BIG DATA INSTITUTE

Li Ka Shing Centre for Health Information and Discovery

An agnostic approach for multiple sclerosis disease states and prognosis using artificial intelligence: four clinical states allow description of the disease and its probabilities for evolution

Presenting author: Habib Ganjgahi,
Senior research fellow,
Statistics department, Big data institute, university of Oxford

Habib Ganjgahi*, Dieter A. Häring*, Gordon Graham, Yang Sun, Stephen Gardiner, Wendy Su, Bernd C. Kieseier, Thomas E. Nichols, Douglas L. Arnold, Robert A. Bermel, Heinz Wiendl^, Chris C. Holmes^.

***Co-First Authors, ^Co-Last Authors**

Scan to download a copy of this presentation

Disclosures

- **D.A.** has received personal compensation for serving as a Consultant for Alexion, Biogen, Celgene, Eli Lilly, EMD Serono, Frequency Therapeutics, Genentech, Merck, Novartis, Roche, Sanofi, and Shionogi, and holds an equity interest in NeuroRx.
- **H.W.** has received honoraria for acting as a member of scientific advisory boards for Biogen, Evgen, Genzyme, MedDay Pharmaceuticals, Merck Serono, Novartis, Roche Pharma AG, and Sanofi-Aventis, as well as speaker honoraria and travel support from Alexion, Biogen, Cognomed, F. Hoffmann-La Roche Ltd., Gemeinnützige Hertie-Stiftung, Merck Serono, Novartis, Roche Pharma AG, Genzyme, Teva, and WebMD Global. H.W is acting as a paid consultant for AbbVie, Actelion, Biogen, IGES, Johnson & Johnson, Novartis, Roche, Sanofi-Aventis, and the Swiss Multiple Sclerosis Society. 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 RE Children's Foundation, Biogen, GlaxoSmithKline GmbH, Roche Pharma AG, and Sanofi-Genzyme.
- **R.B.** has served as a consultant for Astra Zeneca, Biogen, EMD Serono, Genzyme, Genentech, Novartis and VielaBio. He receives research support from Biogen, Genentech and Novartis.
- **D.A., P.A., G.G., W.S.,** and **B.K.** are employees of Novartis.
- **H.G., Y.S., S.G., T.N.** and **C.H.** are current employees of the Big Data Institute (Oxford, UK) which received funding from Novartis to collaborate on AI in Medicine including the work presented here. **T.N.** received consulting fees from Perspectum Ltd.
- **Funding source:** This study is based on the Novartis-Oxford BDI collaboration for AI in Medicine, which was funded by Novartis Pharma AG, Basel, Switzerland.

Motivation

Background

Current subtypes (RRMS, SPMS, PPMS) are based on consensus definitions and only on two features (occurrence of relapse and disability progression)[1,2,3]

Goal

An evidence-based characterisation 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

Objectives

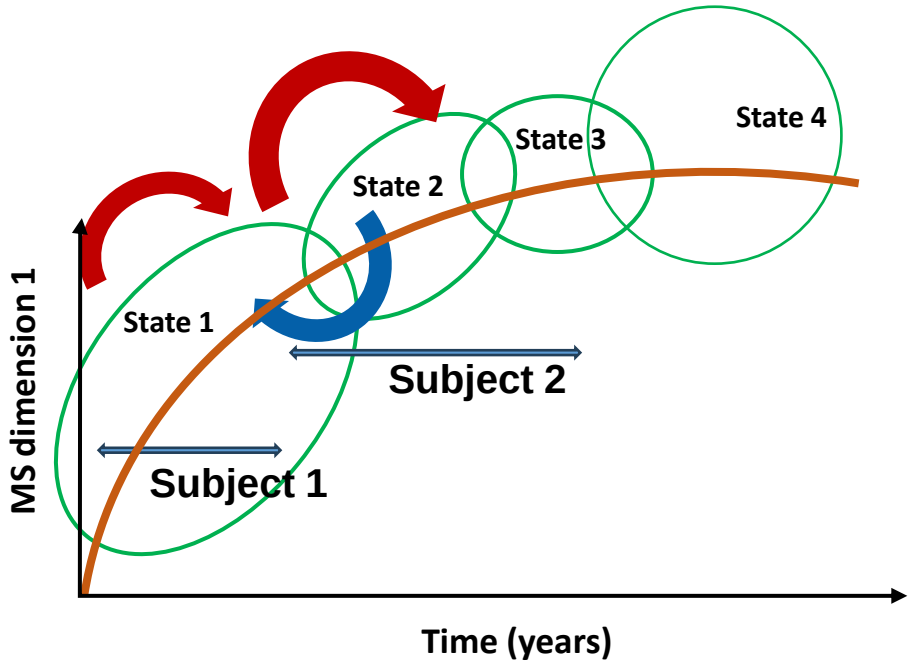
- To identify **key dimensions to describe/characterise MS disease evolution**
- To discover **disease states and probabilities** using multimodal clinical and MRI data
- Evaluate **treatment response and individual disease evolution**

1. Lublin FD, et al. Neurology. 1996;46:907−11. 2. Lublin FD, et al. Neurology. 2014; 83:278−86. 3. Lublin FD, et al. Neurology. 2020; 94:1088−92. MS, multiple sclerosis; PPMS, primary progressive MS; RRMS, relapsing-remitting MS; SPMS, secondary progressive MS.

Method: FAHMM

Factor Analysis followed by Hidden Markov Model

- MS is a **multidimensional disease** with a combination of different clinical and MRI features
- **Dynamic modelling** of dimensions:
  - MS continuum **grouped** into different **states**
  - Quantify **movement** between different groups (states) by **probability**
- Outputs
  - **State means**: phenotypes
  - **Transition probabilities**: progression



Results I: Dimensions to describe MS

- **Data**

  NO.MS[1] 8023 MS patients, 15 years of follow-up, >120,000 visits
  - Discovery: 6419; validation: 1604
  - RRMS (n = 5761), SPMS (n = 1550), or PPMS (n = 712)

- Analysis based on **clinical and MRI data**

- **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

Key dimensions of Multiple Sclerosis

Estimated loading matrix from probabilistic latent variable modelling		Physical disability	Diffuse brain damage	Focal inflammation	
				Relapse	Asymptomatic: Gd+T1 lesions
Clinical	EDSS	**-0.58**	**0.34**	0	0
	Timed 25-foot walk	**-0.66**	0	0	0
	Hand coordination (9HPT)	**-0.57**	0	0	0
	Cognition (PASAT)	0	**-0.35**	0	0
	Relapse (Y/N)	0	0.00	**-0.36**	0
MRI	T2 lesion volume	0	**0.54**	0	**-0.33**
	Normalized brain volume	0	**-0.53**	0	0
	Number of Gd+ T1 lesions	0	0.00	0*	**-0.56**

Variable weight 0 — 1

1. Dahlke F, Arnold DL, Aarden P, et al. Multiple Sclerosis Journal. 2021. 9HPT, 9-hole peg test; EDSS, Expanded Disability Status Scale; FAHMM, factor analysis followed by Hidden Markov Model; Gd+: gadolinium-enhancing; MRI, magnetic resonance imaging; MS, multiple sclerosis; NO.MS, Novartis-Oxford MS; PASAT, paced auditory serial-addition test; PPMS, primary progressive MS; RRMS, relapsing-remitting MS; SPMS, secondary progressive MS. *As a limitation of the NO.MS database, at the time of relapse there is usually no corresponding MRI scan.

Result II: Discovering MS states

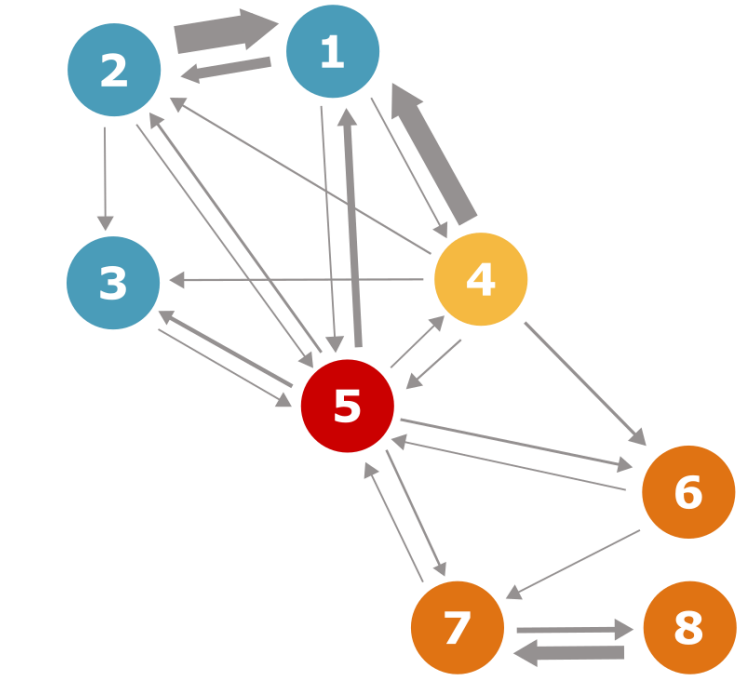
The model suggests 8 statistical states of MS:

- **States 1 – 3:** low levels of physical and cognitive impairment and low levels of cumulative damage – clinically stable and low disease activity
- **States 4 – 5:** transient states of high level of inflammation with or without symptoms
- **States 6 – 8:** high level of impairment, high cumulative damage levels to the CNS, low focal inflammation and low probability to transition to earlier states of MS

	State 1	*State 2*	*State 3*	*State 4*	*State 5*	*State 6*	*State 7*	*State 8*
EDSS	1.9	2.05	3.91	2.52	3.88	4.22	5.66	5.01
T25-FW	5.08	5.05	7.82	6.13	9.49	8.86	19	15.17
Hand coord.	19.78	20.16	23.81	22.04	25.49	28.09	36.77	40.08
Cognition	53.07	50.64	51.2	48.66	47.23	43.15	42.24	34.73
T2 lesion vol.	7332.06	7334.04	1755.11	10926.84	9569.03	12168.27	11207.71	10[illegible]
Brain volume	1.51	1.53	1.5	1.53	1.51	1.45	1.45	1.43
T1 Gd+ lesions	0	0	0	1.33	1.98	0.01	0	0.57
Relapse	0	0	0	0	1	0	0	0

Severe
Moderate
Mild

The table shows state means. For 'Relapse', it is the probability to be in a relapse. For Gd+ T1 lesions it is the number of lesions.



Low probability of transition within 1 month

High probability of transition within 1 month

- In this dynamic model patients can move in any order between states along the arrows.
- Transitions between states not connected by an arrow are highly unlikely

CNS, central nervous system; Coord., coordination; EDSS, Expanded Disability Status Scale; Gd+, gadolinium-enhancing; MS, multiple sclerosis; T25-FW, timed 25-foot walk; vol., volume

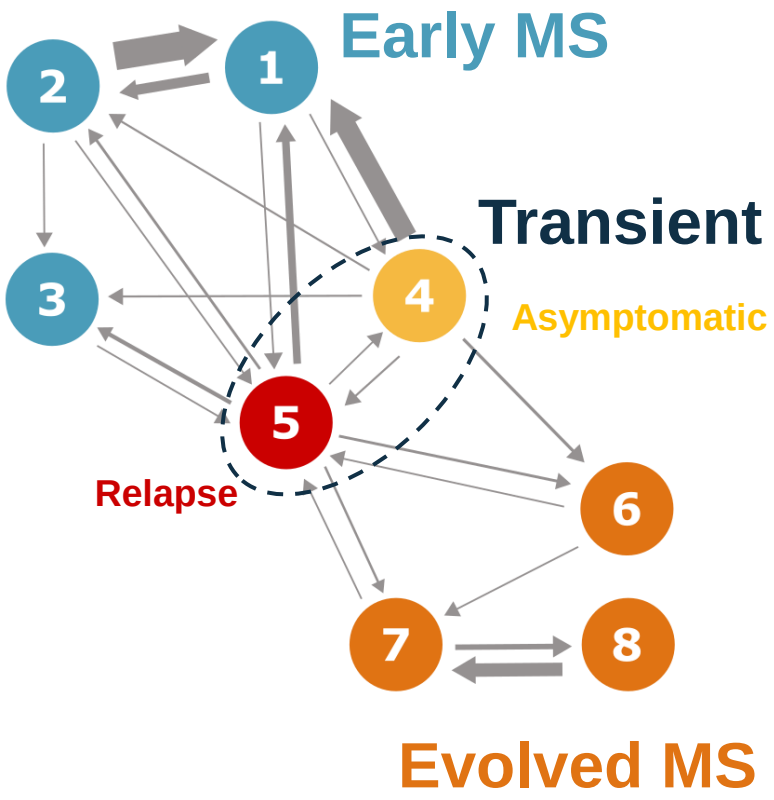
Result II: Four clinical states of MS

	Early MS			Transient		Evolved MS		
Age, median	39	39	47	36	40	47	49	47
Female (%)	69%	70%	68%	70%	71%	60%	58%	60%
Diagnosed subtype								
RRMS (%)	91%	91%	52%	90%	82%	49%	15%	16%
SPMS (%)	6.4%	5.3%	31%	7.4%	17%	38%	69%	73%
PPMS (%)	3.0%	3.8%	16%	2.4%	1.1%	14%	16%	11%
	State 1	*State 2*	*State 3*	*State 4*	*State 5*	*State 6*	*State 7*	*State 8*
EDSS	1.9	2.05	3.91	2.32	3.88	4.22	5.86	6.01
T25-FW	5.08	5.05	7.82	6.13	9.49	8.86	19	38.37
Hand coord.	19.78	20.16	23.81	22.04	25.49	28.09	36.77	49.08
Cognition	53.07	50.84	51.2	48.66	47.23	43.15	42.24	34.92
T2 lesion vol.	7332.06	7334.04	1735.11	10926.84	9569.03	12168.27	11207.71	24017.56
Brain volume	1.51	1.53	1.5	1.53	1.51	1.45	1.45	1.43
T1 Gd+ lesions	0	0	0	3.34	1.98	0.01	0	0.57
Relapse	0	0	0	0	1	0	0	0

Severe

Moderate

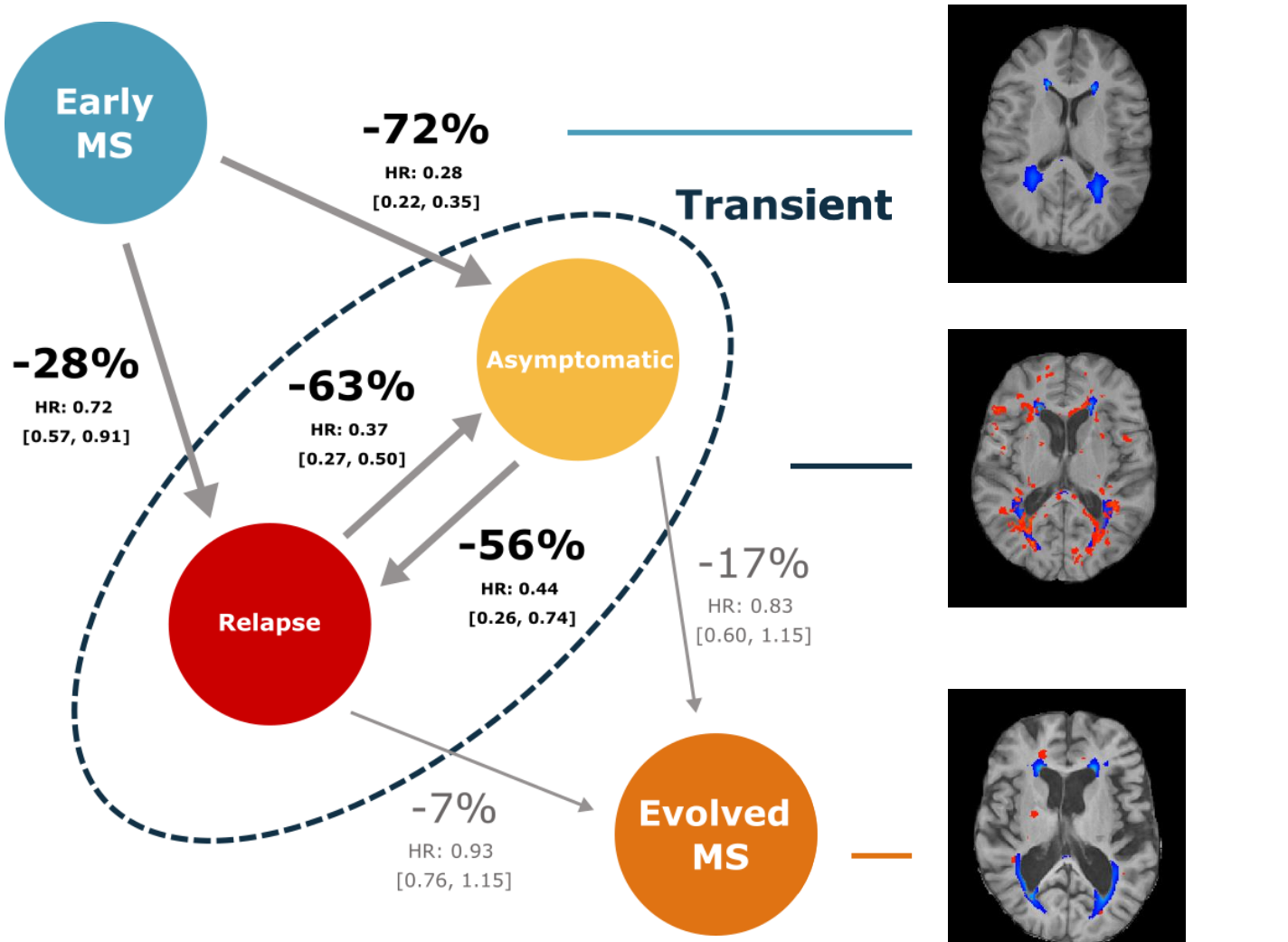
Mild



The table shows state means. For 'Relapse', it is the probability to be in a relapse. For Gd+ T1 lesions it is the number of lesions.

CNS, central nervous system; Coord., coordination; EDSS, Expanded Disability Status Scale; Gd+, gadolinium-enhancing; MS, multiple sclerosis; T25-FW, timed 25-foot walk; vol., volume

Losing reserve capacity in transient /accumulating MS states



T2 lesions Gd+ T1 lesions

Transition

- Patients transition from **Early** states to **Evolved** states through **Transient** states by **acquiring damage** to the CNS; this can occur **asymptomatically** or be accompanied by a **relapse**

Disease modifying therapy

- DMTs **reduce** the **probability** of moving to **Transient** states of MS
- DMTs **enhance** the **chance of remaining** in **Early MS** states

Individual prediction

Time to reach **Evolved MS** from **Early MS**

- Out of sample performance: **C-index 0.82**
- **Brier score: 0.06**

CNS, central nervous system; DMT, disease-modifying therapy; Gd+, gadolinium-enhancing; HR, hazard ratio.

Conclusions

	Clinical phenotypes of MS[1]	FAHMM disease stages[2]
Dimensions to define phenotypes/states	**Two dimensions** 1. Disability progression (mechanism) 2. Relapse	**Dimensions** 1. Physical disability (absolute level) **2. Brain damage (reserve capacity)** 3. Relapse **4. Asymptomatic disease activity**
Modifiers of phenotypes (applicable to all phenotypes)	**Two modifiers** Inflammatory activity (MRI lesions) Clinical progression	**No modifiers** -
Main classification	**Phenotypes** **1. Relapsing remitting MS** **2. Secondary progressive MS** **3. Primary progressive MS**	**Disease continuum** **1. Early MS** **2. Transient : Asymptomatic** **3. Transient: Relapse** **4. Evolved MS**

- Agnostic, evidence-based dynamic description of MS evolution
  - Transition from Early to Evolved MS through accumulation of damage to the brain
  - No distinction between SPMS and PPMS
  - DMT increases chance of remaining in Early MS
  - Ability to better prognosticate individuals

1. Lublin FD, et al. Neurology. 2014. 2. Ganjgahi et al. ECTRIMS 2023. DMT, disease modifying therapy; FAHMM, factor Analysis followed by Hidden Markov Model; MS, multiple sclerosis; PPMS, primary progressive MS; SPMS, secondary progressive MS

Acknowledgments

A. Huffman
Y. Sun
S. Gardiner
T. Nichols
C. Holmes

D. Haering
P. Aarden
M.-C. Mousseau
G. Graham
W. Su
B. Kieseier

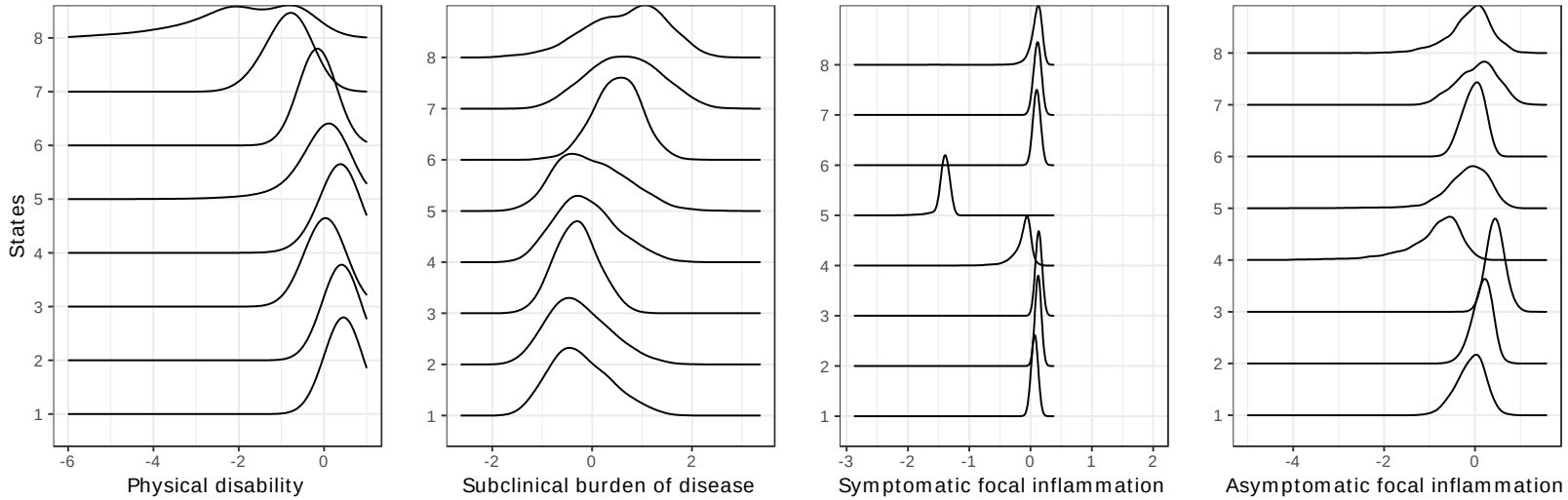
D. Arnold
R. Bermel
H. Wiendl

Physicians and patients who participated in these studies

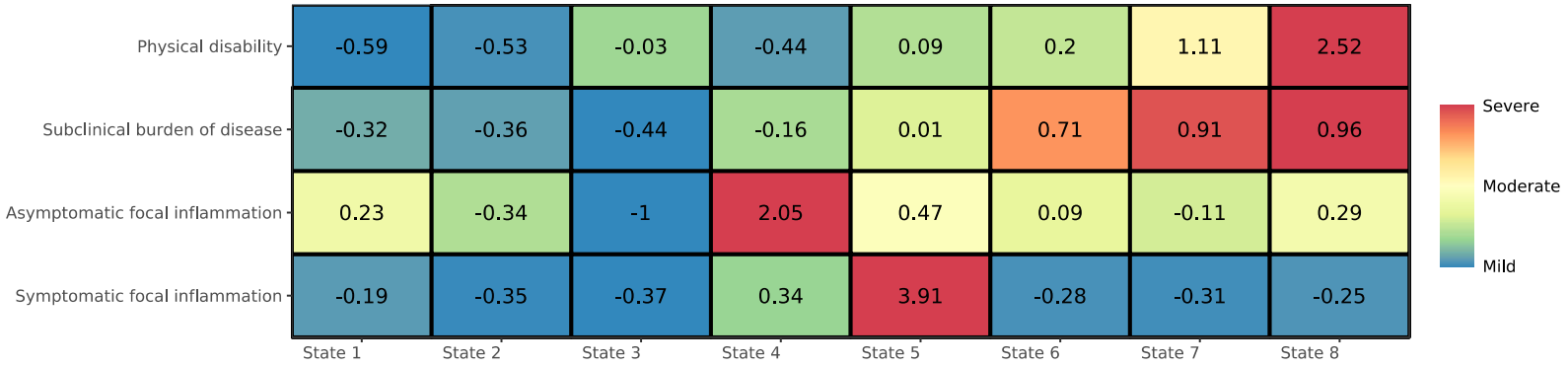
Thank you

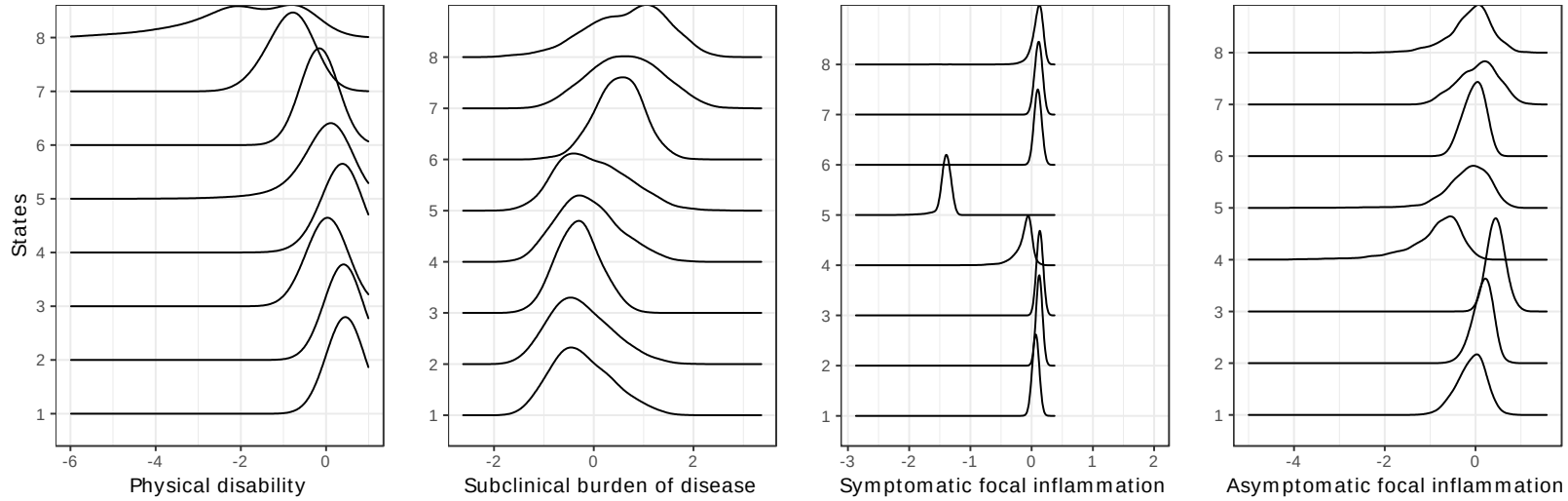
Back up

States


States
8
7
6
5
4
3
2
1
-6
-4
-2
0
Physical disability
-2
0
2
Subclinical burden of disease
-3
-2
-1
0
1
2
Symptomatic focal inflammation
-4
-2
0
Asymptomatic focal inflammation

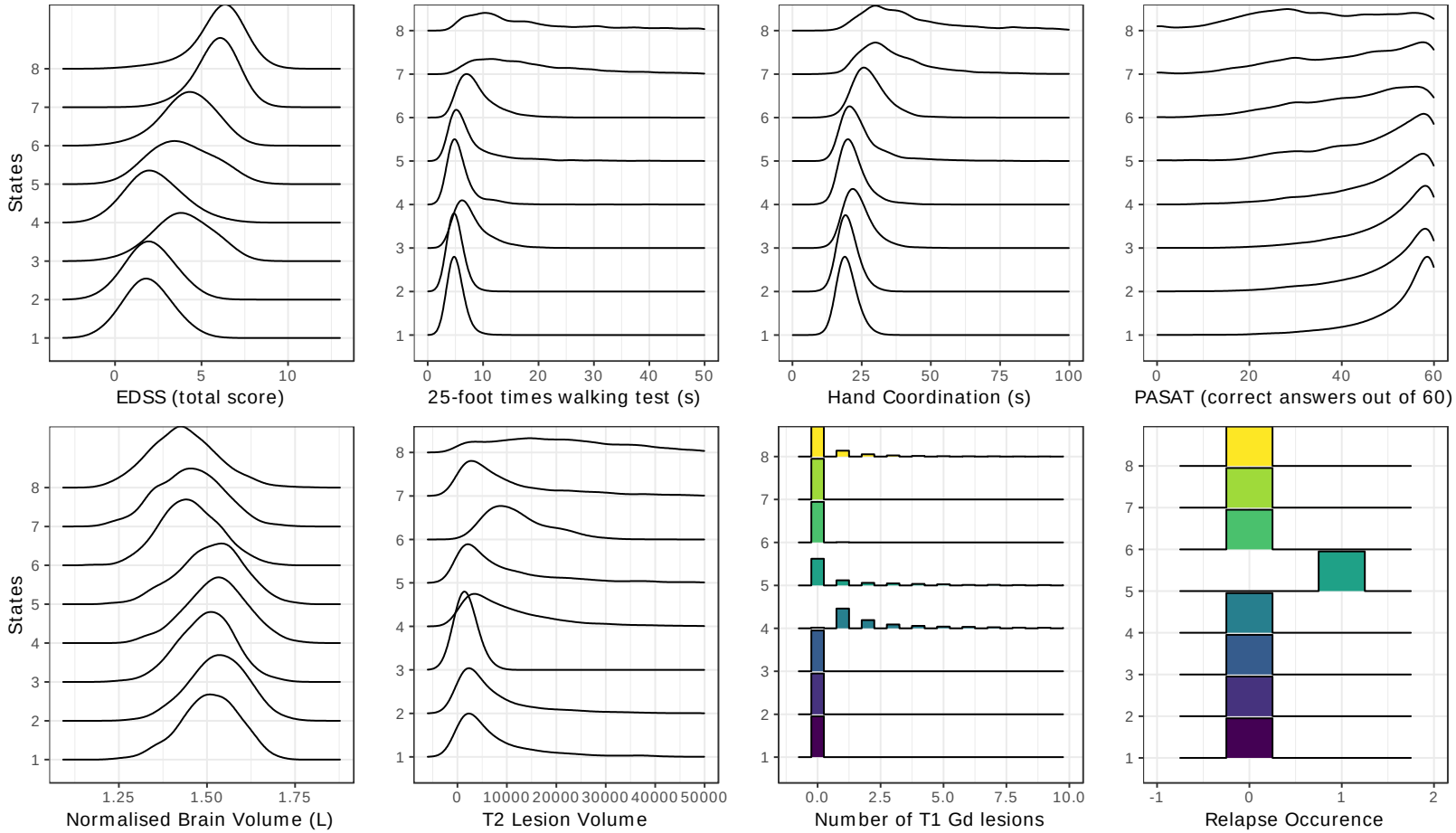
States density: Composite scores



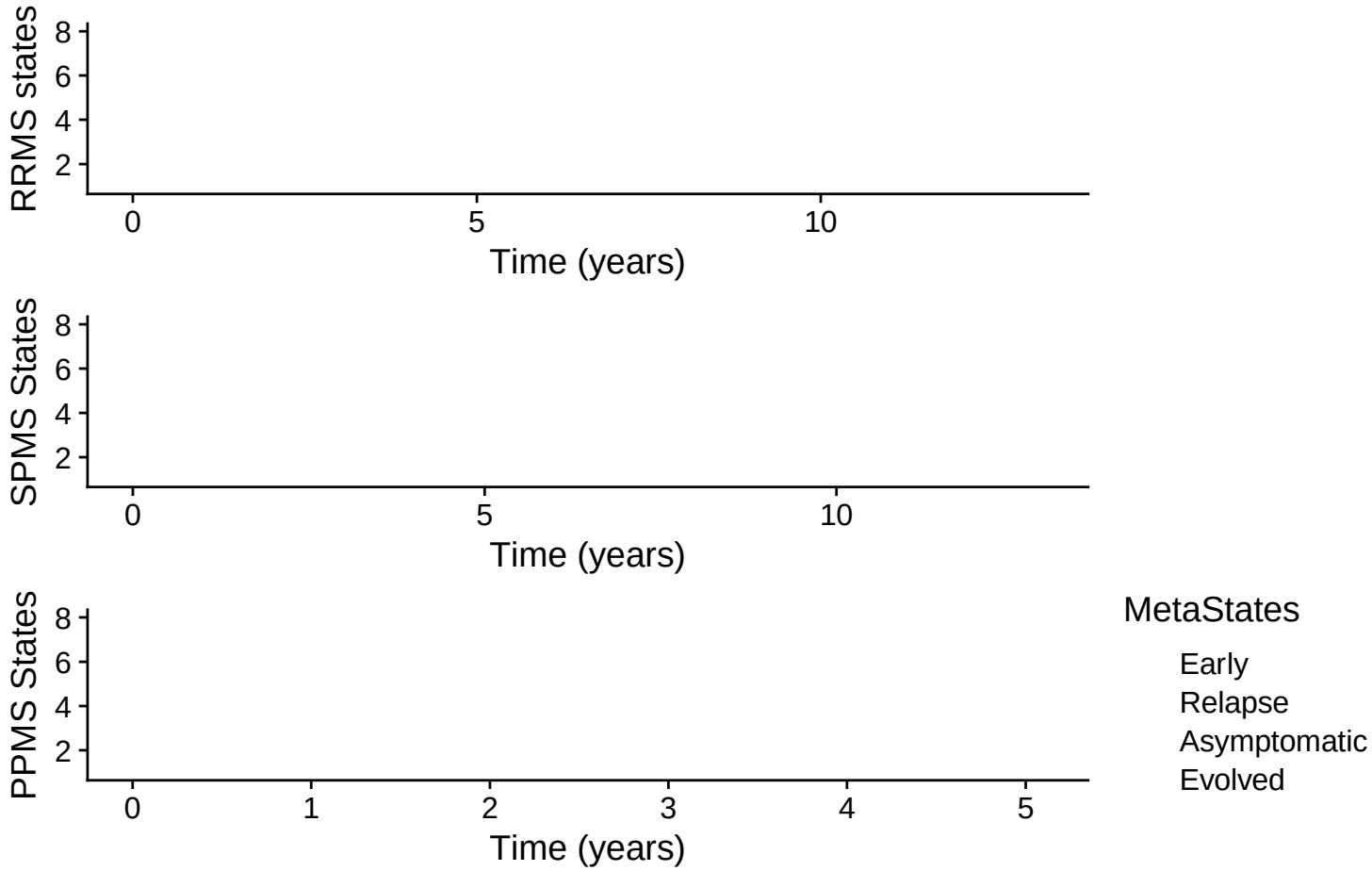


States density: Original features

States do overlap but there are features that separate them: T1GD and Relapse occurrence



States density: Original features



Baseline demography in states

Variable	State 1 N = 33`512	State 2 N = 11`045	State 3 N = 15`253	State 4 N = 3`709	State 5 N = 5`594	State 6 N = 9`802	State 7 N = 11`007	State 8 N = 4`859
Age								
Mean (SD)	39 (9)	39 (9)	47 (9)	37 (9)	40 (10)	46 (9)	48 (9)	47 (9)
Median (IQR)	39 (32, 46)	39 (32, 46)	47 (41, 53)	36 (30, 43)	40 (33, 47)	47 (40, 53)	49 (42, 55)	47 (41, 54)
Gender								
Female	22`991 (69%)	7`702 (70%)	10`336 (68%)	2`597 (70%)	3`984 (71%)	5`924 (60%)	6`361 (58%)	2`892 (60%)
Male	10`521 (31%)	3`343 (30%)	4`917 (32%)	1`112 (30%)	1`610 (29%)	3`878 (40%)	4`646 (42%)	1`967 (40%)
MS Type								
RRMS	30`390 (91%)	10`039 (91%)	7`987 (52%)	3`345 (90%)	4`595 (82%)	4`763 (49%)	1`670 (15%)	782 (16%)
SPMS	2`129 (6.4%)	587 (5.3%)	4`797 (31%)	274 (7.4%)	937 (17%)	3`686 (38%)	7`596 (69%)	3`555 (73%)
PPMS	993 (3.0%)	419 (3.8%)	2`469 (16%)	90 (2.4%)	62 (1.1%)	1`353 (14%)	1`741 (16%)	522 (11%)
Duration since first symptom								
Mean (SD)	11 (8)	10 (8)	14 (9)	10 (8)	13 (9)	16 (9)	18 (9)	19 (9)
Median (IQR)	9 (5, 20)	8 (4, 20)	11 (8, 22)	8 (4, 20)	9 (5, 21)	20 (8, 22)	21 (10, 23)	21 (10, 23)

Endpoints in states

Variable	State 1 N = 33`512	State 2 N = 11`045	State 3 N = 15`253	State 4 N = 3`709	State 5 N = 5`594	State 6 N = 9`802	State 7 N = 11`007	State 8 N = 4`859
EDSS								
Mean (SD)	1.90 (1.07)	2.05 (1.11)	3.91 (1.48)	2.32 (1.40)	3.88 (1.63)	4.22 (1.32)	5.86 (0.90)	6.01 (1.29)
Median (IQR)	2 (1, 2.5)	2 (1.5, 3)	4 (3, 5)	2 (1.5, 3.5)	4 (2.5, 5)	4 (3.5, 5)	6 (5.5, 6.5)	6.5 (6, 6.5)
Timed 25-foot walk (s)								
Mean (SD)	5.08 (1.29)	5.05 (1.26)	7.82 (3.20)	6.13 (3.56)	9.49 (11.03)	8.86 (3.76)	19.00 (10.69)	38.37 (35.77)
Median (IQR)	4.90 (4.20, 5.70)	4.85 (4.15, 5.70)	7.00 (5.60, 9.10)	5.20 (4.35, 6.55)	6.15 (4.81, 9.25)	7.95 (6.35, 10.25)	16.30 (10.95, 24.75)	23.25 (11.50, 57.04)
9-Hole peg test (s)								
Mean (SD)	19.78 (3.33)	20.16 (3.53)	23.81 (5.46)	22.04 (5.45)	25.49 (10.88)	28.09 (6.40)	36.77 (13.42)	49.08 (30.72)
Median (IQR)	19.35 (17.50, 21.70)	19.68 (17.75, 22.08)	22.85 (20.12, 26.38)	21.00 (18.50, 24.15)	22.50 (19.52, 27.12)	27.12 (23.82, 31.40)	33.35 (27.73, 42.50)	38.73 (29.89, 56.16)
PASAT								
Mean (SD)	53.07 (8.38)	50.84 (9.73)	51.20 (9.12)	48.66 (10.88)	47.23 (12.11)	43.15 (12.67)	42.24 (13.92)	34.92 (15.35)
Median (IQR)	56 (50, 59)	54 (47, 58)	54 (47, 58)	52 (43, 57)	51 (40, 57)	46 (34, 54)	45 (31, 54)	35 (24, 48)
Vol. T2 lesions (mL)								
Mean (SD)	7.33 (8.36)	7.33 (8.84)	1.74 (1.44)	10.93 (11.00)	9.57 (11.57)	12.17 (6.18)	11.21 (13.23)	24.02 (17.85)
Median (IQR)	4.34 (1.76, 9.65)	4.08 (1.76, 9.20)	1.33 (0.61, 2.49)	7.58 (3.39, 14.66)	5.38 (1.89, 12.62)	10.93 (7.55, 15.78)	6.45 (2.57, 14.91)	20.75 (11.29, 33.07)
NBV (L)								
Mean (SD)	1.51 (0.08)	1.53 (0.08)	1.50 (0.08)	1.53 (0.09)	1.51 (0.09)	1.45 (0.08)	1.45 (0.09)	1.43 (0.10)
Median (IQR)	1.51 (1.45, 1.57)	1.53 (1.48, 1.59)	1.50 (1.45, 1.55)	1.53 (1.47, 1.58)	1.52 (1.45, 1.57)	1.45 (1.39, 1.51)	1.45 (1.39, 1.51)	1.43 (1.37, 1.49)
PIRA events								
# per year	0.0218	0.0199	0.0860	0.0242	0.0019	0.0852	0.1520	0.1476
CDF at 2 years (%)	4.3%	3.9%	15.8%	4.7%	0.4%	15.7%	26.2%	25.6%

Transition probability matrix of FAHMM

- **Transitions from early states** (1, 2, 3) **to evolved states** (6, 7 ,8) **is low** (<0.02%).
- **Transitions for evolved** states (6, 7 ,8) **to early states** (1, 2, 3) **is also low** (<0.02%).

To states

From states	Early MS 1	Early MS 2	Early MS 3	Transient states 4	Transient states 5	Evolved MS 6	Evolved MS 7	Evolved MS 8
1	0.67	0.24	0	0.04	0.04	0	0	0
2	0.82	0.07	0.03	0.02	0.03	0.02	< 0.02	0
3	0.02	0	0.91	0	0.04	0	0.01	0
4	0.62	0.06	0.06	0.08	0.06	0.09	0.02	0
5	0.22	0.08	0.12	0.04	0.37	0.08	0.07	0.03
6	0.02	0	0	0.02	0.04	0.86	0.05	0
7	0	< 0.02	0	0	0.03	0.02	0.79	0.16
8	0	0	0	0	0.03	0	0.39	0.57

Transition probability matrix of FAHMM

- **FAHMM is a dynamic model of MS:** Displayed in this matrix are the **probabilities of moving from one state to another from one month to the next**

From states \ To states	Early MS 1	Early MS 2	Transient states 3	Transient states 4	Transient states 5	Evolved MS 6	Evolved MS 7	Evolved MS 8
1	0.67	0.24	0	0.04	0.04	0	0	0
2	0.82	0.07	0.03	0.02	0.03	0.02	0	0
3	0.02	0	0.91	0	0.04	0	0.01	0
4	0.62	0.06	0.06	0.08	0.06	0.09	0.02	0
5	0.22	0.08	0.12	0.04	0.37	0.08	0.07	0.03
6	0.02	0	0	0.02	0.04	0.86	0.05	0
7	0	0	0	0	0.03	0.02	0.79	0.16
8	0	0	0	0	0.03	0	0.39	0.57

Baseline demography in four clinical states

Variable	Early MS N = 2`952	Asymptomatic N = 1`622	Relapse N = 561	Evolved MS N = 1`284
Age				
Mean (SD) / Median (IQR)	41 (10) / 42 (34, 48)	37 (9) / 36 (30, 43)	39 (10) / 39 (31, 46)	46 (9) / 47 (40, 53)
Gender				
Female / Male	2`040 (69%) / 912 (31%)	1`103 (68%) / 519 (32%)	409 (73%) / 152 (27%)	751 (58%) / 533 (42%)
MS type				
RRMS / SPMS / PPMS	2`308 (78%) / 319 (11%) / 325 (11%)	1`456 (90%) / 117 (7.2%) / 49 (3.0%)	458 (86%) / 72 (13%) / 4 (0.7%)	366 (29%) / 719 (56%) / 199 (15%)
Duration since 1st symptoms [yrs]				
Mean (SD) / Median (IQR)	10.1 (8.5) / 7.5 (3.5, 20.0)	8.7 (7.8) / 7.5 (3.5, 20.0)	11.0 (9.2) / 7.5 (3.5, 20.0)	15.4 (9.0) / 20.0 (7.5, 20.0)
# of relapses last year before trial				
Mean (SD) / Median (IQR)	1.09 (0.87) / 1.00 (1.00, 1.00)	1.30 (0.85) / 1.00 (1.00, 2.00)	1.49 (1.04) / 1.00 (1.00, 2.00)	0.53 (0.73) / 0.00 (0.00, 1.00)
EDSS				
Mean (SD) / Median (IQR)	2.80 (1.45) / 2.50 (2.00, 4.00)	2.48 (1.41) / 2.00 (1.50, 3.50)	3.12 (1.50) / 3.00 (2.00, 4.00)	5.24 (1.17) / 5.50 (4.50, 6.00)
PASAT				
Mean (SD) / Median (IQR)	48.10 (10.61) / 51.00 (42.00, 56.00)	46.35 (11.64) / 50.00 (40.00, 56.00)	44.94 (13.36) / 49.00 (37.00, 55.25)	35.91 (14.29) / 36.00 (26.00, 48.00)
Volume T2 Lesions [mL]				
Mean (SD) / Median (IQR)	5.42 (7.59) / 2.58 (1.08, 6.35)	11.50 (11.85) / 7.70 (3.30, 15.55)	9.82 (11.83) / 5.50 (2.01, 13.01)	16.59 (15.16) / 12.07 (6.56, 21.53)
Normalized Brain Volume [L]				
Mean (SD) / Median (IQR)	1.53 (0.08) / 1.53 (1.47, 1.59)	1.53 (0.09) / 1.53 (1.47, 1.59)	1.52 (0.09) / 1.53 (1.46, 1.58)	1.46 (0.09) / 1.46 (1.40, 1.52)

Motivation

- **Current Subtypes** (RRMS, SPMS, PPMS): **consensus definitions** solely the basis of **basic clinical features** (occurrence of relapse and physical disability)[1,2,3]
- **Other dimensions**: cognition, damages to different tissues in the brain measured by Magnetic Resonance Imaging (MRI) [REF]
- **Limitations:** predicting individual disease courses [REF], or treatment response [REF]
- **Goal:** An evidence-based characterisation 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
- **Objectives**
  - To identify **key dimensions to describe/characterise MS progression**
- Identify new subtypes using both clinical and MRI data to
  - Characterise disease progression
  - Evaluate treatment response and individual progression

1. Lublin FD, et al. Neurology. 1996;46:907−11. 2. Lublin FD, et al. Neurology. 2014; 83:278−86. 3. Lublin FD, et al. Neurology. 2020; 94:1088−92.