

Clinical and Research Applications of the Electronic Medical Record in Multiple Sclerosis: A Narrative Review of Current Uses and Future Applications

Carol Swetlik, MD, MS; Riley Bove, MD, MSc; and Marisa McGinley, DO, MSc

CE INFORMATION

ACTIVITY AVAILABLE ONLINE: To access the article and evaluation online, go to <https://www.highmarksce.com/mscare>.

TARGET AUDIENCE: The target audience for this activity is physicians, advanced practice clinicians, nursing professionals, pharmacists, mental health professionals, social workers, and other health care providers involved in the research and management of patients with multiple sclerosis (MS).

LEARNING OBJECTIVES:

1. Characterize existing EMR platforms designed specifically for care of people with MS.
2. Describe relevant variables that are captured in the EMR that allow identification of EMR-based cohorts of people with MS.

ACCREDITATION:



JOINT ACCREDITATION
INTERPROFESSIONAL CONTINUING EDUCATION

In support of improving patient care, this activity has been planned and implemented by the Consortium of Multiple Sclerosis Centers (CMSC) and Intellisphere, LLC. The CMSC is jointly accredited by the Accreditation Council for Continuing Medical Education (ACCME), the Accreditation Council for Pharmacy Education (ACPE), and the American Nurses Credentialing Center (ANCC), to provide continuing education for the healthcare team.



IPCE CREDIT™

This activity was planned by and for the healthcare team, and learners will receive .5 Interprofessional Continuing Education (IPCE) credit for learning and change.

PHYSICIANS: Physicians: The CMSC designates this journal-based activity for a maximum of .5 *AMA PRA Category 1 Credit(s)*™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

NURSES: The CMSC designates this enduring material for .5 contact hour of nursing continuing professional development (NCPD) (none in the area of pharmacology).

PHARMACISTS: This knowledge-based activity (UAN JA4008165-9999-22-033-H01-P) qualifies for (.5) contact hour (.05 CEUs) of continuing pharmacy education credit.

PSYCHOLOGISTS: This activity is awarded 0.5 CE credits.

SOCIAL WORKERS: As a Jointly Accredited Organization, the CMSC is approved to offer social work continuing education by the Association of Social Work Boards (ASWB) Approved Continuing Education (ACE) program. Organizations, not individual courses, are approved under this program. State and provincial regulatory boards have the final authority to determine whether an individual course may be accepted for continuing education credit. The CMSC maintains responsibility for this course. Social workers completing this course receive .5 continuing education credits.

DISCLOSURES: It is the policy of the Consortium of Multiple Sclerosis Centers to mitigate all relevant financial disclosures from planners, faculty, and other persons that can affect the content of this CE activity. For this activity, all relevant disclosures have been mitigated.

Francois Bethoux, MD, editor in chief of the *International Journal of MS Care (IJMSC)*, has served as physician planner for this activity. He has disclosed no relevant relationships. **Alissa Mary Willis, MD**, associate editor of *IJMSC*, has disclosed no relevant relationships. Authors **Carol Swetlik, MD**, **Riley Bove, MD**, and **Marisa McGinley, DO**, have disclosed no relevant financial relationships.

The staff at *IJMSC*, CMSC, and Intellisphere, LLC who are in a position to influence content have disclosed no relevant financial relationships. **Laurie Scudder, DNP, NP**, continuing education director CMSC, has served as a planner and reviewer for this activity. She has disclosed no relevant financial relationships.

METHOD OF PARTICIPATION:

Release Date: November 1, 2022; Valid for Credit through: November 1, 2023.

In order to receive CE credit, participants must:

- 1) Review the continuing education information, including learning objectives and author disclosures.
- 2) Study the educational content.
- 3) Complete the evaluation, which is available at <https://www.highmarksce.com/mscare>.

Statements of Credit are awarded upon successful completion of the evaluation. There is no fee to participate in this activity.

DISCLOSURE OF UNLABELED USE: This educational activity may contain discussion of published and/or investigational uses of agents that are not approved by the FDA. The CMSC and Intellisphere, LLC do not recommend the use of any agent outside of the labeled indications. The opinions expressed in the educational activity are those of the faculty and do not necessarily represent the views of the CMSC or Intellisphere, LLC.

DISCLAIMER: Participants have an implied responsibility to use the newly acquired information to enhance patient outcomes and their own professional development. The information presented in this activity is not meant to serve as a guideline for patient management. Any medications, diagnostic procedures, or treatments discussed in this publication should not be used by clinicians or other health care professionals without first evaluating their patients' conditions, considering possible contraindications or risks, reviewing any applicable manufacturer's product information, and comparing any therapeutic approach with the recommendations of other authorities.

ABSTRACT

BACKGROUND: The electronic medical record (EMR) has revolutionized health care workflow and delivery. It has evolved from a clinical adjunct to a multifaceted tool, with uses relevant to patient care and research.

METHODS: A MEDLINE literature review was conducted to identify data regarding the use of EMR for multiple sclerosis (MS) clinical care and research.

RESULTS: Of 282 relevant articles identified, 29 were included. A variety of EMR integrated platforms with features specific to MS have been designed, with options for documenting disease course, disability status, and treatment. Research efforts have focused on early diagnosis identification, relapse prediction, and surrogates for disability status.

CONCLUSIONS: The available platforms and associated research support the utility of harnessing EMR for MS care. The adoption of a core set of discrete EMR elements should be considered to support future research efforts and the ability to harmonize data across institutions.

Int J MS Care. 2022;24(6):287-294. doi:10.7224/1537-2073.2022-066

The electronic medical record (EMR) has changed how information is documented and consumed in all facets of health care. It was originally created for clinical billing, but over time EMR has been refined and is now in daily use, as well as being used in subspecialties and research. An EMR is an electronic version of the medical chart that contains information about a patient's health and condition but often does not transfer between clinicians or health systems. In comparison, an electronic health record is broader, typically viewed and used across multiple providers, with complete information regarding all the patient's medical conditions and health status. In multiple sclerosis (MS) care, EMR tools have become increasingly available, with provider- and patient-reported information readily collected and used for clinical and research applications. In a 2016 survey, 91% of neurology clinicians welcomed the opportunity for MS-specific documentation, and a similar proportion showed interest in extended and interconnected electronic documentation for patients with MS.¹ Platforms specific to MS facilitate patient engagement through patient-reported outcomes and data visualization tools that share relevant health care information. In addition to aiding clinical care, EMR

facilitates research studies when data entry is accurate, consistent, and comprehensive. These high-quality data regarding patient clinical status, symptoms, and quality of life capture longitudinal information that may aid in early diagnosis, monitor response to treatment, anticipate risk of relapse, and identify patients for clinical trial participation.^{2,3} However, clear and current characterization of available tools and their effect on clinical care and research efforts has not been described. This review highlights available EMR tools, their current use in clinical care, and how they may improve our characterization and understanding of MS.

METHODS

Literature Search Strategy

Ovid MEDLINE was searched for articles from inception to April 27, 2022. Keywords and Medical Subject Headings (MeSH) related to MS, EMR, and point-of-care tools were used in combination to search for articles (TABLE S1, available online at IJMSC.org). Additional articles were identified by the authors based on their knowledge of and review of citations in retrieved articles.

Selection Criteria and Data Extraction

The inclusion criteria for the narrative review included description of the data entry tools for the EMR, applications of these tools, and the clinical impact of these tools to evaluate characteristics of patients with MS, including diagnosis, progression, relapse, disability, and comorbidities. Papers were abstracted into a standardized spreadsheet with title, citation, authors, and abstract included. Review of this information was performed independently by 2 authors (C.S. and M.M.). Authors abstracted agreement for inclusion independently, and overlapping articles were included in the final narrative review. Abstracts without full-length accompanying articles were excluded.

RESULTS

Of 282 articles identified, 29 were included that reported prospective validation studies of EMR tools, large observational studies, and/or review articles discussing the use of real-world EMR data.

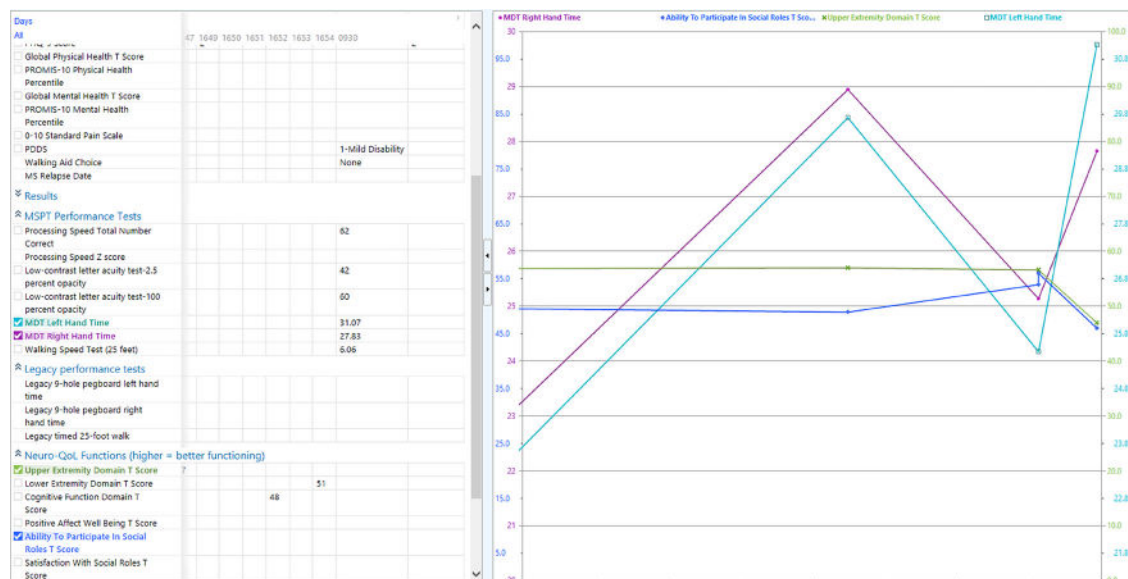
Standardized Data Entry Tools for the EMR

An early innovation in standardized data entry tools was the creation of a nationwide, EMR-embedded MS-based registry associated with a national health care system (Department of Veterans Affairs [VA]). The VA MS Surveillance Registry (MSSR) is one of the oldest

From the Mellen Center for Multiple Sclerosis, Cleveland Clinic, Cleveland, OH, USA (CS, MM); and the UCSF Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA, USA (RB). Correspondence: Marisa McGinley, DO, MSc, Mellen Center U-10, Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44195; email: mcginlm@ccf.org.

Note: Supplementary material for this article is available at IJMSC.org.

© 2022 Consortium of Multiple Sclerosis Centers.

FIGURE 1. Example of MSPT Results in an EMR Incorporating Patient-Reported Outcomes and Neuroperformance Tests

EMR, electronic medical record; MSPT, Multiple Sclerosis Performance Test.

EMR-based cohorts, comprising patients who have been followed since the mid-1990s.⁴

Newer EMR tools have since been designed, including the Cleveland Clinic's Multiple Sclerosis Performance Test (MSPT). The MSPT is a tablet-based battery of patient-reported outcomes and performance tests, including walking speed, manual dexterity, processing speed, and contrast sensitivity.⁵ The MSPT interfaces with a secure, cloud-based system to allow for integration into the EMR. This interface consists of an application programming interface that conforms to Health Level 7 standards using Fast Healthcare Interoperability Resources and Clinical Document Architecture standards. This type of application programming interface uses encryption (ie, AES-256) to allow for bidirectional communication between the MSPT device and the EMR.⁶ This allows the data to instantaneously integrate with the EMR and be inserted into note templates to allow for real-time utilization for all office visits (FIGURE 1). High rates of patient completion overall have been noted, particularly among younger patients and patients with less severe disability.⁵ In addition, the integration of this technology facilitates more consistent data capture of neuroperformance tests and patient-reported outcomes, as completion rates in all testing domains increased after MSPT implementation compared with before MSPT implementation.⁵ After MSPT implementation, follow-up patient completion rates increased significantly over time, from 13.9% of patients completing the MSPT in December 2015 to 77.2% completing in May 2018 ($P < .001$), with an increase in completion of walking and manual dexterity measures in particular. The MSPT is used at all Cleveland Clinic MS outpatient clinical encounters.

The NorthShore Health System in the Chicago (Illinois) area designed a tool kit that includes measures to assess anxiety, disability, fatigue, motor function, and cognition.⁷ In addition to disease-specific data entry, note generation can also be customized. Clinical practice is streamlined through discrete data entry for progress notes rather than free text. In addition, clinicians receive notifications for quality improvement based on collected data. For example, entering a patient who screens positive for symptoms of depression or anxiety will generate an alert to address these symptoms through medication, specialist referral, or deferring. Early evaluation confirms overall high tool kit use by physicians and overall little missing data.⁷

The MS NeuroShare platform at Sutter Health in Northern California is another example of a dual-facing platform, viewable to clinician and patient, that displays a patient's data relevant to MS in the EMR.⁸ Simple displays, such as laboratory values, were found to be used more than those that required additional entered information, such as the Timed 25-Foot Walk test (T25FW).

Research Applications of EMR-Based Tools

The creation of a "learning health system" has been proposed to facilitate research efforts beyond individual clinical care. An example of a learning health system in MS is the MS PATHS network.⁹ Composed of 10 health care institutions under a shared governance model, each institution contributes data captured during routine clinical care, including patient-completed measures from the MSPT, quantitative imaging measures based on brain magnetic resonance imaging, and DNA, RNA, and serum biobanking.

Studies comparing patients with MS identified in the EMR compared with gold standard research data have been conducted and show that patients with MS can be accurately identified via EMR-based searches.¹⁰ A recent validation study compared 4142 patients with MS with 323 research patients who also received clinical care at the University of California San Francisco to assess concordance with disability measures, including the Expanded Disability Status Scale (EDSS), T25FW, and disease subtype. The data captured in the clinical chart generally matched research data.¹⁰

Diagnosis

Recent efforts have focused on characterizing the prevalence and incidence of MS on a national level. Algorithms to accurately identify MS in administrative health care data sets have been developed using *International Classification of Diseases, Ninth (ICD-9)* and *Tenth (ICD-10) Revision* codes and MS disease-modifying therapy (DMT) use within 1 year.¹¹ These algorithms were applied, with physician-adjudicated MS cases as the reference standard, to generate an estimated 2010 prevalence of MS in the US adult population, accumulated over 10 years, of 309.2 per 100,000, with more than 700,000 cases of MS in the United States.¹²

In EMR-based investigations, a prodromal phase before diagnosis of suspected MS-related symptoms has been proposed, particularly in patients with relapsing-remitting MS. In a retrospective, multicenter study, a mean of 3.14 health care contacts per patient per year were found, with a median of 6 prescriptions and 6 ancillary tests in the 5 years before MS diagnosis. Although a comparable age-matched control cohort was not available, back pain, myalgias, joint pain, gastrointestinal symptoms, sensory symptoms, and psychiatric symptoms were all found to present in the 5 years before a classical demyelinating event and subsequent MS diagnosis.¹³ Similarly, natural language processing (NLP) tools have been explored in smaller studies attempting to use computerized programs to identify and mine clinician-entered narrative signs and symptoms in the chart in an effort to diagnosis MS.¹⁴ Based on NLP, as many as 40% of patients with MS could be correctly identified as such before entrance of an ICD-9 code in their chart. Challenges arose in discriminating MS from other known neurologic disorders, including stroke, neuropathy, or migraines, as 90% of patients flagged for potential MS were found to have an alternative diagnosis based on a gold standard ICD-9 code comparator, creating a posttest probability of 10%.

Data from the EMR have been imbedded into the Scalable Precision Medicine Oriented Knowledge Engine (SPOKE) to obtain high-dimensional health status profiles and identify individuals at risk for MS.¹⁵ A knowledge graph that uses 16 nodes of more than 3 million types, SPOKE considers biological mechanisms in the setting of patient-specific health data analytics.

High-dimensional individual health status profiles can then be obtained from SPOKE (SPOKEsigs), allowing a random forest classifier to identify individuals at risk for MS up to 5 years before their documented diagnosis based on data extracted from the University of California San Francisco EMR. Similarly, machine-based learning principles, which rely on systems learning and improving from experience without direct reprogramming, have been applied to promote early identification of approximately half of patients with MS through inclusion of highly relevant variables related to MS, with a specificity of 91.3%.¹⁶ These studies suggest that statistical modeling using EMR data has the potential to aid in the early diagnosis of MS.

Relapses

Relapse activity prediction tools have also been developed to characterize disease activity. Using EMR data linked with research data from the Comprehensive Longitudinal Investigation of Multiple Sclerosis at Brigham and Women's Hospital (CLIMB) cohort, a training set of 1435 patients was created, as well as a validation set of 186 patients with EMR-only data. In the training set, EMR-based demographic and clinical information was extracted, and NLP was applied to free-text clinical narratives, such as outpatient encounters, radiology reports, and discharge summaries. The study's aim was to predict the future risk of relapse within 1 year using previous relapse history based on EMR data without requiring actual relapse history. Consistent with previous literature, the final model used predictors of age, disease duration, and number of relapses in the previous year, avoiding use of DMTs and imaging results, to create a parsimonious and timeless model independent of changing therapeutic strategies.¹⁷ Other clinical tools measuring severity of disability, such as the Multiple Sclerosis Severity Score (MSSS) and radiographic features such as brain parenchymal fraction (BPF), were also extracted from the EMR and compared with the research-grade CLIMB data. Training and testing data sets were used to develop a BPF algorithm, which was then tested with a validation data set. The performance of the BPF algorithm in the validation set was reduced from a mean R^2 of 0.226 to 0.016 when only codified EMR variables (eg, prescriptions, demographics, and billing codes for diagnostic procedures) were included. It was reduced to 0.000760 when only EMR narrative variables from NLP (eg, symptoms, signs, medications, magnetic resonance imaging reports, and the treating neurologist's impression) were included. Thus, neither type of EMR data alone was sufficient to produce an estimate of BPF. When the BPF algorithm included only sex, age at symptom onset, and disease duration, a mean R^2 of 0.2860 was generated, suggesting that the existing EMR variables are not informative enough to be a surrogate measure of BPF. However, EMR-derived MSSS was able to differentiate

patients with relapsing-remitting disease and those with progressive disease. Considering differences in sex, age at symptom onset, and disease duration, patients with primary and secondary progressive MS had a higher mean EMR-derived MSSS than those with relapsing-remitting MS on the validation set (with an observed MSSS of 4.36 compared with a derived MSSS of 3.22 for patients with progressive disease and an observed MSSS of 0.73 compared with a derived MSSS of 1.10 for patients with relapsing-remitting disease).¹⁸

Disability Level

Surrogate measures for the EDSS have been evaluated. In a study of 1599 US-based patients, key symptom terms related to EDSS domains were mapped to EDSS scores and then searched for across the problem list and the vitals tables, including T25FW results and orders for medical equipment. A severity score was assigned to each term based on expert opinion, and ICD codes, procedural codes, and orders for durable medical equipment were mapped to *Current Procedural Terminology* codes with the aim of assessing pharmacy and medical costs based on severity.¹⁹ Ultimately, formal validation of the true patient EDSS scores was unable to be achieved because few patients had an EDSS score in the chart, although stratification of associated medical cost based on estimated severity of disability was confirmed.

Truong et al²⁰ also tried to estimate EDSS scores by mapping Kurtzke Functional Systems Scores to generate EDSS scores with a retrospective cohort of patients with MS from the Intermountain Healthcare Provider-Payer Integrated Delivery Network. This study mapped the components of the Kurtzke Functional Systems Scores to ICD-9-CM codes to generate EDSS scores and calculate the change in EDSS scores to create a tool that could identify patients with MS with disability progression and quantify MS disability using administrative claims. Progression measures were limited by a lack of information regarding patient phenotype (relapsing-remitting vs progressive MS), and true validation against EDSS scores was not performed.

Both these studies used health claims databases as the study cohorts. Application of these approaches to the EMR itself could allow for estimation of disability based on procedures, equipment orders, and visit-associated ICD codes in lieu of formal EDSS testing, while providing access to other data available in the EMR.

Machine learning principles have also been applied to MSBase, an observational, international MS cohort. When a more complete history of progression was included in the model, it showed significant improvement in predicting disability progression, with an area under the curve of 0.85.²¹ MSBase is not an EMR-integrated database, but future efforts to integrate the elements included in this registry with the EMR may facilitate accurate disability and progression status.

PRACTICE POINTS

- » Multiple sclerosis–based tool kits and platforms allow disease-specific documentation of motor, cognitive, and neuropsychiatric symptoms that can inform clinical care over time.
- » Research efforts have used the electronic medical record as a tool for early diagnosis, as well as detection of relapses, disease course, and comorbid conditions.

Natural language processing has also been used to predict EDSS score. In a recent study by Yang et al,²² NLP was used to calculate an EDSS score from the clinic notes about patients with MS. Prediction of total EDSS score was overall superior to prediction of EDSS subscores and, ultimately, performance improved when considering notes with known values of the EDSS subscores.²² Beyond machine learning and NLP efforts to calculate the EDSS, efforts have been made to incorporate validated patient-reported disability outcomes, including Patient-Determined Disease Steps scale and patient-reported EDSS scores, into the EMR.^{23,24} These patient-reported outcomes have the potential to address clinician time constraints and facilitate longitudinal monitoring of disability.

Complications and Associated Comorbidities of MS

Infection risk and accurate identification of infections in patients with MS was investigated via EMR-based algorithms. An algorithm was applied to the charts of a cohort of 6000 patients with MS, with 30,000 age-, sex-, and race-matched controls, to identify patients with MS by ICD-9 code 340 and ICD-10 code G35 and by dispensed DMTs using a previously validated algorithm. Random sample chart abstractions were performed to define discrete infectious episodes, and algorithms with the highest positive predictive value were identified using ICD codes to query for select outpatient infections, including herpes simplex virus, varicella zoster virus, and urinary tract infection.^{11,25} When the positive predictive value of the algorithm was less than 70%, laboratory data and prescriptions were included in the algorithm. Random charts (n = 20) from the patients with MS and the control patients were reviewed by trained professionals to validate the algorithms at each step. This study functioned as a proof-of-concept for applying ICD-based algorithms to identify patient groups and to evaluate outcomes, with a gold standard of chart review as a comparison. The positive predictive values of the algorithms were between 80% and 100% in patients with MS and between 75% and 100% in the general population, with no significant differences between the groups. This method could potentially be applied to other adverse effects of MS treatment and comorbidities.

TABLE 1. Proposed Core Set of Discrete Elements for MS-Tailored EMR Platforms

Category	Discrete elements
Disease characteristics	<ul style="list-style-type: none"> • ICD-10 code • Initial disease course • Present disease course • Relapse history • Relapse treatment
Imaging	<ul style="list-style-type: none"> • Date and type of imaging study • MRI findings: new T2 lesions, T1 gadolinium-enhancing lesions
Disability measures	<ul style="list-style-type: none"> • Walking speed (eg, Timed 25-Foot Walk) • Manual dexterity test (eg, Nine-Hole Peg Test) • Processing speed test (eg, Symbol Digit Modalities Test) • Low-contrast letter acuity (eg, Sloan) • EDSS and/or PDDS scale
DMTs	<ul style="list-style-type: none"> • Past DMTs • Current DMT • Duration of use

DMT, disease-modifying therapy; EDSS, Expanded Disability Status Scale; EMR, electronic medical record; ICD-10, *International Classification of Diseases*, Tenth Revision; MRI, magnetic resonance imaging; MS, multiple sclerosis; PDDS, Patient-Determined Disease Steps.

Suggested Discrete Elements for EMR Collection

As noted previously herein, factors that are discretely collected in the EMR have an advantage over free text or nonstandardized collection. Although free text can permit nuanced documentation, omission of pertinent details and nonstandardized recording are likely. Discrete element collection permits organized data collection and viewing, facilitates patient-centered care, and fuels research efforts. Based on our review of the literature, EMR platforms designed specifically for patients with MS should consider capturing a core set of discrete patient and disease characteristics, including imaging findings, disability measures, and treatment use (TABLE 1). Patient-reported outcomes and symptom screening results should be collected and displayed in tandem with provider-entered characteristics to allow data to be viewed and acted on. All components should be clearly associated with date of collection and, ideally, with options for graphic display of trends over time. Secure and streamlined data exportation should be prioritized to permit research and quality improvement projects.

DISCUSSION

In MS care and research, real-world data have been used to augment the data available from randomized controlled trials. Historically, the data used for observational studies was from large administrative claims databases and MS registries, both distinct from the EMR. Each system came with benefits and limitations. Administrative claims databases leverage large numbers of patients but have limited clinical information and often lack information regarding treatments and outcomes. The MS

registries were created to curate important disease-specific information regarding treatment and clinical outcomes, but they rely on data entry in a separate system from the EMR. The process requires time, and the information is typically available only for research purposes and cannot be used to inform clinical care. The EMR contains both free text and discrete data elements, enabling detailed characterization of patients during clinical visits without duplicative data entry.

The VA MSSR was an early effort to pair a disease state registry with a national health care system to provide registry-level granularity with collection of demographic characteristics, disease course, and treatments in tandem with a large cohort size.²⁶ Integrating the VA MSSR across other MS registries and with other EMRs remains challenging, and continued and standardized data collection of physician and patient-reported outcomes is needed. In such a large system, regular updates to accommodate new best practices and the integration of these updates across EMRs may be cumbersome. Nevertheless, this early integration of the EMR with disease state identification has fueled subsequent efforts. Tool kits can promote the entry and integration of complex tasks specific to MS across domains of function and symptoms. More recent tool kits, such as the MSPT, are designed to facilitate implementation of guidelines and promote structured clinical documentation, more ideal for EMR-based research and promoting best practice care.⁵ Patient-reported data readily viewable to both parties, such as that in MS NeuroShare, have the additional benefits of promoting integration of the patient's perspective, reducing clinician data entry burden, and enhancing a patient's understanding of and involvement in care.⁷ The standardized EMR data collection methods of the NorthShore platform allow for the succinct review of individual patient data during visits while also being easily exported in aggregate for use in observational studies to inform population-based management of patients. These efforts have the potential to facilitate pragmatic trials and improve clinical care through the standardization of office visits.

Challenges can arise when attempting to use disease-specific EMR tools for clinical care. Some of these were identified during the implementation of MS NeuroShare and should be anticipated for future platforms.⁸ A key observation is the importance of workflow for both patients and clinicians. For patients, the request to complete questionnaires before clinical visits is often dismissed or overlooked, leading to missing data or delayed data entry.⁸ For clinicians, streamlined data curation leads to more effective data use. Even when data are more efficiently curated, it requires time to review the information with patients to ensure patients appreciate the value in the data collected and see that it has a meaningful clinical use.

Although the primary aim of the EMR remains patient care, research has evolved rapidly as a secondary gain.

Efforts to validate data collection in the EMR through prospective studies facilitate the generation of real-world evidence. The potential of these data includes characterization of longitudinal disease outcomes, analysis of large sample sizes, integration of clinician- and patient-reported outcomes, validation of outcome measures, creation of mean trajectories according to phenotype, real-world safety monitoring, and health care utilization trends. The platforms in this review, although designed in a clinical context, support these and other research aims. The potential of EMR observational studies is the generation of real-world evidence without requiring maintenance of a large, enrolled, prospective longitudinal cohort. Ideally, given time constraints in outpatient visits, surrogates for motor function, mood, and other domains of function would be identified and allow for imputation of functional status across these domains when more formal testing (eg, EDSS) is not recorded or not possible.

As described previously herein, notable efforts have been made to describe the disease course of MS via chart-based studies to promote early diagnosis, quantify progression of disability, and identify relapses. Some of the challenges are aging cohorts (in which relapse frequency may be expected to decline) and changes in diagnostic criteria, diagnostic approaches, and available treatments. In addition, EMR data extraction methods vary; although manual review of charts may permit more exact collection, it is time-consuming and still error-prone.

Use of chart-based methods, as described in the diagnostic algorithms, often requires patients to have repeated encounters with the health care system, the correct use of ICD codes in their charts, and, often, an already-initiated DMT. These steps may occur relatively late in the diagnostic process, and, therefore, there is a need for methods to accurately capture patients earlier in their disease course or even before its beginning.

Because ICD codes do not typically discriminate among different types of MS, accurate phenotype characterization has also been attempted based on EMR data. In a VA system-based study, NLP was used to search EMR clinical notes and generate phenotype. Unfortunately, although the NLP algorithm was generally successful, phenotype was documented on only one-third of patients' charts, pointing to ongoing challenges in identifying patients with progressive MS and the historical concern that this documentation may potentially limit access to DMTs. In these situations that are challenging even on a clinical basis, medication regimens and patterns of health care utilization may remain important phenotype proxies that NLP may be able to identify.²⁷

A substantial benefit of EMR-based cohorts is that they overcome the enrollment biases of clinical trials, which notoriously overrepresent younger, White individuals with mild disease. Therefore, the clinical insights generated from EMR-based cohorts, although still capable of substantial biases,²⁸ may be more representative of the

diverse population of individuals living with MS.^{6,10} As with all observational data, challenges arise regarding data quality and completeness. If use of the platforms is not required for each patient or not used by each clinician, utility of the data set to answer questions dependent on these platforms decreases. Similarly, if information pertinent to a question is not included in the platform or readily exported from the EMR (eg, imaging findings), timely data collection may be challenging.^{3,29} Finally, especially for data collection across health systems, substantial funding will be required to incentivize harmonized, consistent data collection. Patients who do not seek care or are underserved in clinical outpatient populations will still not be adequately represented, even with use of EMR-based data.

CONCLUSION

In general, although many EMR tools may be implemented with a primary or secondary aim of supporting guideline-based care,⁷ deployment of EMR and its associated tools has preceded design guidelines and best practices. Platform intent and health care system customization with options to integrate across other departments, disease states, and external systems is optimal for data pooling. As described previously herein, successful EMR tools are still relatively novel, many with less than a decade of use, and, therefore, long-term outcomes regarding followed cohorts are not readily available. Older cohort studies are smaller, with registry- or clinical trial-based designs, and may not capture heterogeneity as well or emphasize patient-reported outcomes across all domains of function.

At this time, tools such as NLP show promise regarding facilitating data extraction, but manual data collection still has some additional benefits. Prediction of an MS diagnosis using the EMR shows promise, although capturing discrete EMR elements that can predict disability and relapse risk and correlate with radiographic findings is more challenging. Given the growing number of large databases, ability to discern phenotype and selectively mine discreet groups of patient data may permit more rapid recruitment for clinical trials outside of individual office encounters. Future studies could focus on the emergence of health care disparities not as readily captured in structured research-based environments with similar levels of access to care across participants. EMR-based tools have the potential to lead to better characterization of MS, improve care delivery, improve understanding of MS in underrepresented populations, and assist clinical trial recruitment. □

FINANCIAL DISCLOSURES: Dr Bove is a recipient of the National Multiple Sclerosis Society Harry Weaver Award; has received research support from the National Multiple Sclerosis Society, the Hilton Foundation, the California Initiative to Advance Precision Medicine, the Sherak Foundation, Akili Interactive,

and Roche Genentech; and has received personal compensation for consulting from Alexion, Biogen, EMD Serono, Novartis, Pear Therapeutics, Roche Genentech, and Sanofi. Dr McGinley has served on scientific advisory boards for EMD Serono, Genzyme, and Genentech; consulted for Octave; received research funding from Novartis and Biogen; and receives funding via a KL2 grant (KL2TR002547) from the Clinical and Translational Science Collaborative of Cleveland via the National Center for Advancing Translational Sciences component of the National Institutes of Health. Dr Swetlik declares no conflicts of interest.

FUNDING/SUPPORT: None.

REFERENCES

- Kern R, Haase R, Eisele JC, Thomas K, Ziemssen T. Designing an electronic patient management system for multiple sclerosis: building a next generation multiple sclerosis documentation system. *Interact J Med Res.* 2016;5(1):e2. doi:10.2196/IJMR.4549
- Kalincik T, Butzkueven H. Observational data: understanding the real MS world. *Mult Scler.* 2016;22(13):1642-1648. doi:10.1177/1352458516653667
- Cohen JA, Trojano M, Mowry EM, Uitdehaag BMJ, Reingold SC, Marrie RA. Leveraging real-world data to investigate multiple sclerosis disease behavior, prognosis, and treatment. *Mult Scler J.* 2020;26(1):23-37. doi:10.1177/1352458519892555
- Wallin MT, Whitham R, Maloni H, et al. The Multiple Sclerosis Surveillance Registry: a novel interactive database within the Veterans Health Administration. *Fed Pract.* 2020;37(Suppl 1):S18.
- Macaron G, Moss BP, Li H, et al. Technology-enabled assessments to enhance multiple sclerosis clinical care and research. *Neurol Clin Pract.* 2020;10(3):222. doi:10.1212/CPJ.0000000000000710
- Rhodes JK, Schindler D, Rao SM, et al. Multiple sclerosis performance test: technical development and usability. *Adv Ther.* 2019;36(7):1741-1755. doi:10.1007/S12325-019-00958-X
- Claire Simon K, Hentati A, Rubin S, et al. Successful utilization of the EMR in a multiple sclerosis clinic to support quality improvement and research initiatives at the point of care. *Mult Scler J Exp Transl Clin.* 2018;4(4). doi:10.1177/2055217318813736
- Bove R, Bruce CA, Lunders CK, et al. Electronic health record technology designed for the clinical encounter: MS NeuroShare. *Neurol Clin Pract.* 2021;11(4):318-326. doi:10.1212/CPJ.0000000000000986
- Mowry EM, Bermel RA, Williams JR, et al. Harnessing real-world data to inform decision-making: Multiple Sclerosis Partners Advancing Technology and Health Solutions (MS PATHS). *Front Neurol.* 2020;11. doi:10.3389/FNEUR.2020.00632
- Damotte V, Lizée A, Tremblay M, et al. Harnessing electronic medical records to advance research on multiple sclerosis. *Mult Scler.* 2019;25(3):408-418. doi:10.1177/1352458517747407
- Culpepper WJ, Marrie RA, Langer-Gould A, et al. Validation of an algorithm for identifying MS cases in administrative health claims datasets. *Neurology.* 2019;92(10):e1016-e1028. doi:10.1212/WNL.00000000000007043
- Wallin MT, Culpepper WJ, Campbell JD, et al. The prevalence of MS in the United States. *Neurology.* 2019;92(10):e1029-e1040. doi:10.1212/WNL.00000000000007035
- Jorge A, André A, Rocha AL, et al. Defining the prodromal phase of multiple sclerosis based on healthcare access in a Portuguese population: ProdMS study. *Mult Scler Relat Disord.* 2021;55:103154. doi:10.1016/j.msard.2021.103154
- Chase HS, Mitrani LR, Lu GG, Fulgieri DJ. Early recognition of multiple sclerosis using natural language processing of the electronic health record. *BMC Med Inform Decis Mak.* 2017;17(1):24. doi:10.1186/s12911-017-0418-4
- Nelson CA, Bove R, Butte AJ, Baranzini SE. Embedding electronic health records onto a knowledge network recognizes prodromal features of multiple sclerosis and predicts diagnosis. *J Am Med Informatics Assoc.* 2022;29(3):424-434. doi:10.1093/jamia/ocab270
- Wang R, Luo W, Liu Z, et al. Integration of the Extreme Gradient Boosting model with electronic health records to enable the early diagnosis of multiple sclerosis. *Mult Scler Relat Disord.* 2021;47. doi:10.1016/j.msard.2020.102632
- Ahuja Y, Kim N, Liang L, et al. Leveraging electronic health records data to predict multiple sclerosis disease activity. *Ann Clin Transl Neurol.* 2021;8(4):800-810. doi:10.1002/acn3.51324
- Xia Z, Secor E, Chibnik LB, et al. Modeling disease severity in multiple sclerosis using electronic health records. *PLoS One.* 2013;8(11):e78927. doi:10.1371/journal.pone.0078927
- Berkovich R, Fox E, Okai A, et al. Identifying disability level in multiple sclerosis patients in a U.S.-based health plan claims database. *J Med Econ.* 2021;24(1):46-53. doi:10.1080/13696998.2020.1857257
- Truong CTL, Le HV, Kamau AW, et al. Creating a real-world data, United States healthcare claims-based adaptation of Kurtzke Functional Systems Scores for assessing multiple sclerosis severity and progression. *Adv Ther.* 2021;38(9):4786-4797. doi:10.1007/S12325-021-01858-9
- De Brouwer E, Becker T, Moreau Y, et al. Longitudinal machine learning modeling of MS patient trajectories improves predictions of disability progression. *Comput Methods Programs Biomed.* 2021;208:106180. Published correction appears in *Comput Methods Programs Biomed.* 2022;213:106479.
- Yang Z, Pou-Prom C, Jones A, 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). doi:10.2196/25157
- Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of Patient Determined Disease Steps (PDDS) scale scores in persons with multiple sclerosis. *BMC Neurol.* 2013;13(1):37. doi:10.1186/1471-2377-13-37
- Romeo AR, Rowles WM, Schleimer ES, et al. An electronic, unsupervised patient-reported Expanded Disability Status Scale for multiple sclerosis. *Mult Scler J.* 2021;27(9):1432-1441. doi:10.1177/1352458520968814
- Smith JB, Li BH, Gonzales EG, Langer-Gould A. Validation of algorithms for identifying outpatient infections in MS patients using electronic medical records. *Mult Scler Relat Disord.* 2022;57. doi:10.1016/j.msard.2021.103449
- Flachenecker P, Buckow K, Pugliatti M, et al. Multiple sclerosis registries in Europe: results of a systematic survey. *Mult Scler.* 2014;20(11):1523-1532. doi:10.1177/1352458514528760
- Nelson RE, Butler J, LaFleur J, Knippenberg K, Kamau AWC, DuVall SL. Determining multiple sclerosis phenotype from electronic medical records. *J Manag Care Spec Pharm.* 2016;22(12):1377-1382. doi:10.18553/jmcp.2016.22.12.1377
- Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science.* 2019;366(6464):447-453. doi:10.1126/science.aax2342
- Scott PJ, Dunscombe R, Evans D, Mukherjee M, Wyatt JC. Learning health systems need to bridge the "two cultures" of clinical informatics and data science. *J Innov Health Inform.* 2018;25(2):126-131. doi:10.14236/JHI.V25I2.1062