Missouri University of Science and Technology
Scholars' Mine

Electrical and Computer Engineering Faculty Research & Creative Works

Electrical and Computer Engineering

11 Jan 2023

Subtypes of Relapsing-Remitting Multiple Sclerosis Identified by Network Analysis

Quentin Howlett-Prieto

Chelsea Oommen

Michael D. Carrithers

Donald C. Wunsch
Missouri University of Science and Technology, dwunsch@mst.edu

et. al. For a complete list of authors, see https://scholarsmine.mst.edu/ele_comeng_facwork/4619

Follow this and additional works at: https://scholarsmine.mst.edu/ele_comeng_facwork

Part of the Electrical and Computer Engineering Commons

Recommended Citation

Q. Howlett-Prieto et al., "Subtypes of Relapsing-Remitting Multiple Sclerosis Identified by Network Analysis," *Frontiers in Digital Health*, vol. 4, article no. 1063264, Frontiers Media, Jan 2023.
The definitive version is available at https://doi.org/10.3389/fdgth.2022.1063264

This work is licensed under a Creative Commons Attribution 4.0 License.

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Electrical and Computer Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

frontiers | Frontiers in Digital Health

TYPE Brief Research Report
PUBLISHED 11 January 2023
DOI 10.3389/fdgth.2022.1063264

Check for updates

OPEN ACCESS

EDITED BY
Lina F. Soualmia,
Université de Rouen, France

REVIEWED BY
Leif Simmatis,
University Health Network (UHN), Canada,
A. N. M. Bazlur Rashid,
Edith Cowan University, Australia

*CORRESPONDENCE
Daniel B. Hier
✉ hierd@mst.edu

SPECIALTY SECTION
This article was submitted to Health Informatics, a section of the journal Frontiers in Digital Health

RECEIVED 06 October 2022
ACCEPTED 22 December 2022
PUBLISHED 11 January 2023

CITATION
Howlett-Prieto Q, Oommen C, Carrithers MD, Wunsch II DC and Hier DB (2023) Subtypes of relapsing-remitting multiple sclerosis identified by network analysis.
Front. Digit. Health 4:1063264.
doi: 10.3389/fdgth.2022.1063264

COPYRIGHT
© 2023 Howlett-Prieto, Oommen, Carrithers, Wunsch II and Hier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Subtypes of relapsing-remitting multiple sclerosis identified by network analysis

Quentin Howlett-Prieto[1], Chelsea Oommen[1], Michael D. Carrithers[1], Donald C. Wunsch II[2] and Daniel B. Hier[1,2]*

[1]Department of Neurology and Rehabilitation, University of Illinois at Chicago, Chicago, IL, United States, [2]Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO, United States

We used network analysis to identify subtypes of relapsing-remitting multiple sclerosis subjects based on their cumulative signs and symptoms. The electronic medical records of 113 subjects with relapsing-remitting multiple sclerosis were reviewed, signs and symptoms were mapped to classes in a neuro-ontology, and classes were collapsed into sixteen superclasses by subsumption. After normalization and vectorization of the data, bipartite (subject-feature) and unipartite (subject-subject) network graphs were created using NetworkX and visualized in Gephi. Degree and weighted degree were calculated for each node. Graphs were partitioned into communities using the modularity score. Feature maps visualized differences in features by community. Network analysis of the unipartite graph yielded a higher modularity score (0.49) than the bipartite graph (0.25). The bipartite network was partitioned into five communities which were named fatigue, behavioral, hypertonia/weakness, abnormal gait/sphincter, and sensory, based on feature characteristics. The unipartite network was partitioned into five communities which were named fatigue, pain, cognitive, sensory, and gait/weakness/hypertonia based on features. Although we did not identify pure subtypes (e.g., pure motor, pure sensory, etc.) in this cohort of multiple sclerosis subjects, we demonstrated that network analysis could partition these subjects into different subtype communities. Larger datasets and additional partitioning algorithms are needed to confirm these findings and elucidate their significance. This study contributes to the literature investigating subtypes of multiple sclerosis by combining feature reduction by subsumption with network analysis.

KEYWORDS

multiple sclerosis, phenotype, network analysis, communities, modularity, subtype, feature reduction, subsumption

Introduction

Multiple sclerosis (MS) is one of several immune-mediated demyelinating diseases of the central nervous system that includes transverse myelitis, optic neuritis, neuromyelitis optical, acute disseminated encephalomyelitis, and acute hemorrhagic leukoencephalopathy (1). MS has traditionally been divided into four clinical course phenotypes that include relapsing-remitting multiple sclerosis (RRMS), primary progressive multiple sclerosis (PPMS), secondary progressive multiple sclerosis (SPMS), and relapsing progressive multiple sclerosis

(RPMS) (2). In 2013, the criteria for MS phenotypes were revised to remove RPMS (3–6). More recently, MS has been additionally classified by activity (active or inactive) or phase (relapsing or progressive) (4, 5, 7, 8). Another way of subtyping neurologic diseases is by deep phenotyping where signs and symptoms are recorded in detail and are mapped to a restricted terminology (9–13). Patients can then be subtyped according to patterns of signs and symptoms.

MS may have a variable onset with the diverse symptoms of optic neuritis, facial pain, hemifacial spasm, Lhermitte's sign, transverse myelitis, limb weakness, limb numbness, urinary retention, dysmetria, intention tremor, incoordination, dysarthria, hearing loss, color blindness, gait disturbance, and diplopia (14). MS variably involves the optic nerve (painful loss of vision), the spinal cord (sphincter dysfunction, monoparesis, hemiparesis, hypoesthesia, paresthesia), the brainstem and cerebellum (diplopia, oscillopsia, vertigo, ataxia, tremor, facial weakness), or the cerebral hemispheres (hemiparesis, hemihypoesthesia) (8). Subtypes of multiple sclerosis based on clinical presentation (signs and symptoms) are recognized (15–18) including tremor (19), ataxia (20), visual disturbances (21, 22), sensory symptoms (numbness and paresthesias) (23–26), pyramidal tract findings (weakness, hyperreflexia, spasticity, and hypertonia) (27–29), or spinal cord findings (paraparesis, sphincter dysfunction, and sensory levels) (30, 31). Other MS subjects show cognitive impairment (32, 33), dysarthria (34), dysautonomia (35), depression (36), imbalance (37), paroxysmal symtoms (38), or fatigue (39, 40).

The Kurtzke Functional System Score (FSS) (41) is useful in rating sensory, visual, sphincter, mental, pyramidal, cerebellar, and brainstem dysfunction in MS. However, there is a limited ability to categorize MS subjects based on their predominant clinical presentation. A network analysis of subjects with MS based on their signs and symptoms could assist in identifying clinically significant subtypes of MS.

This paper is organized as follows. We first review prior work on finding subtypes of multiple sclerosis based on signs and symptoms. We then describe our proposed approach to finding subtypes of multiple sclerosis based on deep phenotyping, subsumption of phenotype classes into superclasses, and network analysis. In the Methods section, we describe how deep phenotyping was performed, how the features were collapsed into superclasses, and how the networks were created and partitioned. In the Results section, we report the partitioning of the networks into five communities of MS subjects. In the Discussion section, we discuss the identified communities as possible clinical subtypes of MS. Finally, we discuss the limitations of network analysis as a method of finding MS subtypes.

Prior work

Although network analysis has not been used to identify clinical subtypes of MS, other work is relevant to this undertaking (**Table 1**). Depression and anxiety have been reported in MS in about 27% of patients, but no specific phenotype has been described (42, 43, 62). Zhang et al. (63) examined 13 common symptoms of MS in 1985 MS subjects and found that depression, pain, and walking difficulties were the strongest predictors of impaired quality of life. Cognitive impairment is frequent in MS, possibly affecting 40–80% of subjects (33, 45–50). Pure cognitive subtypes (cognitive impairment without other major neurological signs) have been described in a small minority of MS patients (47, 50). In their review of functional connectivity based on functional MRI, Tahedl et al. (64) suggested that cognitive impairment in MS was associated with disruptions of the default-mode network of the brain, whereas sensory-motor deficits were associated with disruptions of the sensory-motor networks of the brain. Although fatigue is frequent in MS, specific subtypes have not been described (51).

Although uncommon, spinal MS (leg weakness, sphincter dysfunction, sensory levels, spasticity, and hyperreflexia), as well as opticospinal MS (combining spinal MS with optic nerve involvement), are recognized forms of MS (65–67). Opticospinal MS must be differentiated from neuromyelitis optica, a similar but etiologically different disease from MS. Cree et al. (60) have suggested that spinal MS and opticospinal MS may be more common in Blacks than Whites. Nociti et al. (30) reported spinal MS in 2.3% of their cohort of subjects.

Cerebellar and brainstem phenotypes of MS have been reported (44) with prominent ataxia and cranial nerve deficits. Naismith et al. (17) compared 79 Black subjects with MS to 80 White subjects (17) and found more tremor, ataxia, and need for assistive walking devices in the Black MS subjects. They speculated that the optico-spinal, cognitive, and ataxic-spastic phenotypes are more common in Black than White subjects. In a small study, Ayache et al. (19) found tremor in 56% of their cohort of MS subjects but did not identify a specific phenotype.

Sensory symptoms are common in MS, including pain, hypesthesias hyperesthesias, band-like sensations, and paresthesias (26, 56); however, no specific sensory phenotype has been described. Optic neuritis is common in MS but generally recovers fully or partially. Gerbis et al. (61) describe 5 subjects from a cohort of 550 MS who had severe unilateral optic neuritis without recovery, and suggest that these cases may represent a subtype of MS subjects.

Functional Systems Scores (FSS) are a good candidate for identifying clinical subtypes of MS. It is widely used in MS clinical trials and is divided into seven domains (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, and cerebral) (68). An asymmetric distribution of scores in these domains could identify subtypes of MS. Yang et al. (69) used a combination of a convolutional neural network and a rule-based natural language algorithm to accurately predict

TABLE 1 Summarization of prior work relevant to subtyping multiple sclerosis by phenotypic feature.

Author	Domain	*N*	Cite	Findings	Limitations
Donnchadha	Behavioral	33	(42)	27% had anxiety	Small N, no correlations with other features
Koch	Behavior	1,376	(43)	27% depressed, stable over time	No correlation with EDSS
Ganesvaran	Brainstem or cerebellum	20	(44)	80% of 20 MS patients with psoriasis 80% had brainstem or cerebellar lesions	Small N
Naismith	Cerebellar and gait	166	(17)	Compared to Whites, Blacks had more gait and cerebellar deficits	Small N
Rocca	Cognitive	227	(45)	Loss of connectivity predicts memory and attention deficits	Correlation with cognitive impairment
Hancock	Cognitive	1,281	(46)	Cognitive domain impairment: 48% intact, 22% 1-domain, 24% 2-domain, 15% multi-domain cognitive deficits	Only examined cognitive impairment
De Meo	Cognitive	1,212	(33)	5 cognitive subtypes identified Only 19.5% completely normal	Only examined cognitive impairment
Staff	Cognitive	23	(47)	Mayo Clinic reported 23 MS patients with isolated cognitive impairment	Population not reported
Leavitt	Cognitive	128	(48)	43.7% cognitively impaired Memory, Processing speed, or both	Convenience sample
Cabeça	Cognitive	35	(49)	Discriminant analysis found reaction time best discriminator	Study did not identify cognitive subtypes
Zurawski	Cognitive	2,302	(50)	2.6% had pure cognitive phenotype	Only examined cognition
Beckerman	Fatigue	264	(51)	88 with low and 174 with high fatigue physical from mental fatigue correlated	Did not correlate fatigue with other features
Bove	FSS[a]	1,028	(52)	Median and range FSS provided	No classification by subtype
Kalincik	FSS[a]	14,969	(53)	Increased disability with relapse on on pyramidal, cerebellar, sphincter FSS	Not classified by phenotype
Stewart	FSS[a]	19,504	(54)	Pyramidal, cerebellar, sphincter add to disability with relapse	Not classified by phenotype
Scott	FSS[a]	1,173	(55)	On followup, most worsening on pyramidal, sensory, cerebellar, and sphincter FSS scales	Not classified by phenotype
Revil	Pain	112	(56)	40 pain free with normal sensation 44 central pain with hyposensitivity	Only examined pain
Tsantes	Relapse phenotype	199	(57)	47% of relapses recurred at initial optic, spinal, brainstem-cerebellum sites	Small N
Mowry	Relapse phenotype	195	(58)	Relapse more likely to recur at optic nerve spinal cord, or brainstem-cerebellum	Small N
Deen	Relapse phenotype	190	(59)	Relapse recurred at optic nerve brainstem-cerebellum, spinal cord	Small N
Nociti	Spinal cord	563	(30)	13/563 had spinal MS (2.3%)	Retrospective study
Sanders	Sensory	127	(26)	84% had paresthesias	Small N
Cree	Spinal cord	1490	(60)	Compared to White, Black MS subjects more corticospinal and transverse myelitis	
Ayache	Tremor	32	(19)	56% with tremor	Not population-based
Gerbis	Vision	550	(61)	5 of 550 patients had severe optic neuritis that never recovered	Only examined optic neuritis

Author is the first author, N is the number of subjects in the study, cite is the reference number. Studies are sorted by domain.
[a]Functional system score.

Kurtzke Functional System Scores (FSS) from the EHR notes of 4906 multiple sclerosis subjects. SUMMIT (Serially Unified Multicenter Multiple Sclerosis Investigation) is an international effort to create a repository of deeply phenotyped MS subjects utilizing standardized neurological examinations and the Kurtzke FSS (12). However, no subtypes based on FSS have been reported. Similarly, Dahlke et al. (70) examined the clinical course in 34,987 MS patients who had

entered into clinical trials (31,863 with relapsing-remitting MS, 1873 with secondarily progressive MS, and 986 with primary progressive MS) but did not characterize MS subjects further as to clinical phenotype. Other ongoing longitudinal studies have been undertaken to characterize MS clinical phenotypes (16, 71) but they have not yet yielded new subtypes.

The increment in neurological deficits after MS relapses has been investigated (53–55, 57–59). Increasing disability in some subjects has been linked to the accumulation of pyramidal, sensory, cerebellar, and sphincter abnormalities (53–55). Furthermore, in some subjects relapses tend to occur at the same anatomical site as previous attack, and this is especially so for the optic nerve, spinal cord, brainstem, and cerebellum sites (57–59), suggesting that neurological signs and symptoms could accumulate at those affected areas. If relapses recur at sites of the previous attacks, this could foster subtypes of MS based on repeated relapses at the same anatomic site.

A network (also called a graph) is an assembly of nodes that are interconnected by edges (52). When all connected nodes come from the same class, the graph is unipartite. When each node is connected to a node of a second class, the graph is bipartite (72). Networks can be partitioned into communities of like nodes (also called clusters) (73, 74). Barabási (75) defines a community as "a locally dense connected subgraph in a network (page 325)," and that "… we expect nodes that belong to a community to have a higher probability of linking to other members of that community than to nodes that do not belong to the same community …." Some of the partitioning algorithms depend upon the maximization of modularity which measures how well each community is separated from other communities.

Network analysis has proven useful in visualizing complex relationships between the phenotypes, genes, proteins, and metabolic pathways that underlie human diseases (76–81). Network analysis has provided important insights in into brain connectivity, and neuroimaging (82, 83). Network analysis has identified potential genetic causes of autism (84) and has clustered autism subjects by phenotype (85, 86). Network analysis has been used to identify genes that govern MS susceptibility (87–89), proteins implicated in the etiology of MS (90), as well as brain areas that undergo disconnection in MS (45, 91–93).

Proposed approach

The review of prior work suggested that there is a gap in identifying subtypes of MS based on signs and symptoms. Our goal was to identify clinical subtypes of RRMS using network analysis after feature reduction. We found 244 unique neurologic signs and symptoms in a cohort of 113 subjects with relapsing-remitting MS, mapped them to classes in a neuro-ontology, and then collapsed the classes into sixteen superclasses (**Figures 1A,B**). For each subject, the count of signs and symptoms in each superclass was normalized. A bipartite graph was created using NetworkX, with each subject node connected to one of sixteen superclass nodes by an edge proportional to the normalized count of signs and symptoms. Distances between subjects were calculated by the cosine similarity of their signs and symptoms. A unipartite graph was created in NetworkX where the nodes were subjects, and the edges were inter-subject distances. The unipartite and bipartite graphs were visualized in Gephi and partitioned into communities based on the Louvain algorithm (94). Modularity scores were used to evaluate the quality of the partitions. We used feature maps to characterize the communities. This approach could lead to classifying MS patients by clinical phenotype and supplement the phenotyping of MS subjects by disease course.

Methods

Subjects

One hundred and twenty MS subjects followed at the University of Illinois-Neuroscience Center were enrolled in the University of Illinois at Chicago (UIC) Neuroimmunology Biobank between August 2018 and March 2020 (mean age 42.7 $\pm$ 12.8 years, 73% female, 27% male, 58% Black, 42% White). The Biobank is approved by the Institutional Review Board (IRB) of the University of Illinois College of Medicine. All subjects provided informed written consent at enrollment. Subjects were between 18-80 years old and had a diagnosis of RRMS based on the 2017 McDonald criteria (95). Subjects had been recruited for a study of blood biomarkers in MS where RRMS was an inclusion criterion and progressive MS was an exclusion criterion. Seven subjects with normal neurological examinations were excluded from the analysis leaving a final study sample of 113 subjects.

Neuro-phenotyping

The neurological progress notes from the electronic health record of all subjects were reviewed, and neurological signs and symptoms were recorded (11). The cumulative signs and symptoms (both active and resolved) of each subject were recorded and mapped to concepts in a neuro-ontology with 1,600 possible concepts (96). Subjects had 13.2 $\mp$ 9.2 signs and symptoms (mean $\mp$ standard deviation). The 113 subjects had 1,453 total signs and symptoms (244 unique signs and symptoms). Subsumption (97) was used to collapse the signs and symptoms (**Figure 1A**) into 16 superclasses (**Figure 1B**) that included *behavior*, *cognitive*, *cranial nerve*,

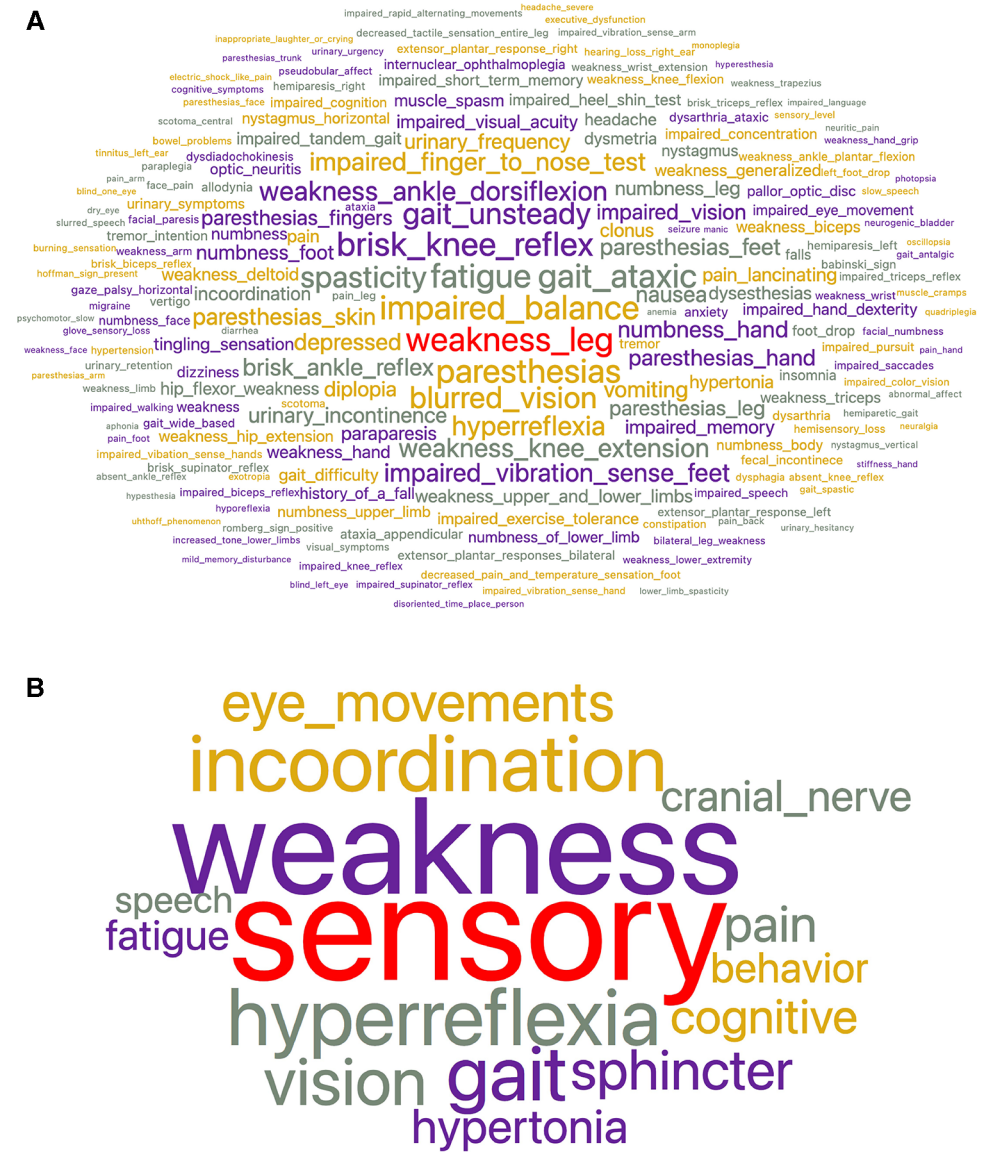

FIGURE 1
(**A**) Word cloud representing the frequency of signs and symptoms in the entire MS cohort before subsumption. Word size is proportional to frequency. There were 244 unique signs and symptoms. The most frequent signs and symptoms were leg weakness, impaired balance, fatigue, and paresthesias. Supporting files available on the project's GitHub site. (**B**) Word cloud representing the frequency of signs and symptoms in the entire MS cohort after subsumption into 16 superclasses. Word size is proportional to frequency. The largest superclasses are sensory, weakness, hyperreflexia, and incoordination. Supporting files available on the project GitHub site.

eye movement, *fatigue*, *gait*, *hyperreflexia*, *hypertonia*, *incoordination*, *pain*, *sensory*, *speech*, *sphincter*, *tremor*, *vision*, and *weakness*. The largest superclasses were weakness, sensory, incoordination, and hyperreflexia. Each subject was represented as a 17-dimension vector where the first element of the vector was the case identification label, and the subsequent sixteen elements were the count for each of the sixteen superclasses (**Figure 2A**). Counts were normalized over the interval [0, 1] using the *continuize* widget in Orange 3.32.0 (98) (**Figure 2B**). We chose to normalize counts because counts varied significantly between superclasses. For supporting data, see the project GitHub site.

Network analysis, distance metrics, feature maps

Network analyses were performed on normalized 113 × 17 data arrays (89, 98, 99). NetworkX (100) converted the data arrays to GraphML files compatible with Gephi. Bipartite

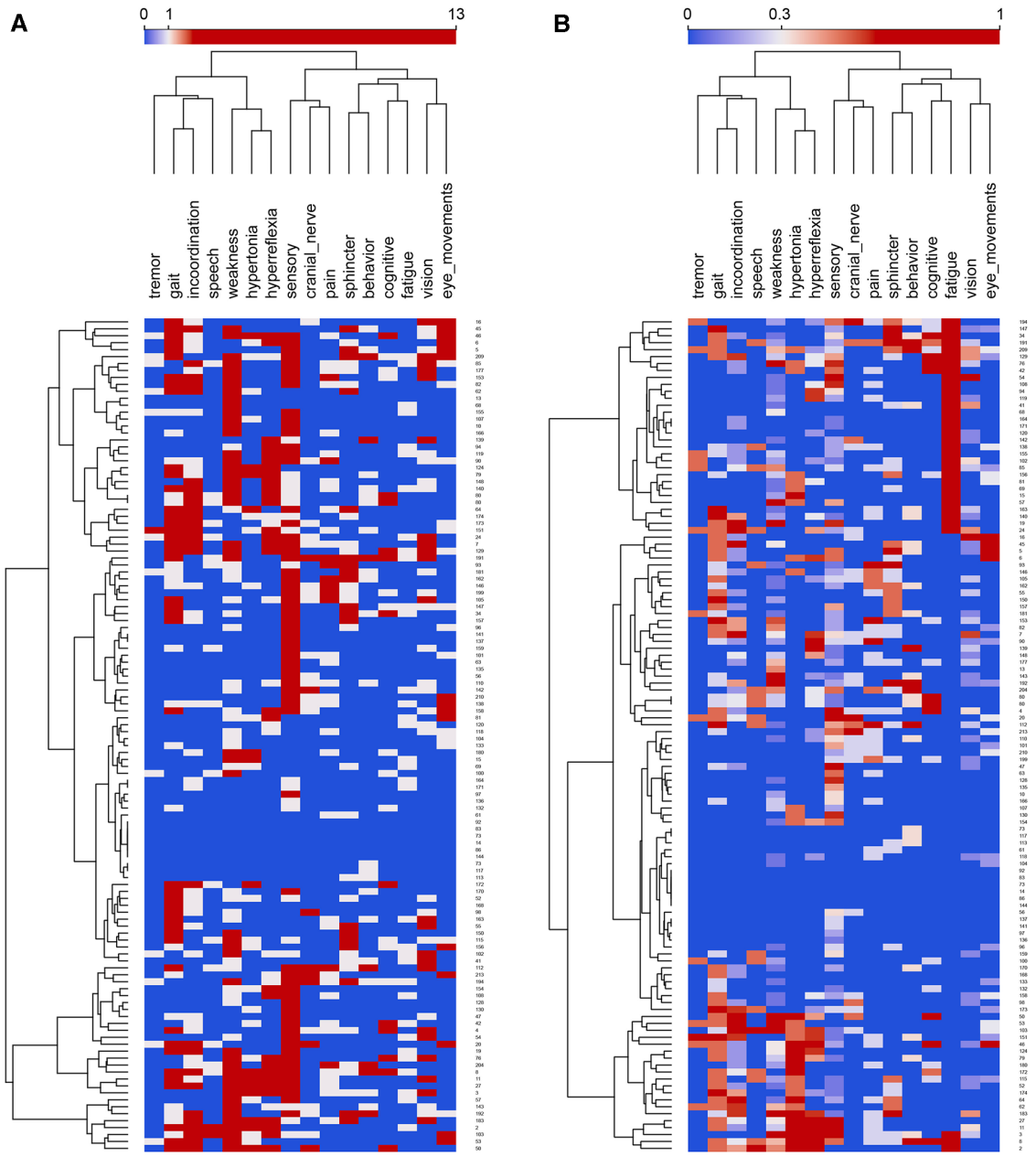

FIGURE 2
(A) Feature map of the entire cohort of MS patients before normalization. Rows are subjects, and columns are superclasses. Normalized feature counts in the columns range between 0 to 13 and the color scale is centered on 3 features. Rows and columns are clustered hierarchically with Ward linkage. Column distances by Pearson correlation coefficient; row distances are Euclidean. (B) Feature map of the entire cohort of MS patients after normalization. Rows are subjects, and columns are superclasses. Normalized feature counts in the columns range between 0 to 1 and the color scale is centered on 0.3 features. Rows and columns are clustered hierarchically with Ward linkage. Column distances by Pearson correlation coefficient; row distances are Euclidean.

networks were visualized in Gephi 0.9.7 using a variety of layouts, with the final analysis using the Force Atlas layout with a repulsion force of 10,000. Visual inspection showed Force Atlas to have the optimal spacing of nodes and clarity of visualization. The bipartite network contained nodes of subjects and features (signs and symptoms) as nodes with a magnitude of the edges connecting subjects to features equal to the normalized feature score for each subject. In the bipartite networks, there were no direct subject-subject or feature-feature edges. Node sizes were proportional to the average weighted degree of each node. Communities were named based on their predominant features. Nodes were colored by their community membership, and colors were used consistently across graphs based on feature

predominance. Edge widths were proportional to edge weight for the bipartite graphs. The unipartite networks were based on distances between subjects derived from the feature vectors for each subject. Distances were calculated in Orange using the *distances* widget for Pearson, Euclidean, and cosine distances. Visual inspection of the network graphs showed that the cosine-based graphs were superior to those based on the Pearson or Euclidean distances. Only the cosine distances were retained for further analysis (101). For the unipartite graphs, all nodes were subjects, and the edges were subject similarity based on the cosine distances. Node size was proportional to the degree (number of edges for each node). The edge width was fixed. Gephi was used to partition the unipartite and bipartite networks into communities based on the Louvain algorithm (94). The Louvain algorithm maximizes modularity (a measure of community separation). Modularity rises from 0.0 as the number of intra-community edges increases relative to inter-community edges. Larger values of modularity reflect a more robust separation of the communities. The degree, average degree, and modularity class for each node were calculated by Gephi. Modularity resolution was set to 1.0 for the unipartite graph and 1.15 for the bipartite graph. For the unipartite graph, two subjects were excluded from the final analysis as they formed communities with only one node. For the normalized unipartite graph, a cosine distance threshold of 0.4 was used to exclude weak edges. Feature means for each community were calculated by SPSS 28 (IBM, Chicago, IL). Differences between community feature means were tested by one-way ANOVA (SPSS). Feature maps were created with the *heat map* widget from Orange. The word cloud was created with the *word cloud* widget from Orange. The concordance for set membership between communities was measured by the Jaccard Index (102) where J is the Jaccard Index, and A and B are the set memberships of two communities:

$$J = \frac{A \cap B}{A \cup B}.$$

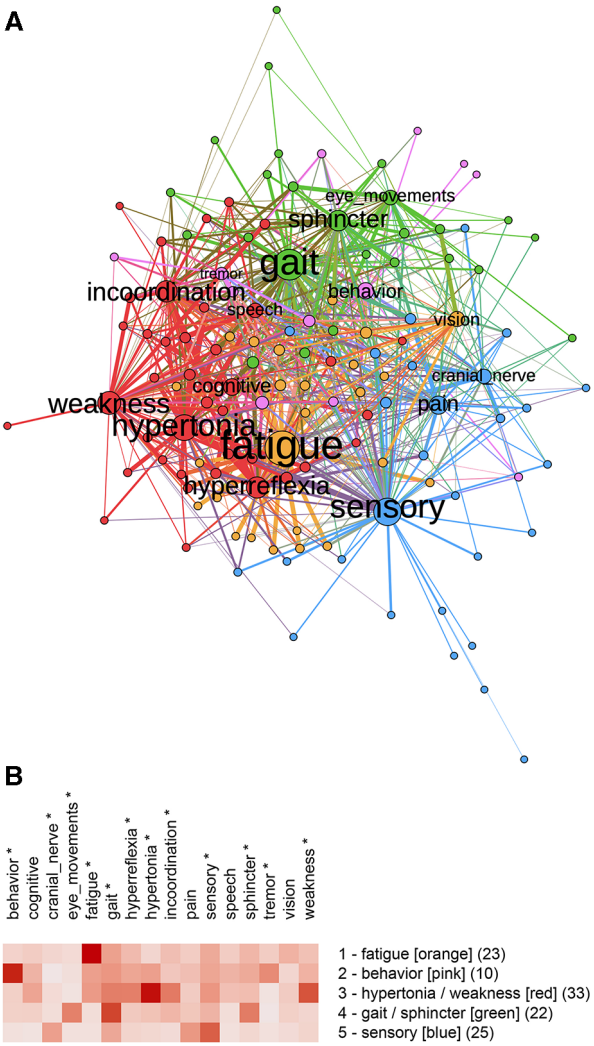

FIGURE 3
(A) Bipartite graph of normalized features. Labeled nodes are features; unlabeled nodes are subjects. Weighted connections (edges) are between features and subjects. (B) Feature map of the five communities identified by partitioning the bipartite graph with unnormalized features. Modularity analysis of the bipartite graph of normalized data showed five communities. Community 1 was predominantly fatigue, Community 2 was predominantly behavioral, Community 3 was weakness and hypertonia, Community 4 was gait and sphincter, and Community 5 was predominantly sensory features indicating features that differed significantly by the community (One-way ANOVA, $p < 0.05$, $df = 4$).

Results

The largest superclasses of signs and symptoms in this cohort of MS subjects were sensory, weakness, incoordination, and hyperreflexia (**Figure 1B**). To prevent the superclasses of weakness and sensory from dominating the network analysis, the signs and symptoms were normalized on the interval [0, 1] before network analysis and partitioning. Visual inspection of the feature map of the MS cohort suggested some clustering of subjects on signs and symptoms (**Figures 2A,B**) and that a network analysis to identify distinct communities would be fruitful.

The bipartite graph was partitioned into five communities (**Figure 3A**) with a modularity score of 0.25. Communities were named and color-coded by the one or two features with the highest community means as **fatigue** ($n = 23$), **behavior** ($n = 10$), **hypertonia/weakness** ($n = 33$), **gait/sphincter** ($n = 22$), and **sensory** ($n = 25$) (**Figure 3B**). ANOVA showed significant differences between communities for behavior ($p < .001$), cranial nerve ($p = .008$), eye movements ($p < .001$), fatigue ($p < .001$), gait ($p = .029$), hyperreflexia

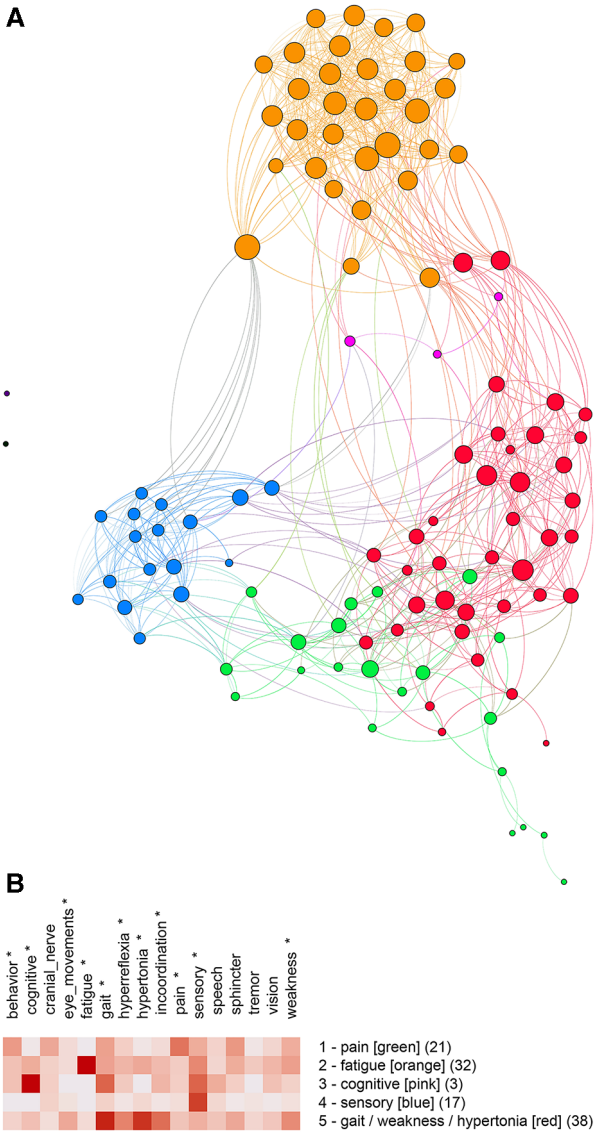

FIGURE 4
(**A**) Unipartite graph based on normalized features. Nodes are subjects, and node size is proportional to the number of edges. The largest communities are gait/weakness/hypertonia (red, $n = 38$) and fatigue (orange, $n = 32$). Note the small cognitive community (pink, $n = 3$). (**B**) Feature map of the five communities identified by partitioning the unipartite network graph based on normalized features. Asterisks indicate features that differed significantly by the community (One-way ANOVA, $p < 0.05$, $df = 4$).

($p = .031$), hypertonia ($p < .001$), incoordination ($p = .002$), sensory ($p = .018$), sphincter ($p = .021$), tremor ($p = .006$), and weakness ($p = .001$).

The unipartite graph was partitioned into five communities (Figure 4A) with a modularity score of 0.49. Communities were named by their predominant features: **pain**, **fatigue**, **cognitive**, **sensory**, and **weakness/gait/hypertonia** (Figure 4B). ANOVA analysis showed significant differences between communities for behavior ($p = .033$), cognitive ($p < .001$), eye movements ($p = .047$), fatigue ($p < .001$), gait ($p < .001$), hyperreflexia ($p = .014$), hypertonia ($p < .001$), incoordination ($p < .001$), pain ($p < .001$), sensory ($p < .001$), and weakness ($p = .037$).

Although partitioning the bipartite and unipartite graphs produced somewhat different communities, similarities between community membership and graphs were notable. We used the Jaccard Index (a set similarity measure) to assess the similarity between communities. Membership for the **fatigue** ($J = 0.72$) and **sensory** ($J = 0.56$) communities was similar for the unipartite and bipartite graphs. The unipartite graph community **gait/weakness/hypertonia** showed similarity to the bipartite graph communities **hypertonia/weakness** ($J = 0.36$) and **gait/sphincter** ($J = 0.36$). A complete table of Jaccard Index values is available on the project's GitHub site.

Discussion

Multiple sclerosis can present as sensory loss, weakness, incoordination, sphincter disturbance, diplopia, visual loss, cognitive impairment, fatigue, or even pain. We have used network analysis to identify distinct clinical subtypes of multiple sclerosis based on signs and symptoms. We first mapped the signs and symptoms of a cohort of multiple sclerosis subjects to concepts from neuro-ontology. We then created a bipartite graph, where subjects and their signs and symptoms were nodes in a graph (Figure 3A). In a bipartite graph, subjects are connected to signs and symptoms and not to other subjects. When the signs and symptoms of a subject are converted to vectors, distances between subjects can be calculated so that subject nodes can be connected to other subjects to form a unipartite graph (Figure 4A). Network analysis allowed us to identify communities of multiple sclerosis subjects who shared signs and symptoms in common. Partitioning of the unipartite and bipartite graphs based on modularity score identified communities with strong fatigue and sensory feature predominance. Both partitions had communities characterized by weakness combined with hypertonia or gait findings. Partitioning of the bipartite graph produced a small community with behavioral changes (depression, anxiety, etc.) and a gait/sphincter community. Partitioning of the unipartite graph produced a small community with cognitive findings and a medium-sized community with pain (Figure 4B).

Partitions of the unipartite graph yielded higher modularity scores than the bipartite graph, suggesting that the partitioning of the unipartite graph was more robust. The named communities for Figure 4B (**pain**, **fatigue**, **cognitive**, **sensory**, and **gait/weakness/hypertonia**) deserve special consideration as potentially identifiable multiple sclerosis subtypes. We found a strong overlap between the fatigue and sensory communities across both graphs as measured by the Jaccard Index. Significant overlap between the **gait/weakness/hypertonia** community from the unipartite graph with the **gait/sphincter** and **hypertonia/**

weakness communities from the bipartite graph was noted. Although the partitioning of networks based on features suggests that identifiable MS subtypes may exist, variability across partitions does not permit a definitive characterization of subtypes.

Although we did not correlate community features with MRI findings, the communities detected may reflect the anatomic location of MS lesions (103, 104). Of particular interest is the tendency for relapses to occur at the sites of previous MS attacks (57–59). Recurring relapses at the same anatomic site could lead to increased symptoms in certain domains (e.g., weakness, incoordination, sphincter, visual, etc.) and make subtypes of MS more discernible. On the other hand, "pure" subtypes of MS (i.e., pure motor, pure sensory, pure cognitive) are uncommon; nearly all MS patients in our cohort have signs and symptoms in multiple symptomatic domains (see, for example, **Figures 2A,B**).

Two strengths of this study should be mentioned. First, community detection was done by network analysis which offers an alternative to unsupervised machine learning algorithms based on cluster analysis (105, 106). Second, we used subsumption and the hierarchical organization of signs and symptoms in an ontology to reduce the number of features used in the analysis (97). The current study demonstrates that subsumption can successfully group signs and symptoms of MS subjects into superclasses (**Figures 1A, B**). These superclasses can be used to characterize the clinical features of communities identified by network analysis.

The current study has several limitations. The sample size was small ($N = 113$). The small sample size could cause a selection bias that influenced the communities found by network analysis. Network analysis of larger sample sizes may detect more robust communities with a different profile of predominant features. In particular, we did not identify communities of MS subjects with predominant vision, cranial nerve, or incoordination signs and symptoms, although such communities likely exist (15, 20–22). Another limitation was that we evaluated only one partitioning algorithm (Louvain). A limitation of the Louvain algorithm is that it does not exhaustively examine all possible partitions, so partitioning is non-deterministic, and partitions may change with each run (73, 107, 108). Other partitioning algorithms are available and might yield different results. We used subsumption to reduce the number of clinical features from 244 to sixteen. Different subsumption strategies would likely yield different results. We calculated distances between subjects using the cosine distance metric; other distance metrics are available and may have resulted in different results. Although the modularity scores of the partitions are comparable to those obtained on standard datasets like the Karate Club (73, 108), they are still modest (0.25–0.49). Another limitation was that subjects in the study were diagnosed with the RRMS phenotype. Without further analysis, our data cannot be extrapolated to other disease course phenotypes. Our analysis did not consider the race or sex of the subjects, which could influence clinical subtype (60, 109, 110). Finally, we partitioned MS subjects based on their accumulated signs and symptoms. Examining networks based on signs and symptoms at a single time would be instructive.

Conclusions

MS phenotypes based on the clinical course are well-established. Clinical subtypes of MS based on clinical presentation are increasingly recognized. After mapping the signs and symptoms of a cohort of MS patients to classes in neuro-ontology and then collapsing these classes into sixteen superclasses, we used network analysis to identify clinical subtypes of MS based on signs and symptoms. Feature maps (**Figures 3B**, **4B**) suggest that identifiable subtypes of MS with predominant signs and symptoms related to weakness, sensation, behavior, cognition, pain, and fatigue deserve further investigation. The clinical subtyping of MS subjects could supplement phenotyping by disease course. Additional studies may reveal that MS subtypes correlate with epigenetic, radiological, immunologic, or protein biomarkers.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by The Biobank is approved by the Institutional Review Board (IRB) of the University of Illinois College of Medicine. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Design and concept by DBH and QHP. Data acquisition by MDC, DBH, QHP, and CO. Computations by DBH and QHP. Data analysis, interpretation, and writing by DBH, QHP, MDC, CO, and DCWII. Revisions and approval by DBH, QHP, MDC, CO, and DCWII. All authors contributed to the article and approved the submitted version.

Funding

MDC acknowledges research funding from the Department of Veterans Affairs (BLR&D Merit Award BX000467) and prior support from Biogen. DCW II acknowledges research funding

support from the Mary K. Finley Endowment at the Missouri University of Science and Technology.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Kerr DA. The lumping, splitting of inflammatory CNS diseases. *Neurology*. (2006) 66(10):1466–7. doi: 10.1212/01.wnl.0000221747.37657.c6

2. Lublin FD, Reingold SC, SC for the National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology*. (1996) 46(4):907–11. doi: 10.1212/WNL.46.4.907

3. Marcus JF, Waubant EL. Updates on clinically isolated syndrome, diagnostic criteria for multiple sclerosis. *Neurohospitalist*. (2013) 3(2):65–80. doi: 10.1177/1941874412457183

4. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sørensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*. (2014) 83(3):278–86. doi: 10.1212/WNL.0000000000000560

5. Lublin FD. New multiple sclerosis phenotypic classification. *Eur Neurol*. (2014) 72(1):Suppl. 1–5. doi: 10.1159/000367614

6. Engelhard J, Oleske DM, Schmitting S, Wells KE, Talapala S, Barbato LM. Multiple sclerosis by phenotype in Germany. *Mult Scler Relat Disord*. (2022) 57:103326. doi: 10.1016/j.msard.2021.103326

7. Kantarci OH. Phases and phenotypes of multiple sclerosis. *Continuum*. (2019) 25(3):636–54. doi: 10.1212/CON.0000000000000737

8. Oh J, Vidal-Jordana A, Montalban X. Multiple sclerosis: clinical aspects. *Curr Opin Neurol*. (2018) 31(6):752–9. doi: 10.1097/WCO.0000000000000622

9. Robinson PN. Deep phenotyping for precision medicine. *Hum Mutat*. (2012) 33(5):777–80. doi: 10.1002/humu.22080

10. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping, genomic data. *Nature*. (2018) 562 (7726):203–9. doi: 10.1038/s41586-018-0579-z

11. Hier DB, Yelugam R, Azizi S, Wunsch III DC. A focused review of deep phenotyping with examples from neurology. *Eur Sci J*. (2022) 9:4–19. doi: /10.19044/esj.2022.v18n4p4

12. Bove R, Chitnis T, Cree BA, Tintoré M, Naegelin Y, Uitdehaag BM, et al. SUMMIT (serially unified multicenter multiple sclerosis investigation): creating a repository of deeply phenotyped contemporary multiple sclerosis cohorts. *Mult Scler J*. (2018) 24(11):1485–98. doi: 10.1177/1352458517726657

13. Delude CM. Deep phenotyping: the details of disease. *Nature*. (2015) 527 (7576):S14–5. doi: 10.1038/527S14a

14. Poser C. Onset symptoms of multiple sclerosis. *J Neurol Neurosurg Psychiatr*. (1995) 58(2):253. doi: 10.1136/jnnp.58.2.253-a

15. Ford H. Clinical presentation, diagnosis of multiple sclerosis. *Clin Med*. (2020) 20(4):380. doi: 10.7861/clinmed.2020-0292

16. Loonstra FC, De Ruiter LR, Doesburg D, Lam KH, Van Lierop ZY, Moraal B, et al. Project Y: the search for clues explaining phenotype variability in MS. *Mult Scler Relat Disord*. (2022) 57:103337. doi: 10.1016/j.msard.2021.103337

17. Naismith RT, Trinkaus K, Cross A. Phenotype and prognosis in African-Americans with multiple sclerosis: a retrospective chart review. *Mult Scler J*. (2006) 12(6):775–81. doi: 10.1177/1352458506070923

18. Cree BA, Reich DE, Khan O, De Jager PL, Nakashima I, Takahashi T, et al. Modification of multiple sclerosis phenotypes by African ancestry at HLA. *Arch Neurol*. (2009) 66(2):226–33. doi: 10.1001/archneurol.2008.541

19. Ayache SS, Chalah MA, Al-Ani T, Farhat WH, Zouari HG, Créange A, et al. Tremor in multiple sclerosis: the intriguing role of the cerebellum. *J Neurol Sci*. (2015) 358(1-2):351–6. doi: 10.1016/j.jns.2015.09.360

20. Mills RJ, Yap L, Young CA. Treatment for ataxia in multiple sclerosis. *Cochrane Database Syst Rev*. (2007) 1. doi: 10.1002/14651858.CD005029.pub2

21. Jacobs DA, Galetta SL. Multiple sclerosis, the visual system. *Ophthalmol Clin North Am*. (2004) 17(3):265–73. doi: 10.1016/j.ohc.2004.05.011

22. Costello F. Vision disturbances in multiple sclerosis. *Semin Neurol*. (2016) 36 (02):185–95. doi: 10.1055/s-0036-1579692

23. Rae-Grant AD, Eckert NJ, Bartz S, Reed JF. Sensory symptoms of multiple sclerosis: a hidden reservoir of morbidity. *Mult Scler J*. (1999) 5(3):179–83. doi: 10.1177/135245859900500307

24. O'Connor AB, Schwid SR, Herrmann DN, Markman JD, Dworkin RH. Pain associated with multiple sclerosis: systematic review, proposed classification. *Pain*. (2008) 137(1):96–111. doi: 10.1016/j.pain.2007.08.024

25. Kratz AL, Whibley D, Alschuler KN, Ehde DM, Williams DA, Clauw DJ, et al. Characterizing chronic pain phenotypes in multiple sclerosis: a nationwide survey study. *Pain*. (2021) 162(5):1426. doi: 10.1097/j.pain.0000000000002136

26. Sanders E, Arts R. Paraesthesiae in multiple sclerosis. *J Neurol Sci*. (1986) 74 (2-3):297–305. doi: 10.1016/0022-510X(86)90115-2

27. Rizzo M, Hadjimichael O, Preiningerova J, Vollmer T. Prevalence and treatment of spasticity reported by multiple sclerosis patients. *Mult Scler J*. (2004) 10(5):589–95. doi: 10.1191/1352458504ms1085oa

28. Hoang PD, Gandevia SC, Herbert RD. Prevalence of joint contractures and muscle weakness in people with multiple sclerosis. *Disabil Rehabil*. (2014) 36 (19):1588–93. doi: 10.3109/09638288.2013.854841

29. Cordani C, Hidalgo de la Cruz M, Meani A, Valsasina P, Esposito F, Pagani E, et al. MRI correlates of clinical disability and hand-motor performance in multiple sclerosis phenotypes. *Mult Scler J*. (2021) 27(8):1205–21. doi: 10.1177/1352458520958356

30. Nociti V, Cianfoni A, Mirabella M, Caggiula M, Frisullo G, Patanella AK, et al. Clinical characteristics, course and prognosis of spinal multiple sclerosis. *Spinal Cord*. (2005) 43(12):731–4. doi: 10.1038/sj.sc.3101798

31. Wiesel PH, Norton C, Glickman S, Kamm MA. Pathophysiology and management of bowel dysfunction in multiple sclerosis. *Eur J Gastroenterol Hepatol*. (2001) 13(4):441–8. doi: 10.1097/00042737-200104000-00025

32. Amato MP, Zipoli V, Portaccio E. Cognitive changes in multiple sclerosis. *Expert Rev Neurother*. (2008) 8(10):1585–96. doi: 10.1586/14737175.8.10.1585

33. De Meo E, Portaccio E, Giorgio A, Ruano L, Goretti B, Niccolai C, et al. Identifying the distinct cognitive phenotypes in multiple sclerosis. *JAMA Neurol*. (2021) 78(4):414–25. doi: 10.1001/jamaneurol.2020.4920

34. Noffs G, Perera T, Kolbe SC, Shanahan CJ, Boonstra FM, Evans A, et al. What speech can tell us: a systematic review of dysarthria characteristics in Multiple Sclerosis. *Autoimmun Rev*. (2018) 17(12):1202–9. doi: 10.1016/j.autrev.2018.06.010

35. Miglis MG, Muppidi S. Autonomic dysfunction in multiple sclerosis and other updates on recent autonomic research. *Clin Auton Res*. (2018) 28 (4):391–3. doi: 10.1007/s10286-018-0548-5

36. Sá MJ. Psychological aspects of multiple sclerosis. *Clin Neurol Neurosurg*. (2008) 110(9):868–77. doi: 10.1016/j.clineuro.2007.10.001

37. Soyuer F, Mirza M, Erkorkmaz Ü. Balance performance in three forms of multiple sclerosis. *Neurol Res*. (2006) 28(5):555–62. doi: 10.1179/016164105X49373

38. Pop R, Kipfer S. Paroxysmal kinesigenic dyskinesia–like phenotype in multiple sclerosis. *Mult Scler J*. (2017) 23(13):1795–7. doi: 10.1177/1352458517702535

39. Induruwa I, Constantinescu CS, Gran B. Fatigue in multiple sclerosis—a brief review. *J Neurol Sci*. (2012) 323(1–2):9–15. doi: 10.1016/j.jns.2012.08.007

40. Kos D, Kerckhofs E, Nagels G, D'hooghe M, Ilsbroukx S. Origin of fatigue in multiple sclerosis: review of the literature. *Neurorehabil Neural Repair*. (2008) 22 (1):91–100. doi: 10.1177/1545968306298934

41. Amato MP, Fratiglioni L, Groppi C, Siracusa G, Amaducci L. Interrater reliability in assessing functional systems and disability on the Kurtzke scale in multiple sclerosis. *Arch Neurol.* (1988) 45(7):746–8. doi: 10.1001/archneur.1988.00520310052017

42. Ó Donnchadha S, Burke T, Bramham J, O'Brien MC, Whelan R, Reilly R, et al. Symptom overlap in anxiety and multiple sclerosis. *Mult Scler J.* (2013) 19 (10):1349–54. doi: 10.1177/1352458513476742

43. Koch MW, Patten S, Berzins S, Zhornitsky S, Greenfield J, Wall W, et al. Depression in multiple sclerosis: a long-term longitudinal study. *Mult Scler J.* (2015) 21(1):76–82. doi: 10.1177/1352458514536086

44. Ganesvaran G, Greer J, Pender M. Prominent brainstem and cerebellar involvement in multiple sclerosis with psoriasis. *Mult Scler J.* (2009) 15 (6):763–6. doi: 10.1177/1352458509103612

45. Rocca MA, Valsasina P, Meani A, Falini A, Comi G, Filippi M. Impaired functional integration in multiple sclerosis: a graph theory study. *Brain Struct Funct.* (2016) 221(1):115–31. doi: 10.1007/s00429-014-0896-4

46. Hancock LM, Galioto R, Samsonov A, Busch RM, Hermann B, Matias-Guiu JA. A proposed new taxonomy of cognitive phenotypes in multiple sclerosis: the international classification of cognitive disorders in MS (IC-CoDiMS). *Mult Scler J.* (2022):13524585221127941.

47. Staff NP, Lucchinetti CF, Keegan BM. Multiple sclerosis with predominant, severe cognitive impairment. *Arch Neurol.* (2009) 66(9):1139–43. doi: 10.1001/archneurol.2009.190

48. Leavitt VM, Tosto G, Riley CS. Cognitive phenotypes in multiple sclerosis. *J Neurol.* (2018) 265(3):562–6. doi: 10.1007/s00415-018-8747-5

49. Cabeça HLS, Rocha LC, Sabbá AF, Tomás AM, Bento-Torres NVO, Anthony DC, et al. The subtleties of cognitive decline in multiple sclerosis: an exploratory study using hierarchichal cluster analysis of CANTAB results. *BMC Neurol.* (2018) 18(1):140. doi: 10.1186/s12883-018-1141-1

50. Zurawski J, Healy B, Ratajska A, Barker L, Glanz B, Houtchens M. Identification of a predominant cognitive phenotype in patients with multiple sclerosis. *Eur J Neurol.* (2020) 27(6):1083–8. doi: 10.1111/ene.14186

51. Beckerman H, Eijssen IC, van Meeteren J, Verhulsdonck MC, de Groot V. Fatigue profiles in patients with multiple sclerosis are based on severity of fatigue and not on dimensions of fatigue. *Sci Rep.* (2020) 10(1):1–10. doi: 10.1038/s41598-020-61076-1

52. Borsboom D, Deserno MK, Rhemtulla M, Epskamp S, Fried EI, McNally RJ, et al. Network analysis of multivariate data in psychological science. *Nat Rev Dis Primers.* (2021) 1(1):1–18.

53. Kalincik T, Buzzard K, Jokubaitis V, Trojano M, Duquette P, Izquierdo G, et al. Risk of relapse phenotype recurrence in multiple sclerosis. *Mult Scler J.* (2014) 20(11):1511–22. doi: 10.1177/1352458514528762

54. Stewart T, Spelman T, Havrdova E, Horakova D, Trojano M, Izquierdo G, et al. Contribution of different relapse phenotypes to disability in multiple sclerosis. *Mult Scler J.* (2017) 23(2):266–76. doi: 10.1177/1352458516643392

55. Scott T, Wang P, You X, Mann M, Sperling B. Relationship between sustained disability progression and functional system scores in relapsing-remitting multiple sclerosis: analysis of placebo data from four randomized clinical trials. *Neuroepidemiology.* (2015) 44(1):16–23. doi: 10.1159/000369621

56. Rivel M, Achiron A, Dolev M, Stern Y, Zeilig G, Defrin R. Unique features of central neuropathic pain in multiple sclerosis: results of a cluster analysis. *Eur J Pain.* (2022) 26(5):1107–22. doi: 10.1002/ejp.1934

57. Tsantes E, Leone MA, Curti E, Cantello R, Vecchio D, Granella F. Location of first attack predicts the site of subsequent relapses in multiple sclerosis. *J Clin Neurosci.* (2020) 74:175–9. doi: 10.1016/j.jocn.2020.02.017

58. Mowry EM, Deen S, Malikova I, Pelletier J, Bacchetti P, Waubant E. The onset location of multiple sclerosis predicts the location of subsequent relapses. *J Neurol Neurosurg Psychiatry.* (2009) 80(4):400–3. doi: 10.1136/jnnp.2008.157305

59. Deen S, Bacchetti P, High A, Waubant E. Predictors of the location of multiple sclerosis relapse. *J Neurol Neurosurg Psychiatry.* (2008) 79(10):1190–3. doi: 10.1136/jnnp.2007.136440

60. Cree B, Khan O, Bourdette D, Goodin D, Cohen J, Marrie R, et al. Clinical characteristics of African Americans vs Caucasian Americans with multiple sclerosis. *Neurology.* (2004) 63(11):2039–45. doi: 10.1212/01.WNL.0000145762.60562.5D

61. Gerbis N, Parratt J. Severe unilateral optic neuritis in multiple sclerosis. *J Neurol Neurosurg Psychiatry.* (2018) 89:A41. doi: 10.1136/jnnp-2018-ANZAN.103

62. Diaz-Olavarrieta C, Cummings JL, Velazquez J, Garcia de al Cadena C. Neuropsychiatric manifestations of multiple sclerosis. *J Neuropsychiatry Clin Neurosci.* (1999) 11(1):51–7. doi: 10.1176/jnp.11.1.51

63. Zhang Y, Taylor BV, Simpson Jr S, Blizzard L, Campbell JA, Palmer AJ, et al. Feelings of depression, pain, and walking difficulties have the largest impact on the quality of life of people with multiple sclerosis, irrespective of clinical phenotype. *Mult Scler J.* (2021) 27(8):1262–75. doi: 10.1177/1352458520958369

64. Tahedl M, Levine SM, Greenlee MW, Weissert R, Schwarzbach JV. Functional connectivity in multiple sclerosis: recent findings and future directions. *Front Neurol.* (2018) 9:828. doi: 10.3389/fneur.2018.00828

65. Schee JP, Viswanathan S. Pure spinal multiple sclerosis: a possible novel entity within the multiple sclerosis disease spectrum. *Mult Scler J.* (2019) 25 (8):1189–95. doi: 10.1177/1352458518775912

66. Kira Ji. Neuromyelitis optica and opticospinal multiple sclerosis: mechanisms and pathogenesis. *Pathophysiology.* (2011) 18(1):69–79. doi: 10.1016/j.pathophys.2010.04.008

67. Takeuchi W, Fujimori J, Nakashima I. Multiple sclerosis limited to spinal cord lesions. *Clin Exp Neuroimmunol.* (2021) 12(2):111–5. doi: 10.1111/cen3.12635

68. Noseworthy J, Vandervoort M, Wong C, Ebers G. Interrater variability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical trial. *Neurology.* (1990) 40(6):971. doi: 10.1212/WNL.40.6.971

69. Yang Z, Pou-Prom C, Jones A, Banning M, Dai D, Mamdani M, et al. Assessment of natural language processing methods for ascertaining the expanded disability status scale score from the electronic health records of patients with multiple sclerosis: algorithm development and validation study. *JMIR Med Inform.* (2022) 10(1):e25157. doi: 10.2196/25157

70. Dahlke F, Arnold DL, Aarden P, Ganjgahi H, Häring DA, Čuklina J, et al. Characterisation of MS phenotypes across the age span using a novel data set integrating 34 clinical trials (NO, MS cohort): age is a key contributor to presentation. *Mult Scler J.* (2021) 27(13):2062–76. doi: 10.1177/1352458520988637

71. Bergamaschi R, Beghi E, Bosetti C, Ponzio M, Santucci C, Lepore V, et al. Do patients' and referral centers' characteristics influence multiple sclerosis phenotypes? Results from the Italian multiple sclerosis and related disorders register. *Neurol Sci.* (2022):1–11. doi: 10.1007/s10072-022-06169-7

72. Pavlopoulos GA, Kontou PI, Pavlopoulou A, Bouyioukos C, Markou E, Bagos PG. Bipartite graphs in systems biology and medicine: a survey of methods and applications. *GigaScience.* (2018) 7(4):giy014. doi: 10.1093/gigascience/giy014

73. Newman ME. Modularity and community structure in networks. *Proc Natl Acad Sci.* (2006) 103(23):8577–82. doi: 10.1073/pnas.0601602103

74. Kramer J, Boone L, Clifford T, Bruce J, Matta J. Analysis of medical data using community detection on inferred networks. *IEEE J Biomed Health Inform.* (2020) 24(11):3136–43. doi: 10.1109/JBHI.2020.3003827

75. Barabási AL. *Network science*. Cambridge, UK: Cambridge University Press (2016).

76. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* (2011) 12(1):56–68. doi: 10.1038/nrg2918

77. Carter H, Hofree M, Ideker T. Genotype to phenotype via network analysis. *Curr Opin Genet Dev.* (2013) 23(6):611–21. doi: 10.1016/j.gde.2013.10.003

78. Bertolero MA, Yeo BT, D'Esposito M. The diverse club. *Nat Commun.* (2017) 8(1):1–11. doi: 10.1038/s41467-017-01189-w

79. Chen Y, Zhang X, Zhang GQ, Xu R. Comparative analysis of a novel disease phenotype network based on clinical manifestations. *J Biomed Inform.* (2015) 53:113–20. doi: 10.1016/j.jbi.2014.09.007

80. Gosak M, Markovič R, Dolenšek J, Rupnik MS, Marhl M, Stožer A, et al. Network science of biological systems at different scales: a review. *Phys Life Rev.* (2018) 24:118–35. doi: 10.1016/j.plrev.2017.11.003

81. Ren X, Wang S, Huang T. Decipher the connections between proteins and phenotypes. *Biochim Biophys Acta Proteins Proteom.* (2020) 1868(11):140503. doi: 10.1016/j.bbapap.2020.140503

82. Betzel RF, Bertolero MA, Gordon EM, Gratton C, Dosenbach NU, Bassett DS. The community structure of functional brain networks exhibits scale-specific patterns of inter-and intra-subject variability. *Neuroimage.* (2019) 202:115990. doi: 10.1016/j.neuroimage.2019.07.003

83. Bassett DS, Zurn P, Gold JI. On the nature and use of models in network neuroscience. *Nat Rev Neurosci.* (2018) 19(9):566–78. doi: 10.1038/s41583-018-0038-8

84. Emberti Gialloreti L, Enea R, Di Micco V, Di Giovanni D, Curatolo P. Clustering analysis supports the detection of biological processes related to autism spectrum disorder. *Genes.* (2020) 11(12):1476. doi: 10.3390/genes11121476

85. Matta J, Zhao J, Ercal G, Obafemi-Ajayi T. Applications of node-based resilience graph theoretic framework to clustering autism spectrum disorders phenotypes. *Appl Netw Sci.* (2018) 3(1):1–22. doi: 10.1007/s41109-018-0093-0

86. Matta J, Dobrino D, Yeboah D, Howard S, Yasser EM, Obafemi-Ajayi T. Connecting phenotype to genotype: PheWAS-inspired analysis of autism spectrum disorder. *Front Hum Neurosci*. (2022) 16:1–16. doi: 10.3389/fnhum.2022.960991

87. Baranzini SE, Khankhanian P, Patsopoulos NA, Li M, Stankovich J, Cotsapas C, et al. Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. *Am J Hum Genet*. (2013) 92(6):854–65. doi: 10.1016/j.ajhg.2013.04.019

88. Slim L, Chatelain C, Foucauld Hd, Azencott CA. A systematic analysis of gene–gene interaction in multiple sclerosis. *BMC Med Genomics*. (2022) 15 (1):1–14. doi: 10.1186/s12920-022-01247-3

89. Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, function using NetworkX. In: Varoquaux G, Vaught T, Millman J, editors. *Proceedings of the 7th Python in Science*; Pasadena, CA, USA. Los Alamos, NM: Los Alamos National Laboratory (LANL); 2008.

90. Cervantes-Gracia K, Husi H. Integrative analysis of multiple sclerosis using a systems biology approach. *Sci Rep*. (2018) 8(1):1–14. doi: 10.1038/s41598-018-24032-8

91. Schiavi S, Azzari A, Mensi A, Graziano N, Daducci A, Bicego M, et al. Classification of multiple sclerosis patients based on structural disconnection: a robust feature selection approach. *J Neuroimaging*. (2022) 32:647–55. doi: 10.1111/jon.12991

92. Schoonheim MM, Meijer KA, Geurts JJ. Network collapse, cognitive impairment in multiple sclerosis. *Front Neurol*. (2015) 6:82. doi: 10.3389/fneur.2015.00082

93. Fleischer V, Radetz A, Ciolac D, Muthuraman M, Gonzalez-Escamilla G, Zipp F, et al. Graph theoretical framework of brain networks in multiple sclerosis: a review of concepts. *Neuroscience*. (2019) 403:35–53. doi: 10.1016/j.neuroscience.2017.10.033

94. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech: Theory Exp*. (2008) 2008(10): P10008. doi: 10.1088/1742-5468/2008/10/P10008

95. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. (2018) 17(2):162–73. doi: 10.1016/S1474-4422(17)30470-2

96. Hier DB, Brint SU. A Neuro-ontology for the neurological examination. *BMC Med Inform Decis Mak*. (2020) 20(1):1–9. doi: 10.1186/s12911-020-1066-7

97. Wunsch III DC, Hier DB. Subsumption is a novel feature reduction strategy for high dimensionality datasets. *Eur Sci J*. 2022;18:20–36. Accessed August 12, 2022. doi: 10.190444/esj.2022.v18n4p20

98. Demsar J, Zupan B. Orange: data mining fruitful and fun-A historical perspective. *Informatica*. (2013) 37(1):55.

99. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 3. Burnaby, BC Canada: Public Knowledge Project (2009). p. 361–362. doi: 10.1609/icwsm.v3i1.13937

100. Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using networkX. In: Varoquaux G, Vaught T, Millman J, editors. *Proceedings of the 7th Python in Science Conference*; Pasadena, CA, USA; scipy.org 2008. p. 11–15.

101. Hier DB, Kopel J, Brint SU, Wunsch DC, Olbricht GR, Azizi S, et al. Evaluation of standard and semantically-augmented distance metrics for neurology patients. *BMC Med Inform Decis Mak*. (2020) 20(1):1–15. doi: 10.1186/s12911-020-01217-8

102. Jaccard P. Distribution de la flore alpine dans le Bassin des Drouces et dans quelques régions voisines. *Bull Soc Vaud Sci Nat*. (1901) 37(140):241–72.

103. Huber SJ, Paulson GW, Chakeres D, Pakalnis A, Brogan M, Phillips BL, et al. Magnetic resonance imaging and clinical correlations in multiple sclerosis. *J Neurol Sci*. (1988) 86(1):1–12. doi: 10.1016/0022-510X(88)90002-0

104. Stevens JC, Farlow MR, Edwards MK, Yu Pl. Magnetic resonance imaging: clinical correlation in 64 patients with multiple sclerosis. *Arch Neurol*. (1986) 43 (11):1145–8. doi: 10.1001/archneur.1986.00520110039011

105. Xu R, Wunsch D. Survey of clustering algorithms. *IEEE Trans Neural Netw*. (2005) 16(3):645–78. doi: 10.1109/TNN.2005.845141

106. Xu R, Wunsch DC. Clustering algorithms in biomedical research: a review. *IEEE Rev Biomed Eng*. (2010) 3:120–54. doi: 10.1109/RBME.2010.2083647

107. Fortunato S, Hric D. Community detection in networks: a user guide. *Phys Rep*. (2016) 659:1–44. doi: 10.1016/j.physrep.2016.09.002

108. Motschnig N, Ramharter A, Schweiger O, Zabka P, Foerster KT. On comparing and enhancing common approaches to network community detection [Preprint] (2021). Available at: https://doi.org/10.48550/arXiv.2108.13482.

109. Weinstock-Guttman B, Jacobs L, Brownscheidle C, Baier M, Rea D, Apatoff B, et al. Multiple sclerosis characteristics in African American patients in the New York State Multiple Sclerosis Consortium. *Mult Scler J*. (2003) 9(3):293–8. doi: 10.1191/1352458503ms909oa

110. Caldito NG, Saidha S, Sotirchos ES, Dewey BE, Cowley NJ, Glaister J, et al. Brain and retinal atrophy in African-Americans versus Caucasian-Americans with multiple sclerosis: a longitudinal study. *Brain*. (2018) 141(11):3115–29. doi: 10.1093/brain/awy245