Progression of clinical markers in prodromal Parkinson's disease and dementia with Lewy bodies: a multicentre study

Stephen Joza,[1] Michele T. Hu,[2] Ki-Young Jung,[3] Dieter Kunz,[4] Ambra Stefani,[5] Petr Dušek,[6] Michele Terzaghi,[7,8] Dario Arnaldi,[9,10] Aleksandar Videnovic,[11] Mya C. Schiess,[12] Wiebke Hermann,[13] Jee-Young Lee,[14] Luigi Ferini-Strambi,[15] Simon J. G. Lewis,[16] Laurène Leclair-Visonneau,[17,18] Wolfgang H. Oertel,[19,20] Elena Antelmi,[21] Friederike Sixel-Döring,[19,22] Valérie Cochen De Cock,[23,24] Claudio Liguori,[25,26] Jun Liu,[27] Federica Provini,[28,29] Monica Puligheddu,[30] Alessandra Nicoletti,[31] Claudio L. A. Bassetti,[32] Jitka Bušková,[33] Yves Dauvilliers,[34] Raffaele Ferri,[35] Jacques Y. Montplaisir,[36,37] Michael Lawton,[38] Han-Joon Kim,[3] Frederik Bes,[4] Birgit Högl,[5] Karel Šonka[6], Giuseppe Fiamingo,[7,8] Pietro Mattioli,[9] Maria Lorena Lavadia,[11] Jessika Suescun,[12] Kyung Ah Woo,[14] Sara Marelli,[15] Kaylena Ehgoetz Martens,[39] Annette Janzen,[19] Giuseppe Plazzi,[29,40] Brit Mollenhauer,[22,41] Mariana Fernandes,[25] Yuanyuan Li,[27] Pietro Cortelli,[28,29] Michela Figorilli,[30] Calogero Edoardo Cicero,[31] Carolin Schaefer,[32] Lily Guiraud,[34] Giuseppe Lanza,[35,42] Jean-François Gagnon,[36,37] Jun-Sang Sunwoo,[43] Abubaker Ibrahim,[5] Nicola Girtler,[9,10] Claudia Trenkwalder,[22,41] Luca Baldelli,[28,29] Amelie Pelletier,[1,36] and Ronald B. Postuma[1,36] for the International REM Sleep Behavior Disorder Study Group

Abstract

The neurodegenerative synucleinopathies, including Parkinson's disease and dementia with Lewy bodies, are characterized by a typically lengthy prodromal period of progressive subclinical motor and non-motor manifestations. Among these, idiopathic REM sleep behavior disorder (iRBD) is a powerful early predictor of eventual phenoconversion, and therefore represents a critical opportunity to intervene with neuroprotective therapy. To inform the design of randomized trials, it is essential to study the natural progression of clinical markers during the prodromal stages of disease in order to establish optimal clinical endpoints.

In this study, we combined prospective follow-up data from 28 centers of the International REM Sleep Behavior Disorder Study Group representing 12 countries. Polysomnogram-confirmed REM sleep behavior disorder subjects were assessed for prodromal Parkinson's disease using the Movement Disorder Society criteria and underwent periodic structured sleep, motor, cognitive,

autonomic and olfactory testing. We used linear mixed-effect modelling to estimate annual rates of clinical marker progression stratified by disease subtype, including prodromal Parkinson's disease and prodromal dementia with Lewy bodies. In addition, we calculated sample size requirements to demonstrate slowing of progression under different anticipated treatment effects.

Overall, 1160 subjects were followed over an average of 3.3±2.2 years. Among clinical variables assessed continuously, motor variables tended to progress faster and required the lowest sample sizes, ranging from 151-560 per group (at 50% drug efficacy and 2-year follow-up). By contrast, cognitive, olfactory, and autonomic variables showed modest progression with higher variability, resulting in high sample sizes. The most efficient design was a time-to-event analysis using combined milestones of motor and cognitive decline, estimating 117 per group at 50% drug efficacy and 2-year trial duration. Finally, while phenoconverters showed overall greater progression than non-converters in motor, olfactory, cognitive, and certain autonomic markers, the only robust difference in progression between Parkinson's disease and dementia with Lewy bodies phenoconverters was in cognitive testing.

This large multicenter study demonstrates the evolution of motor and non-motor manifestations in prodromal synucleinopathy. These findings provide optimized clinical endpoints and sample size estimates to inform future neuroprotective trials.

Author affiliations:

1 Department of Neurology, McGill University, Montreal General Hospital, Montreal, Canada

2 Nuffield Department of Clinical Neurosciences, Division of Neurology and Oxford Parkinson's Disease Centre, University of Oxford, UK

3 Department of Neurology, Seoul National University College of Medicine, Seoul National University Hospital, Seoul, Republic of Korea

4 Clinic for Sleep & Chronomedicine, St. Hedwig-Krankenhaus, Berlin, Germany

5 Medical University Innsbruck, Department of Neurology, Innsbruck, Austria

6 Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czechia

7 Sleep Medicine and Epilepsy Unit, IRCCS Mondino Foundation, Pavia, Italy

8 Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

9 Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Clinical Neurology, University of Genoa

10 IRCCS Ospedale Policlinico San Martino, Genoa, Italy

11 Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

12 Department of Neurology, The University of Texas Health Science Center at Houston, Texas, USA

13 Department of Neurology, University of Rostock, Rostock, Germany

14 Department of Neurology, Seoul National University College of Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, South Korea

15 Sleep Disorders Center, Vita-Salute San Raffaele University, Milan, Italy

16 ForeFront Parkinson's Disease Research Clinic, Brain and Mind Centre, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia

17 Department of Clinical Neurophysiology, CHU de Nantes, Nantes, France

18 Nantes Université, Inserm, TENS, The Enteric Nervous System in Gut and Brain Diseases, Nantes, France

19 Department of Neurology and Section on Clinical Neuroscience, Philipps University Marburg, Marburg, Germany

20 Institute for Neurogenomics, Helmholtz Center for Health and Environment, München-Neuherberg, Germany

21 Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy

22 Paracelsus Elena Klinik, Centre for Movement Disorders, Kassel, Germany

23 EuroMov Digital Health in Motion, University of Montpellier, IMT Mines Ales, 34090 Montpellier, France

24 Department of Neurology and Sleep, Beau Soleil Clinic, Montpellier, France

25 Sleep Medicine Center, Department of Systems Medicine, University of Rome Tor Vergata, Rome, Italy

26 Neurology Unit, University Hospital of Rome Tor Vergata, Rome, Italy

27 Department of Neurology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

28 Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy

29 IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy

30 Sleep Center, Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy

31 Department of Medical, Surgical Sciences and Advanced Technologies, GF Ingrassia, University of Catania, Catania, Italy

32 Department of Neurology, University Hospital Bern (Inselspital), 3010, Bern, Switzerland

33 National Institute of Mental Health, Klecany, Third Faculty of Medicine, Charles University, Prague, Czech Republic

34 Sleep Unit, Department of Neurology, Hôpital Gui de Chauliac, Montpellier, INSERM U1061, Montpellier, F-34093 Cedex 5 France

35 Clinical Neurophysiology Research Unit, Oasi Research Institute-IRCCS. Via Conte Ruggero 73, 94018, Troina, Italy

36 Centre d'Études Avancées en Médecine du Sommeil, Hôpital du Sacré-Cœur de Montréal, Montréal, Canada

37 Department of Psychology, Université du Québec à Montréal, Montréal, Quebec, Canada

38 School of Social and Community Medicine, Bristol Medical School, University of Bristol, UK

39 Department of Kinesiology and Health Sciences, University of Waterloo, Waterloo, Ontario, Canada

40 Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio-Emilia, Modena, Italy

41 Department of Neurosurgery, University Medical Center Goettingen, Göttingen, Germany

42 Department of Surgery and Medical-Surgical Specialties, University of Catania. Via Santa Sofia 78, 95123, Catania, Italy

43 Department of Neurology, Kangbuk Samsung Hospital, Seoul, Republic of Korea

Correspondence to: Ronald B. Postuma MD, MSc

Montreal General Hospital, 1650 Cedar Ave., H3G 1A4 Montreal, QC, Canada

E-mail: ron.postuma@mcgill.ca

Running title: Progression in prodromal PD and DLB

Keywords: REM sleep behavior disorder; Parkinson's disease; dementia with Lewy bodies; prodromal stage; evolution

Introduction

Despite much promise, no therapeutic intervention has been able to alter the progression of the neurodegenerative synucleinopathies,[1–3] which include Parkinson's disease (PD), dementia with Lewy bodies (DLB), and multiple system atrophy (MSA). Aside from drug inefficacy, the lack of benefit could also reflect the possibility that the underlying neurodegenerative process has already progressed to a point beyond which no intervention would benefit. Therefore, targeting the prodromal stages of disease, when time still remains to prevent irreversible degeneration, could be the critical point at which to intervene.[4]

Synucleinopathies are distinctive for both a typically long prodromal period prior to phenoconversion to the overt stages of disease and for the involvement of multiple clinical domains, including motor and cognitive abnormalities, olfactory dysfunction, constipation, dysautonomia, and sleep disorders.[5] Among these, idiopathic REM sleep behavior disorder (iRBD), a parasomnia characterized by loss of REM atonia and consequent dream-enactment behavior, is common in all synucleinopathies.[6] It is also a powerful predictor of phenoconversion: the vast majority (>80%) of individuals with iRBD will ultimately develop an overt degenerative synucleinopathy, with a phenoconversion rate of approximately 6-8% per year.[7,8]

iRBD subjects are therefore ideal candidates for neuroprotective trials. However, optimal endpoints to assess drug efficacy have yet to be established and are required to ensure that future trials are optimally designed. It is unclear to what degree different prodromal markers progress in the early stages of disease. Moreover, it remains to be established how a given clinical marker's progression is affected by disease subtype (e.g., prodromal PD vs. prodromal DLB)[9].

Although previous longitudinal multicenter studies have measured the degree to which clinical markers are predictive of phenoconversion in iRBD,[7,8,10] a systematic approach quantifying the progression of each marker over time has not been performed. Those studies that have longitudinally and systematically assessed marker progression in iRBD have been from single centers,[9,11] or required the use of expensive or sophisticated biomarker analyses that may not be suitable as primary outcome measures in Phase 3 trials.[12,13]

In the present study, we combined the prospective results of 28 centers of the International RBD Study Group (IRBDSG) to: (i) assess the progression of clinical motor and non-motor markers in iRBD subjects over 5 years of follow-up; (ii) determine to what degree this progression differs depending on phenoconversion type; and (iii) calculate required sample sizes to inform the design of randomized neuroprotective trials for prodromal synucleinopathies.

Materials and methods

Study subjects

All study subjects had polysomnogram-confirmed iRBD according to standard criteria[14] and were without parkinsonism or dementia at baseline. Data were collected between 2003 – 2021, with the majority of subjects (80.0%) recruited after 2014 (**Supplementary Fig. 1**). Subjects were systematically assessed at baseline visit, and for inclusion, were required to have at least one follow-up examination. To reflect the situation of a clinical trial, in the primary analysis, subjects were required to meet MDS research criteria for probable prodromal PD, defined according to the criteria as having at least an 80% probability of prodromal Parkinson's disease[5] (using all information available at each center). For subjects that did not meet criteria at baseline but did in subsequent years (13.1% of all subjects), the baseline year was set to the first year in which criteria were met. Ethics approval was obtained from the local institutional boards of each center with subject consent in accordance with the Declaration of Helsinki.

Study procedures

Subjects underwent periodic structured sleep, motor, cognitive, autonomic, and olfactory testing on an approximately annual basis. For inclusion, we did not require that each marker was tested in each patient; rather, centers sent results for all markers that were systematically assessed (detailed in **Supplementary Table 1**). To be analyzed, each marker of interest needed to be systematically assessed by at least two centers and in at least 100 subjects at baseline. Markers included:

i. Standardized motor examination: tested with the MDS-UPDRS-III. For the primary analysis, we combined both the 2008 and 1987 versions of the UDPRS. When the 1987 UPDRS-III was used (36% of subjects at baseline), scores were adjusted by multiplying by a weighting factor of 1.2[15]; an intercept term (i.e. the addition of 2.3) was not used since the calibration was originally developed for early PD, rather than prodromal PD, and would have lead to inaccurately inflated baseline MDS-UPDRS-III scores (e.g. a minimum score of 2.3 for a completely normal UPDRS-III).

ii. Standardized motor symptoms: MDS-UPDRS-II. If the 1987 UPDRS-II was used, scores were adjusted by multiplying by a weighting factor of 1.1 and adding an intercept of 0.2.[15]

iii. Standardized non-motor symptoms: MDS-UPDRS-I.

iv. Quantitative motor testing: Timed-up-and-go (TUG)[16] and Purdue Pegboard (scores reported are the 30 second task involving both hands).[17] Since one center (Houston) used a longer distance TUG (14 meters, rather than 6 meters), scores were additionally standardized to TUG velocity in meters per second (m/s) by dividing the distance of the task by time.

v. Olfaction: 40-item University of Pennsylvania Smell Identification Test (UPSIT), 12-item Cross-Cultural Smell Identification Test (CCSIT), or the 12- or 16-item Sniffin' Sticks (SS) tests. To harmonize results, z-scores were created for each test stratified by sex and/or age using published normatives and averaged.[18–21]

vi. Sleep: Epworth Sleepiness Scale (ESS),[22] Insomnia Severity Index (ISI),[23] Pittsburgh Sleep Quality Index (PSQI),[24] and the REM Sleep Behavior Disorder Screening Questionnaire (RBDSQ).[25]

vii. Office-based cognitive testing: Mini-Mental State Examination (MMSE)[26] and Montreal Cognitive Assessment (MoCA).[27]

viii. Autonomic symptoms: Scales for Outcomes in Parkinson's disease - Autonomic Dysfunction (SCOPA-AUT) scale.[28]

ix. Orthostatic blood pressure: assessed supine and after 1-3 minutes standing. Since the timing and number of standing measurements varied between centers, postural scores from 1-3 minutes were averaged together.

x. Psychiatric symptoms: Beck Depression Inventory,[29] Beck Anxiety Inventory,[30] 30-item Geriatric Depression Scale (GDS),[31] and the Hospital Anxiety and Depression Scale (HADS).[32] To harmonize scores, z-scores were created for each test using the mean and standard deviation at baseline. Individual test z-scores were then averaged to create overall z-scores for depression and anxiety.

Statistical Analysis

Statistical analyses were performed using *R* (version 4.1.2) and Stata (version 13.0).

Outcomes

The progression of variables of interest are described using annual mean and standardized response mean (SRM), which is computed by dividing the mean change from baseline of each individual patient by the standard deviation of the change of the total cohort (allowing diverse measures to be compared to one another). Linear mixed-effect modeling (LMEM)[33] was used to estimate the yearly progression rate of each variable of interest with subject (random slopes) and study center (random intercepts) as random effects and baseline age and follow-up year as fixed effects. Visual inspection of residual plots for each variable did not reveal obvious deviations from homoscedasticity or normality (**Supplementary Fig. 2**). Estimates of the annual progression rates were subdivided by phenoconversion status (PD-phenoconverters, DLB-phenoconverters, and those not known to have phenoconverted during 5 years of follow-up) and are displayed along with the overall estimated progression rate for the total cohort. MSA-phenoconverters were included as part of the total cohort analysis, but the progression of MSA-phenoconverters, specifically, could not be accurately calculated due to low numbers. Rates of progression between different sub-groups were compared using interaction terms between follow-up year and phenoconversion status; *p*-values were obtained by likelihood ratio tests of the full model with the interaction term against the model without the interaction term. Survival analysis for subjects that phenoconverted to a defined neurodegenerative disease was performed using Kaplan-Meier analysis to estimate annual phenoconversion risk.

Secondary analyses examining progression rates stratified by baseline age and by sex were performed. For age analysis, we excluded subjects over the age of 79 years at baseline since too few were studied to allow reliable estimates (**Supplementary Fig. 1**; also note that subjects of advanced age might be excluded from enrollment in a neuroprotective clinical trial).

Missing data

Imputation by linear interpolation[34] was used if data was missing in a single follow-up year between two other data points. Since data were not collected in years following a subject's phenoconversion, and since subsequent treatment could reduce the estimation of a marker's

progression, values were imputed in these years by adding the mean change of the whole group during that year to the last measured value (i.e., at phenoconversion).[11]

Sample size calculations

Sample size estimates for a hypothetical intervention to slow disease progression of each variable of interest were estimated by comparison of slopes between LMEMs for treated and untreated groups.[35] Sample size estimates were also calculated for time-to-event analyses[36] for a hypothetical trial in which phenoconversion is the primary outcome. Additional time-to-event analyses for significant motor decline (defined as a sustained increase in MDS-UPDRS-III of ≥ 4 points),[37] a significant cognitive decline (defined as a sustained reduction in MoCA ≥ 3 points, i.e., an effect size ≈ 1 according to the baseline MoCA standard deviation), or a combined milestone of cognitive and/or motor decline. Similarly, a significant increase in the combined MDS-UPDRS-I+II+III score was defined as a sustained increase ≥ 12 points, based on the baseline standard deviation. A sustained change was defined as a change in score that was observed in two consecutive years. Sample sizes are presented for a 2-arm parallel trial in which treatment is expected to reduce the rate of progression by a constant amount throughout follow-up. Presented are required sample sizes to detect 30% or 50% treatment effects for a 2- or 3-year trial with periodic 6 month-follow-up (for continuous variable analysis) specifying 80% power and 2-sided alpha=0.05.

Data availability

De-identified subject data used in this study are available upon reasonable request from the corresponding author (R.B.P.).

Results

Subjects

Detailed baseline demographics for each center are shown in **Supplementary Table 1** and summarized in **Table 1**. Data were collected from a total of 1647 subjects from 28 centers in 12 countries, from which 210 were excluded for having only a single baseline visit, and 1 was excluded due to a diagnosis of PD at baseline. From the remaining 1436 subjects, 1160 (80.8%) met MDS prodromal PD criteria and were included in the primary analysis. Since only 10% of

subjects had follow-up data beyond five years, the majority of whom were followed by a single center (Montreal; **Fig. 1A**), analyses of variable progression and sample size calculations were limited to data from the first five years of follow-up. Mean age at baseline was 68.5 ± 7.0 years, 78.4% were male, time from iRBD diagnosis was 1.28 ± 2.3 years, and time from self-reported iRBD symptom onset was 6.4 ± 6.4 years. The mean follow-up time (i.e., the duration between baseline and last examination or time of phenoconversion) was 3.3 ± 2.2 years, translating to 3828 total person-years of follow-up.

During five years of follow-up, 220 subjects were known to have phenoconverted to a defined neurodegenerative disease, including 129 (58.6%) who developed parkinsonism as the first disease manifestation (of whom 11 were eventually diagnosed with MSA) and 41.4% who developed dementia first. Using Kaplan-Meier analysis, this corresponded to a phenoconversion rate of 4.4% at 1 year, 18.2% at 3 years, and 31.7% at 5 years (**Fig. 1B**). Baseline characteristics of subjects who phenoconverted within 5 years are summarized in **Supplementary Tables 2-3**. DLB-phenoconverters were significantly older than both PD-phenoconverters and non-converters (DLB=72.9±6.5, PD=68.8±7.2, non-converters=68.1±6.9 years; $p<0.001$ for all comparisons).

Progression of clinical markers

The progression of clinical markers for the total cohort and subdivided by phenoconversion status over five years of follow-up are illustrated in **Fig. 2, Fig. 3** and **Supplementary Fig. 3**. Annual change as assessed by SRMs and estimated annual progression rate for each marker is detailed in **Table 2**, **Supplementary Table 4**, and **Fig. 4**, while estimated annual progression rate subdivided by phenoconversion status is detailed in **Table 3** and **Supplementary Table 5**. Progression rates for the entire cohort (without stratifying by MDS prodromal criteria) are shown in **Supplementary Tables 6-7**.

Motor markers

Motor symptoms and motor signs showed the greatest degree of progression over time (**Fig. 2 and Table 2**). For example, MDS-UPDRS-III (excluding action tremor, which does not progress in iRBD)[9] had an estimated yearly progression rate of 1.73 points, with SRM=0.30 after 1 year and 0.97 after 5 years. Similarly, annual decline in Purdue Pegboard score was estimated to be -0.81 pegs, with SRM -0.35 and -1.15 at 1- and 5-year follow-up. More modest rates of progression were observed with MDS-UPDRS-II (SRM 0.2 and 0.8 at 1- and 5-year follow-up) and TUG velocity (SRM -0.09 and -0.67 at 1- and 5-year follow-up). A combined MDS-UPDRS-I+II+III score progressed by 2.81 points per year, with SRM=0.35 after 1 year and 1.20 after 5 years.

Phenoconverters had significantly greater annual progression rates in all motor variables compared with non-converters (**Table 3**), with the greatest distinction found in the MDS-UPDRS-III without action tremor score (annual progression in DLB=4.02, PD=3.69, non-converters=0.61 points; $p<0.001$). When comparing between PD- and DLB-phenoconverters, a slight but statistically significant increased slope in PD-phenoconverters was observed in the MDS-UPDRS-II and MDS-UPDRS-III scores ($p=0.037$ and $p=0.008$, respectively), although baseline MDS-UPDRS-III scores were significantly higher in DLB-phenoconverters (**Supplementary Table 3**, $p=0.028$).

Cognitive markers

Within the total cohort, both MoCA and MMSE demonstrated slow progression in the average score over time (**Fig. 3 and Table 2**), with an estimated annual decline of -0.07 and -0.25 points, respectively. These were associated with 1- and 5-year SRMs of 0.03 to -0.22 and -0.07 to -0.58.

A more dramatic decline was seen in phenoconverters compared with non-converters (**Table 3**), with annual decline in MMSE score of -0.09 points in non-converters *vs*. -0.42 in PD-phenoconverters and -0.81 DLB-phenoconverters ($p<0.001$). Estimated annual progression in

MoCA score in fact slightly increased in non-converters compared with a decline in phenoconverters (DLB=-0.73, PD=-0.09, non-converters=0.06 points; $p<0.001$). Rates of progression in both MMSE and MoCA were significantly different when comparing between PD- and DLB-phenoconverters ($p<0.001$ for both), with greater decline in DLB-phenoconverters.

Autonomic symptoms and signs

Autonomic symptoms as assessed by SCOPA-AUT total score increased slightly over time (**Fig. 3**, **Supplemental Fig. 3**, **Table 2**, and **Supplementary Table 4**), with estimated annual progression rate of 0.36 and 1- and 5-year SRMs of 0.13 and 0.31, respectively. Autonomic signs as assessed by orthostatic blood pressure showed mild increase in systolic pressure drop over time, with an estimated annual progression rate of 1.44 mmHg (1- and 5-year SRMs of 0.08 to 0.36).

Although PD-phenoconverters had a similarly modest annual rate of progression in total SCOPA-AUT score compared with non-convertors, DLB-phenoconverters had a significantly increased rate (DLB =1.57, PD=0.15, non-converters=0.20; $p<0.001$). This was driven by increased annual rates of progression in SCOPA-urinary and SCOPA-cardiovascular sub-scores (**Supplementary Fig. 3** and **Supplementary Table 5**), which also individually differed significantly from PD- phenoconverters ($p=0.004$ and $p<0.001$, respectively). When comparing the progression of postural systolic drop, although phenoconverters had a significantly increased rate of progression relative to non-phenoconverters ($p=0.002$), no significant difference was observed between PD- and DLB-phenoconverters ($p=0.553$).

Olfactory function

Olfactory z-scores slightly decreased over time in the total cohort (**Fig. 3 and Table 2**), with an estimated yearly progression rate of -0.09 and SRMs at 1- and 5-year follow-up of -0.07 and -0.64, respectively. The estimated yearly progression rate was significantly greater in PD- and

DLB-phenoconverters (-0.28in both) compared with non-converters (-0.06, $p<0.001$), without any significant difference between PD and DLB-phenoconverters ($p=0.958$).

Sleep symptoms

Sleep quality, as assessed by ESS, ISI, RBDSQ, and PSQI, paradoxically showed slight improvement in scores over time (**Fig. 3 and Table 2**), with SRMs ranging from -0.05 to -0.27 at 1-year follow-up and -0.17 to -0.52 at 5-year follow-up. When comparing non-converters and phenoconverters, a significant difference was seen only in ISI score (DLB=-0.99, PD=-0.77, non-converters=-0.43; $p=0.006$).

Psychiatric symptoms

Both depression and anxiety z-scores progressed only minimally or not at all, with SRMs ranging from -0.02 to 0.20 during the five years of follow-up (**Fig. 3 and Table 2**). No significant difference in the annual progression rate between phenoconverters and non-phenoconverters was observed.

Progression rates stratified by baseline age and by sex

Age at baseline followed a roughly normal distribution, with a median age of 68.8 ± 7.0 years (**Supplementary Fig. 1**). The results of clinical marker progression stratified by decade are shown in **Supplementary Fig. 5** and **Supplementary Table 8**. In general, clinical markers progressed along similar trajectories, with faster rates of decline in motor and cognitive scores among older participants (e.g. MDS-UPDRS-III progression at ages 50-59=1.08, ages 60-69=1.45, ages 70-79=1.78 points). With respect to sex, clinical markers progressed at similar rates between sexes, except for olfactory loss, which did not progress in females, and RBDSQ and PSQI, which worsened in females (**Supplementary Fig. 6** and **Supplementary Table 9**).

Sample size calculations

Using the estimated yearly progression rate of each variable, we calculated the required sample sizes for an interventional 1:1 placebo-controlled trial at different treatment efficacies (30% or 50% reduction in clinical progression) for different study lengths (**Table 4** and **Supplementary Fig. 4**). For example, assuming a treatment efficacy of 50% reduction in the progression of MDS-UPDRS-III (excluding action tremor) with 6-month follow-up periods, the required sample size at 80% power would be 213 subjects per group for a 2-year study. Using a combined MDS-UPDRS score (i.e. the sum of parts I, II, and III) would require slightly fewer subjects at 183 per group for a 2-year study. Under similar assumptions, based on time-to-event analysis to reduce the rate of phenoconversion by 50%, we estimated that 409 subjects per arm would need to be enrolled in a 2-year trial. The most efficient trial design was found to be a combined motor and cognitive endpoint of a sustained increase in MDS-UPDRS-III (excluding action tremor) score $\geq$ 4 and/or a sustained decrease in MoCA score $\geq$ 3; this provided an estimated sample size of 117 subjects per arm in a 2-year study and 88 subjects in a 3-year study (with 389 and 294 subjects for an agent with 30% efficacy).

Sample sizes were also calculated for the entire cohort, including subjects that did not meet MDS prodromal PD criteria (**Supplementary Table 10**). This increased sample size requirements for the majority of continuous motor variables or event milestones by approximately 10-30%.

Aside from increasing the assumed treatment effect and stratifying by MDS prodromal PD criteria, the other driver of required sample sizes was the extent of follow-up duration (**Supplementary Fig. 4**). Increasing the follow-up time from 1 to 2 years resulted in greater sample size reductions in all variables tested than any increases beyond 2 years. For example, a 1-year trial targeting a 50% reduction of the combined motor and cognitive endpoint required 229 subjects, versus 117 subjects in a 2-year trial, or 88 subjects for a 3-year trial.

Discussion

This international longitudinal prospective study represents the largest and most comprehensive systematic assessment of clinical marker progression in iRBD that has been performed. We demonstrate several important insights, including: (i) motor assessment using the MDS-UPDRS-III and quantitative motor testing shows the greatest degree of progression over time; (ii) there is moderate progression of other non-motor markers, particularly the MDS-UPDRS-II, MMSE, and olfactory scores, and limited to no progression in psychiatric and some autonomic measures; (iii) while phenoconverters showed overall greater progression than non-converters in motor, olfactory, cognitive, and certain autonomic markers, the only robust difference in progression between PD and DLB-phenoconverters was in cognitive testing; and (iv) the most efficient trial design for future randomized trials was a combined endpoint of a sustained increase in MDS-UPDRS-III and/or a sustained decrease in MoCA score, while stratifying by MDS prodromal PD criteria and extending trial duration from 1 to 2 years yielded the largest reductions in sample size.

Clinical marker progression

Quantitative motor assessment by standardized clinical exam or simple office-based motor testing showed clear progression over the study period, in keeping with prior studies.[7,9,11] Unsurprisingly, given that motor function is the primary means of defining parkinsonism, phenoconverters had significantly increased rates of progression compared with non-converters.

With respect to non-motor markers, although cognitive function showed moderate decline overall, scores remained stable in non-converters but dramatically declined among phenoconverters. This bimodal distribution likely explains the large difference in sample size requirements when using MoCA as a continuous variable (which includes the stable scores of non-converters, and which could be confounded by practice effects in cognitively-spared subjects) rather than as a milestone of sustained decrease (which dichotomizes into phenoconverters and non-converters).

Olfactory and autonomic dysfunction only mildly progressed when assessed in the total cohort, as previously observed,[9,38] and is in keeping with being among the earliest markers of prodromal disease. Indeed, the inclusion of subjects not meeting MDS prodromal PD criteria (i.e. those likeliest to have more olfactory and autonomic "reserve" to lose) paradoxically decreased the sample size requirements for these variables. Although olfactory dysfunction in phenoconverters appeared to decline more rapidly, this could reflect progressive cognitive dysfunction (i.e. olfactory memory) rather than continued olfactory loss alone.[39] Increasing postural systolic drop was also observed in phenoconverters, which is recognized to be predictive of eventual phenoconversion.[40]

Psychiatric symptoms and sleep symptoms were generally stable over time, in keeping with prior studies.[11,41] In phenoconverters, insomnia scores in fact significantly improved over time relative to non-converters, which could reflect a general subthreshold increase in sleep drive without overt daytime somnolence as patients approach a defined neurodegenerative disease. Alternatively, these trends could be resultant from treatment for sleep or psychiatric disorders.

Secondary analyses stratifying clinical marker progression by baseline age demonstrated somewhat faster rates of decline in motor and cognitive measures in older subjects. By contrast, there were minimal differences when stratifying by sex.

Phenoconversion rate

We found that phenoconversion rates were slightly lower than expected compared to two recent large IRBDSG studies, despite similar baseline ages.[7,10] Our 3-year phenoconversion risk was found to be 18.2% *vs*. 17.9% and 24.2% in the other studies, despite the fact that this study selected subjects that met prodromal PD criteria. Several explanations likely account of this. First, a lower phenoconversion rate was observed in a single large center (Berlin) which had no phenoconversions at all over a 2.7-year follow-up; removal of this center increased the 3-year risk to 20.1%. Second, although there is some overlap in the patient populations with the prior studies, this study includes 8 new centers contributing 155 subjects (13.4% of included subjects),

while several large centers with higher phenoconversion rates that were included in the prior studies were unable to contribute to this one. However, newer centers did not have lower rates of phenoconversion (3-year risk: 19.8%). Third, the inclusion criteria may have enriched toward an overall healthier population than the previous studies. By design, subjects were required to attend periodic and structured assessments longitudinally (whereas only a follow-up clinical examination was required in the other studies), which may have discouraged subjects with mobility or cognitive issues (i.e. those most likely to phenoconvert) from being enrolled.[9] This would be consistent with the unusually low phenoconversion rate in the first year (4.4%) vs. an average annual conversion rate of 6.1% in years 2-5 (a rate consistent with prior studies). In any event, although this study population had lower rates of phenoconversion than expected, longitudinal patient retention is a critical aspect of any proposed therapeutic trial. Therefore, the subjects included in this study are probably representative of those likeliest to be enrolled in a future trial.

Prodromal Parkinson's disease versus prodromal dementia with Lewy bodies

When classified according to the initial phenoconversion event (parkinsonism-first vs. dementia-first), PD- and DLB-phenoconverters showed remarkably similar age-adjusted rates of progression. For example, among motor signs, only MDS-UPDRS-III showed a slightly increased rate in PD-phenoconverters, with the difference possibly explained by the higher baseline MDS-UPDRS-III score in DLB-phenoconverters. This is concordant with a recent single-center study in which no significant between-group difference in motor trajectories was observed.[9] An increased rate of progression in SCOPA-AUT was also observed in DLB-phenoconverters. This was primarily driven by an increased cardiovascular subscore, which largely reflects orthostatic hypotension symptoms; nevertheless, no difference in orthostatic blood pressure was seen between PD- and DLB-phenoconverters, in agreement with studies with more precise orthostatic testing.[42]

Overall, the only robust differentiating clinical marker between PD- and DLB-first phenoconverters was the higher rate of cognitive decline in DLB, as would be expected by definition. This is in agreement with two recent IRBDSG studies (with approximately half of subjects overlapping between them), which observed that baseline cognitive function was the only clear differentiating clinical predictor between PD and DLB phenoconversion.[7,13] Thus, while clear differences in the progression of clinical variables are apparent between those at higher and lower risk of phenoconversion (i.e. phenoconverters and non-converters in this study), the subtypes of prodromal synucleinopathies appear to follow very similar clinical courses. The underlying pathological substrate that accounts for this remains unclear. This could reflect either alternate pathways of synuclein spread or coexistent amyloid or tau pathology driving earlier cortical neurodegeneration.[43,44] It is important to note that all subjects in this study were iRBD patients, who generally have a more diffuse burden of synucleinopathy, and consequently more non-motor manifestations.[6] iRBD identifies subtypes of PD and DLB that are associated with greater progression of motor and non-motor symptoms, diffuse and severe deposition of synuclein at autopsy, enhanced patterns of atrophy earlier in the disease course, and overall poorer prognosis.[45,46] This PD subtype is therefore characterized by a different speed and anatomical pattern of progression than PD subjects without RBD. Therefore, it is not clear to what degree the findings in this study are translatable to prodromal subtypes that do not have iRBD.

Sample size

We calculated sample size estimates for neuroprotective trials using both the progression of continuous clinical variables and categorical events (phenoconversion and motor and cognitive decline milestones) as endpoints. Importantly, we first stratified by MDS prodromal criteria, which retained >80% of subjects; this reduces sample sizes by approximately 10-30% for most motor clinical markers or events of interest. For continuous motor variables, sample sizes for a 2-year trial with HR=0.5 ranged from 151-560 subjects per arm, while substantially higher numbers were required for non-motor variables. Under similar assumptions, sample size estimates using the sum of MDS-UPDRS-I, -II, and -III sub-scores resulted in 183 subjects per arm. The most efficient trial design was a combined motor and cognitive endpoint of a sustained

increase in MDS-UPDRS-III and/or a sustained decrease in MoCA score, which required only 117 subjects for a 2-year study at HR=0.5. These samples size estimates are broadly similar to those calculated in a recent single-center study of clinical markers.[11] They are also similar to the sample sizes calculated in a recent single-center study that assessed serial DAT-PET imaging (i.e. sample size=94 for standard DAT-PET analysis).[47] Notably, using the milestone of phenoconversion to overt disease required substantially larger numbers. Finally, aside from increasing the assumed treatment effect and stratifying by MDS prodromal criteria, sample sizes could also be substantially reduced by increasing the follow-up time from 1 to 2 years, whereas lesser reductions were observed if trials were extended to 3 years or beyond.

Strengths and limitations

Strengths of this study include a large study population prospectively followed over a period of 5 years. Clinical variables representative of most of the critical predictors of phenoconversion were systematically measured, including the motor, cognitive, olfactory, autonomic, psychiatric, and sleep domains. However, several limitations should be discussed. Since each of the 28 centers used their own study protocol, which varied in predictors assessed, methods of assessment, and follow-up frequency, a pragmatic approach was taken with respect to data collection, in which different clinical tests were harmonized across centers in order to maximize recruitment and simplify the analysis. Although different methods of measuring a clinical marker undoubtably vary in sensitivity and statistical power, they have all been shown to have similar performance in PD.[15,18,48,49] Moreover, in this study, all scores were adjusted by center in the LMEMs and followed a broadly similar trend when SRMs were evaluated individually (data not shown). Second, some clinical markers that have been shown to have excellent predictive value were not included in the analysis since they were only performed in sufficient numbers by a single center (e.g. alternate tap test, color-vision testing, etc.).[9,11] The IRBDSG is currently planning a recommended minimal core data collection protocol that will be essential for standardization between centers in the future. Additionally, longitudinal assessment of imaging[13,50] and fluid[51] biomarkers to evaluate neuropathological changes as complementary measures of progression are needed. Third is the use of a generally conservative method of imputation to estimate progression in subjects after phenoconversion, particularly since certain markers can increase

exponentially closer to the time of phenoconversion.[9] Notably, a similar issue would exist in any real-life therapeutic trial, since it would be unethical to withhold symptomatic treatment in phenoconverted subjects. Fourth, medication use could impact upon the progression of markers. Although medication use was not longitudinally collected, the use of either melatonin, clonazepam, or antidepressants at baseline showed only a statistically significant effect of clonazepam on annual decline in MoCA score (clonazepam use=-0.19 points vs non-use=0.012 points, p=0.026; and data not shown). Fifth, subjects destined to convert to a parkinsonism-first vs. dementia-first phenotype cannot be reliably distinguished at time of iRBD diagnosis. If the underlying pathomechanisms that drive neurodegeneration are substantially different between the two,[43,44] a neuroprotective therapy targeting a single pathomechanism may inadequately slow progression in a substantial subgroup of the population, although this could be mitigated by baseline neurocognitive testing.[52,53] Similarly, the 5-10% of subjects expected to phenoconvert to MSA are likely to progress very differently, although this could be mitigated by screening subjects for olfactory loss.[7] Finally, an assumption of LMEMs is linearity over time. Previous studies have demonstrated heterogeneity in the pattern of emergence among prodromal features: some features emerge early and subsequently remain fairly stable over time (e.g., constipation), whereas other features emerge late and increase quickly in the last few years before clinical diagnosis (e.g. motor signs).[9] Consequently, the current results may overestimate the rate of progression of early prodromal features during the last years of the prodromal phase, and conversely underestimate the rate of progression of late-emerging prodromal features. In keeping with this, those phenoconverting within 3-5 years had faster rates of progression in motor and cognitive measures and generally less progression in markers known to have longer latencies. Assuming that a future neuroprotective trial would not run longer than 3 years, using a 5-year window for the LMEMs was felt to be a compromise between the robust inclusion of datapoints for model precision versus achieving an accuracy that reflects the reality of recruiting a patient in whom the time until phenoconversion to overt disease will be unknown.

Conclusion

To conclude, we confirmed patterns of clinical marker progression in prodromal synucleinopathy and demonstrated predicted sample sizes to inform future neuroprotective trials.

Funding

SJ was supported by an Edmond J. Safra Fellowship in Movement Disorders from the Michael J. Fox Foundation.

The Oxford Discovery cohort (lead M.T.Hu) is supported by Parkinson's UK and the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC), UK.

KY Jung received research fund supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (2017M3C7A1029688, 2017R1A2B2012280) and NRF-2022R1H1A2092329.

WO is a Hertie-Senior Research Professor supported by the Charitable Hertie Foundation, Frankfurt/Main, Germany. AJ and WO are supported by the ParkinsonFonds Deutschland.

JFG was funded by a grant from the CIHR and he holds a Canada Research Chair in Cognitive Decline in Pathological Aging.

RBP was funded by a grant from the Canadian Institutes of Health Research (CIHR).

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Lang AE, Siderowf AD, Macklin EA, et al. Trial of Cinpanemab in Early Parkinson's Disease. *https://doi.org/101056/NEJMoa2203395*. 2022;387(5):408-420. doi:10.1056/NEJMOA2203395
2. Pagano G, Taylor KI, Anzures-Cabrera J, et al. Trial of Prasinezumab in Early-Stage Parkinson's Disease. *N Engl J Med*. 2022;387(5):421-432. doi:10.1056/NEJMOA2202867/SUPPL_FILE/NEJMOA2202867_DATA-SHARING.PDF
3. Lang FM, Kwon DY, Aarsland D, et al. An international, randomized, placebo-controlled, phase 2b clinical trial of intepirdine for dementia with Lewy bodies (HEADWAY-DLB). Published online 2021. doi:10.1002/trc2.12171
4. Postuma RB. Neuroprotective Trials in REM Sleep Behavior Disorder The Way Forward Becomes Clearer. Published online 2022. doi:10.1212/WNL.0000000000200235
5. Heinzel S, Berg D, Gasser T, et al. Update of the MDS research criteria for prodromal Parkinson's disease. *Movement Disorders*. 2019;34(10):1464-1470. doi:10.1002/MDS.27802
6. Berg D, Borghammer P, Fereshtehnejad SM, et al. Prodromal Parkinson disease subtypes — key to understanding heterogeneity. *Nat Rev Neurol*. 2021;17(6):349-361. doi:10.1038/s41582-021-00486-9
7. Postuma RB, Iranzo A, Hu M, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain*. 2019;142(3):744-759. doi:10.1093/BRAIN/AWZ030
8. Postuma RB, Iranzo A, Hogl B, et al. Risk factors for neurodegeneration in idiopathic rapid eye movement sleep behavior disorder: A multicenter study. *Ann Neurol*. 2015;77(5):830-839. doi:10.1002/ANA.24385
9. Fereshtehnejad SM, Yao C, Pelletier A, Montplaisir JY, Gagnon JF, Postuma RB. Evolution of prodromal Parkinson's disease and dementia with Lewy bodies: a prospective study. *Brain*. 2019;142(7):2051-2067. doi:10.1093/BRAIN/AWZ111
10. Zhang H, Iranzo A, Högl B, et al. Risk Factors for Phenoconversion in Rapid Eye Movement Sleep Behavior Disorder. *Ann Neurol*. 2022;91(3):404-416. doi:10.1002/ANA.26298

11. Alotaibi F, Pelletier A, Gagnon JF, Montplaisir JY, Postuma RB. Prodromal Marker Progression in Idiopathic Rapid Eye Movement Sleep Behavior Disorder: Sample Size for Clinical Trials. *Movement Disorders*. 2019;34(12):1914-1919. doi:10.1002/MDS.27869

12. Kogan R v., Janzen A, Meles SK, et al. Four-Year Follow-up of [18F]Fluorodeoxyglucose Positron Emission Tomography–Based Parkinson's Disease–Related Pattern Expression in 20 Patients with Isolated Rapid Eye Movement Sleep Behavior Disorder Shows Prodromal Progression. *Movement Disorders*. 2021;36(1):230-235. doi:10.1002/MDS.28260

13. Arnaldi D, Chincarini A, Hu MT, et al. Dopaminergic imaging and clinical predictors for phenoconversion of REM sleep behaviour disorder. *Brain*. 2021;144(1):278-287. doi:10.1093/BRAIN/AWAA365

14. Sateia MJ. International Classification of Sleep Disorders-Third Edition. *Chest*. 2014;146(5):1387-1394. doi:10.1378/CHEST.14-0970

15. Goetz CG, Stebbins GT, Tilley BC. Calibration of unified Parkinson's disease rating scale scores to Movement Disorder Society-unified Parkinson's disease rating scale scores. *Movement Disorders*. 2012;27(10):1239-1242. doi:10.1002/MDS.25122

16. Podsiadlo D, Richardson S. The Timed "Up & Go": A Test of Basic Functional Mobility for Frail Elderly Persons. *J Am Geriatr Soc*. 1991;39(2):142-148. doi:10.1111/J.1532-5415.1991.TB01616.X

17. Postuma RB, Lang AE, Massicotte-Marquez J, Montplaisir J. Potential early markers of Parkinson disease in idiopathic REM sleep behavior disorder. *Neurology*. 2006;66(6):845-851. doi:10.1212/01.WNL.0000203648.80727.5B

18. Lawton M, Hu MTM, Baig F, et al. Equating scores of the University of Pennsylvania Smell Identification Test and Sniffin' Sticks test in patients with Parkinson's disease. *Parkinsonism Relat Disord*. 2016;33:96-101. doi:10.1016/j.parkreldis.2016.09.023

19. Oleszkiewicz A, Schriever VA, Croy I, Hähner A, Hummel T. Updated Sniffin' Sticks normative data based on an extended sample of 9139 subjects. *European Archives of Oto-Rhino-Laryngology*. 2019;276(3):719-728. doi:10.1007/S00405-018-5248-1/TABLES/2

20. Hummel T, Konnerth CG, Rosenheim K, Kobal G. Screening of olfactory function with a four-minute odor identification test: Reliability, normative data, and investigations in patients with

olfactory loss. *Annals of Otology, Rhinology and Laryngology*. 2001;110(10):976-981. doi:10.1177/000348940111001015

21. Menon C, Westervelt HJ, Jahn DR, Dressel JA, O'Bryant SE. Normative Performance on the Brief Smell Identification Test (BSIT) in a Multi-Ethnic Bilingual Cohort: A Project FRONTIER Study 1. *http://dx.doi.org/101080/138540462013796406*. 2013;27(6):946-961. doi:10.1080/13854046.2013.796406

22. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-545. doi:10.1093/SLEEP/14.6.540

23. Bastien CH, Vallières A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med*. 2001;2(4):297-307. doi:10.1016/S1389-9457(00)00065-4

24. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193-213. doi:10.1016/0165-1781(89)90047-4

25. Stiasny-Kolster K, Mayer G, Schäfer S, Möller JC, Heinzel-Gutenbrunner M, Oertel WH. The REM sleep behavior disorder screening questionnaire--a new diagnostic instrument. *Mov Disord*. 2007;22(16):2386-2393. doi:10.1002/MDS.21740

26. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6

27. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc*. 2005;53(4):695-699. doi:10.1111/J.1532-5415.2005.53221.X

28. Visser M, Marinus J, Stiggelbout AM, van Hilten JJ. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. *Mov Disord*. 2004;19(11):1306-1312. doi:10.1002/MDS.20153

29. Beck A, Steer R, Brown G. Beck depression inventory–II. *psycnet.apa.org*. Published online 1996. Accessed August 15, 2022. https://psycnet.apa.org/doiLanding?doi=10.1037/t00742-000

30. Beck AT, Epstein N, Brown G, Steer RA. An Inventory for Measuring Clinical Anxiety: Psychometric Properties. *J Consult Clin Psychol*. 1988;56(6):893-897. doi:10.1037/0022-006X.56.6.893

31. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1982;17(1):37-49. doi:10.1016/0022-3956(82)90033-4

32. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67(6):361-370. doi:10.1111/J.1600-0447.1983.TB09716.X

33. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw*. 2017;82(13):1-26. doi:10.18637/JSS.V082.I13

34. Moritz S, Bartz-Beielstein T. Time Series Missing Value Imputation [R package imputeTS version 3.2]. *R Journal*. 2021;9(1):207-218. doi:10.32614/RJ-2017-009

35. Nash S, Morgan KE, Frost C, Mulick A. Power and sample-size calculations for trials that compare slopes over time: Introducing the slopepower command: *https://doi.org/101177/1536867X211045512*. 2021;21(3):575-601. doi:10.1177/1536867X211045512

36. Anderson K. gsDesign. Published 2022. Accessed August 15, 2022. https://cran.r-project.org/web/packages/gsDesign/index.html

37. Horváth K, Aschermann Z, Ács P, et al. Minimal clinically important difference on the Motor Examination part of MDS-UPDRS. *Parkinsonism Relat Disord*. 2015;21(12):1421-1426. doi:10.1016/J.PARKRELDIS.2015.10.006

38. Iranzo A, Serradell M, Vilaseca I, et al. Longitudinal assessment of olfactory function in idiopathic REM sleep behavior disorder. *Parkinsonism Relat Disord*. 2013;19(6):600-604. doi:10.1016/J.PARKRELDIS.2013.02.009

39. Shin C, Lee JY, Kim YK, et al. Cognitive decline in association with hyposmia in idiopathic rapid eye movement sleep behavior disorder: a prospective 2-year follow-up study. *Eur J Neurol*. 2019;26(11):1417-1420. doi:10.1111/ENE.14025

40. Gibbons CH, Freeman R. Clinical implications of delayed orthostatic hypotension A 10-year follow-up study. Published online 2015.

41. Postuma RB, Gagnon JF, Pelletier A, Montplaisir JY. Insomnia and somnolence in idiopathic RBD: a prospective cohort study. *NPJ Parkinsons Dis*. 2017;3(1). doi:10.1038/S41531-017-0011-7

42. McCarter SJ, Gehrking TL, st. Louis EK, et al. Autonomic dysfunction and phenoconversion in idiopathic REM sleep behavior disorder. *Clinical Autonomic Research*. 2020;30(3):207-213. doi:10.1007/S10286-020-00674-5/TABLES/3

43. Foffani G, Obeso JA. A Cortical Pathogenic Theory of Parkinson's Disease. *Neuron*. 2018;99(6):1116-1128. doi:10.1016/J.NEURON.2018.07.028

44. Adler CH, Beach TG. Neuropathological basis of nonmotor manifestations of Parkinson's disease. *Movement Disorders*. 2016;31(8):1114-1119. doi:10.1002/MDS.26605

45. Fereshtehnejad SM, Zeighami Y, Dagher A, Postuma RB. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. *Brain*. 2017;140(7):1959-1976. doi:10.1093/BRAIN/AWX118

46. Postuma RB, Adler CH, Dugger BN, et al. REM sleep behavior disorder and neuropathology in Parkinson's disease. *Movement Disorders*. 2015;30(10):1413-1417. doi:10.1002/MDS.26347

47. Shin JH, Lee JY, Kim YK, et al. Longitudinal change in dopamine transporter availability in idiopathic REM sleep behavior disorder. *Neurology*. 2020;95(23):e3081-e3092. doi:10.1212/WNL.0000000000010942

48. Williams JR, Hirsch ES, Anderson K, et al. A comparison of nine scales to detect depression in Parkinson disease: Which scale to use? *Neurology*. 2012;78(13):998. doi:10.1212/WNL.0B013E31824D587F

49. Leentjens AFG, Dujardin K, Marsh L, Richard IH, Starkstein SE, Martinez-Martin P. Anxiety rating scales in Parkinson's disease: a validation study of the Hamilton anxiety rating scale, the Beck anxiety inventory, and the hospital anxiety and depression scale. *Mov Disord*. 2011;26(3):407-415. doi:10.1002/MDS.23184

50. Rahayel S, Postuma RB, Montplaisir J, et al. A Prodromal Brain-Clinical Pattern of Cognition in Synucleinopathies. *Ann Neurol*. 2021;89(2):341-357. doi:10.1002/ANA.25962

51. Iranzo A, Fairfoul G, Ayudhaya ACN, et al. Detection of α-synuclein in CSF by RT-QuIC in patients with isolated rapid-eye-movement sleep behaviour disorder: a longitudinal observational study. *Lancet Neurol*. 2021;20(3):203-212. doi:10.1016/S1474-4422(20)30449-X

52. Marchand DG, Montplaisir J, Postuma RB, Rahayel S, Gagnon JF. Detecting the Cognitive Prodrome of Dementia with Lewy Bodies: A Prospective Study of REM Sleep Behavior Disorder. *Sleep*. 2017;40(1). doi:10.1093/SLEEP/ZSW014

53. Marchand DG, Postuma RB, Escudier F, et al. How does dementia with Lewy bodies start? prodromal cognitive changes in REM sleep behavior disorder. *Ann Neurol*. 2018;83(5):1016-1026. doi:10.1002/ANA.25239

Figure legends

Figure 1 Study profile. (**A**) Subjects enrolled in the study grouped by country of origin over time. More subjects were seen at 2-year follow-up than 1-year follow-up since some centers tended to have longer follow-up times (e.g., 18 months to 2 years). (**B**) Kaplan-Meier survival plot of disease-free survival (i.e. free of phenoconversion) with 95% confidence intervals shaded.

Figure 2 Motor outcome measures over 5 years of follow-up for the total cohort and by phenoconversion status. Individual dots represent each subject; solid lines represent estimated progression by linear mixed-effect modeling.

Figure 3 Non-motor outcome measures over 5 years of follow-up for the total cohort and by phenoconversion status. Individual dots represent each subject; solid lines represent estimated progression by linear mixed-effect modeling.

Figure 4 Normalized motor and non-motor outcome measures over 5 years of follow-up. Results were normalized for comparison between variables by standardized response means (SRM), which is computed by dividing the mean change from baseline of each individual patient by the standard deviation of the change of the total cohort.

Table 1 Baseline demographics and phenoconversion outcomes from baseline to 5-year follow-up

	Baseline (*n*=1160)	1-year follow-up (*n*=767)	2-year follow-up (*n*=783)	3-year follow-up (*n*=477)	4-year follow-up (*n*=311)	5-year follow-up (*n*=228)
Demographics						
Age, years	68.5 ± 7.0	69.5 ± 7.1	70.3 ± 6.8	70.8 ± 6.7	72.3 ± 6.5	73.0 ± 6.4
Sex, % male	78.4	78.5	80.5	82.2	80.7	83.3
Handedness, % right	90.6	92.8	90.4	90.9	90.4	87.4
RBD course						
Years from diagnosis	1.28 ± 2.3	2.2 ± 2.3	3.1 ± 2.2	4.4 ± 2.4	5.1 ± 2.0	6.4 ± 2.2
Years from symptom onset	6.4 ± 6.4	7.84 ± 8.0	8.8 ± 7.9	10.0 ± 8.7	11.0 ± 9.4	12.5 ± 10.4
Years from baseline visit	-	1.1 ± 0.4	1.98 ± 0.4	3.0 ± 0.4	4.0 ± 0.4	5.1 ± 0.5
Phenoconversion outcomes						
Phenoconverted, %	-	4.4	11.5	18.2	25.3	31.7
Phenoconverted, *n*	-	51	69	45	33	23
PD	-	29	35	23	20	11
DLB	-	18	31	18	11	12
MSA	-	4	3	3	2	0

Data are presented as mean ± SD. Yearly phenoconverted percentages were calculated by Kaplan-Meier survival analysis. More subjects were seen at 2-year follow-up than 1-year follow-up since some centers tended to have longer follow-up times (18 months to 2 years).
DLB=dementia with Lewy bodies; MSA=multiple system atrophy; PD=Parkinson's disease.

Table 2 Annual marker outcomes and estimated progression rates

Marker	Baseline		1-year follow-up		2-year follow-up		3-year follow-up		5-year follow-up		Yearly progression
	Centers, *n*	Mean ± SD (*n*)	Mean ± SD (*n*)	SRM	Mean ± SD (*n*)	SRM	Mean ± SD (*n*)	SRM	Mean ± SD (*n*)	SRM	Estimate [95 % CI]
MDS-UPDRS											
MDS-UPDRS-I	15	7.67 ± 6.01 (482)	8.14 ± 5.80 (431)	0.13	8.45 ± 5.90 (359)	0.14	8.62 ± 5.55 (240)	0.20	10.48 ± 5.98 (185)	0.54	0.48 [0.34, 0.61]
MDS-UPDRS-II	18	2.31 ± 3.48 (740)	2.81 ± 3.88 (665)	0.20	3.3 ± 4.37 (537)	0.26	4.31 ± 5.1 (378)	0.44	7.02 ± 6.71 (294)	0.80	0.65 [0.55, 0.75]
MDS-UPDRS-III	27	4.02 ± 5.03 (1095)	5.3 ± 7.04 (989)	0.26	6.9 ± 8.98 (751)	0.37	10.0 ± 10.7 (521)	0.58	18.6 ± 15.2 (371)	0.99	1.59 [1.41, 1.76]
MDS-UPDRS-III (no action tremor)	20	3.74 ± 4.92 (805)	5.27 ± 7.01 (722)	0.30	6.95 ± 8.93 (559)	0.40	9.6 ± 10.57 (408)	0.59	18.01 ± 15.37 (302)	0.97	1.73 [1.53, 1.93]
MDS-UPDRS-I+II+III	15	15.2 ± 12.0 (472)	17.7 ± 13.3 (413)	0.35	20.0 ± 14.8 (347)	0.39	24.9 ± 16.0 (230)	0.72	37.9 ± 24.3 (173)	1.20	2.81 [2.38, 3.23]
Quantitative motor[a]											
Timed Up & Go (s)	2	8.04 ± 2.86 (346)	8.06 ± 2.67 (298)	0.08	8.17 ± 3.4 (243)	0.07	8.91 ± 6.58 (183)	0.19	9.05 ± 4.05 (141)	0.42	0.32 [0.15, 0.49]
Purdue Peg Board	2	10.59 ± 4.09 (271)	9.93 ± 3.51 (234)	−0.35	9.53 ± 3.98 (178)	−0.29	8.34 ± 3.6 (129)	−0.70	5.31 ± 4.86 (106)	−1.15	−0.81 [−0.98, −0.64]
Autonomic[a]											
Postural Systolic Drop	6	10.1 ± 16.2 (383)	10.6 ± 15.6 (332)	0.08	11.9 ± 15.6 (259)	0.13	15.1 ± 16.4 (195)	0.22	18.9 ± 17.0 (149)	0.36	1.44 [1.01, 1.87]
SCOPA-AUT Total	10	10.95 ± 7.46 (213)	11.87 ± 7.86 (184)	0.13	12.03 ± 7.54 (140)	0.04	11.61 ± 6.81 (97)	0.11	14.04 ± 7.14 (57)	0.31	0.36 [0.05, 0.66]
Olfactory											
Olfaction z-score	14	−2.28 ± 1.8 (564)	−2.23 ± 1.84 (373)	−0.07	−2.29 ± 2.03 (287)	−0.07	−2.59 ± 2.07 (178)	−0.25	−3.39 ± 2.45 (139)	−0.64	−0.09 [−0.14, −0.05]
Cognitive											
MoCA	21	25.3 ± 3.2 (788)	25.4 ± 3.3 (694)	0.03	25.2 ± 3.6 (523)	0.01	24.8 ± 3.9 (388)	−0.08	24.1 ± 4.4 (273)	−0.22	−0.07 [−0.13, −0.01]
MMSE	15	27.7 ± 2.3 (706)	27.6 ± 2.3 (584)	−0.07	27.2 ± 2.8 (441)	−0.18	26.8 ± 3.1 (312)	−0.29	25.6 ± 3.7 (247)	−0.58	−0.25 [−0.32, −0.19]
Psychiatric symptoms											
Depression z-score	17	0.01 ± 0.98 (684)	0 ± 0.96 (562)	−0.02	0.01 ± 1.01 (437)	0.06	0.09 ± 0.99 (296)	0.11	0.13 ± 0.93 (199)	0.20	0.02 [0, 0.04]
Anxiety z-score	8	0.01 ± 1 (395)	0.01 ± 0.98 (316)	0.11	−0.04 ± 1.03 (257)	0.03	−0.07 ± 0.91 (190)	0.01	0.01 ± 1.03 (136)	0.18	0.02 [−0.01, 0.04]
Sleep symptoms											
ESS	11	6.76 ± 4.49 (583)	6.38 ± 4 (518)	−0.15	6.28 ± 4.18 (374)	−0.16	6.16 ± 4 (249)	−0.13	5.79 ± 4.21 (176)	−0.25	−0.25 [−0.33, −0.16]
ISI	5	9.29 ± 6.35 (310)	8.07 ± 5.44 (271)	−0.27	8.25 ± 5.86 (181)	−0.19	8.01 ± 5.29 (133)	−0.33	6.73 ± 5.99 (107)	−0.52	−0.61 [−0.78, −0.43]
PSQI	2	7.14 ± 4.01 (162)	6.54 ± 3.49 (154)	−0.15	7.15 ± 3.47 (94)	0.07	6.56 ± 2.88 (52)	−0.10	8.02 ± 3.45 (28)	0.14	−0.01 [−0.22, 0.2]
RBDSQ	5	9.43 ± 2.56 (247)	9.26 ± 2.57 (225)	−0.05	9.28 ± 2.76 (184)	−0.05	9.18 ± 2.93 (113)	−0.16	9.23 ± 2.8 (67)	−0.17	−0.09 [−0.2, 0.02]

The progression of variables of interest are described using annual mean ± SD, standardized response mean (SRM), and estimated annual progression rate by LMEM. ESS=Epworth Sleepiness Scale; ISI=Insomnia Severity Index; MMSE=Mini-Mental State Examination; MoCA=Montreal Cognitive Assessment; PSQI=Pittsburgh Sleep Quality Inventory; RBDSQ=REM Behavior Disorder Sleep Questionnaire; SCOPA-AUT= Scales for Outcomes in Parkinson's Disease - Autonomic Dysfunction.

[a]Full results that include all clinical markers and 4-year follow-up data can be found in **Supplementary Table 4**.

Table 3 Estimated progression rates subdivided by phenoconversion state

	Total Cohort	Still Unconverted	PD	DLB	*p*-value	
Variable of interest	**Estimate [95 % CI]**	**Estimate [95% CI]**	**Estimate [95% CI]**	**Estimate [95% CI]**	**Unconverted vs. Phenoconverted**	**PD- vs DLB-phenoconverted**
MDS-UPDRS						
MDS-UPDRS-I	0.48 [0.34, 0.61]	0.22 [0.11, 0.33]	0.95 [0.76, 1.13]	1.13 [0.9, 1.36]	**<0.001**	0.192
MDS-UPDRS-II	0.65 [0.55, 0.75]	0.26 [0.2, 0.31]	1.6 [1.46, 1.74]	1.38 [1.23, 1.53]	**<0.001**	**0.037**
MDS-UPDRS-III	1.59 [1.41, 1.76]	0.59 [0.51, 0.67]	4.41 [4.15, 4.66]	3.86 [3.56, 4.17]	**<0.001**	**0.008**
MDS-UPDRS-III (no action tremor)	1.73 [1.53, 1.93]	0.61 [0.52, 0.71]	4.44 [4.15, 4.73]	4.02 [3.69, 4.36]	**<0.001**	0.070
MDS-UPDRS-I+II+III	2.81 [2.38, 3.23]	1.47 [1.24, 1.7]	7.68 [6.99, 8.36]	6.75 [6.12, 7.38]	**<0.001**	0.082
Quantitative Motor						
Timed Up & Go (s)	0.32 [0.15, 0.49]	0.18 [0.07, 0.29]	0.44 [0.36, 0.52]	0.75 [0.41, 1.08]	**<0.001**	0.069
Timed Up & Go (m/s)	−0.02 [−0.02, −0.01]	−0.01 [−0.02, −0.01]	−0.04 [−0.04, −0.03]	−0.04 [−0.05, −0.03]	**<0.001**	0.221
Purdue Peg Board	−0.81 [−0.98, −0.64]	−0.64 [−0.77, −0.51]	−1.95 [−2.2, −1.7]	−1.59 [−1.86, −1.32]	**<0.001**	0.100
Autonomic[a]						
Postural Systolic Drop	1.44 [1.01, 1.87]	1.02 [0.56, 1.48]	2.05 [1.3, 2.81]	2.38 [1.65, 3.11]	**0.002**	0.553
Postural Diastolic Drop	0.79 [0.48, 1.11]	0.59 [0.28, 0.9]	1.07 [0.6, 1.54]	1.35 [0.75, 1.96]	**0.020**	0.405
SCOPA-AUT Total	0.36 [0.05, 0.66]	0.20 [−0.06, 0.46]	0.15 [−0.25, 0.54]	1.57 [0.97, 2.23]	0.073	**<0.001**
Olfactory						
Olfaction z-score	−0.09 [−0.14, −0.05]	−0.06 [−0.1, −0.02]	−0.28 [−0.36, −0.2]	−0.28 [−0.36, −0.20]	**<0.001**	0.958
Cognitive						
MoCA	−0.07 [−0.13, −0.01]	0.06 [0.01, 0.11]	−0.09 [−0.18, −0.01]	−0.73 [−0.87, −0.59]	**<0.001**	**<0.001**
MMSE	−0.25 [−0.32, −0.19]	−0.09 [−0.14, −0.04]	−0.42 [−0.49, −0.36]	−0.81 [−0.91, −0.7]	**<0.001**	**<0.001**
Psychiatric symptoms						
Depression z-score	0.02 [0, 0.04]	0.02 [0, 0.04]	0.04 [0.02, 0.07]	0.04 [−0.02, 0.09]	0.100	0.854
Anxiety z-score	0.02 [−0.01, 0.04]	0.02 [0, 0.04]	0.01 [−0.03, 0.05]	0.04 [−0.01, 0.1]	0.981	0.504
Sleep Symptoms						
ESS	−0.25 [−0.33, −0.16]	−0.22 [−0.29, −0.14]	−0.19 [−0.31, −0.06]	−0.19 [−0.4, 0.02]	0.643	0.978
ISI	−0.61 [−0.78, −0.43]	−0.43 [−0.6, −0.25]	−0.77 [−1.05, −0.5]	−0.99 [−1.35, −0.64]	**0.006**	0.314
PSQI	−0.01 [−0.22, 0.2]	−0.03 [−0.24, 0.17]	0.28 [−0.05, 0.62]	0.22 [−0.17, 0.6]	0.058	0.712
RBDSQ	−0.09 [−0.2, 0.02]	−0.07 [−0.16, 0.01]	0.06 [−0.07, 0.2]	−0.14 [−0.26, −0.03]	0.394	0.136

Progression is described using estimated annual progression rate by LMEM. *p*-values were obtained by likelihood ratio tests of the full model with the interaction term against the model without the interaction term. ESS=Epworth Sleepiness Scale; ISI=Insomnia Severity Index; MMSE=Mini-Mental State Examination; MoCA=Montreal Cognitive Assessment; PSQI=Pittsburgh Sleep Quality Inventory; RBDSQ=REM Behavior Disorder Sleep Questionnaire; SCOPA-AUT= Scales for Outcomes in Parkinson's Disease - Autonomic Dysfunction. [a]Results of all autonomic symptoms/signs can be found in **Supplementary Table 5**.

Table 4 Calculated sample size estimates to detect differences in marker progression at 50% and 30% drug efficacy

	50% Drug Effectiveness Sample size per group		30% Drug Effectiveness Sample size per group	
	2-year study	**3-year study**	**2-year study**	**3-year study**
Continuous Variable Analysis				
MDS-UPDRS-I	657	445	1825	1236
MDS-UPDRS-II	355	255	986	708
MDS-UPDRS-III	244	175	678	486
MDS-UPDRS-III (without action tremor)	213	153	592	425
MDS-UPDRS-I+II+III	183	141	507	392
Timed Up & Go (s)	1496	1123	1013	10678
Timed Up & Go (m/s)	560	319	1556	886
Purdue Pegboard	151	98	419	272
Postural Systolic Drop	1026	453	2850	1258
SCOPA-Total	2459	1448	6831	4022
Olfaction z-score	2046	1076	5683	2989
MoCA	22007	12930	61131	35917
MMSE	870	612	2417	1700
Depression z-score	7404	3802	20567	10561
Anxiety z-score	11398	6601	31661	18336
Event-based Analysis (time to event)				
Purdue Pegboard increase ≥ 4	273	164	896	540
MDS-UPDRS-III increase ≥ 4	167	108	551	362
MoCA decrease ≤ 3	497	304	1622	997
MDS-UPDRS-III ≥ 4 or MoCA ≤ 3	117	88	389	294
MDS-UPDRS I+II+III ≥ 12	226	121	742	403
Phenoconversion	409	265	1337	869

Sample sizes for a 2-arm parallel trial in which treatment is expected to reduce the rate of progression by a constant amount throughout follow-up. Presented are required sample sizes to detect 30% or 50% treatment effects for a 2- or 3-year trial with periodic 6 month-follow-up (for continuous variable analysis) specifying 80% power and 2-sided alpha=0.05. Sleep symptoms are not included since scores paradoxically improved over time. MMSE=Mini-Mental State Examination; MoCA=Montreal Cognitive Assessment; SCOPA-AUT= Scales for Outcomes in Parkinson's Disease - Autonomic Dysfunction.

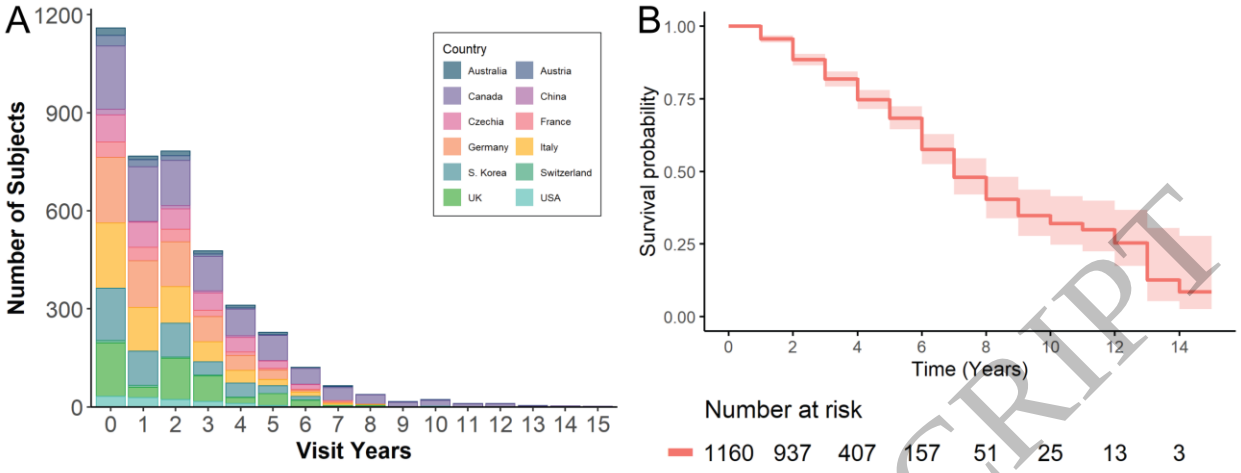

Figure 1
165x64 mm (.46 x DPI)

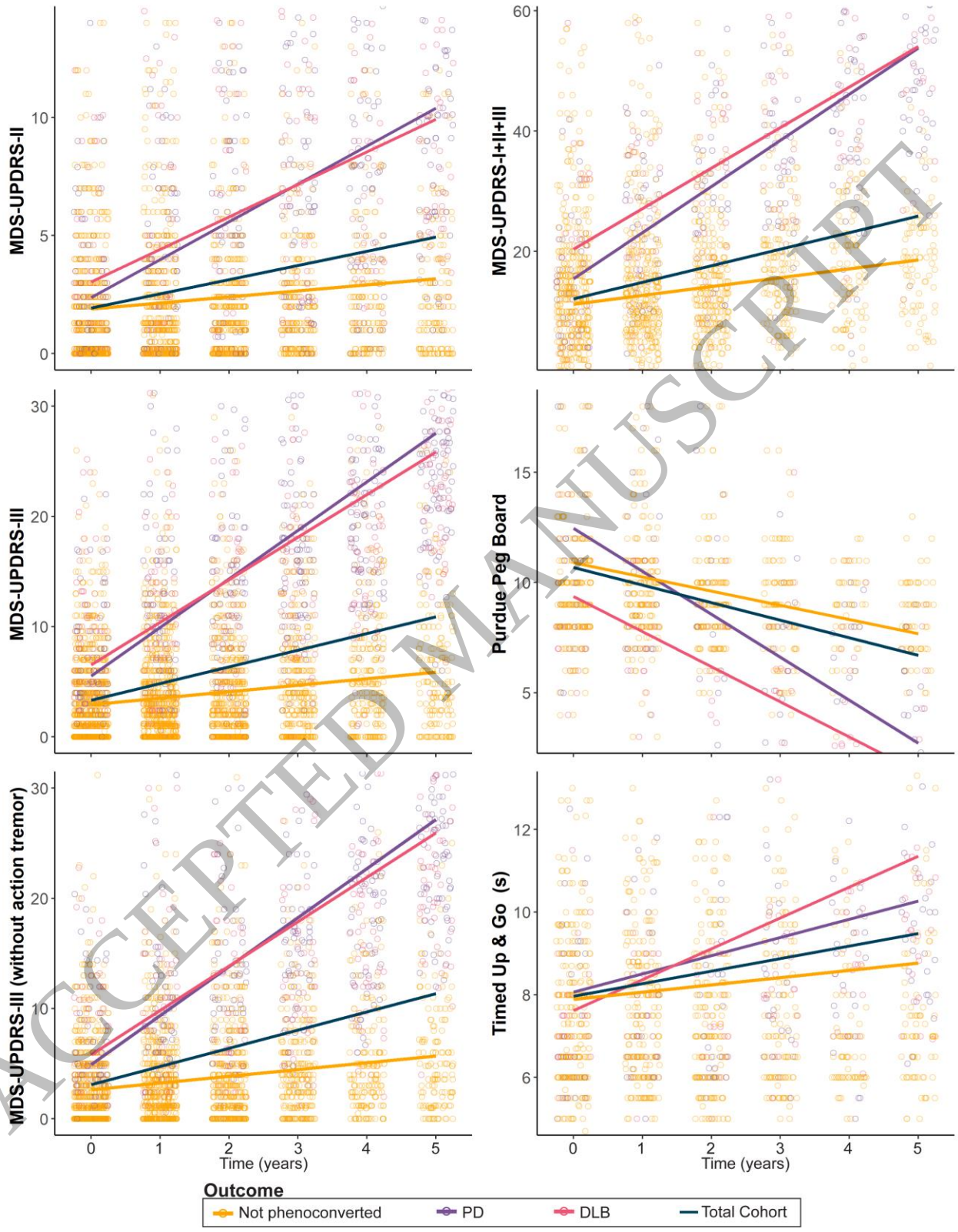

Figure 2
165x211 mm (.46 x DPI)

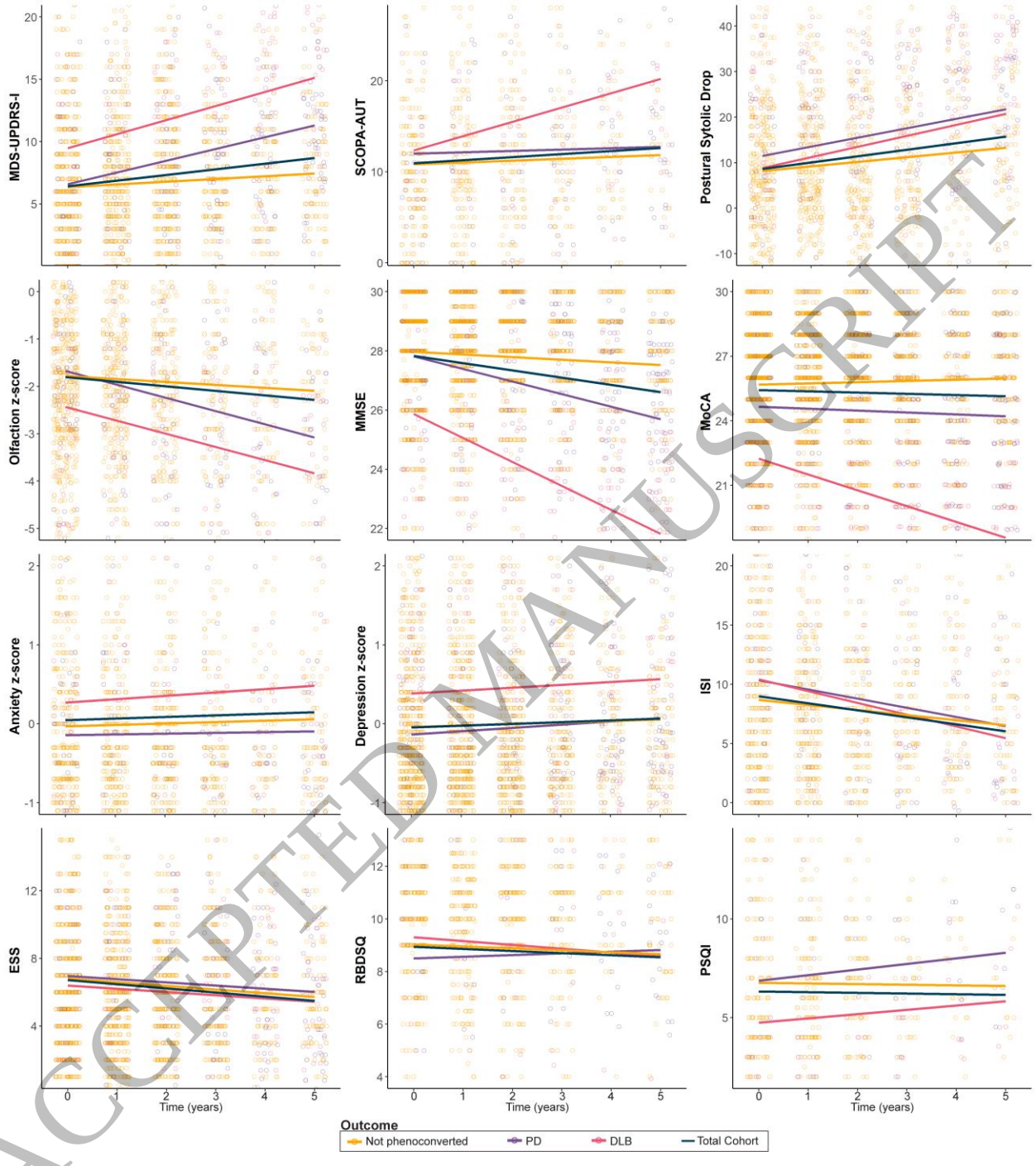

Figure 3
165x184 mm (.46 x DPI)

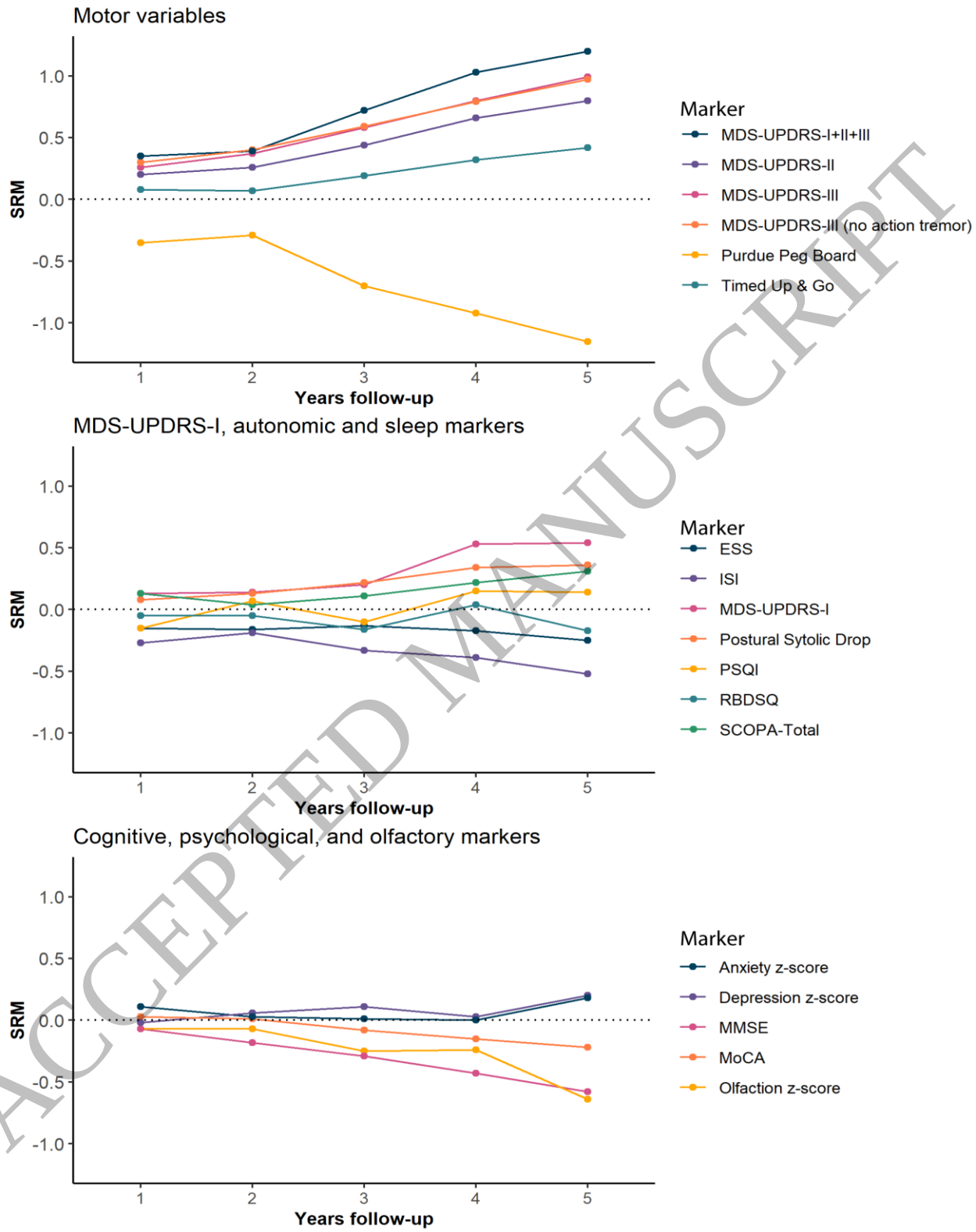

Figure 4
165x222 mm (.46 x DPI)