Table S1. Selected examples of other real-world data cataloguing efforts

Initiative name	Characteristics of the initiative	References and/or links
European Register for Multiple Sclerosis (EuReMS)	● MS specific ● Coverage: Europe ● Provided an overview of data collected, governance, operational and other aspects, which provides a useful basis to analyse large data sets from PwMS	3-6 https://emsp.org/resources/euremss-report/
Landscaping of MS patient cohorts and registries workshop 2018	● MS specific ● Coverage: global ● Provides strategic overviews of many emerging and existing initiatives on e.g., accessibility, size, geographical catchment	2
European medical information framework (EMIF)	● Organised around “communities”,e.g., electronic health record data, Alzheimer's disease cohorts, vaccine benefit-risk data, psychopathology, multiple sclerosis, ... ● Coverage: global, with a focus on Europe ● Allows an infrastructure to easily host different catalogues based on a community-specific “fingerprint”	10, 11 https://emif-catalogue.eu
ConcePTION	● Pregnant and breast-feeding women (not MS-specific) ● Coverage: Europe ● Creating the first Europe-wide breast milk biobank	12 https://www.imi.europa.eu/projects-results/project-factsheets/conception
European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Resources Database	● Disease agnostic ● Coverage: Europe ● Aims, amongst others, to facilitate access of electronic databases in Europe to researchers, by inviting data custodians to provide descriptions of their core data	13, 14 https://www.ema.europa.eu/en/partners-networks/networks/european-network-centres-pharmacoepidemiology-pharmacovigilance-encepp
Secondary Progressive MS Research Collaboration Network (SPMS RCN) cataloguing effort	● RRMS/SPMS patients ● collaboration between 8 European MS Registries ● Aim: generate data for RWE studies on SPMS, address challenges and improve SPMS data collection	7, 8
Maelstrom Research Catalogue	● Different networks and data collection efforts, including MS ● Coverage: US focused MS cohorts ● Contains comprehensive information about epidemiological research networks and studies, and the data they have collected. It also provides information about harmonized data generated by these research networks	9, 15 https://www.maelstrom-research.org/network/msmdc https://www.maelstrom-research.org/page/catalogue
HealthData.gov	● Disease agnostic ● Coverage: US ● Data on a wide range of topics, including environmental health, medical devices, Medicare & Medicaid, social services, community health, mental health, and substance abuse	16 https://healthdata.gov/

Note: Reference information for the numbered citations can be found in the reference list of the article.

Table S2. **Domains currently included in the questionnaire**

Domain	**Covering**
1. Organizational Information	Name of registry, Name and type of organization keeping the registry, Contact person, Financing
2. Background / Purpose	Initiation year, Aim/Purpose, Target population, Sources of data collection
3. Inclusion criteria	Patient and centre inclusion criteria
4. Documentation process	Person performing the documentation, Data collection manner, Longitudinal surveillance supported, Language of registry
5. Data that are collected:	● Personal data ● Basic disease data (e.g. Disease course, Time of disease onset...) ● Relapses (e.g. Date of relapse, relapse symptoms...) ● Disability ● Cognition scales ● Treatments (relapse therapy, DMTs, symptomatic) ● MRI ● Paraclinical measures ● Patient-derived measurements ● Depression ● Fatigue ● Comorbidities ● Socioeconomic status ● Societal services ● Healthcare utilization ● e-health technologies ● COVID-19 & MS (do you collect data on COVID-19&MS? Patient- or clinician-reported?)
6. Quality Control	Quality control mechanisms, triggers of data entry
7. Governance	Informed consent, access to data
8. Status of the registry	Estimated number of currently registered patients and visits

Abbreviations: COVID-19: coronavirus disease 2019; DMTs: disease-modifying therapies; MS: multiple sclerosis

Table S3. **Overview of collected data and its collection coverage by the 38 registries that are part of the MSDA Catalogue**

Coverage is shown per variable and the average coverage per category

Key: 

The coverage (0%-100%) is shown by using a color scale (red to green) and the length of the colored bars gives an impression of the percentage of the coverage.

Abbreviations: COVID-19: Coronavirus disease 2019; DMTs: disease-modifying therapies; MRI: magnetic resonance imaging; MS: multiple sclerosis

[See next page for Table S3]

Category	Variable	Coverage (%)	Average coverage (%) /category
Personal data	Date of birth	100	99
	Sex	97	
Basic disease data	Disease course	89	83
	Time of disease onset	97	
	Time of diagnosis	95	
	Symptoms at onset	76	
	Past disease activity	71	
	Diagnostic accuracy (McDonald / Poser)	68	
Relapses	Date of relapse	82	71
	Relapse symptoms (please specify)	63	
	Steroid treatment of relapse	76	
	Completeness of recovery	63	
Disability	Expanded Disability Status Score (EDSS)	87	53
	Functional System Score (FSS)	53	
	9-Hole Peg Test (9-HPT)	42	
	Timed 25-Foot Walk (T25-FW)	42	
	MS Functional Composite (MSFC)	42	
Cognition scales	Brief interview for Mental Status (BIMS)	3	14
	Cognitive Performance Scale (CPS)	3	
	Mini-Mental State Examination (MMSE)	11	
	Paced Auditory Serial Addition Test (PASAT3)	29	
	Brain Fog Scale	3	
	Symbol Digital Modality Test (SDMT)	34	
	Brief International Cognitive Assessment for MS (BICAMS)	16	
Treatments	Relapse therapy	82	62
Treatments: DMTs	Past disease-modifying therapies	84	
	Start and end dates of past treatments	76	
	Current disease-modifying therapies	95	
	Start date of current treatment	87	
	Reasons for discontinuation of DMTs	68	
	Treatment satisfaction - patient-reported	21	
	Non-serious Drug side effects, ("AEs")	63	
	Serious drug side effect ("SAEs")	61	
	Treatment adherence	29	
Treatments: Symptomatic treatment	Current symptomatic therapies, medical	61	
	Current symptomatic therapies, non-medical	37	
	Complementary / alternative therapy	37	
MRI	MRI (Y/N)	76	52
	T2/FLAIR lesions	61	
	Number of T2/FLAIR lesions	47	
	Number of new T2/FLAIR since previous scan	45	
	T1 with Gadolinium	61	
	Number of enhancing lesions	50	
	Number of new enhancing lesions since previous scan	45	
	MRI: brain volume	29	
Paraclinical measures	Cerebrospinal fluid (CSF) analysis of cells and intrathecal IgG production	63	34
	CSF biomarkers, other (please provide details)	24	
	Evoked potentials / visual evoked potentials	42	
	Optical coherence tomography (OCT)	18	
	Blood biomarkers (please provide details)	24	
Patient-derived measures	EuroQoL Five Dimensions Questionnaire (EQ-5D)	34	16
	36-Item Short Form Health Survey (SF-36)	18	
	Patient Health Questionnaire (PHQ-9)	8	
	MSIS-29 (MS Impact Scale)	24	
	MS Quality of Life-54 (MSQoL-54)	18	
	MS Quality of Life Inventory (MSQLI)	5	
	Guys Neurological Disability Scale (GNDS)	8	
	Patient derived disease steps (PDDS) please specify	16	
	Patient diary	11	
Depression	Depression please specify	45	45
Fatigue	Fatigue Impact Scale (U-FIS)	13	12
	Daily Fatigue Impact Scale (D-FIS)	3	
	Dutch Fatigue Scale (DuFS)	3	
	Krupp's Fatigue Severity Scale (FSS)	24	
	Modified Fatigue Impact Scale (MFIS)	18	
	Visual Analogue Scale for Fatigue (VAS-F)	11	
Co-morbidities	Chronic diseases (please specify)	79	70
	Co-medication (please specify)	61	
Socio-economic data	Occupation	58	59
	Employment	71	
	Education	61	
	Other socio-economic data (please specify)	45	
Societal services	Access to carer/support person	32	22
	Provision of aids (i. e. walking sticks, wheelchair etc.)	21	
	Other societal services data (please specify)	13	
Healthcare utilization	Hospitalization due to MS	45	30
	Any hospitalizations	29	
	Emergency room visits due to MS	21	
	Any emergency room visits	18	
	Outpatient visits for MS	37	
	Any outpatient visits	32	
e-Health technologies	Collecting data using e-Health technologies = yes	29	29
	Social media	*n=2*	*not taken into account for calculation of the average*
	Mobile apps	*n=8*	
	Wearable devices	*n=5*	
	Technical tools for neurorehabilitation	*n=1*	
COVID-19 and MS	Collecting data on COVID-19&MS	68	68
	Patient-reported	*n=8*	*not taken into account for calculation of the average*
	Clinician-reported	*n=8*	
	Both	*n=10*	
	Part of GDSI = Yes	*n=15*	

Table S4. Overview of selected relevant MS and non-MS initiatives and organizations that are or have been focusing on developing guidelines or recommendations to optimize collaborative research using real-world data sets

Initiative	Additional information	MS-specific
European Medicines Agency (EMA)-Guideline on registry-based studies[19]	Aim: optimize use of registry-based studies as a source of real-world evidence. Recommendations are drafted (comments could be given until December 31, 2020), addressing methodological as well as legal and operational aspects in context of registry-based studies used to support regulatory decision making.	No
EMA Patient Registry Initiative[20] (including MS-specific workshop)[21]	Aim: optimize use of existing registries and facilitate establishment of high-quality new registries. MS-specific workshop of 2017: core common data elements for emerging and existing MS registries were discussed.	Partially
Big MS Data (BMSD) Network core protocol post-authorization safety studies (PASS)[22] (based on EMA guidelines)	Core protocol outlines principles of PASS for MS disease-modifying therapies. Aim: provide a basis for registries to qualify and contribute to improvement of knowledge and treatment of MS, with a specific focus on assessing long-term safety effects.	Yes
Heads of Medicines Agencies (HMA)-EMA Joint Big Data Taskforce[23]	Recommendations to address what are well-recognized challenges if big data is to deliver evidence of suitable strength to support decision making across multiple stakeholders.	No
EMA - European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP)[13,14]	ENCePP Code of Conduct[24] (2011): the Code sets out rules and principles for studies, primarily pharmacoepidemiology and pharmacovigilance studies, with an emphasis on noninterventional PASS.	No
PARENT (PAtient REgistries iNiTiative)[25]	Joint Action cofunded by the European Commission (May 2012-October 2014). Aim: support European Union member states in rationalizing development and governance of interoperable patient registries.	No
PROMS (Patient Reported Outcomes for MS) Initiative[26]	Joint initiative of European Charcot Foundation (ECF) and MS International Federation. Aim: maximize impact of science with and of patient input on health, health care, and quality of life of people affected by MS. Represent a unified view on patient-reported outcomes for MS to people affected by MS, health care providers, regulatory agencies, and Healthcare Technology Assessment Agencies.	Yes
European Health Data Space[27]	Aim: common European Health Data Space will promote better exchange and access to different types of health data. The European Health Data Space will be built on 3 main pillars: a strong system of data governance and rules for data exchange, data quality, and strong infrastructure and interoperability.	No
Maelstrom Research	Aim: provide help to challenges in data documentation, cataloguing,[9] harmonization, integration, and co-analysis. MS-specific effort: US National Society and Consortium of Multiple Sclerosis Centers (CMSC) started a cataloguing effort focusing on MS data initiatives in Northern US region using Maelstrom Research pipeline.[15]	Partially
Multiple Sclerosis Outcome Assessments Consortium (MSOAC)[28]	Formed by National MS Society. Aim: developed a Clinical Data Interchange Standards Consortium (CDISC) standard for MS.	Yes
Magnetic Resonance Imaging in MS (MAGNIMS) study group[29]	Aim: formulate recommendations for implementation of brain and spinal cord atrophy measures in clinical management of people with MS and on directions of future research to improve knowledge in this field.	Yes
European Network for Health Technology Assessment (EUnetHTA)[30]	Aim: support collaboration between European HTA organizations that brings added value to health care systems at European, national, and regional levels. Provides a science-based platform for HTA agencies. Promotes transparency, objectivity, independence of expertise.	No