Title
Estimating the Evolution of Disease in the Parkinson's Progression Markers Initiative

Journal
Neurodegenerative Diseases, 18(4)

ISSN
1660-2854

Authors
Iddi, Samuel
Li, Dan
Aisen, Paul S
et al.

Publication Date
2018

DOI
10.1159/000488780

Peer reviewed

Original Paper

Neuro-degenerative Diseases

Neurodegener Dis 2018;18:173–190
DOI: 10.1159/000488780

Received: February 13, 2018
Accepted after revision: March 23, 2018
Published online: August 8, 2018

Estimating the Evolution of Disease in the Parkinson's Progression Markers Initiative

Samuel Iddi[a, b] Dan Li[a] Paul S. Aisen[a] Michael S. Rafii[a] Irene Litvan[d] Wesley K. Thompson[c] Michael C. Donohue[a]

[a]Alzheimer's Therapeutic Research Institute, University of Southern California, San Diego, CA, USA; [b]Department of Statistics, University of Ghana, Accra, Ghana; [c]Department of Neurosciences, University of California, San Diego, CA, USA; [d]Department of Family Medicine and Public Health, University of California, San Diego, CA, USA

Keywords

Biomarkers · Disease trajectories · Clinical diagnosis · Multilevel Bayesian models · Joint mixed-effects models · Latent time shift · Multicohort longitudinal data · Parkinson's disease

Abstract

Parkinson's disease is the second most common neurological disease and affects about 1% of persons over the age of 60 years. Due to the lack of approved surrogate markers, confirmation of the disease still requires postmortem examination. Identifying and validating biomarkers are essential steps toward improving clinical diagnosis and accelerating the search for therapeutic drugs to ameliorate disease symptoms. Until recently, statistical analysis of multicohort longitudinal studies of neurodegenerative diseases has usually been restricted to a single analysis per outcome with simple comparisons between diagnostic groups. However, an important methodological consideration is to allow the modeling framework to handle multiple outcomes simultaneously and consider the transitions between diagnostic groups. This enables researchers to monitor multiple trajectories, correctly account for the correlation among biomarkers, and assess how these associations may jointly change over the long-term course of disease. In this study, we apply a latent time joint mixed-effects model to study biomarker progression and disease dynamics in the Parkinson's Progression Markers Initiative (PPMI) and examine which markers might be most informative in the earliest phases of disease. The results reveal that, even though diagnostic category was not included in the model, it seems to accurately reflect the temporal ordering of the disease state consistent with diagnosis categorization at baseline. In addition, results indicated that the specific binding ratio on striatum and the total Unified Parkinson's Disease Rating Scale (UPDRS) show high discriminability between disease stages. An extended latent time joint mixed-effects model with heterogeneous latent time variance also showed improvement in model fit in a simulation study and when applied to real data.

Introduction

Parkinson's disease (PD) affects nearly 1% of persons over the age of 60 years. Degeneration and death of neurons in the substantia nigra of the brain occurs much ear-

Michael C. Donohue
Alzheimer's Therapeutic Research Institute
University of Southern California
9860 Mesa Rim Rd., San Diego, CA 92121 (USA)
E-Mail mdonohue@usc.edu

lier than the onset of motor symptoms such as slow movement, tremor, and rigidity. Disease progression is characterized by the transition from normal, through preclinical, to clinical stages of the disease. Confirmation of the disease is only possible through an autopsy. The presence of nonmotor symptoms such as olfactory, sleep, and depression help classify patients as prodromal since there is currently no available surrogate marker to confirm disease diagnosis. Currently, the FDA approves DaTSCAN, a specialized imaging technique of the brain that measures levels of dopamine in the substantia nigra, to evaluate suspected PD. Nonetheless, diagnosis of the onset and progression of the disease is often marred by errors due to the unavailability of reliable validated diagnostic markers. Identifying and validating biomarkers are crucial steps in clinical diagnosis and the search for treatment to ameliorate disease symptoms and slow down the progression of PD. Reliable and cost-effective biomarkers for the diagnosis of neurodegenerative diseases are more likely to be discovered if multiple clinical, biological, and imaging assessments are studied simultaneously to provide complementary information.

In studies of neurodegenerative diseases such as PD, clinical, biological, and imaging biomarkers are often collected longitudinally over a short period of time on individuals at different stages of disease. Such studies are aimed at accurately characterizing disease trajectories and identifying important diagnosis and prognosis biomarkers, and ultimately help to accelerate drug discovery. Although studies generally collect multiple outcomes over time, it is very common to find studies where researchers focus on the analysis of changes in a single outcome. In this regard, the linear mixed-effects model is ubiquitous in studies of neurodegenerative diseases. For example, Kennedy et al. [1] applied linear mixed-effects models to estimate the rate of decline of Alzheimer's Disease Assessment Scale Cognitive score (ADASCog) among Alzheimer's disease patients based on the Mini-Mental State Examination (MMSE). Guerrero et al. [2] reported studies which employ mixed-effects models for the analysis of longitudinal Alzheimer's disease markers. Mixed-effects models are also commonly applied in the PD literature. Sturkenboom et al. [3] assessed the effect of occupational therapy on the Canadian Occupational Performance Measure in improving the daily activity of PD patients. Klotsche et al. [4] analyzed changes in health-related quality of life in a longitudinal cohort study of PD patients. Analysis of single longitudinal outcomes can be inefficient for studying the progression of the disease since outcomes may be correlated and hence may provide complementary insights. Single-outcome analyses may also contribute to the lack of strong evidence of symptomatic or disease-modifying treatment effects on clinical outcomes. Inference may be improved if all outcomes are modeled simultaneously, taking into account the intra- and inter-subject variability and accommodating between outcome associations.

Joint models that simultaneously model 2 or more outcomes have received relatively less attention in the neurodegenerative disease literature. This neglect is likely due to problems in interpretation of resulting parameter estimates, or to the computational complexities that arise when the number of outcomes grows or when the outcomes are of mixed types, such as binary, count, continuous, or time-to-event. Moreover, in most studies which do employ multivariate models, only 2 outcomes are modeled at a time [5, 6]. Nevertheless, the multivariate modeling framework offers several potential advantages, such as accommodating all sources of variation and correlation among outcomes. With the rapid emergence and growing popularity of Bayesian estimation methods to handle complex and computationally burdensome models, multivariate modeling techniques have become an active area of research in neurodegenerative studies. Luo and He [7] analyzed longitudinal outcomes and a time-to-event outcome to assess the effect of tocopherol on patients with early PD using a Bayesian Markov Chain Monte Carlo (MCMC) method for the efficient estimation of model parameters. Their modeling framework applied a multilevel item response theory model for multiple longitudinal outcomes and a Cox proportional hazard model for the survival outcome. They demonstrated that their joint model led to a better fit to their data than single analyses per outcome. Luo [8] provides an excellent review of recent developments and issues of joint modeling of longitudinal and survival data. Li et al. [9] proposed a latent time joint mixed-effects model (LTJMM), an extension of the multivariate linear mixed-effects model for longitudinal data, which accommodates a large number of outcomes to be analyzed together and permits unbalanced time measurement of outcomes. A distinctive feature of this model is the inclusion of a latent time shift parameter, which captures the degree of an individual's disease progression. The subject-specific latent times shared across outcomes are assumed to have a homogeneous variance across all subjects.

In this study, we extended the LTJMM by relaxing the homogeneity assumption on the latent times. Heterogeneity was introduced by allowing the latent time to be influenced by subject-level covariates. To achieve this, we modeled the latent time variance in terms of covariates

through an exponential function. The heterogeneous LTJMM was used to estimate trajectories of biomarkers and determine the sequence of biomarker abnormality using data from the Parkinson's Progression Markers Initiative (PPMI) study. Estimation of model parameters was via a Bayesian MCMC framework.

Materials and Methods

Data

The PPMI is a prospective multicenter study of patients at different stages of PD with healthy patients as controls. Roughly 35 centers across North America, Europe, Israel, and Australia are involved in this ongoing study that has collected data over a period of 6 years. All study sites received institutional review board approval before initiating the study, and all study participants provided written informed consent for research. Detailed study design, inclusion criteria, standard protocols, registration, and consent procedures can be found on the study website (www.ppmi-info.org). The study is aimed at identifying novel clinical, imaging, and biological markers of PD and to assess their progression in patients. Discovery and validation of new biomarkers would be beneficial for use in clinical trials of disease-modifying drugs. Patients are classified by their disease stage, namely, healthy controls (HC), PD, scan without evidence of dopaminergic deficit, prodromal, genetic cohort, and genetic registry patients. The genetic cohort and registry subjects may or may not be affected with Parkinson's symptoms. The steps involved in diagnosing and classifying participants into disease categories are reported in the PPMI study protocol.

In this study, we focused on subjects in the HC, PD, and prodromal groups. There were 475 PD, 238 HC, and 187 prodromal individuals representing 52.78, 26.44, and 20.78% of the sample, respectively. Of these participants, 317 (35.22%) were female and 583 (64.78%) were male. The proportion of females in the HC, PD, and prodromal groups was 34.87, 35.58, and 34.76% respectively. Visit times were not balanced, as the schedule of assessments depended on the group. HC were followed every 6 months for the first year and every 12 months thereafter. Before the first 12 months, PD and prodromal patients were followed every 3 months and every 6 months thereafter.

We considered 17 outcomes and modeled 8 of them using age and gender as covariates. Clinical assessments included tremor, postural instability and gait difficulty (PIGD), REM sleep behavior disorder (RBD), University of Pennsylvania Smell Identification Test (UPSIT), Montreal Cognitive Assessment (MOCA), Hopkins Verbal Learning Test (HVLT), Geriatric Depression Scale (GDS), Semantic Fluency Test (SFT), Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS), total Scale for Outcomes in Parkinson's (SCOPA), Line Orientation Test (LINEORT), Montreal Cognitive Assessment (MOCA) and State Trait Anxiety Index (STAI). UPDRS total includes the 3 subtests; UPDRS I gauges mentation, behavior, and mood, UPDRS II assesses activities of daily living, and UPDRS III examines motor. To complement these clinical outcomes, we also investigated biological markers such as cerebrospinal fluid (CSF) α-synuclein, CSF amyloid peptide 42 (Aβ42), CSF phosphorylated tau 181p (p-tau 181)

Table 1. Simulation study results

Model parameters	True values	LTJMM-H model		LTJMM model	
		bias	coverage	bias	coverage
β_{01}	−31.7	1.6929	0.86	1.7119	0.86
β_{11}	−2.06	0.1139	0.94	0.0957	0.98
β_{21}	0.12	−0.0314	0.9	−0.0312	0.88
β_{02}	−30.70	0.8346	0.93	0.8408	0.93
β_{12}	0.52	0.0138	0.97	0.0072	0.95
β_{22}	0.06	−0.0156	0.93	−0.0154	0.92
β_{03}	−7.60	1.3147	0.97	1.3778	0.98
β_{13}	3.90	0.3746	0.95	0.3227	0.98
β_{23}	0.28	−0.031	0.98	−0.0304	0.96
β_{04}	17.20	−0.5776	1.00	−0.399	1.00
β_{14}	−1.56	1.1918	0.98	1.0105	0.98
β_{24}	0.08	−0.0145	1.00	−0.0114	1.00
β_{05}	−36.20	8.2043	0.74	8.2752	0.73
β_{15}	−1.40	0.3505	0.97	0.2837	0.97
β_{25}	0.18	−0.1564	0.68	−0.1553	0.74
γ_1	0.15	0.0106	0.96	0.0105	0.97
γ_2	0.09	0.0064	0.98	0.0063	0.98
γ_3	0.60	0.0461	0.96	0.0457	0.94
γ_4	2.20	0.1588	0.97	0.1576	0.95
γ_5	0.80	0.0577	0.98	0.0573	0.97
$\sigma\alpha_{01}$	3.20	−0.0548	0.96	−0.0574	0.96
$\sigma\alpha_{02}$	1.60	−0.021	0.97	−0.0203	0.97
$\sigma\alpha_{03}$	5.60	−0.0224	0.95	−0.0218	0.95
$\sigma\alpha_{04}$	7.10	0.0800	0.96	0.0781	0.95
$\sigma\alpha_{11}$	0.60	−0.007	0.90	−0.0054	0.9
$\sigma\alpha_{12}$	0.47	0.0209	0.93	0.0214	0.94
$\sigma\alpha_{13}$	1.17	0.0022	0.96	0.0033	0.97
$\sigma\alpha_{14}$	3.70	−0.1244	0.92	−0.1267	0.91
$\sigma\alpha_{15}$	5.70	0.0657	0.96	0.0657	0.96
σ_1	3.10	0.0198	0.93	0.0200	0.93
σ_2	1.60	0.0229	0.97	0.0226	0.95
σ_3	4.80	0.0556	0.95	0.0552	0.97
σ_4	5.80	0.0601	0.94	0.0602	0.94
σ_5	1.60	−6e−04	0.96	−2e−04	0.97
τ_0	3.60	−0.1314	0.95	2.1513	0.00
τ_1	−1.50	0.0131	0.97	–	–
τ_2	0.05	2e−04	0.95	–	–
% best WAIC		61%			
% best LOOIC		63%			

One hundred simulated data sets were generated under model LTJMM (independent random effects). Both LTJMM-H (heterogeneity in the latent time) and LTJMM models with correlated random effects were fit to all simulated data sets. The best performance was determined by the lowest WAIC and LOOIC between the two models for each data and summarized as "% best" over the 100 simulations. Bias and 95% coverage probabilities are also reported for each parameter.

Table 2. Summary measures at baseline

Diagnostic category	Variable	*n*	Min	Max	Mean	SE
HC	Age	199	30.54	83.87	60.79	0.80
	CSF Aβ42	190	88.80	879.50	377.62	8.22
	CSF α-synuclein	190	592.56	8,608.91	2,201.26	78.86
	CSF p-tau 181	190	5.10	73.30	18.22	0.85
	CSF total tau	188	18.40	223.10	52.48	1.98
	GDS short	199	0.00	15.00	1.29	0.15
	HVLT	198	15.00	35.00	26.02	0.32
	LINEORNT	198	4.00	15.00	13.14	0.14
	PIGD	198	0.00	0.80	0.02	0.01
	REM sleep	198	0.00	11.00	2.86	0.17
	SCOPA	197	0.00	35.00	8.10	0.52
	SFT	198	22.00	80.00	51.90	0.80
	STAI	196	40.00	105.00	57.08	1.00
	Tremor	197	0.00	0.64	0.03	0.01
	UPDRS total	197	0.00	20.00	4.63	0.32
	UPSIT	199	10.00	40.00	33.89	0.36
Prodromal	Age	64	58.37	82.12	68.75	0.73
	HVLT	63	9.00	33.00	21.79	0.67
	LINEORNT	62	3.00	15.00	11.97	0.29
	PIGD	63	0.00	0.60	0.10	0.02
	REM sleep	63	1.00	14.00	7.54	0.48
	SCOPA	64	1.00	39.00	15.50	1.22
	SFT	63	26.00	75.00	45.00	1.37
	Tremor	63	0.00	0.45	0.08	0.02
	UPDRS total	63	0.00	31.00	12.32	0.99
	UPSIT	61	7.00	35.00	17.18	0.84
PD	Age	430	33.63	84.71	61.57	0.47
	CSF Aβ42	416	129.20	796.50	370.82	4.91
	CSF α-synuclein	416	332.93	6,694.55	1,847.87	38.72
	CSF p-tau 181	414	4.70	94.10	15.74	0.50
	CSF total tau	412	14.40	121.00	44.80	0.91
	GDS short	428	0.00	14.00	2.32	0.12
	HVLT	426	9.00	36.00	24.47	0.24
	LINEORNT	426	5.00	15.00	12.78	0.10
	PIGD	411	0.00	1.40	0.23	0.01
	REM sleep	416	0.00	13.00	4.50	0.14
	SCOPA	430	0.00	71.00	12.55	0.45
	SFT	426	20.00	103.00	48.79	0.57
	STAI	425	40.00	137.00	65.25	0.89
	Tremor	411	0.00	1.82	0.50	0.02
	UPDRS total	406	7.00	76.00	32.52	0.67
	UPSIT	427	1.00	40.00	22.29	0.40

PIGD, posture instability and gait difficulty; RBD, REM sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; MOCA, Montreal Cognitive Assessment; HVLT, Hopkins Verbal Learning Test; GDS, Geriatric Depression Scale; SFT, Semantic Fluency Test; SCOPA, Scale for Outcomes in Parkinson's; LINEORT, Line Orientation Test; UPDRS, Unified Parkinson's Disease Rating Scale; STAI, State Trait Anxiety Index; PD, Parkinson's disease; HC, healthy control; SE, standard error.

Table 3. Percentiles and percentile ranks of outcomes

Outcomes	Min	25th	50th	75th	Max
Tremor	0.00	0.05	0.12	0.43	2.36
PIGD	0.00	0.07	0.14	0.25	4.00
UPDRS I	0.00	1.96	4.35	7.73	36.00
UPDRS II	0.00	0.67	2.12	6.17	40.00
UPDRS III	0.00	1.00	6.21	19.22	86.00
UPDRS total	0.00	4.64	12.53	30.21	156.00
MOCA	2.00	24.70	26.76	28.30	30.00
UPSIT	1.00	15.26	23.82	33.26	40.00
SBR striatum	0.10	1.21	1.86	2.44	4.24
CSF α-synuclein	332.93	1,372.50	1,865.90	2,450.67	8,608.91
CSF Aβ42	88.80	312.24	372.85	435.31	879.50
CSF p-tau 181	4.70	9.84	13.35	20.59	94.10
CSF total tau	14.40	33.50	42.35	55.23	223.10
HVLT	4.00	19.63	23.62	27.92	36.00
REM sleep	0.00	1.52	3.54	7.12	17.00
GDS short	0.00	0.52	1.10	2.73	15.00
LINEORNT	0.00	10.55	12.52	13.81	15.00
SFT	7.00	39.79	47.83	55.92	103.00
STAI	40.00	47.70	58.07	71.06	150.00
SCOPA	0.00	5.12	9.79	19.51	83.00

PIGD, posture instability and gait difficulty; RBD, REM sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; MOCA, Montreal Cognitive Assessment; HVLT, Hopkins Verbal Learning Test; GDS, Geriatric Depression Scale; SFT, Semantic Fluency Test; SCOPA, Scale for Outcomes in Parkinson's; LINEORT, Line Orientation Test; UPDRS, Unified Parkinson's Disease Rating Scale; STAI, State Trait Anxiety Index.

and CSF total tau (t-tau). These outcomes have long been known to be associated with the development and progression of PD. We also included DaTSCAN summaries of the striatum.

Statistical Methods

We employed the LTJMM to assess the relationship between CSF measures, imaging biomarkers, and the clinical outcomes and further explored its use to study progression among diagnostic groups. The LTJMM allows for multiple outcomes, measured longitudinally for each patient potentially at different visit times. Suppose y_{ijk} represents the outcome k ($k = 1, 2, \ldots, p$) observed at time j ($j = 1,\ldots, q$) for each individual, i ($i = 1, 2,\ldots, n$), t_{ij} is the time of measurement, and x_{ijk} is a set of covariates for the ith individual at time j. The model is given by

$$y_{ijk} = x^t_{ijk}\beta_k + y_k(t_{ij} + \delta_i) + \alpha_{0ik} + \alpha_{1ik}t_{ij} + \varepsilon_{ijk}, \quad (1)$$

where β_k and y_k are coefficients corresponding to covariates and time, respectively, α_{0ik} and α_{1ik} are subject-specific random intercept and slope components specific to each outcome and assumed to follow a multivariate normal distribution with mean 0 and variance-covariance matrix, D. A random or "latent" time shift, δ_i, specific to each subject but shared across all outcomes is introduced to quantify the disease progression of an individual relative to the population. The δ_i are assumed to follow a normal distribution with mean 0 and variance σ^2_δ. Finally, the independent random error term ε_{ijk} is assumed to follow a normal distribution with mean 0 and variance σ^2. Additionally, δ_i is assumed to be independent of the subject-specific random effects and the pure random errors. However, the subject-specific random intercepts and slopes are allowed to be correlated to reflect the dependence among outcomes. Heterogeneity in the latent time (LTJMM-H) is introduced by modeling σ^2_δ as follows:

$$\sigma^2_{\delta i} = \exp(Z^t_i\tau), \quad (2)$$

where Z_i represents a set of covariates for individual assumed to be a subset of X_i and τ is a vector of parameters corresponding to the covariates. With this, we have allowed the variability of the latent time to vary across subjects.

To efficiently estimate the model parameters, an MCMC approach was adopted. We assign a weakly informative normal prior with zero mean and variance 100 on the regression and y_k parameters. For the variance components of the time shift and random error term, a weakly informative half-Cauchy (0, 2.5) distribution was imposed. For the subject-specific random effects, the variance-covariance matrix was first decomposed into variance and correlation components. A Cholesky decomposition was then applied to the correlation matrix to ensure efficiency and stability. A half-Cauchy (0, 25) prior was placed on the variance part and the LKJ prior on the correlation matrix as recommended by Lewandowski et al. [10]. To ensure identifiability of the model, we constrained the random intercepts for each subject to sum to zero (i.e., $\Sigma^p_{k=1}\alpha_{0ik} = 0$).

Simulation Study

A limited simulation study was conducted to study the effect of assuming a homogeneous latent time variance when in fact there is between-subject variation in latent time variance. We generated data from the LTJMM-H model with uncorrelated random effects as follows:

$$y_{ijk} = \beta_{0k} + \beta_{1k}x_{1ijk} + \beta_{2k}x_{2ijk} + y_k(t_{ij} + \delta_i) + \alpha_{0ik} + \alpha_{1iktij} + \varepsilon_{ijk},$$

where $i = 1, 2,\ldots, 5$, $x_{1ijk} \sim \text{bin}(0.5)$, $x_{2ijk} \sim N(55, 5^2)$, α_{0ik}, and α_{1ik} are random intercept and slope, respectively, and assumed to be independent. Further,

$$\log(\sigma^2_{\delta i}) = \tau_0 + \tau_1 x_{1i} + \tau_2 x_{2i}.$$

The sample size was set to 100 with each individual time, $t_{ijk} \sim$ Uniform(0, 6).

Simulation parameters were set by fitting the LTJMM model to 5 of the original outcomes in the PPMI study. Heterogeneity was introduced by allowing a small effect of x_{1i} and x_{2i} in the submodel. Actual parameter values for both fixed effects and variance components can be seen in Table 1. We then simulated 100 datasets and fitted both the LTJMM and the LTJMM-H with correlated random

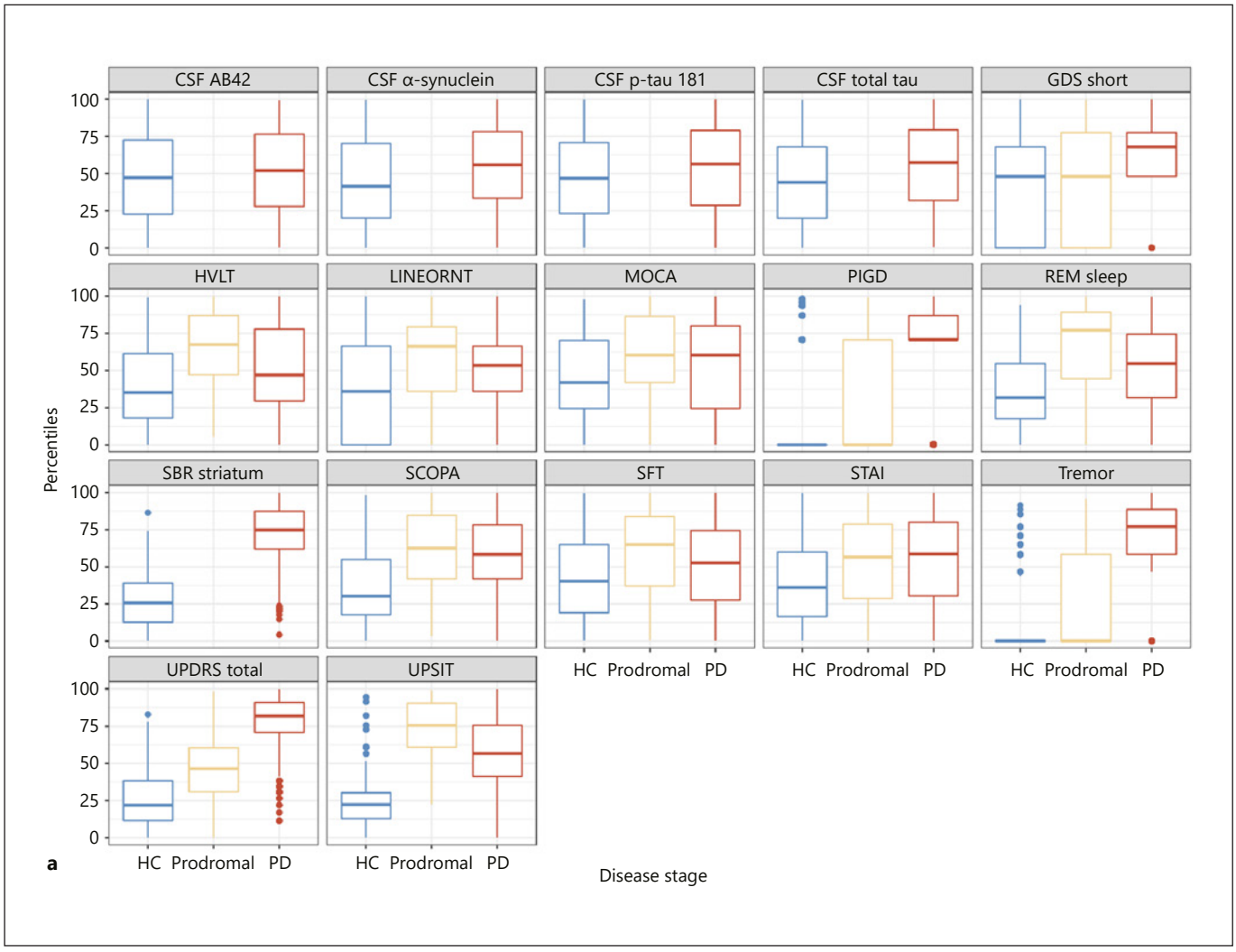


Fig. 1. a Box plots of the observed quantiles of each outcome by disease category. **b** Spearman's correlation between the outcomes. PIGD, posture instability and gait difficulty; RBD, REM sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; MOCA, Montreal Cognitive Assessment; HVLT, Hopkins Verbal Learning Test; GDS, Geriatric Depression Scale; SFT, Semantic Fluency Test; SCOPA, Scale for Outcomes in Parkinson's; LINEORT, Line Orientation Test; UPDRS, Unified Parkinson's Disease Rating Scale; STAI, State Trait Anxiety Index; PD, Parkinson's disease; HC, healthy control.

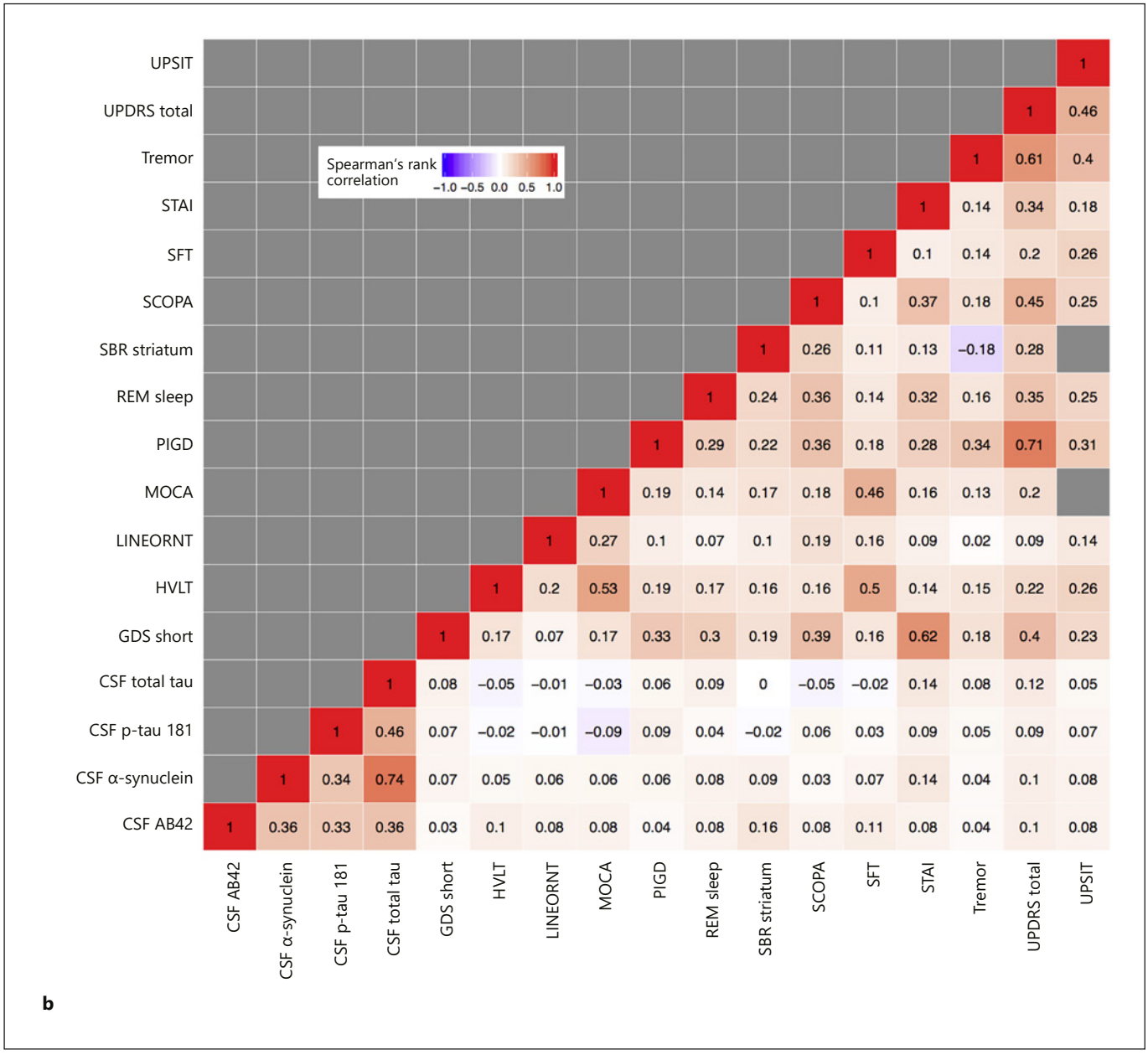
(Figure continued on next page.)

effects to the data. To apply the MCMC algorithm, we specified two Markov chains, each run for 2,000 iterations including 1,000 warm-up iterations, which were discarded. Simulation results are presented in Table 1. Both models consistently estimated fixed parameters, variance parameters of the random effects, and the random error with high coverage probability. However, the LTJMM performed poorly in the estimation of the latent time variance resulting in severe bias and zero coverage probability for the true parameters. Also, the LTJMM-H model resulted in lower widely applicable information criterion (WAIC) and leave-one-out cross-validation information criterion (LOOIC) values in many of our simulations, demonstrating better model fit compared to LTJMM.

Analysis of PPMI

Exploratory Analysis

Table 2 summarizes patient characteristics and outcomes at baseline in the PPMI dataset. Using ANOVA, the diagnostic groups at baseline yielded a significant dif-



Color version available online

1

ference in means, with all p values <0.05. The average age was higher in the prodromal group compared to HC and PD categories. Tremor, one of the main clinical manifestations of PD, was unsurprisingly higher on average among PD patients relative to the other diagnostic groups.

To compare outcomes on a common scale, raw scores were transformed into percentiles using a weighted empirical cumulative distribution function. The inverse of the sample proportion of diagnostic category for each outcome was used as weights. Table 3 shows percentile scores corresponding to selected percentile ranks. Figure 1a shows the box plot of the percentile ranks, representing disease severity, for each outcome in each of the three diagnostic categories. From this figure, we observe broad variation in percentile ranks of the outcomes for each disease group. There were outlying observations, particularly for the motor outcomes, tremor, PIGD, and UPDRS total and olfactory measures indexed by UPSIT. The median ranks differ among diagnostic groups and were lowest in the HC group. The Spearman rank correlations be-

Table 4. Results of posterior estimates of parameters and corresponding 95% credible intervals for model with 8 outcomes from the Parkinson's Progression Markers Initiative (PPMI)

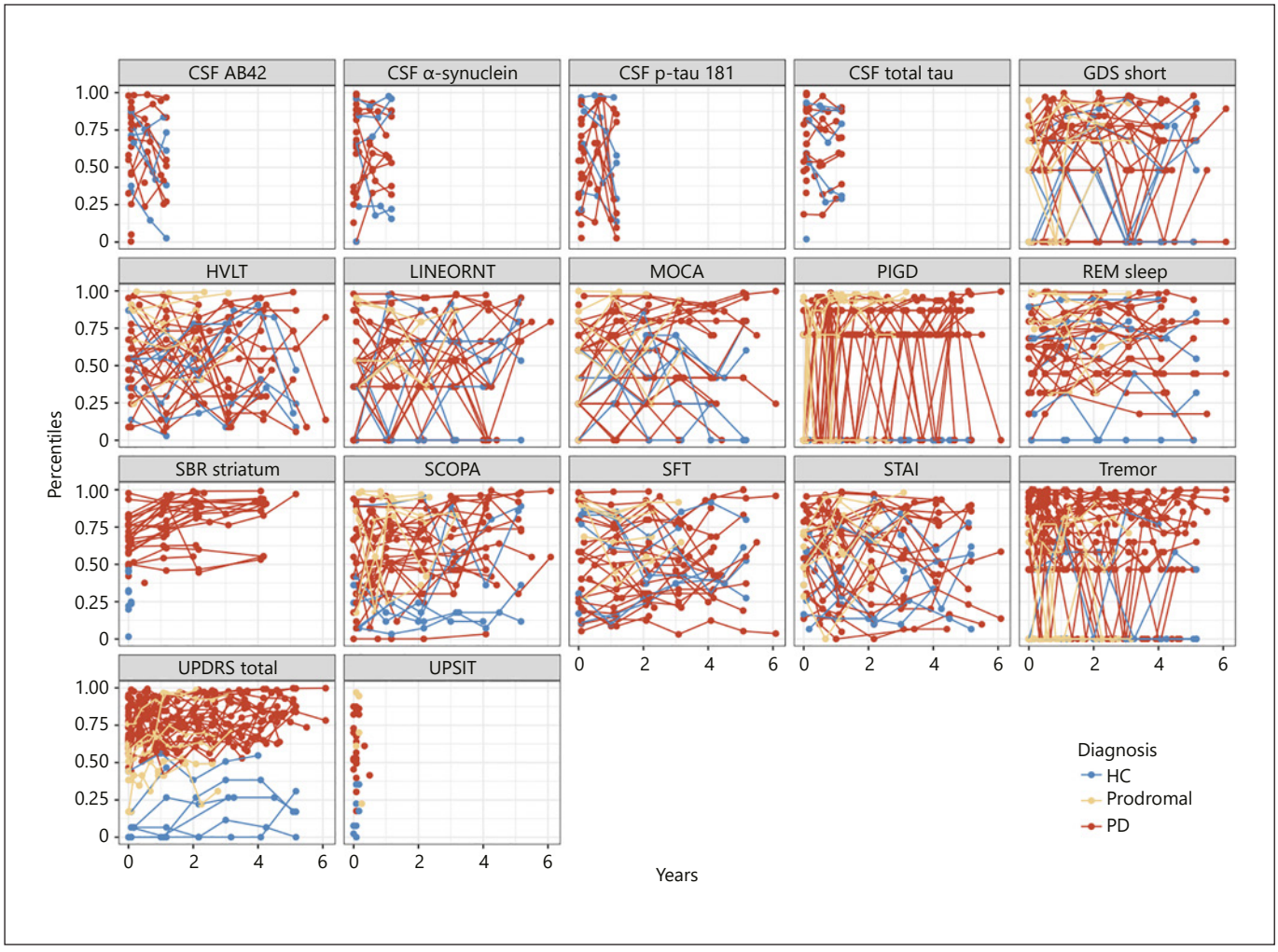
Parameter	Posterior mean	95% credible interval
CSF α-synuclein		
Intercept	0.8976	(0.4341, 1.3793)
Age	–0.0123	(–0.0200, –0.0050)
Female	–0.1160	(–0.2745, 0.0401)
Latent time, γ_1	0.0274	(0.0190, 0.0363)
Error variance, σ_1	0.4288	(0.4023, 0.4552)
PIGD		
Intercept	–1.9367	(–2.6663, –1.1232)
Age	0.0114	(–5e-04, 0.0226)
Female	0.0866	(–0.1579, 0.3315)
Latent time, γ_2	0.1415	(0.1225, 0.1621)
Error variance, σ_2	1.1573	(1.1347, 1.1788)
REM sleep		
Intercept	0.2397	(–0.1764, 0.6504)
Age	–0.0053	(–0.0118, 8e-04)
Female	–0.1052	(–0.2437, 0.0491)
Latent time, γ_3	0.0446	(0.0364, 0.053)
Error variance, σ_3	0.5806	(0.5646, 0.5968)
SBR striatum		
Intercept	–0.8843	(–1.2600, –0.5038)
Age	0.0157	(0.0098, 0.0215)
Female	–0.0677	(–0.2017, 0.0604)
Latent time, γ_4	0.0599	(0.0502, 0.0697)
Error variance, σ_4	0.2118	(0.2002, 0.2237)
Heterogeneous parameters		
τ_1	5.1320	(4.4178, 5.8806)
τ_2	–0.0071	(–0.018, 0.0033)
τ_3	–0.0184	(–0.2247, 0.1919)
SCOPA		
Intercept	–2.3726	(–2.7089, –2.0074)
Age	0.0336	(0.0279, 0.0387)
Female	0.5104	(0.3944, 0.6296)
Latent time, γ_5	0.0406	(0.0338, 0.0486)
Error variance, σ_5	0.4683	(0.4559, 0.4819)
Tremor		
Intercept	0.3527	(–0.3759, 1.1549)
Age	–0.0174	(–0.0295, –0.0064)
Female	–0.1660	(–0.3967, 0.0813)
Latent time, γ_6	0.1236	(0.1084, 0.1404)
Error variance, σ_6	0.8108	(0.7962, 0.827)
UPDRS total		
Intercept	–0.2535	(–0.6679, 0.176)
Age	0.0059	(–8e-04, 0.0123)
Female	–0.0364	(–0.1758, 0.0979)
Latent time, γ_7	0.0880	(0.0769, 0.0999)
Error variance, σ_7	0.3494	(0.343, 0.3559)
UPSIT		
Intercept	–1.2072	(–1.6254, –0.7688)
Age	0.0206	(0.0138, 0.0275)
Female	–0.2175	(–0.348, –0.0731)
Latent time, γ_8	0.0511	(0.0418, 0.0615)
Error variance, σ_8	0.3688	(0.0688, 0.6602)
Model criteria		
WAIC	61054.80	
LOOIC	62158.01	

tween outcomes are displayed in a heat map in Figure 1b. Total tau was highly correlated with CSF α-synuclein (r = 0.74). PIGD was also strongly associated with UPDRS total (r = 0.71). Generally, motor outcomes were highly correlated. CSF biomarkers were also generally correlated with each other but weakly associated with clinical outcomes. A cognitive outcome, MOCA, was moderately positively correlated with verbal working memory assessment indexed by the HVLT (r = 0.53), and depression (GDS) was moderately correlated with anxiety (STAI) with r = 0.62.

Figure 2 displays the individual profiles for each outcome for 50 randomly selected participants. Tremor, PIGD, UPDRS total, MOCA, specific binding ratio (SBR) on striatum, GDS, and STAI were first measured prior to baseline (at screening). For all participants, the UPSIT score was taken only at baseline because it does not change with disease progression. While there was considerable variation between patients, individual trajectories also appear to show a slight progression (increase) over time.

Statistical Analysis

Figure 1a demonstrates that some of the outcomes more clearly distinguished between HC and PD than others. These measures were likely to carry more information related to latent time. Hence, our statistical analysis was limited to 8 of these outcomes, namely, PIGD, SBR striatum, SCOPA, tremor, UPDRS, UPSIT, CSF α-synuclein and REM sleep behavior disorder. An analysis with all 17 outcomes is included in the online supplementary Appendix (for all online suppl. material, see www.karger.com/doi/10.1159/000488780).



Color version available online

Fig. 2. Observed subject-specific profile showing the profiles for 50 randomly selected subjects per outcome and by disease category. PIGD, posture instability and gait difficulty; RBD, REM sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; MOCA, Montreal Cognitive Assessment; HVLT, Hopkins Verbal Learning Test; GDS, Geriatric Depression Scale; SFT, Semantic Fluency Test; SCOPA, Scale for Outcomes in Parkinson's; LINEORT, Line Orientation Test; UPDRS, Unified Parkinson's Disease Rating Scale; STAI, State Trait Anxiety Index; PD, Parkinson's disease; HC, healthy control.

We analyzed the data using the model described in section 3 with three Markov chains, each run for 8,000 iterations, including a warm-up of 4,000 iterations, which was discarded. The LTJMM and the proposed LTJMM-H were implemented and the best model was selected according to two model selection criteria, the WAIC and the LOOIC. The LTJMM-H was chosen over the LTJMM since it produced lower WAIC and LOOIC values of 61,054.8 and 62,158.01 compared to 61,203.94 and 62,290.43 for the homogeneous version, respectively. For the analysis of all 17 outcomes, WAIC and LOOIC for the LTJMM-H were 117,860 and 120,166.7, which were lower compared to the LTJMM values of 143,431.9 and 145,055.2, respectively.

Posterior means with corresponding 95% posterior credible intervals are reported in Table 4 for the selected model. Figure 3 shows the density plot of estimated random latent time shifts by diagnostic categories. Although diagnostic category was not included in the model, the model seemed to accurately reveal the temporal ordering of disease state consistent with the diagnosis groups and biology of PD. The distribution of latent time overlapped much more between the HC and prodromal groups compared to the HC and PD groups. This was not surprising

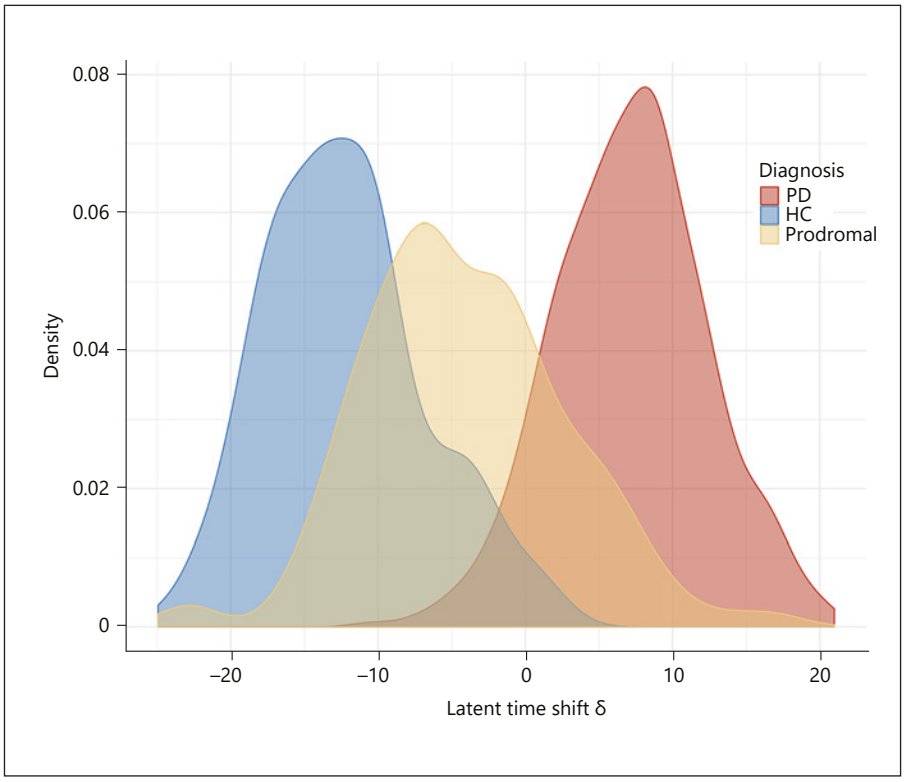


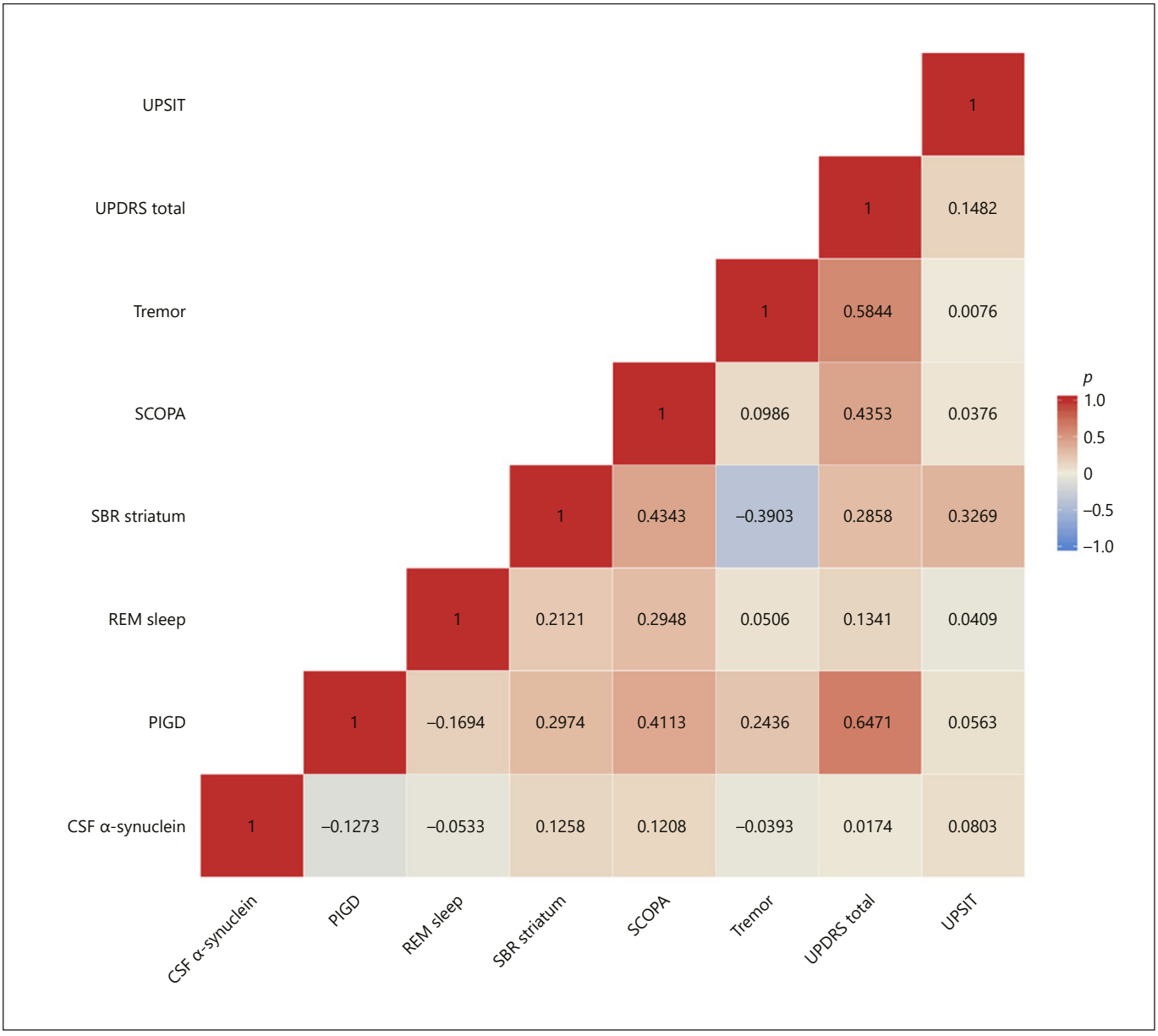
Fig. 3. Distribution of estimate of time shift by disease status. PD, Parkinson's disease; HC, healthy control.

since HC shared some nonmotor features (sleep disturbances, constipation, olfactory deficits, depression) with both prodromal and PD, but HC shared fewer features with PD than prodromal. Thus, the latent time shift estimates provided a continuous alternative to diagnosis, which was objectively derived from a comprehensive joint model of longitudinal measures of disease progression [9]. Comparing results from the two sets of outcomes, Table 4 and online supplementary Table A.1 show that the model was generally robust to the inclusion of less informative outcome measures, but density plots, in Figure 3 and online supplementary Figure A.1, show that latent time was more successful in parsing out the diagnosis groups when the less informative outcomes were omitted.

Figure 4 (and online suppl. Fig. A.2 for all outcomes) shows the correlation of random slopes between pairs of outcomes. Consistent with the observed correlation matrix in Figure 1b, we observed that the slopes for PIGD were correlated with UPDRS total. Similarly, GDS short and STAI showed moderate association. In general, we observed that the biological markers shared a stronger pair-wise association compared to their association with clinical markers.

Figure 5 shows the subject-level predicted severity versus age for 10 randomly selected participants. The model reasonably predicted the trend and direction of individual trajectories, albeit with some variation. For subjects with more time points, the predicted trajectories appeared to perform better, with some reduction in the difference between the observed and predicted trajectories compared to subjects with fewer time points. Placed side-by-side in Figure 6a and b are the observed and predicted subject-level trajectories against age. The two sets of trajectories appeared similar, giving an indication that the model was predicting reasonably well. In addition, SBR striatum and total UPDRS in Figure 6a and b clearly showed the discriminability of these measures between disease stages in the observed and predicted trajectories.

Figure 7 displays the average long-term population trajectories for females with progressive disease. To generate these plots, it was necessary to "calibrate" the independent variables age and latent time by assigning the latent time for any given age. For the purposes of these plots, we assumed the estimated median latent time among HC (−13 years) at the average age for the entire sample (61 years). This figure indicates that PIGD and

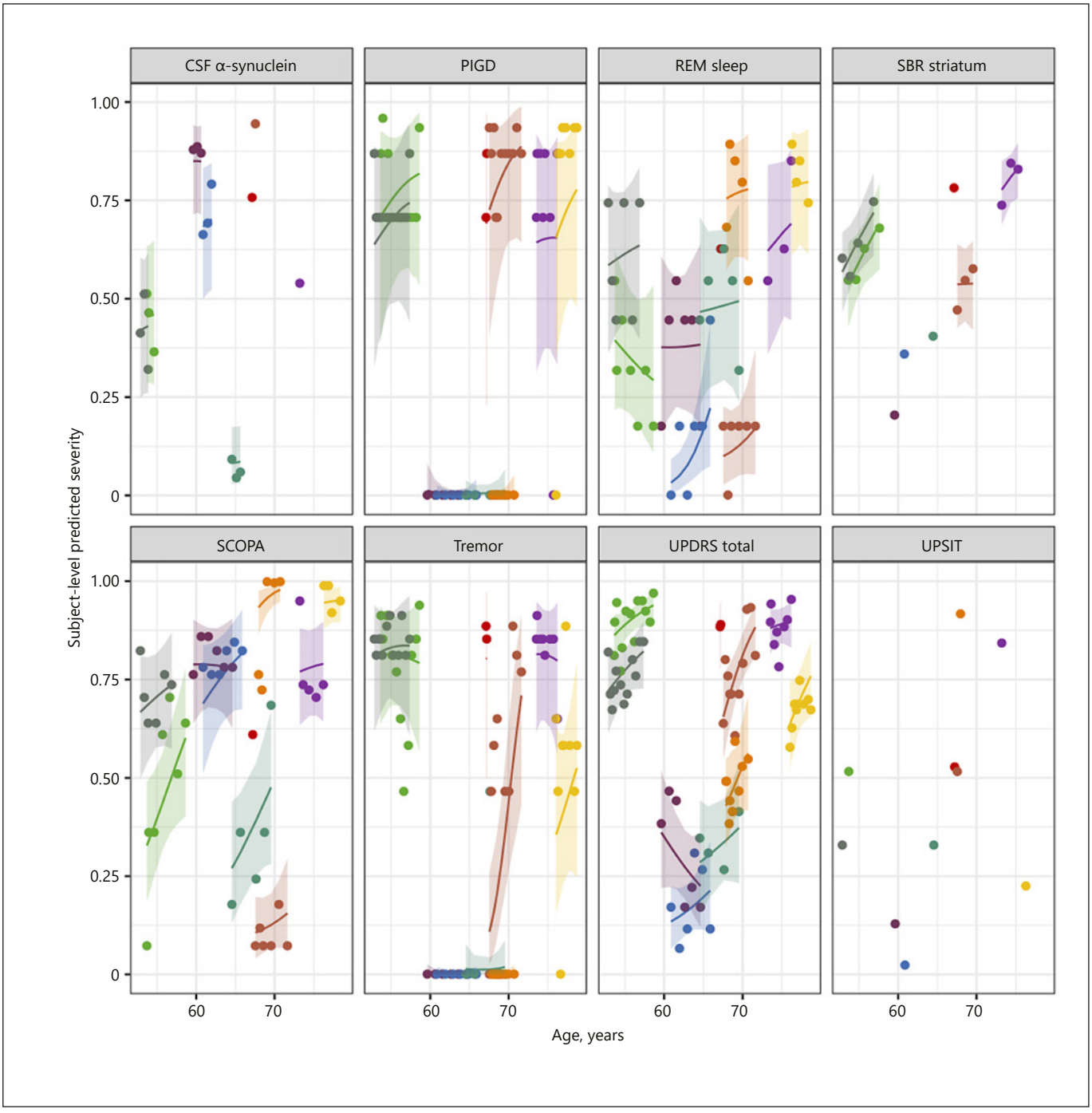


Color version available online

Fig. 4. Correlation between random slopes for pair of outcomes. PIGD, posture instability and gait difficulty; RBD, REM sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; SCOPA, Scale for Outcomes in Parkinson's; UPDRS, Unified Parkinson's Disease Rating Scale; PD, Parkinson's disease; HC, healthy control.

tremor reached the 50th percentile later in the disease progression compared to UPDRS, which occurred much earlier. We also fitted the model treating the UPDRS subtests I, II, and III as separate outcomes (see online suppl. Fig. A.3b). We found that UPDRS I (mental) reached the 50th percentile level first on average, followed by UPDRS III (motor) and II (activities of daily living) in close succession. Due to the strong gender effect, as can be observed from the table of estimates, a similar plot for males will position SCOPA differently in the ordering (see Fig. 8).

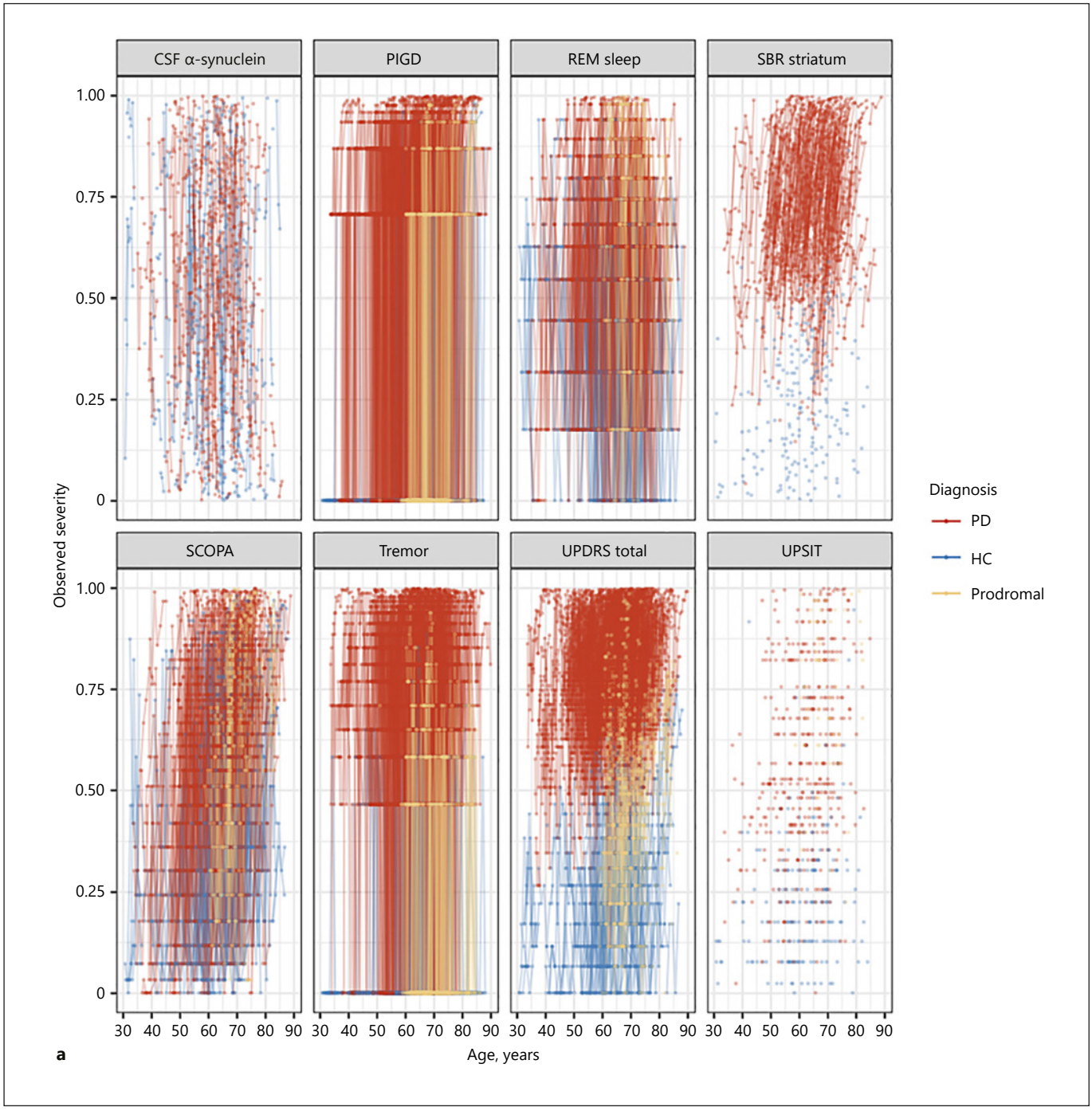
Figure 8, a positional variance diagram, shows the estimated variation in the ordering of marker abnormalities for females and males. Figure 8 was derived by slicing



Color version available online

Fig. 5. Subject-level prediction for random subjects. Each color represents the predicted profile of a randomly selected individual. PIGD, posture instability and gait difficulty; REM sleep, rapid eye movement sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; SCOPA, Scale for Outcomes in Parkinson's; UPDRS, Unified Parkinson's Disease Rating Scale; PD, Parkinson's disease; HC, healthy control.

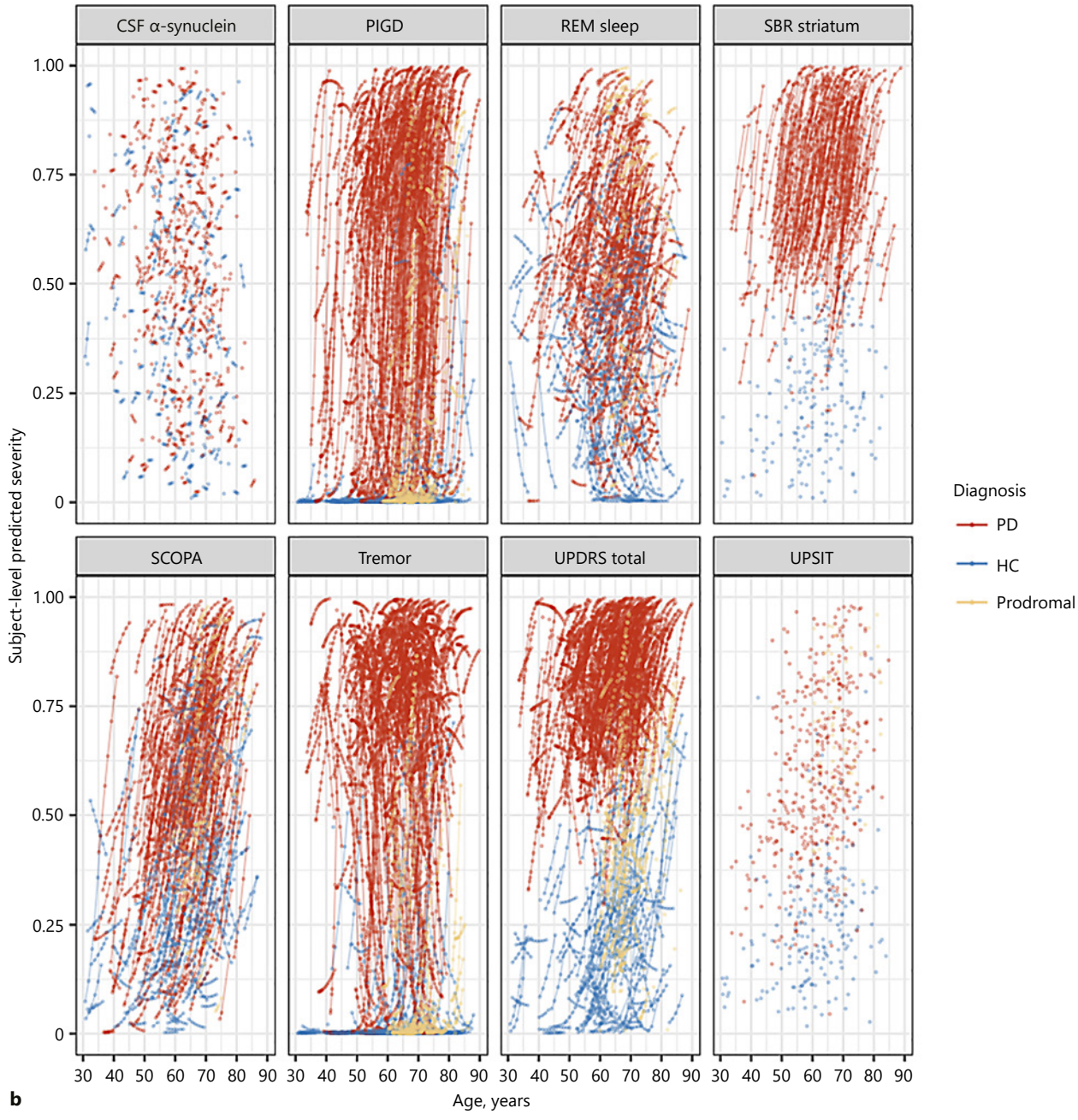
Neurodegener Dis 2018;18:173–190
DOI: 10.1159/000488780



Color version available online

Fig. 6. a Observed trajectories. **b** The predicted profile for all subjects by outcome. PIGD, posture instability and gait difficulty; REM sleep, rapid eye movement sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; SCOPA, Scale for Outcomes in Parkinson's; UPDRS, Unified Parkinson's Disease Rating Scale; PD, Parkinson's disease; HC, healthy control.

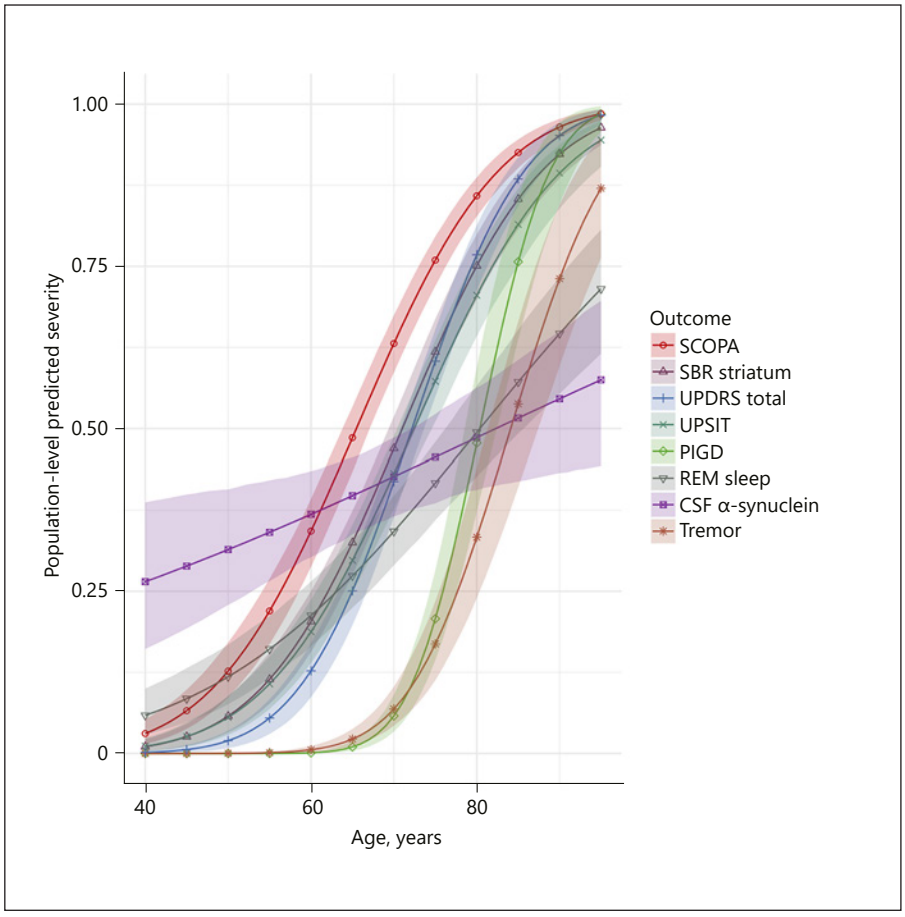
(Figure continued on next page.)



6

Color version available online

Neurodegener Dis 2018;18:173–190
DOI: 10.1159/000488780



Color version available online

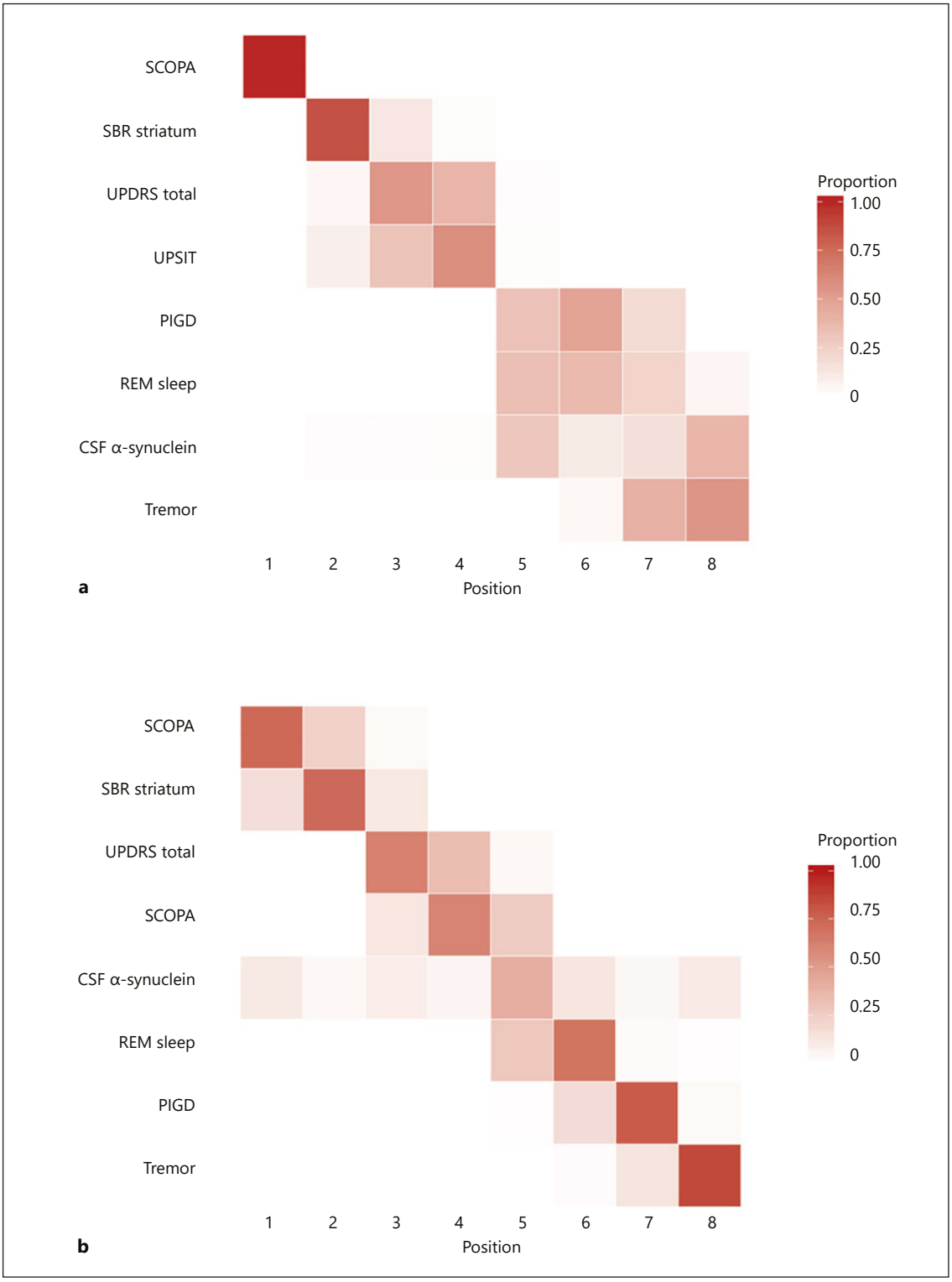
Fig. 7. Population-level predicted severity for females with mean age and median latent time of a healthy control. The legend is ordered by the age at which the predicted severity level for each outcome is 0.5 (sliced horizontally). PIGD, posture instability and gait difficulty; REM sleep, rapid eye movement sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; SCOPA, Scale for Outcomes in Parkinson's; UPDRS, Unified Parkinson's Disease Rating Scale.

Figure 7 horizontally at the 50th percentile and observing the order of markers as they appear from left to right. We used the posterior distribution to evaluate the uncertainty in these estimated orderings. For each draw from the posterior distribution, we generated curves similar to Figure 7 and sliced horizontally at the 50th percentile to determine the ordering of marker abnormalities for that posterior sample. The diagram represents the distribution (proportion of MCMC samples) of the outcome orderings. The bolder the color of the cell, the higher the certainty of the ordering for that outcome. In Figure 9, we display the positional diagram based on subject-level predictions. For each subject in the sample, we ordered the outcomes based on the age at which the 50th percentile of the outcome is attained. As one might expect, there was greater variation in the orderings from subject to subject (Fig. 9) compared to the average ordering for the sample (Fig. 8).

Discussion

In this paper, we applied a joint mixed-effects model with latent time to analyze the various stages of PD included in the PPMI study. We found that although diagnostic category was not included in the model, it seems to accurately reveal the temporal ordering of disease state consistent with diagnosis categorization at baseline. In other words, the distribution of the estimates of the latent time reflected the subjective assessment of patient disease status at baseline.

In addition, the association between estimated random effects revealed that biological markers share a stronger pair-wise association but a weaker association with clinical markers. The biomarkers not being very informative (greater variance and lower signal-to-noise ratio) probably explains this weak association with clinical outcomes. It is worth mentioning that the role of β-amyloid and tau protein as biomarkers in PD is not well established, but a number of publications report altered



Color version available online

Fig. 8. Positional variance diagrams of the central ordering. **a** Female. **b** Male. The *x* axis is the event position. Difference in ordering may be attributed to the strong effect of gender on some of these outcomes. PIGD, posture instability and gait difficulty; REM sleep, rapid eye movement sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; SCOPA, Scale for Outcomes in Parkinson's; UPDRS, Unified Parkinson's Disease Rating Scale; PD, Parkinson's disease; HC, healthy control.

Neurodegener Dis 2018;18:173–190
DOI: 10.1159/000488780

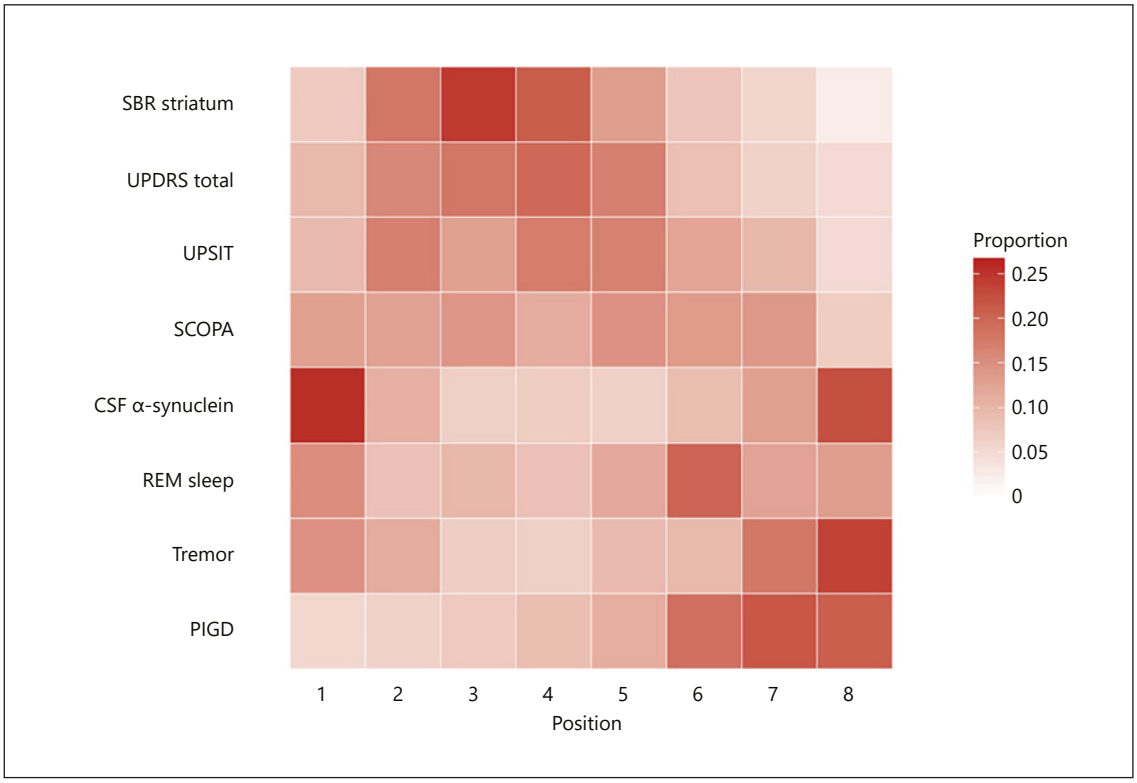


Fig. 9. Positional variance diagram of the central ordering. PIGD, posture instability and gait difficulty; REM sleep, rapid eye movement sleep behavior disorder; UPSIT, University of Pennsylvania Smell Identification Test; SCOPA, Scale for Outcomes in Parkinson's; UPDRS, Unified Parkinson's Disease Rating Scale; PD, Parkinson's disease; HC, healthy control.

CSF Aβ1–42, total tau, or p-tau in patients with PD with or without dementia compared with HC [11]. Clinical markers also tend to be correlated among each other, with PIGD and total UPDRS, and GDS short and STAI sharing strong to moderate associations. Inspection of observed and predicted severity shows that SBR striatum and total UPDRS can provide better discrimination between disease stages.

Long-term population-level trajectories were also derived from the proposed model. The results indicated late manifestation of PIGD and tremor but earlier total SCOPA and total UPDRS abnormalities. The SCOPA and UPDRS include nonmotor features that appear at prodromal disease stages (earlier in the course of the illness). The positional variance diagram provided an ordering consistent with that from the population trajectories. Similar trajectories can be obtained for any subgroup of the population with ease using our modeling approach. It should be kept in mind that the data used to develop the models are restricted to short-term follow-up after enrollment (median <4 years), which might limit predictive accuracy when evaluated over longer periods of disease progression.

We also demonstrated through a limited simulation study that allowing heterogeneity in the variance of latent time can improve LTJMM fit. The variance of the latent time was modeled in terms of observed baseline covariates. Particularly, the simulation indicated that estimation of parameters is consistent with the LTJMM model except that the between-subject latent time variance is severely biased. Given the importance of the subject-specific latent time as a data-driven alternative to categorical disease status, it is critical that the latent time variance is accurately estimated. Violation of the homogeneous latent time can, therefore, be detrimental. We encourage the exploration of different submodels to determine which baseline covariates are significant in explaining the between-subject latent time variation. In the application of the two models to the PPMI study, both WAIC and LOOIC favored the extended LTJMM model with heterogeneous latent time.

Acknowledgments

We are grateful to the PPMI study volunteers and their families. This work was supported by a Biomarkers Across Neurodegenerative Disease (BAND-14-338179) grant from the Alzheimer's Association, Michael J. Fox Foundation, and Weston Brain Institute, and National Institute on Aging grant R01-AG049750. Data used in the preparation of this article were obtained from the PPMI database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org.

Disclosure Statement

The authors have no conflict of interest to declare.

References

1 Kennedy RE, Cutter GR, Wang G, Schneider LS: Using baseline cognitive severity for enriching Alzheimer's disease clinical trials: how does Mini-Mental State Examination predict rate of change? Alzheimers Dement (NY) 2015;1:46–52.
2 Guerrero R, Schmidt-Richbergh A, Ledig C, Tong T, Wolz R, Rueckert D: Instantiated mixed effects modeling of Alzheimer's disease markers. Neuroimage 2016;142:113–125.
3 Sturkenboom IH, Graff MJ, Hendriks JC, Veenhuizen Y, Munneke M, Bloem BR, Sanden MW: Efficacy of occupational therapy for patients with Parkinson's disease: a randomized controlled trial. Lancet Neurol 2014;13:557–566.
4 Klotsche J, Reese JP, Winter Y, Oertel WH, Irving H, Wittchen H, Rehm J, Dodel R: Trajectory classes of decline in health-related quality of life in Parkinson's disease: a pilot study. Value Health 2011;14:329–338.
5 Iddi S, Molenberghs G: A joint marginalized multilevel model for longitudinal outcomes. J Appl Stat 2012;39:24132430.
6 Tsiatis AA, Davidian M: A joint modeling of longitudinal and time-to-event data: an overview. Stat Sin 2004;14:809–834.
7 Luo S, He B: Joint modeling of multivariate longitudinal measurements and survival data with applications to Parkinson's disease. Stat Methods Med Res 2016;25:1346–1358.
8 Luo S: A Bayesian approach to joint analysis of multivariate longitudinal data and parametric accelerated failure time. Stat Med 2014;33:580–594.
9 Li D, Iddi S, Thompson WK, Donohue MC: Bayesian latent time joint mixed effect models for multicohort longitudinal data. Stat Methods Med Res 2017, DOI: 10. 1177/0962280217737566 (Epub ahead of print).
10 Lewandowski D, Kurowicka D, Joe H: Generating random correlation matrices based on vines and extended onion method. J Multivar Anal 2009;100:1989–2001.
11 Brockmann K, Schulte C, Deuschle C, Hauser A-K, Heger T, Gasser T, Maetzler W, Berg D: Neurodegenerative CSF markers in genetic and sporadic PD: classification and prediction in a longitudinal study. Parkinsonism Relat Disord 2015;21;1427–1434.