Multiple Sclerosis Disease States as Identified by Unsupervised Machine Learning on Multimodal Longitudinal Patient Trajectories

P1208

Habib Ganjgahi[1]*, Dieter A. Häring[2]*, Gordon Graham[2], Yang Sun[3], Stephen Gardiner[3], Wendy Su[2], Bernd C. Kieseier[2], Thomas E. Nichols[3], Douglas L. Arnold[4], Robert A. Bermel[5], Heinz Wiendl[6,7]^, Chris C. Holmes[1]^

*Co-first authors; ^co-last authors

[1]Department of Statistics, University of Oxford, Oxford, UK; [2]Novartis Pharma AG, Basel, Switzerland; [3]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK; [4]Brain Imaging Centre, Montréal Neurological Institute and Hospital, McGill University, Montréal, QC, Canada; [5]Department of Neurology, Mellen MS Center, Cleveland Clinic, Cleveland, OH, USA; [6]Department of Neurology, University Hospital Münster, Münster, Germany; [7]Brain and Mind Institute, University of Sydney, Sydney, Australia.

Introduction

- Phenotypes of multiple sclerosis (MS) have been defined[1] and revised[2] based on consensus definitions to describe the clinical disease course, help with communication, prognosticate, design and recruit for clinical trials, and make decisions on treatment
- However, the multifaceted patient experience of MS is not well reflected in the current classification of MS phenotypes, which is based on only two dimensions, i.e., the presence or absence of clinical relapse and disability progression.[1,3] The classical phenotypes (clinically isolated syndrome, relapsing remitting MS [RRMS], secondary progressive MS [SPMS], primary progressive MS [PPMS]) may not have distinct pathophysiological properties and reveal limitations in prognosticating individual disease courses, or treatment response
- Unsupervised machine learning techniques can be used to analyse multidimensional longitudinal patient trajectories from thousands of patients and visits in an unbiased and data-driven way to discover key dimensions describing the disease and to identify homogenous states of MS

Goal

- An evidence-based characterization of MS based on an analysis of multivariate clinical and radiological disease trajectories of patients with MS across the entire disease spectrum using unsupervised machine learning

Objective

- To identify **key dimensions to describe MS** based on an analysis of the covariation of standardized clinical and radiological features
- To discover **MS disease states**, i.e. phases in which the patient trajectories from different patients resemble each other
- To quantify the **transition probabilities** between MS disease states and the effect of disease-modyfying therapy (DMTs) on these transition probabilities

Methods

Patient data

- A total of 8052 MS patients (discovery: 6444; validation: 1608) from the Novartis-Oxford MS (NO-MS) clinical trial database[4], with up to 15 years of follow-up, >120000 visits, three magnetic resonance imaging (MRI) features, and a diagnosis at study entry of either RRMS (n=5794), SPMS (n=1548), or PPMS (n=710) were included in this analysis

Data Analysis

- A scalable unsupervised machine learning method (Factor Analysis followed by Hidden Markov Model; FAHMM) was developed to analyse patient trajectories using longitudinal data from the NO–MS dataset, including:
  - **Clinical data:** expanded disability status scale (EDSS), timed 25-foot walk test (T25FWT), 9-hole peg test (9HPT), paced auditory serial addition test (PASAT) and occurrence of relapse
  - **MRI data:** number of gadolinium (Gd)-enhancing T1 lesions, volume of T2 lesions and normalised brain volume derived from percentage brain volume change
- The FAHMM is agnostic to the classical phenotypes, i.e. the diagnosed subtype of MS (RRMS, SPMS or PPMS) is not used in the modelling
- Dimensions to describe MS were identified using a probabilistic latent variable analysis to exploit shared information between measured variables (i.e., linear combination of the originally measured clinical and MRI variables)
- The FAHMM model was used to discover the number and identity of disease states of MS (Bayesian information criterion). Within a disease state the multimodal feature trajectories are similar for different patients; patients can stay for a period in a specific disease state or they can move back and forth between disease states
- The transition probability between disease states and the impact of treatment (any DMT vs no treatment or placebo) were extracted from the transition matrix of the FAHMM model

Results

The FAHMM identified:

- Three reproducible key dimensions of MS:

Physical disability: associated with EDSS, T25FWT, 9HPT

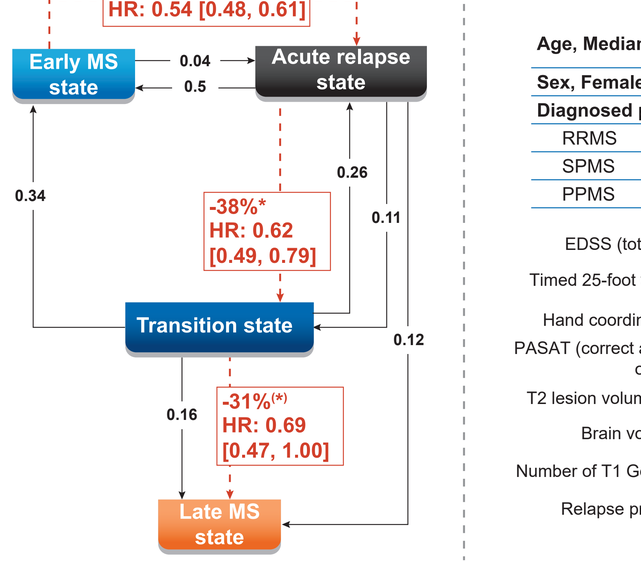
Subclinical burden of disease & cognition: high T2 lesion volume, low normalised brain volume, associated cognitive deficits

Focal inflammation: symptomatic (lesions and relapse), or asymptomatic (lesions but no relapse)

- Nine distinct MS disease states, each differing in terms of the key dimensions of MS, and which can be grouped into four interpretable meta-states based on clinical review (**Figure 1** and **Table 1**)

Figure 1. Multiple sclerosis disease states and impact of treatment on patient transitioning between states

A. Nine states of MS and the transition probabilities between them, grouped in four meta-states



Early MS state
- Patients with early MS can transition back and forth between early states, or into the acute relapse state
- Limited physical disability
- Low subclinical disease burden
- Ongoing subclinical inflammation (but clinically stable)
- Patients with early MS can transition into the acute relapse state

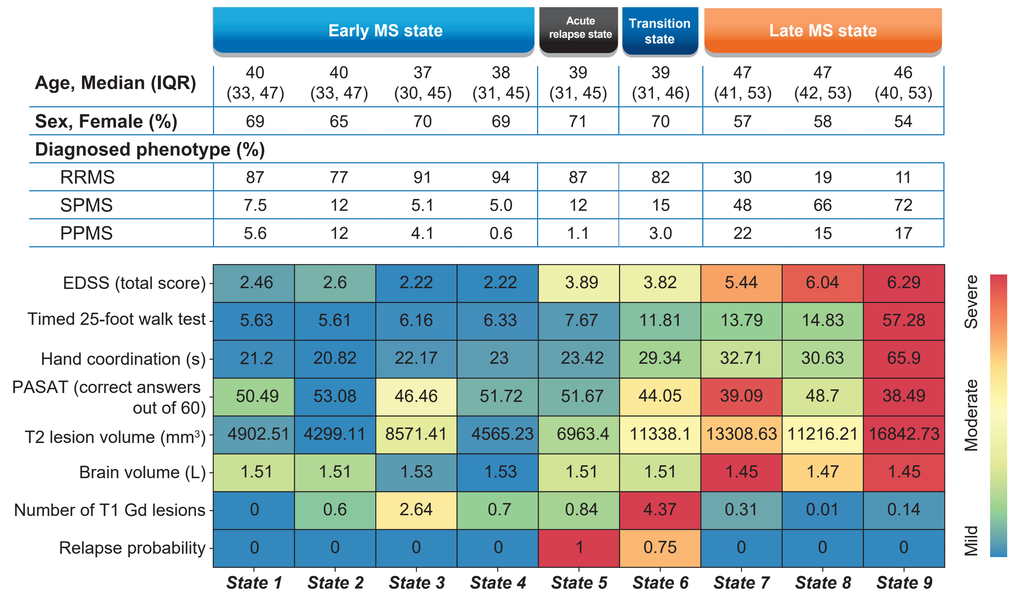
Acute relapse state
- 100% probability to be in a relapse
- High level of subclinical inflammation (MRI lesions)
- Patients can recover to one of the early MS states, or to the transition state, or to late MS

Transition state
- (e.g. PMS patients with superimposed relapses, or RRMS patients with progression)
- Moderate disability (mean EDSS 3.8)
- High subclinical disease burden, often accompanied by cognitive deficits
- High level of focal inflammation (MRI lesions and/or relapses)
- Patients may recover to one of the early states of MS, they may relapse, or transition to late MS

Late MS state
- High level of physical disability
- Cognitive deficits
- Low level of inflammation
- Patients with late MS have low probability of recovery to earlier states of MS

B. The impact of treatment on the transitions between clinically interpretable meta-states of MS

C. The nine states of MS based on the empirical means of the original variables



	Early MS state				Acute relapse state	Transition state	Late MS state		
Age, Median (IQR)	40 (33, 47)	40 (33, 47)	37 (30, 45)	38 (31, 45)	39 (31, 45)	39 (31, 46)	47 (41, 53)	47 (42, 53)	46 (40, 53)
Sex, Female (%)	69	65	70	69	71	70	57	58	54
Diagnosed phenotype (%)									
RRMS	87	77	91	94	87	82	30	19	11
SPMS	7.5	12	5.1	5.0	12	15	48	66	72
PPMS	5.6	12	4.1	0.6	1.1	3.0	22	15	17
EDSS (total score)	2.46	2.6	2.22	2.22	3.89	3.82	5.44	6.04	6.29
Timed 25-foot walk test	5.63	5.61	6.16	6.33	7.67	11.81	13.79	14.83	57.28
Hand coordination (s)	21.2	20.82	22.17	23	23.42	29.34	32.71	30.63	65.9
PASAT (correct answers out of 60)	50.49	53.08	46.46	51.72	51.67	44.05	39.09	48.7	38.49
T2 lesion volume (mm^3)	4902.51	4299.11	8571.41	4565.23	6963.4	11338.1	13308.63	11216.21	16842.73
Brain volume (L)	1.51	1.51	1.53	1.53	1.51	1.51	1.45	1.47	1.45
Number of T1 Gd lesions	0	0.6	2.64	0.7	0.84	4.37	0.31	0.01	0.14
Relapse probability	0	0	0	0	1	0.75	0	0	0
	State 1	*State 2*	*State 3*	*State 4*	*State 5*	*State 6*	*State 7*	*State 8*	*State 9*

Severe / Moderate / Mild

EDSS, expanded disability status scale; FAHMM, factor analysis followed by hidden markov model; Gd, gadolinium; HR, hazard ratio; MS, multiple sclerosis; PASAT, paced auditory serial addition test; PMS, progressive MS; RRMS, relapsing remitting MS; SPMS, secondary progressive MS

A. An illustration of the nine states of MS, grouped into four clinically interpretable meta-states, and the transition probabilities between these states as proposed by the FAHMM (coloured circles and light grey arrows), arrow boldness increases with increased transitioning probability, transition probabilities are thresholded, i.e. only those >8% are shown, for time steps of 1 month; **B.** Black arrows show the probability of transitioning between clinically interpretable meta-states, i.e. to any of the states within the meta-state. Transition probabilities flagged with an asterisk are significant at the 5% level, those asterix within bracket are significant at the 10% level. Treatments were analysed as 'any DMT' (interferon beta-1a, glatiramer acetate, teriflunomide, fingolimod, siponimod or ofatumumab) vs none (i.e. either placebo treatment or no treatment), red dashed arrows (and hazard ratios) show the impact of treatment on transitioning from one state to another; **C.** For ease of interpretation the states were sorted and labelled 1–9 by physical disability score. Subclinical disease burden/damage: disease related small brain volume and high T2 lesion volume.

Table 1. Comparison of current clinical phenotypes of MS versus Factor Analysis followed by Hidden Markov Model (FAHMM) phenotypes of MS

	Clinical phenotypes of MS (Lublin et al., 2014)	FAHMM disease stages (Ganjgahi et al., current poster)
Dimensions to define phenotypes/states	**2** Disability progression (mechanism) Relapse	**3+** Disability (absolute level) Subclinical damage & cognitive deficits Inflammation (T1 Gd lesions, relapses)
Modifiers of phenotypes (applicable to all phenotypes)	**2** Inflammatory activity (MRI lesions) Clinical progression	**0** (none)
Main classification	**Relapsing remitting MS** **Secondary progressive MS** **Primary progressive MS**	**Early MS** **Acute relapse** **Transition state** **Late MS**

Conclusions

- The FAHMM identified three key dimensions to characterize MS: (1) the level of disability, (2) the level of subclinical damage and cognitive impairment, and (3) the level of focal inflammation (relapses and lesions)
- The FAHMM distinguished nine states which can be grouped into four meta-states: four "*early MS*", one "*acute relapse*", one "*transition*", and three "*late MS*" states
- Transition from "*early*" to "*late*" MS only occurs when there is an accumulation of subclinical damage, reducing the patient reserve capacity. Subclinical disease activity should be prevented from the start. The model confirms that DMTs significantly lower the probability of patients to transition from "early" to "late MS"
- Within "late MS" FAHMM found no distinction between SPMS and PPMS; low inflammatory progressive disease is a single meta-state of MS

References

1. Lublin FD, et al. *Neurology*. 996;46:907−11.
2. Lublin FD, et al. *Neurology*. 014;83:278−86.
3. Larochelle C, et al. *Trends Neurosci*. 2016; 39:325−339.
4. Mallon A-M, et al. *BMC Medical Research Methodology*. 2021;21:250.

Acknowledgments

The authors acknowledge the following Novartis employees: **Rohit Bhandari** and **Marie-Catherine Mousseau** for medical writing assistance and coordinating author reviews, and **Rupa De** for creative design assistance. The final responsibility for the content lies with the authors.

Disclosures

The study was funded by Novartis Pharma AG, Basel, Switzerland. **HB:** Nothing to disclose. **DAH:** Consulting fees from Albert Charitable Trust, Alexion Pharma, Biogen, Celgene, Frequency Therapeutics, Genentech, Med-Ex Learning, Merck, Novartis, Population Council, Receptos, Roche and Sanofi-Aventis; grants from Biogen, Immunotec, an equity interest in NeuroRx and is a full-time employee of Novartis Pharma AG, Basel, Switzerland. **GG:** Full-time employee of Novartis Pharma AG, Basel, Switzerland. **YS:** Nothing to disclose. **SG:** Nothing to disclose. **Wendy Su:** Full-time employee of Novartis Pharma AG, Basel, Switzerland. **BCK:** Full-time employee of Novartis Pharma AG, Basel, Switzerland. **TEN:** Nothing to disclose. **DLA:** Consulting fees from Biogen, Celgene, Frequency Therapeutics, Genentech, Merck, Novartis, Race to Erase MS, Roche, and Sanofi-Aventis, Shionogi, Xfacto Communications; grants from Immunotec and Novartis, and an equity interest in NeuroRx. **RAB:** served as a consultant for AstraZeneca, Biogen, EMD Serono, Genzyme, Genentech, Novartis and VielaBio. He receives research support from Biogen, Genentech, and Novartis. **HW:** received honoraria for acting as a member of Scientific Advisory Boards for Biogen, Genzyme, Merck Serono, Novartis, Roche Pharma AG, Sanofi-Aventis and UCB; as well as speaker honoraria and travel support from Alexion, Biogen, Biologix, Cognomed, F. Hoffmann-La Roche Ltd., Gemeinnützige Hertie Stiftung, Merck, Novartis, Roche Pharma AG, Genzyme, TEVA and WebMD Global. Prof. Wiendl is acting as a paid consultant for Actelion, Argenx, Biogen, Bristol Myers Squibb, EMD Serono, Idorsia, IGES, Immunic, Immunovant, Janssen, Johnson & Johnson, Novartis, Roche, Sanofi, the Swiss Multiple Sclerosis Society and UCB. His research is funded by the German Ministry for Education and Research (BMBF), Deutsche Forschungsgemeinschaft (DFG), Else Kröner Fresenius Foundation, Fresenius Foundation, the European Union, Hertie Foundation, NRW Ministry of Education and Research, Interdisciplinary Center for Clinical Studies (IZKF) Muenster and Biogen, GlaxoSmithKline, Roche Pharma AG, and Sanofi-Genzyme.
Chris C. Holmes: Nothing to disclose.

Poster presented at the 38th Congress of the European Committee for Treatment and Research in Multiple Sclerosis, 26–28 October 2022

Visit the web at: https://bit.ly/ectrims2022

Presenter email address: habib.ganjgahi@bdi.ox.ac.uk

Scan this QR code to download a copy Poster