differences and high incidence. *Acta Neurol Scand* 2019;139:353–9.
7 Urru SA, Antonelli A, Sechi GM, *et al*. Prevalence of multiple sclerosis in Sardinia: A systematic cross-sectional multi-source survey. *Mult Scler* 2020;26:372–80.
8 Inojosa H, Proschmann U, Akgün K, *et al*. A focus on secondary progressive multiple sclerosis (Spms): Challenges in diagnosis and definition. *J Neurol* 2021;268:1210–21.
9 Krajnc N, Bsteh G, Berger T. Clinical and Paraclinical biomarkers and the Hitches to assess conversion to secondary progressive multiple sclerosis: A systematic review. *Front Neurol* 2021;12.
10 Horáková D, Boster A, Bertolotto A, *et al*. Proportion of Alemtuzumab-treated patients converting from Relapsing-Remitting multiple sclerosis to secondary progressive multiple sclerosis over 6 years. *Mult Scler J Exp Transl Clin* 2020;6.
11 Chiang GC. Atrophied T2 lesion volume in multiple sclerosis: Reducing the focus on lesion counting. *Radiology* 2019;293:434–5.
12 Genovese AV, Hagemeier J, Bergsland N, *et al*. Atrophied brain T2 lesion volume at MRI is associated with disability progression and conversion to secondary progressive multiple sclerosis. *Radiology* 2019;293:424–33.
13 Scalfari A, Romualdi C, Nicholas RS, *et al*. The cortical damage, early Relapses, and onset of the progressive phase in multiple sclerosis. *Neurology* 2018;90:e2107–18.
14 Calabrese M, Favaretto A, Martini V, *et al*. Grey matter lesions in MS: From histology to clinical implications. *Prion* 2013;7:20–7.
15 Tavazzi E, Zivadinov R, Dwyer MG, *et al*. Mri biomarkers of disease progression and conversion to secondary-progressive multiple sclerosis. *Expert Rev Neurother* 2020;20:821–34.
16 Wattjes MP, Rovira À, Miller D, *et al*. Magnims consensus guidelines on the use of MRI in multiple sclerosis–establishing disease prognosis and monitoring patients. *Nat Rev Neurol* 2015;11:597–606.
17 Chard D, Trip SA. Resolving the Clinico-radiological paradox in multiple sclerosis. *F1000Res* 2017;6:1828.
18 Barkhof F. The Clinico-radiological paradox in multiple sclerosis Revisited. *Curr Opin Neurol* 2002;15:239–45.
19 Aoki S, Fujimori J, Mikami R, *et al*. Assisting the diagnosis of multiple sclerosis using a set of regional brain volumes: A classification model for patients and healthy controls. *Informatics in Medicine Unlocked* 2021;26:100766.
20 Ntoskou K, Messinis L, Nasios G, *et al*. Cognitive and language deficits in multiple sclerosis: Comparison of Relapsing Remitting and secondary progressive subtypes. *Open Neurol J* 2018;12:19–30.
21 Filippi M, Rocca MA. MR imaging of multiple sclerosis. *Radiology* 2011;259:659–81.
22 McNicholas N, Hutchinson M, McGuigan C, *et al*. Mcdonald diagnostic criteria: A review of the evidence. *Mult Scler Relat Disord* 2018;24:48–54.
23 Yousefi M, Panahali M, Azarkhish K, *et al*. Diagnostic value of DWI-MRI for the detection of acute plaques in the relapse phase of multiple sclerosis. *Ro J Neurol* 2021;20:35–40. 10.37897/RJN.2021.1.5 Available: https://rjn.com.ro/rjn-vol-xx-no-1-year-2021/
24 Barquero G, La Rosa F, Kebiri H, *et al*. Rimnet: A deep 3D Multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis. *Neuroimage Clin* 2020;28.
25 Cetin O, Seymen V, Sakoglu U. Multiple sclerosis lesion detection in Multimodal MRI using simple clustering-based Segmentation and classification. *Informatics in Medicine Unlocked* 2020;20.
26 Campion T, Smith RJP, Altmann DR, *et al*. FLAIR* to visualize veins in white matter lesions: A new tool for the diagnosis of multiple sclerosis? *Eur Radiol* 2017;27:4257–63.
27 George IC, Sati P, Absinta M, *et al*. Clinical 3-Tesla flair* MRI improves diagnostic accuracy in multiple sclerosis. *Mult Scler* 2016;22:1578–86.
28 Hesseltine SM, Law M, Babb J, *et al*. Diffusion Tensor imaging in multiple sclerosis: Assessment of regional differences in the axial plane within normal-appearing Cervical spinal cord. *AJNR Am J Neuroradiol* 2006;27:1189–93.
29 Zurita M, Montalba C, Labbé T, *et al*. Characterization of Relapsing-Remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion MRI data. *Neuroimage Clin* 2018;20:724–30.
30 La Rosa F, Barquero G, Al-Louzi O, *et al*. Automated assessment of paramagnetic rim lesions in multiple sclerosis patients with 3t and 7t Mp2Rage. Proceedings of the 2021 ISMRM & SMRT Annual Meeting & Eshibition. An Online Experience. (ISMRM); 2021
31 Seiler A, Nöth U, Hok P, *et al*. Multiparametric quantitative MRI in neurological diseases. *Front Neurol* 2021;12:640239.
32 Politis M, Giannetti P, Su P, *et al*. Increased Pk11195 PET binding in the cortex of patients with MS correlates with disability. *Neurology* 2012;79:523–30.
33 Kebir S, Rauschenbach L, Weber M, *et al*. Machine learning-based differentiation between multiple sclerosis and glioma WHO II°-IV° using O-(2-[18F] Fluoroethyl)-L-tyrosine positron emission tomography. *J Neurooncol* 2021;152:325–32.
34 Bauckneht M, Capitanio S, Raffa S, *et al*. Molecular imaging of multiple sclerosis: From the clinical demand to novel Radiotracers. *EJNMMI Radiopharm Chem* 2019;4:6.
35 Leocani L, Comi G. Neurophysiological investigations in multiple sclerosis. *Curr Opin Neurol* 2000;13:255–61.
36 Karaca BK, Akşahin MF, Öcal R. Detection of multiple sclerosis from Photic stimulation EEG signals. *Biomedical Signal Processing and Control* 2021;67.
37 Tataroglu C, Genc A, Idiman E, *et al*. Cortical silent period and motor evoked potentials in patients with multiple sclerosis. *Clin Neurol Neurosurg* 2003;105:105–10.
38 Kale N, Agaoglu J, Onder G, *et al*. Correlation between disability and transcranial magnetic stimulation abnormalities in patients with multiple sclerosis. *J Clin Neurosci* 2009;16:1439–42.
39 de Santiago L, Sánchez Morla EM, Ortiz M, *et al*. A computer-aided diagnosis of multiple sclerosis based on mfVEP recordings. *PLoS One* 2019;14.
40 Laron M, Cheng H, Zhang B, *et al*. Comparison of multifocal visual evoked potential, standard automated Perimetry and optical coherence tomography in assessing visual pathway in multiple sclerosis patients. *Mult Scler* 2010;16:412–26.
41 Harirchian MH, Karimi N, Nafisi S, *et al*. Vestibular evoked Myogenic potential for diagnoses of multiple sclerosis: Is it beneficial? *Med Glas (Zenica)* 2013;10:321–6.
42 Selmaj I, Cichalewska M, Namiecinska M, *et al*. Global Exosome Transcriptome profiling reveals biomarkers for multiple sclerosis. *Ann Neurol* 2017;81:703–17.
43 Ebrahimkhani S, Vafaee F, Young PE, *et al*. Exosomal Microrna signatures in multiple sclerosis reflect disease status. *Sci Rep* 2017;7:14293.
44 Vistbakka J, Elovaara I, Lehtimäki T, *et al*. Circulating Micrornas as biomarkers in Progressive multiple sclerosis. *Mult Scler* 2017;23:403–12.
45 Gelibter S, Pisa M, Croese T, *et al*. Spinal fluid myeloid Microvesicles predict disease course in multiple sclerosis. *Ann Neurol* 2021;90:253–65.
46 Graner M, Pointon T, Manton S, *et al*. Oligoclonal IgG antibodies in multiple sclerosis target patient-specific peptides. *PLoS One* 2020;15.
47 Williams T, Tur C, Eshaghi A, *et al*. Serum Neurofilament light and MRI predictors of cognitive decline in patients with secondary progressive multiple sclerosis: Analysis from the MS-STAT randomised controlled trial. *Mult Scler* 2022;28:1913–26.
48 Kapoor R, Smith KE, Allegretta M, *et al*. Serum Neurofilament light as a biomarker in Progressive multiple sclerosis. *Neurology* 2020;95:436–44.
49 Gaetani L, Schoonheim MM. Serum Neurofilament light chain predicts cognitive worsening in secondary progressive multiple sclerosis better than brain MRI measures. *Mult Scler* 2022;28:1831–3.

50 DeLuca GC, Yates RL, Beale H, *et al*. Cognitive impairment in multiple sclerosis: Clinical, Radiologic and pathologic insights. *Brain Pathol* 2015;25:79–98. 10.1111/bpa.12220 Available: http://doi.wiley.com/10.1111/bpa.2015.25.issue-1
51 Arnett PA, Strober LB. Cognitive and neurobehavioral features in multiple sclerosis. *Expert Rev Neurother* 2011;11:411–24.
52 Akbar N, Honarmand K, Kou N, *et al*. Validity of a computerized version of the symbol digit modalities test in multiple sclerosis. *J Neurol* 2011;258:373–9.
53 Friedova L, Rusz J, Motyl J, *et al*. Slowed articulation rate is associated with information processing speed decline in multiple sclerosis: A pilot study. *J Clin Neurosci* 2019;65:28–33.
54 Singla J, Grover D, Bhandari A. Medical expert systems for diagnosis of various diseases. *IJCA* 2014;93:36–43.
55 Hosseini A, Asadi F, Arani LA. Development of a knowledge-based clinical decision support system for multiple sclerosis diagnosis. *J Med Life* 2020;13:612–23.
56 Filippi M, Preziosa P, Pagani E, *et al*. Microstructural magnetic resonance imaging of cortical lesions in multiple sclerosis. *Mult Scler* 2013;19:418–26.
57 Airas L, Rissanen E, Rinne J. Imaging of Microglial activation in MS using PET: Research use and potential future clinical application. *Mult Scler* 2017;23:496–504.
58 Rissanen E, Tuisku J, Rokka J, *et al*. In vivo detection of diffuse inflammation in secondary progressive multiple sclerosis using PET imaging and the Radioligand ^{11}C-Pk11195 . *J Nucl Med* 2014;55:939–44.
59 Sucksdorff M, Matilainen M, Tuisku J, *et al*. Brain TSPO-PET predicts later disease progression independent of Relapses in multiple sclerosis. *Brain* 2020;143:3318–30.
60 Misin O, Matilainen M, Nylund M, *et al*. Innate immune cell–related pathology in the thalamus signals a risk for disability progression in multiple sclerosis. *Neurol Neuroimmunol Neuroinflamm* 2022;9.
61 Högel H, Rissanen E, Barro C, *et al*. Serum glial fibrillary acidic protein correlates with multiple sclerosis disease severity. *Mult Scler* 2020;26:210–9.
62 Belvisi D, Tartaglia M, Borriello G, *et al*. Are neurophysiological biomarkers able to discriminate multiple sclerosis clinical subtypes? *Biomedicines* 2022;10.
63 National Heart L, Institute B. *Study quality assessment tools*. 2013.
64 Gong X, Xu Q, Xu Z, *et al*. Real-time Elastography for the differentiation of benign and malignant breast lesions: A meta-analysis [breast cancer research and treatment 130]. *Breast Cancer Res Treat* 2011;130:11–8.
65 Qu X, Huang X, Wu L, *et al*. Comparison of virtual Cystoscopy and Ultrasonography for bladder cancer detection: A meta-analysis. *Eur J Radiol* 2011;80:188–97.
66 Rdrrio. Metafor: Meta-analysis package for R. 2022.
67 Github. Metafor: Meta-analysis package for R: Funnel plots. 2022.
68 Steinhauser S, Schumacher M, Rücker G. Determining optimal cut-offs in the meta-analysis of diagnostic test accuracy studies [Ph.D. thesis]. Universität, 2016
69 rdrrio. Diagmeta: Meta-analysis of diagnostic accuracy studies with several Cutpoints. 2022.
70 Steinhauser S, Schumacher M, Rücker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. *BMC Med Res Methodol* 2016;16:97.
71 Shapiro SS, Wilk MB. An analysis of variance test for Normality (complete samples). *Biometrika* 1965;52:591–611.
72 ByteCompile T, LazyData T. Package 'Metafor.
73 Weisstein EW. Bonferroni correction. 2004. Available: https://mathworld. wolfram. com/
74 Higgins JPT, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
75 Harrer M, Cuijpers P, Furukawa TA, *et al*. Doing meta-analysis with R: A hands-on guide [CRC press]. 2021.
76 Metafor project. Forest plot with subgroups. 2022.
77 Documentation R. Meta-analysis of diagnostic tests. 2022.
78 Doebler P, Holling H. Meta-analysis of diagnostic accuracy with Mada. *R Packag* 2015;1:15.
79 Arends LR, Hamza TH, van Houwelingen JC, *et al*. Bivariate random effects meta-analysis of Roc curves. *Med Decis Making* 2008;28:621–38.