doi:10.1093/brain/awz030 BRAIN 2019: 142; 744–759 | 744

BRAIN
A JOURNAL OF NEUROLOGY

# Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study

**Ronald B. Postuma,[1,2] Alex Iranzo,[3] Michele Hu,[4] Birgit Högl,[5] Bradley F. Boeve,[6] Raffaele Manni,[7] Wolfgang H. Oertel,[8] Isabelle Arnulf,[9] Luigi Ferini-Strambi,[10] Monica Puligheddu,[11] Elena Antelmi,[12,13] Valerie Cochen De Cock,[14] Dario Arnaldi,[15] Brit Mollenhauer,[16] Aleksandar Videnovic,[17] Karel Sonka,[18] Ki-Young Jung,[19] Dieter Kunz,[20] Yves Dauvilliers,[21] Federica Provini,[22,23] Simon J. Lewis,[24] Jitka Buskova,[25] Milena Pavlova,[26] Anna Heidbreder,[27] Jacques Y. Montplaisir,[2] Joan Santamaria,[14] Thomas R. Barber,[4] Ambra Stefani,[5] Erik K. St.Louis,[6] Michele Terzaghi,[7] Annette Janzen,[8] Smandra Leu-Semenescu,[9] Guiseppe Plazzi,[12,13] Flavio Nobili,[15] Friederike Sixel-Doering,[16] Petr Dusek,[18] Frederik Bes,[20] Pietro Cortelli,[22,23] Kaylena Ehgoetz Martens,[24] Jean-Francois Gagnon,[28] Carles Gaig,[3] Marco Zucconi,[10] Claudia Trenkwalder,[15] Ziv Gan-Or,[29,30] Christine Lo,[4] Michal Rolinski,[4] Philip Mahlknecht,[5] Evi Holzknecht,[5] Angel R. Boeve,[6] Luke N. Teigen,[6] Gianpaolo Toscano,[7] Geert Mayer,[31] Silvia Morbelli,[32] Benjamin Dawson,[1] Amelie Pelletier[1,2] and the International REM Sleep Behavior Disorder Study Group**

See Morris and Weil (doi:10.1093/brain/awz014) for a scientific commentary on this article.

Idiopathic REM sleep behaviour disorder (iRBD) is a powerful early sign of Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy. This provides an unprecedented opportunity to directly observe prodromal neurodegenerative states, and potentially intervene with neuroprotective therapy. For future neuroprotective trials, it is essential to accurately estimate phenoconversion rate and identify potential predictors of phenoconversion. This study assessed the neurodegenerative disease risk and predictors of neurodegeneration in a large multicentre cohort of iRBD. We combined prospective follow-up data from 24 centres of the International RBD Study Group. At baseline, patients with polysomnographically-confirmed iRBD without parkinsonism or dementia underwent sleep, motor, cognitive, autonomic and special sensory testing. Patients were then prospectively followed, during which risk of dementia and parkinsonsim were assessed. The risk of dementia and parkinsonism was estimated with Kaplan-Meier analysis. Predictors of phenoconversion were assessed with Cox proportional hazards analysis, adjusting for age, sex, and centre. Sample size estimates for disease-modifying trials were calculated using a time-to-event analysis. Overall, 1280 patients were recruited. The average age was 66.3 ± 8.4 and 82.5% were male. Average follow-up was 4.6 years (range = 1–19 years). The overall conversion rate from iRBD to an overt neurodegenerative syndrome was 6.3% per year, with 73.5% converting after 12-year follow-up. The rate of phenoconversion was significantly increased with abnormal quantitative motor testing [hazard ratio (HR) = 3.16], objective motor examination (HR = 3.03), olfactory deficit (HR = 2.62), mild cognitive impairment (HR = 1.91–2.37), erectile dysfunction (HR = 2.13), motor symptoms (HR = 2.11), an abnormal DAT scan (HR = 1.98), colour vision abnormalities (HR = 1.69), constipation (HR = 1.67), REM atonia loss (HR = 1.54), and age (HR = 1.54). There was no significant predictive value of sex, daytime somnolence, insomnia, restless legs syndrome, sleep apnoea, urinary dysfunction, orthostatic symptoms, depression, anxiety, or hyperechogenicity on substantia nigra ultrasound. Among predictive markers, only cognitive

Received June 13, 2018. Revised November 6, 2018. Accepted November 9, 2018

**variables were different at baseline between those converting to primary dementia versus parkinsonism. Sample size estimates for definitive neuroprotective trials ranged from 142 to 366 patients per arm. This large multicentre study documents the high phenoconversion rate from iRBD to an overt neurodegenerative syndrome. Our findings provide estimates of the relative predictive value of prodromal markers, which can be used to stratify patients for neuroprotective trials.**

1 Department of Neurology, McGill University, Montreal General Hospital, Montreal, Canada
2 Centre d'Études Avancées en Médecine du Sommeil, Hôpital du Sacré-Cœur de Montréal, Montréal, Canada
3 Neurology Service, Hospital Clinic de Barcelona, IDIBAPS, CIBERNED, Barcelona, Spain
4 Oxford Parkinson's Disease Centre (OPDC) and Oxford University, Oxford, UK
5 Department of Neurology, Innsbruck Medical University, Innsbruck, Austria
6 Mayo Clinic, Rochester, MN, USA
7 Unit of Sleep Medicine and Epilepsy, IRCCS, C.Mondino Foundation, Pavia, Italy
8 Department of Neurology, Philipps-Universität, Marburg, Germany
9 Sleep disorders unit, Pitie-Salpetriere Hospital, IHU@ICM and Sorbonne University, Paris, France
10 Sleep Disorders Center, Department of Neurology, Scientific Institute Ospedale San Raffaele, Vita-Salute University, Milan, Italy
11 Sleep Center, Department of Cardiovascular and Neurological Sciences, University of Cagliari, Italy
12 Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy
13 IRCCS Institute of the Neurological Sciences, Ospedale Bellaria, ASL di Bologna, Bologna, Italy
14 Sleep and Neurology Unit, Beau Soleil Clinic, Montpellier, France; EuroMov, University of Montpellier, Montpellier, France
15 Clinical Neurology, Dept. of Neuroscience (DINOGMI), University of Genoa, and Polyclinic San Martino Hospital, Genoa, Italy
16 Department of Neurosurgery (C.T.) University Medical Center, Göttingen; Paracelsus-Elena-Klinik (B.M., C.T. F. S-D.), Kassel, Germany
17 Movement Disorders Unit and Division of Sleep Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, USA
18 Department of Neurology and Centre of Clinical Neurosciences of the First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic
19 Neuroscience Research Institute, Seoul National University College of Medicine, Department of Neurology, Seoul National University Hospital, Seoul, Korea
20 Institute of Physiology Charité-Universitätsmedizin Berlin. Germany
21 Sleep Unit, Department of Neurology, Hôpital Gui de Chauliac, Montpellier, INSERM U1061, Montpellier, F-34093 Cedex 5 France
22 Department of Biomedical and Neuromotor Sciences, Bellaria Hospital, University of Bologna, Bologna, Italy
23 IRCCS Institute of Neurological Sciences of Bologna, Bellaria Hospital, Bologna, Italy
24 Brain and Mind Centre University of Sydney, Camperdown, Australia
25 National Institute of Mental Health, Klecany, Third Faculty of Medicine, Charles Unviersity, Prague, Czech Republic
26 Department of Neurology, Brigham and Women's Hospital, Boston; Harvard Medical School, Boston, USA
27 Institute for Sleep Medicine and Neuromuscular Disorders, University Hospital Muenster, Muenster, Germany
28 Department of Psychology, Université du Québec à Montréal, Montreal, Quebec, Canada
29 Department of Human Genetics, McGill University, Montreal, Canada
30 Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Canada
31 Department of Neurology, Hephata Klinik, Schwalmstadt-Treysa, Germany
32 Nuclear Medicine, Department of Health Sciences (DISSAL), University of Genoa and Polyclinic San Martino Hospital, Genoa, Italy

Correspondence to: Ronald Postuma
Montreal General Hospital - McGill University
1650 Cedar Avenue Montreal Quebec H3G 1A4
Canada
E-mail: ron.postuma@mcgill.ca

**Keywords:** REM sleep behaviour disorder; Parkinson's disease; dementia with Lewy bodies; multiple system atrophy

**Abbreviations:** iRBD = idiopathic REM sleep behaviour disorder; MCI = mild cognitive impairment; MDS-UPDRS = Movement Disorders Society Unified Parkinson Disease Rating Scale; MSA = multiple system atrophy

# Introduction

The neurodegenerative synuclein aggregation disorders, namely Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy (MSA), all have a prodromal interval; that is, a period during which neurodegenerative symptoms/signs are present, but full clinical disease has not yet developed (Berg *et al.*, 2015). In the synucleinopathies,

this interval is notably long, often exceeding a decade (Berg *et al.*, 2015). This provides an unprecedented opportunity to provide potential neuroprotective therapy early, perhaps even preventing the development of parkinsonism and dementia.

Unlike many neurological diseases, whose prodromal states are predominantly identified by abnormalities in the same domain [e.g. mild cognitive impairment (MCI) is the primary prodromal marker of Alzheimer's disease], prodromal synucleinopathy markers are notably diverse. In addition to subtle motor signs, the potential prodromal markers include autonomic abnormalities, olfactory loss, cognitive changes, depression, anxiety, etc. (Goldman and Postuma, 2014). Most are relatively non-specific, such that the large majority of marker-positive subjects will never develop disease. However, a notable exception is idiopathic REM sleep behaviour disorder (iRBD).

RBD is a parasomnia in which the normal paralysis of REM sleep is lost, such that patients 'act out' their dreams (Schenck *et al.*, 2013*b*; Hogl *et al.*, 2018). Idiopathic RBD [alternatively termed 'isolated' (Hogl *et al.*, 2018) or 'cryptogenic' RBD] has a prevalence of ~1% over age 60 (Kang *et al.*, 2013; Haba-Rubio *et al.*, 2018; Pujol *et al.*, 2017), although most do not present to medical attention. Observational studies, generally from single centres, have suggested that most patients with iRBD will eventually develop a defined neurodegenerative disease, almost always diagnosed as synucleinopathy (Wing *et al.*, 2012; Schenck *et al.*, 2013*a*; Iranzo *et al.*, 2014; Arnulf *et al.*, 2015; Mahlknecht *et al.*, 2015; Postuma *et al.*, 2015*a*, *d*; Li *et al.*, 2017). In this context, RBD is likely related to neurodegeneration in the pontine or medullary areas associated with control of REM atonia (Valencia Garcia *et al.*, 2018). The latency from symptom onset to disease phenoconversion (i.e. conversion from iRBD to defined dementia with Lewy bodies, Parkinson's disease, or MSA) averages over 10 years (Schenck *et al.*, 2013*b*). Therefore, this implies that 1% of the elderly population have a readily-diagnosable but often-undetected early-stage neurodegenerative syndrome.

So far, most studies of phenoconversion risk and predictors came from single centres, so whether this is seen across different countries and different contexts remains unclear. In this study, we combined the prospective experience of 24 centres from the International RBD Study Group, to quantify the risk of phenoconversion to defined parkinsonism/dementia and to test 21 potential predictors of phenoconversion.

# Materials and methods

## Subjects

For inclusion, all subjects had to have iRBD confirmed on polysomnogram according to American Academy of Sleep Medicine Criteria (American Academy of Sleep Medicine and Hauri, 2007), and be free of parkinsonism or dementia on baseline neurological examination. Each patient had at least one follow-up examination during which systematic assessment for parkinsonism and dementia was performed. All patients gave written informed consent according to the Declaration of Helsinki, and ethics approval was obtained from the local institutional boards.

## Baseline variables

Centres collected all available information on baseline variables, then followed patients prospectively. We did not require that each variable be tested in each patient; rather, centres sent results for all those variables that they systematically assessed. Neither did we require that all variables be assessed with the same technique, as centres had different testing protocols for prodromal markers. For the analyses of hazard ratio (HR) in with tests were categorized as abnormal or normal, each centre defined each variable as abnormal/normal according to their own testing protocols, unless otherwise stated below. Detailed numbers of patients assessed with each variable is provided in Supplementary Table 1. Variables of interest and the assessment methods used included:

(i) Standardized motor examination: tested with the Movement Disorders Society Unified Parkinson Disease Rating Scale (MDS-UPDRS) (Goetz *et al.*, 2008). Either the 1987 UPDRS or 2008 MDS-UPDRS version could be used. For stratification purposes, the cut-off score was $>3$ excluding action tremor (Postuma *et al.*, 2012).

(ii) Standardized motor symptoms: UPDRS/MDS-UPDRS Part II (Fahn *et al.*, 1987; Goetz *et al.*, 2008).

(iii) Quantitative motor testing: tests included the alternate-tap test (Nutt *et al.*, 2000; Postuma *et al.*, 2015*c*), Purdue PegBoard (Desrosiers *et al.*, 1995; Postuma *et al.*, 2015*c*), 3-Metre Timed-Up-and-Go (Podsiadlo and Richardson, 1991; Postuma *et al.*, 2015*c*), or Flamingo balance test (Barber *et al.*, 2017). If multiple tests were conducted in one centre, the majority had to be abnormal to classify the testing as abnormal.

(iv) Olfaction: 12- or 40-item University of Pennsylvania Smell Identification Test (Doty *et al.*, 1984) or Sniffin Sticks (Hummel *et al.*, 1997; Mahlknecht *et al.*, 2015).

(v) Colour vision: Farnsworth-Munsell 100-Hue test (Farnsworth, 1943).

(vi) Physician-documented insomnia: Insomnia Severity Index (Bastien *et al.*, 2001), Athens Insomnia Scale (Soldatos *et al.*, 2000), or clinical interview.

(vii) Excessive daytime somnolence: Epworth Sleepiness scale (Johns, 1991; Valencia Garcia *et al.*, 2018) or clinical interview.

(viii) Restless legs syndrome: diagnosed using clinical interview.

(ix) Sleep apnoea: apnoea-hypopnoea index cut-off $\geqslant 15$/h (secondary analysis was also performed using cut-off $\geqslant 5$/h).

(x) REM sleep without atonia: scored as % tonic and phasic chin REM on the polysomnographic trace, using either Montreal scoring (Montplaisir *et al.*, 2010), or % 'any' tone using SINBAR scoring, chin $\pm$ arm (Frauscher *et al.*, 2012). For combined stratification, we divided each individual's score by the mean estimate from their centre.

(xi) Constipation: Unified MSA Rating Scale (UMSARS) (Wenning *et al.*, 2004), SCOPA-AUT (Visser *et al.*, 2004), Rome Criteria (Higgins and Johanson, 2004), or clinical interview.

(xii) Urinary symptoms: UMSARS, SCOPA-AUT, or clinical interview.

(xiii) Erectile dysfunction: UMSARS, SCOPA-AUT, or clinical interview.
(xiv) Orthostatic symptoms: UMSARS, SCOPA-AUT, PD-NMS-Quest (Chaudhuri *et al.*, 2006), or clinical interview.
(xv) Orthostatic blood pressure: assessed lying and after 1–3 min standing. For illustration/stratification purposes only, a cut-off systolic drop of >10 mmHg was used.
(xvi) Cognition, neuropsychological testing: MCI/neurocognitive disorder diagnosed as abnormal neuropsychological testing (generally two or more tests abnormal in one or more domain, adjusted for age and education), plus subjective cognitive complaint, and preserved activities of daily living. We also assessed predictive value of abnormal cognitive testing, regardless of reported cognitive symptoms.
(xvii) Cognition, office-based diagnosis: Folstein Mini-Mental State Examination (MMSE) (Folstein *et al.*, 1975) or Montreal Cognitive Assessment (MoCA) (Nasreddine *et al.*, 2005). For stratification, an education-adjusted MoCA < 26 and MMSE < 28 were defined as abnormal (for combined analysis, MoCA given priority, as it has a validated MCI cut-off) (Nasreddine *et al.*, 2005). For definition of office-based possible MCI, we also required cognitive complaint/symptoms.
(xviii) Depression: Beck Depression Inventory (Beck *et al.*, 1961), Geriatric Depression Scale (Yesavage *et al.*, 1982), Patient Health Questionnaire-9 (Kroenke *et al.*, 2001) or clinical interview.
(xix) Anxiety: Beck Anxiety Inventory (Beck *et al.*, 1988), Neuropsychiatric Inventory (Cummings *et al.*, 1994), State-Trait Anxiety Inventory (Gaudry *et al.*, 1975), Hospital Anxiety and Depression scale (Zigmond and Snaith, 1983), Leeds Anxiety Scale (Snaith *et al.*, 1976), and NMS-Quest (Chaudhuri *et al.*, 2006), Generalized Anxiety Disorder-7 (Spitzer *et al.*, 2006).
(xx) Dopamine-transporter single photon emission tomography (DAT-SPECT), focusing on the putamen as region of interest.
(xxi) Substantia nigra pars compacta hyperechogenicity measured by transcranial ultrasound.

## Follow-up and disease conversion

All centres prospectively followed patients with in-person evaluation to diagnose phenoconversion to defined parkinsonism [defined as bradykinesia plus at least one of rigidity or rest tremor (Postuma *et al.*, 2015*b*)] or dementia [defined as cognitive impairment on standardized testing with functional impairment (Dubois *et al.*, 2007)]. For patients with parkinsonism as the primary disease manifestation, the primary diagnosis (Parkinson's disease/MSA) was made according to the treating neurologist. This differential diagnosis incorporated all available follow-up information (i.e. any patient who was initially diagnosed with Parkinson's disease at phenoconversion but who was subsequently found to have MSA would be included as MSA). For dementia conversions, all patients had polysomnogram-diagnosed RBD; therefore, they met 2017 criteria for probable dementia with Lewy bodies with a clinical core symptom plus biomarker loss of REM atonia (McKeith *et al.*, 2017). One centre (Barcelona) excluded patients with MCI from their cohort at baseline and delineated *de novo* MCI as a phenoconversion. To prevent any resulting bias in conversion risk estimates, we delineated phenoconversion as new parkinsonism or neuropsychological-examination-diagnosed MCI for this centre only (and also conducted sensitivity analysis removing that centre). Survival analysis was done using Kaplan-Meier analysis to estimate disease risk. For overall analysis, time = 0 was the first baseline in-person evaluation. In analysis of individual variables, not all variables might have been measured at the same baseline visit. Therefore, we calculated the interval for each individual variable (i.e. time = 0 was the first evaluation of that specific variable). For the stratification analysis (and for illustration in figures), we defined predictive markers binarily (normal/abnormal). For those variables without definable cut-offs for abnormality within RBD (e.g. REM atonia, age), we stratified as above versus below mean values (Table 2). For testing potential prodromal markers, the primary analysis was Cox proportional hazards analysis adjusting for baseline age, sex, and centre. Each prodromal marker was analysed both as continuous and categorical variables. To facilitate comparison between variables, we present the categorical analysis (stratified as discussed above); note that in all cases, statistical significance ($P < 0.05$ threshold) was the same for continuous and categorical analyses. Finally, we assessed the predictive value of selected markers in combination. To increase precision and reliability only combinations that could be tested in at least three centres, with >50 patients in each possible combination (i.e. none, one, or both variables present) were eligible for combined analysis.

On secondary analysis, among convertors who were diagnosed with Lewy body disease (i.e. excluding MSA), we compared those who developed dementia as the first disease manifestation versus parkinsonism-first conversions (if both were diagnosed on the same visit, the patient was classified as dementia-first).

Finally, we estimated sample size requirements for a future neuroprotective trial. This assumed a categorical definitive end-point (defined disease phenoconversion), with two groups (placebo versus a single-dose of active treatment), two-sided alpha = 0.05, and 80% power. We used time-to-event analysis (http://www.quesgen.com/SSSurvival.php), for a 2-year trial, assuming an agent that reduces phenoconversion with HR = 0.5. We calculated sample size for the population as a whole, and using stratification by prodromal marker testing, using directly-observed conversion rates, and also by using the hazard ratio from the current study estimates (i.e. adjusting for centre effects by recalculating the conversion rate in each single analysis to equal the median conversion rate in the entire group). For assessment of MDS prodromal criteria, we included only patients who had sufficient testing to reasonably estimate their % probability, which was defined as four or more prodromal variables including at least one of the three highest-specificity variables (olfaction, objective motor examination/quantitative testing, DAT-SPECT); for all calculations, the likelihood ratio of RBD (130) was included.

## Data availability

The original database from the study can be obtained by contacting the first author (R.B.P.).

# Results

## Participants

A total of 1280 patients from 24 centres were included in this study. Recruitment data from each centre are

## Table 1 Recruitment data and patient follow-up

| Cente | *n* | Follow-up duration, y, mean | Total patient-years | UPDRS /MDS-UPDRS Part III | Quantitative motor testing | UPDRS /MDS UPDRS Part II | Olfaction | Colour vision | Other sleep symptoms | PSG - apnoea | PSG - REM atonia % | Autonomic symptoms | Orthostatic blood pressure | Office-based cognitive tests | Neuropsychological examination | Depression | Anxiety | Substantia nigra ultrasound | DAT-SPECT | Total patient/ year/ variables |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Montreal | 154 | 4.3 | 664 | X | X | X | X | X | ± | X | X | X | X | X | X | X | X | | | 10 789 |
| Barcelona | 202 | 5.4 | 1091 | | | | | | | X | X | ± | | | | X | | | | 3926 |
| Oxford | 120 | 2.2 | 266 | X | X | X | X | | X | X | | X | X | X | | X | X | | | 3790 |
| Innsbruck | 98 | 4.5 | 461 | ± | | ± | ± | | X | X | | X | | | | | | | | 3559 |
| Rochester, Minnesota | 96 | 4.4 | 424 | X | | ± | X | ± | ± | ± | | ± | ± | X | X | ± | X | | | 3434 |
| Pavia | 74 | 5.4 | 400 | X | | | | | ± | X | | ± | ± | X | X | ± | ± | ± | ± | 3240 |
| Marburg | 43 | 3.4 | 145 | X | | X | X | X | | | | X | | X | | X | X | X | X | 1697 |
| Paris | 97 | 2.9 | 278 | ± | | ± | ± | | ± | X | X | ± | ± | X | ± | ± | | | ± | 1585 |
| Milan | 51 | 2.4 | 123 | X | | X | X | | ± | X | ± | X | X | X | X | X | X | | | 1583 |
| Cagliari | 37 | 2.9 | 108 | X | | X | | | X | X | ± | X | | X | X | ± | ± | | ± | 1295 |
| Bologna (Antelmi) | 24 | 4.2 | 100 | | | X | | | X | X | X | ± | X | X | ± | X | ± | | | 990 |
| Montpellier (Cochen) | 34 | 1.9 | 65 | X | | X | ± | | X | X | X | X | | X | | X | ± | | X | 916 |
| Genoa | 31 | 2.7 | 84 | X | | X | X | ± | ± | ± | X | ± | | X | X | X | ± | | X | 856 |
| Kassel | 22 | 2.3 | 50 | X | | X | X | ± | X | X | X | X | X | X | ± | X | | X | X | 795 |
| Boston (MGH) | 16 | 3.4 | 54 | X | | X | | | X | X | X | X | | X | | X | X | | | 680 |
| Prague (Sonka) | 31 | 1.5 | 46 | X | | | X | X | X | X | | X | | X | | X | X | X | X | 662 |
| Seoul | 28 | 2.1 | 58 | | | | ± | | X | X | | ± | | X | X | X | | | X | 557 |
| Berlin | 47 | 1.2 | 55 | X | | | X | | X | ± | ± | | | X | X | X | X | | X | 528 |
| Montpellier (Dauvilliers) | 17 | 3.5 | 60 | ± | | | | | X | X | ± | | ± | | | | | | ± | 281 |
| Bologna (Provini) | 9 | 2.3 | 21 | X | X | | | | X | X | X | X | X | X | X | ± | ± | | | 277 |
| Sydney | 13 | 2.2 | 29 | X | | X | ± | ± | | | X | | | X | | X | X | ± | | 200 |
| Prague (Buskova) | 9 | 1.3 | 12 | X | | X | X | | X | X | X | X | | X | | X | X | | | 162 |
| Boston (Brigham) | 6 | 4.0 | 24 | ± | | ± | | | X | ± | | ± | ± | | | ± | | | | 130 |
| Muenster | 19 | 1.2 | 23 | | | | X | | ± | X | X | ± | | | | ± | | ± | | 124 |
| Total | 1280 | 72.7 | 4890 | | | | | | | | | | | | | | | | | 42 698 |

X = systematically assessed and available for large majority (i.e. ⩾67%) of participants; ± = assessed in a subset (10–66%).

**Table 2 Baseline predictors of neurodegenerative phenoconversion in iRBD**

| | Developed disease *n* = 353 | Still disease-free *n* = 927 | Unadjusted HR (95%CI) | HR, adjusted age/sex/centre (95% CI) |
|---|---|---|---|---|
| Age | 67.6 ± 6.9 | 65.9 ± 8.8 | **1.52 (1.23–1.88)** | **1.54 (1.23–1.91)** |
| Sex, % male | 83.9 | 82.0 | 0.98 (0.73–1.30) | 0.93 (0.70–1.24) |
| UPDRS Part III | | | | |
| Combined: abnormal | 62.1% | 29.9% | **2.70 (2.03–3.60)** | **3.03 (2.21–4.15)** |
| 1987 UPDRS | 5.84 ± 4.72 (*n* = 142) | 2.84 ± 3.36 (*n* = 279) | **2.46 (1.75–3.45)** | **2.75 (1.89–4.01)** |
| MDS-UPDRS | 6.26 ± 4.94 (*n* = 57) | 3.12 ± 3.82 (*n* = 299) | **3.48 (2.03–5.97)** | **3.77 (2.11–6.77)** |
| Quantitative motor abnormal | 62.7% (*n* = 75) | 22.7% (*n* = 198) | **3.46 (2.16–5.56)** | **3.16 (1.86–5.37)** |
| UPDRS Part II | | | | |
| Combined, above mean | 56.0% | 32.2% | **1.62 (1.12–2.36)** | **2.11 (1.35–3.32)** |
| 1987 UPDRS | 1.38 ± 1.79 (*n* = 72) | 1.13 ± 1.85 (*n* = 157) | 1.23 (0.76–1.99) | 1.29 (0.75–2.22) |
| MDS-UPDRS | 3.33 ± 4.21 (*n* = 51) | 1.87 ± 3.84 (*n* = 233) | **2.94 (1.66–5.20)** | **4.75 (2.33–9.66)** |
| Olfaction abnormal | 78.7% (*n* = 127) | 63.5% (*n* = 501) | **2.33 (1.52–3.58)** | **2.62 (1.67–4.12)** |
| Olfaction, excluding MSA | 80.3% (*n* = 122) | 63.5% (*n* = 501) | **2.54 (1.62–3.98)** | **2.91 (1.81–4.67)** |
| Colour vision abnormal | 52.9% (*n* = 70) | 32.4% (*n* = 170) | **1.62 (1.01–2.56)** | **1.69 (1.01–2.78)** |
| Insomnia | 31.6% (*n* = 79) | 29.9% (*n* = 328) | 0.86 (0.54–1.39) | 0.90 (0.54–1.52) |
| Daytime somnolence | 34.6% (*n* = 263) | 29.3% (*n* = 755) | 1.06 (0.82–1.38) | 1.16 (0.89–1.51) |
| Restless legs syndrome | 17.2% (*n* = 169) | 17.9% (*n* = 504) | 0.92 (0.61–1.40) | 1.06 (0.67–1.68) |
| Apnoea (AHI ⩾ 15) | 27.7% (*n* = 271) | 27.4% (*n* = 811) | 1.06 (0.81–1.38) | 0.92 (0.70–1.23) |
| REM %[a]: above mean | 60.0% | 47.5% | **1.65 (1.13–2.42)** | **1.54 (1.05–2.27)** |
| Tonic REM % (MTL) | 52.6 ± 29.1 (*n* = 104) | 47.3 ± 29.2 (*n* = 255) | 1.41 (0.95–2.08) | 1.38 (0.93–2.05) |
| Phasic REM % (MTL) | 32.5 ± 18.3 (*n* = 80) | 30.0 ± 21.5 (*n* = 170) | 1.18 (0.76–1.84) | 1.37 (0.84–2.26) |
| % Any (SINBAR) | 65.0 ± 21.1 (*n* = 18) | 59.1 ± 23.7 (*n* = 91) | 2.69 (0.62–11.7) | 3.40 (0.75–15.1) |
| Constipation | 56.4% (*n* = 202) | 38.7% (*n* = 628) | **1.69 (1.27–2.23)** | **1.67 (1.24–2.24)** |
| Urinary dysfunction | 34.3 (*n* = 143) | 30.5% (*n* = 544) | 1.20 (0.85–1.69) | 1.06 (0.73–1.54) |
| Erectile dysfunction | 63.1% (*n* = 65) | 36.5% (*n* = 211) | **1.89 (1.13–3.21)** | **2.13 (1.10–4.13)** |
| Orthostatic symptoms | 33.6% (*n* = 119) | 28.2% (*n* = 412) | 1.29 (0.88–1.89) | 1.41 (0.93–2.13) |
| Systolic blood pressure drop | 14.4 ± 18.6 (*n* = 87) | 6.5 ± 13.7 (*n* = 267) | **1.55 (1.04–2.29)** | 1.37 (0.90–2.08) |
| Abnormal office: cognitive test (regardless of complaint) | 53.0% (*n* = 185) | 34.4% (*n* = 591) | **1.63 (1.22–2.18)** | **1.55 (1.15–2.11)** |
| MoCA <26 | 24.7 ± 3.2 (*n* = 84) | 25.8 ± 2.9 (*n* = 346) | 1.47 (0.95–2.27) | 1.47 (0.93–2.32) |
| MMSE <28 | 26.9 ± 3.4 (*n* = 132) | 28.1 ± 1.8 (*n* = 375) | **1.69 (1.20–2.38)** | **1.58 (1.10–2.28)** |
| Neuropsychological abnormal (regardless of complaint) | 60.9% (*n* = 138) | 25.9% (*n* = 328) | **2.09 (1.48–2.94)** | **1.89 (1.22–2.94)** |
| Mild cognitive impairment[b] | | | | |
| Neuropsychological testing | 55.4% (*n* = 121) | 16.4% (*n* = 299) | **2.53 (1.77–3.62)** | **2.37 (1.45–3.88)** |
| MoCA/MMSE | 41.7% (*n* = 151) | 17.4% (*n* = 477) | **1.98 (1.43–2.74)** | **1.91 (1.34–2.73)** |
| Depression | 28.8% (*n* = 226) | 25.6% (*n* = 632) | 1.17 (0.87–1.56) | 1.20 (0.88–1.63) |
| Anxiety | 22.6% (*n* = 116) | 17.7% (*n* = 429) | 1.46 (0.93–2.27) | 1.44 (0.88–2.35) |
| Substantia nigra ultrasound | 64.3% (*n* = 14) | 64.5% (*n* = 65) | 1.14 (0.35–3.72) | 1.19 (0.29–4.82) |
| DAT scan (putamen) abnormal | 69.2% (*n* = 52) | 37.3% (*n* = 193) | **2.22 (1.22–4.05)** | **1.98 (1.05–3.73)** |
| MDS prodromal criteria | 92.7% (*n* = 150) | 71.1% (*n* = 440) | **4.52 (2.44–8.35)** | **5.37 (2.77–10.4)** |

Continuous variables are presented as mean ± standard deviation (*n*). To allow direct comparisons between markers, all continuous variables are stratified to normal versus abnormal; for values with no defined abnormal cut-off above (e.g. age) results were stratified as above or below mean values. Hazard ratios are presented according to Cox proportional hazards analysis performed with logistic regression adjusting for age, sex, and centre.

[a]Each result is stratified to above or below mean values for that centre. The combined analysis combines tonic/phasic/any tone. Measures for which the confidence intervals do not cross one (i.e. $P < 0.05$) are highlighted in bold.

[b]Reference group is normal cognitive testing (regardless of cognitive complaint). Diagnosis of MDS prodromal criteria includes the likelihood ratio of RBD.

AHI = Apnoea–Hypopnea Index; MoCA = Montreal Cognitive Assessment; MMSE = Mini Mental State Examination; MTL = Montreal; SINBAR = Sleep Innsbruck Barcelona.

summarized on Table 1. Mean age at baseline was 66.3 ± 8.4 and 82.5% were male. The mean follow-up duration (between first baseline examination and last contact or disease conversion) was 3.6 years (maximum = 19 years), translating to 4890 total person-years of follow-up.

## Overall outcome

During follow-up, 352 (28%) converted to an overt neurodegenerative syndrome (Fig. 1). The mean interval between baseline evaluation and phenoconversion was 4.6 ± 3.5 years. The median time to phenoconversion was 8.0 years,

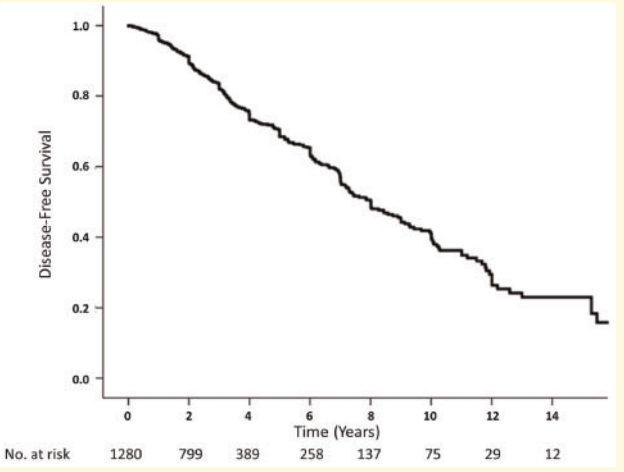


**Figure 1 Kaplan-Meier plot of disease-free survival (i.e. free of parkinsonism or dementia) among patients with iRBD.**

with an overall phenoconversion rate of 6.25% per year. The risk of phenoconversion on Kaplan-Meier analysis was 10.6% after 2 years, 17.9% after 3 years, 31.3% after 5 years, 51.4% after 8 years, 60.2% after 10 years, and 73.5% after 12 years. With regards to disease classifications, 199 (56.5%) developed parkinsonism as the first disease manifestation [of whom 16 (4.5%) were diagnosed with probable MSA], and 153 (43.5%) developed dementia first.

## Predictors of outcome

Kaplan-Meier analysis of selected predictors is illustrated on Fig. 2. On Cox proportional hazards analysis, adjusting for age, sex, and centre, numerous measures significantly predicted outcome (Table 2 and Figs 2–4). These included: (i) quantitative motor testing (HR = 3.16); (ii) standardized motor examination [HR = 3.03 overall, higher for MDS-UPDRS (3.77) than UPDRS-III (2.75)]; (iii) olfaction (HR = 2.62). Predictive value also improved when excluding MSA patients (HR = 2.91); (iv) MCI, with better prediction using neuropsychological examination (HR = 2.37) than with office-based testing (HR = 1.91); (v) erectile dysfunction (HR = 2.13); (vi) motor symptoms: HR = 2.11, with better prediction for the MDS-UPDRS-II (HR = 4.75) than the 1987 UPDRS-II (HR = 1.29); (vii) DAT-SPECT (HR = 1.98); (viii) neuropsychological testing (regardless of cognitive complaint) (HR = 1.89); (ix) colour vision (HR = 1.69); (x) constipation (HR = 1.67); (xi) REM sleep without atonia (HR = 1.54, on combined analysis only); (xii) brief office-based cognitive tests (regardless of cognitive complaint) (MMSE/MoCA combined HR = 1.55); and (xiii) age (HR = 1.54 for above versus below mean).

In addition, systolic blood pressure drop at a cut-off of 10 mm (HR = 1.55) predicted outcome on unadjusted analysis, but not after adjusting for age, sex, and centre (HR = 1.37) (using a cut-off of 20 mm, the unadjusted HR was 1.37 (0.88–2.15) and adjusted HR was 1.20 (0.74–1.91). The MDS prodromal criteria (which combines numerous variables) predicted outcome with the highest hazard ratio (HR = 5.37).

By contrast, we saw no significant predictive differences according to sex, insomnia symptoms, daytime somnolence, restless legs syndrome, apnoea, urinary dysfunction, orthostatic symptoms, depression, anxiety, or substantia nigra ultrasound.

## Secondary and sensitivity analyses

Among the 336 patients diagnosed with Lewy Body disease (i.e. excluding MSA), there were relatively few differences between patients who converted to dementia first versus parkinsonism first (Table 3). Age and sex were similar. All motor measures were similar except for quantitative motor testing, which was more likely to be abnormal in those developing dementia first (82.4%) than parkinsonism first (47.2%). Olfaction was similar in both groups, as were all sleep symptoms and polysomnographic variables. Autonomic symptoms were similar, as was orthostatic blood pressure drop, depression or anxiety. Although power was limited, we also saw no differences in proportion of patients with abnormal DAT-SPECT or substantia nigra ultrasound. The only variables that differed strongly (all $P < 0.001$) were those that tested cognition, including office based cognitive testing, neuropsychological examination, and colour vision testing which predicted only dementia [note that colour vision predominantly tests visuoperceptual cognition in Parkinson's disease (Bertrand *et al.*, 2012)].

Excluding results from centres that already published data on these predictors did not substantially affect the hazard ratio. For example, the hazard ratio of UPDRS excluding Montreal (Postuma *et al.*, 2012) was 3.04, versus 3.03 for entire group. The hazard ratio of olfaction excluding both Montreal (Postuma *et al.*, 2011) and Innsbruck (Mahlknecht *et al.*, 2015) was 2.53, versus 2.62.

## Sample size calculations

Based on the time-to-event analysis, we estimated that 366 patients per arm would need to be recruited into a 2-year trial to have 80% power to find a 50% reduction in disease phenoconversion (i.e. 65 phenoconversion events; Table 4). Adjusting the study duration altered sample sizes roughly proportionally to the proportion in duration (e.g. 4-year trial = 192 per group, 1-year trial = 709 per group). Testing different effectiveness assumptions, a drug providing 80% reduction in phenoconversion would require 84 patients per group (12 phenoconversion events) while a 30% reduction would require 959 (190 phenoconversion events).

The most powerful single selection procedure (abnormal quantitative motor testing) reduced sample size to 166–197 patients; however, only 34% of the iRBD population had abnormal testing and so would be included in such a study.

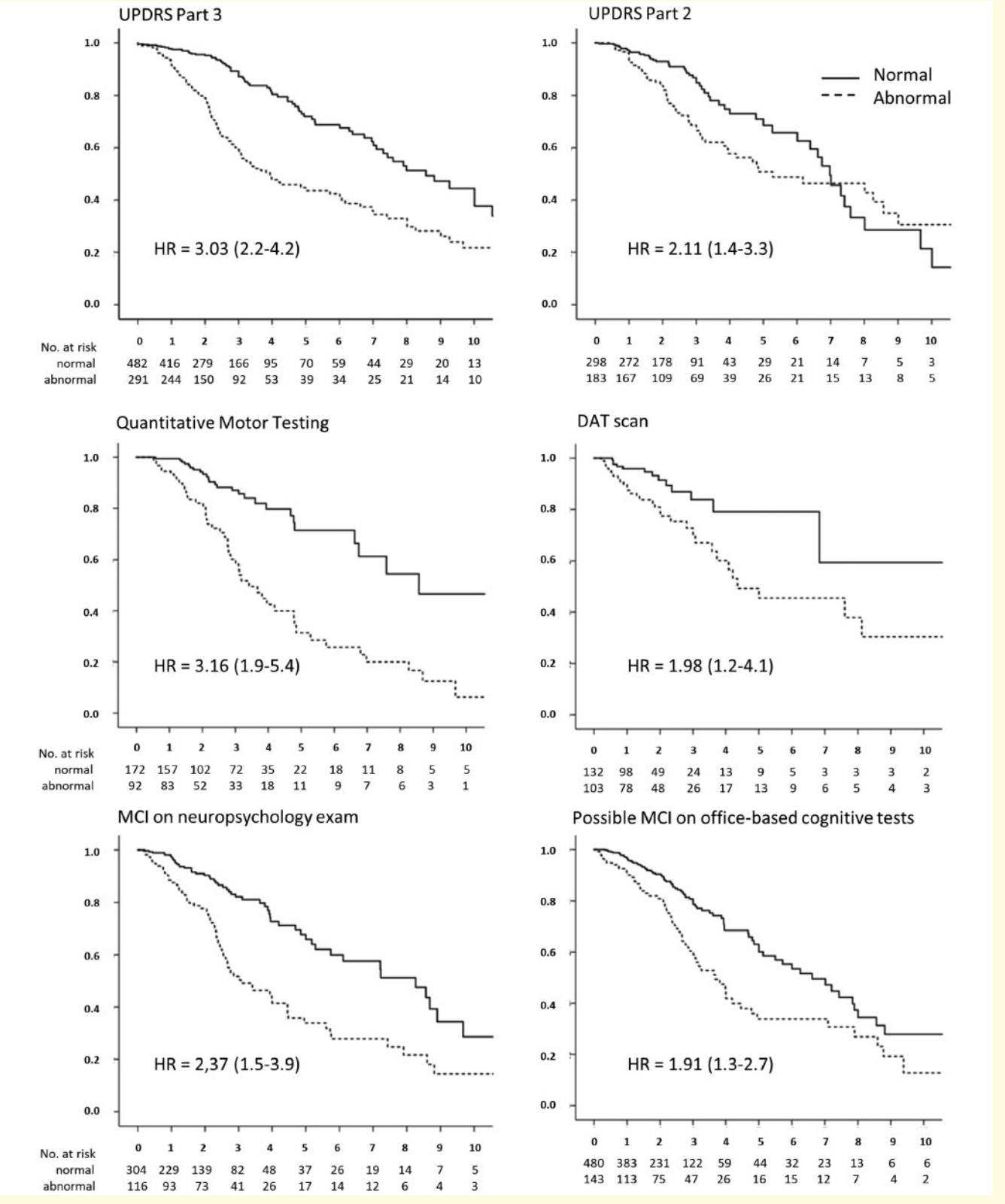


**Figure 2 Kaplan-Meier plot of disease-free survival of patients with iRBD stratified according to presence of motor and cognitive markers.** Results are presented according to baseline assessment (i.e. patients who develop a *de novo* marker abnormality over the course of the follow-up remain in the 'marker-free' group). Solid line indicates patients with normal values, dashed line abnormal values. Hazard ratios (HRs) are with Cox proportional hazards, adjusting for age, sex, and centre, with 95% confidence intervals in parentheses.

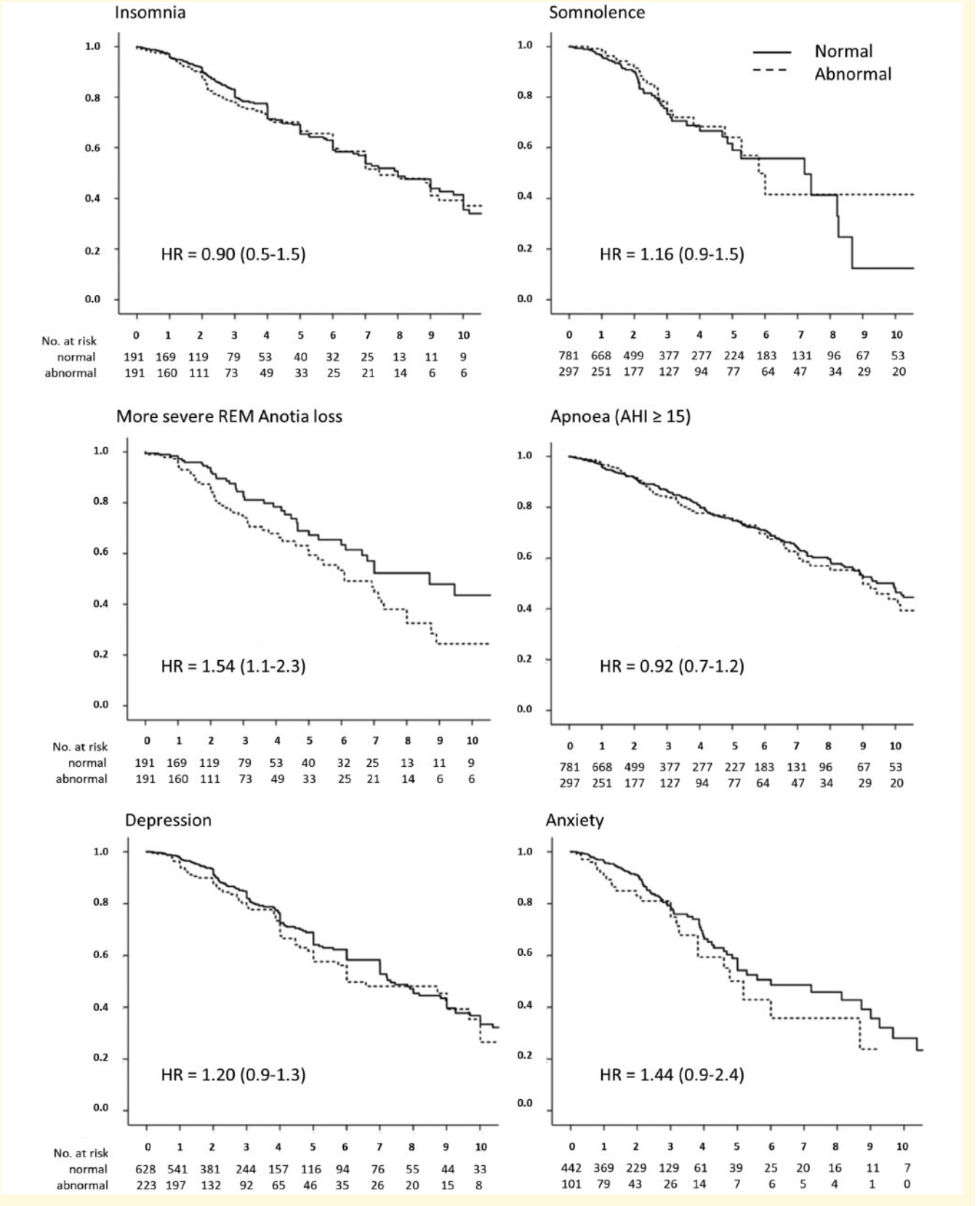


**Figure 3 Kaplan-Meier plot of disease-free survival of patients with iRBD stratified according to presence of sleep and psychiatric markers.** Solid line indicates patients with normal values, dashed line abnormal values. Hazard ratios (HRs) are with Cox proportional hazards, adjusting for age, sex, and centre, with 95% confidence intervals in parentheses.

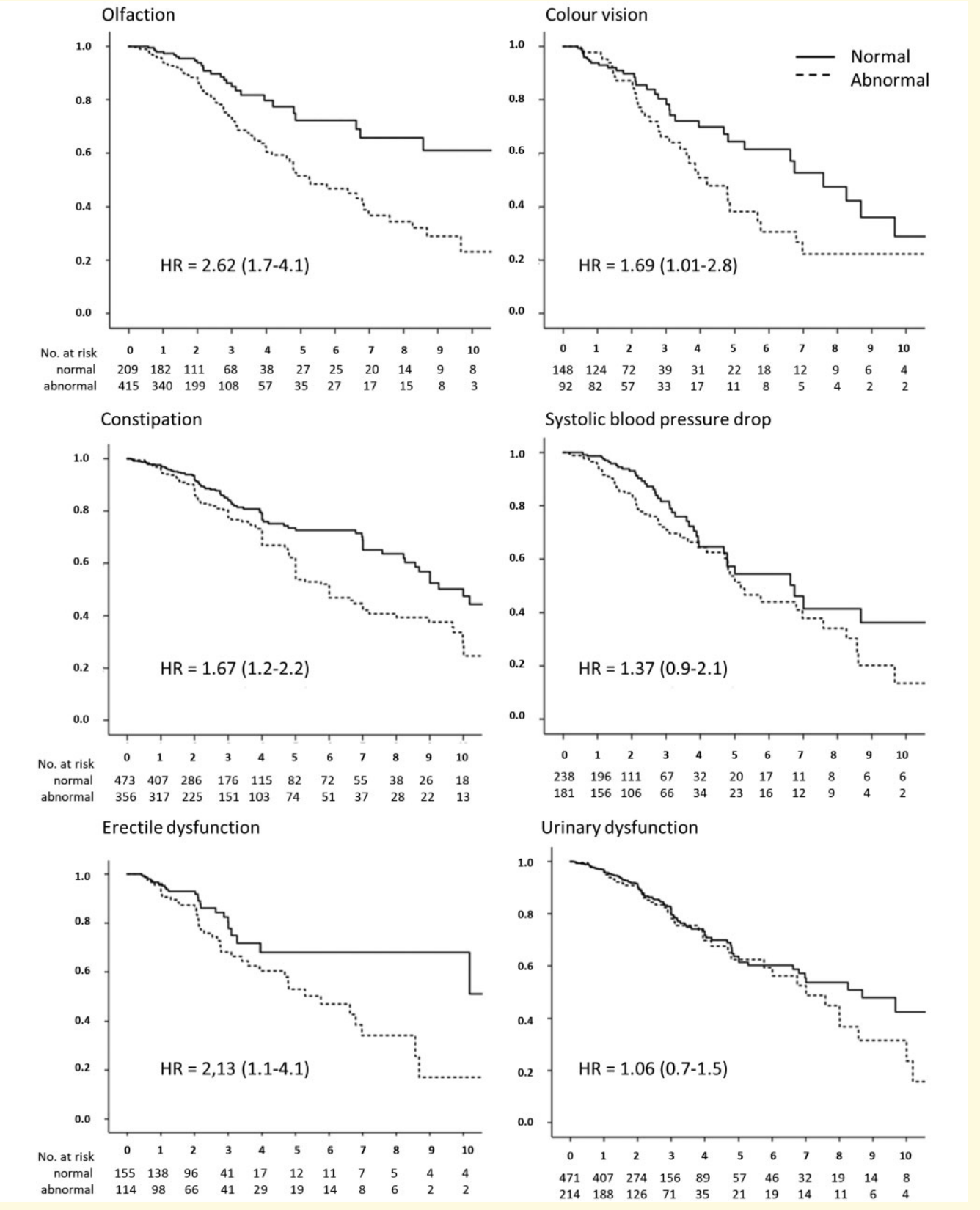


**Figure 4 Kaplan-Meier plot of disease-free survival of patients with iRBD stratified according to presence of special sensory and autonomic markers.** Solid line indicates patients with normal values, dashed line abnormal values. Hazard ratios (HRs) are with Cox proportional hazards, adjusting for age, sex, and centre, with 95% confidence intervals in parentheses.

**Table 3 Diagnosed Lewy body disease, divided into parkinsonism versus dementia-first**

| | Parkinsonism-first *n* = 184 | Dementia-first *n* = 146 | *P*-value |
|---|---|---|---|
| Age | 67.4 ± 6.6 | 68.3 ± 7.1 | 0.23 |
| Sex, % male | 81.0 | 88.4 | 0.068 |
| UPDRS Part III | | | |
| Combined: abnormal | 60.4% | 63.7% | 0.64 |
| 1987 UPDRS | 5.40 ± 4.38 (*n* = 60) | 6.17 ± 4.96 (*n* = 77) | 0.34 |
| MDS-UPDRS | 5.56 ± 5.08 (*n* = 41) | 6.36 ± 3.69 (*n* = 14) | 0.53 |
| Quantitative Motor Abnormal | 47.2% (*n* = 36) | 82.4% (*n* = 34) | 0.002 |
| UPDRS Part II | | | |
| Combined, above mean | 50.0% | 61.7% | 0.22 |
| 1987 UPDRS | 1.44 ± 1.84 (*n* = 35) | 1.10 ± 1.46 (*n* = 34) | 0.51 |
| MDS-UPDRS | 2.38 ± 2.75 (*n* = 34) | 5.60 ± 6.12 (*n* = 15) | 0.27 |
| Olfaction abnormal | 75.7% (*n* = 70) | 86.5% (*n* = 52) | 0.13 |
| Colour vision abnormal | 30.3% (*n* = 33) | 73.5% (*n* = 34) | <0.001 |
| Insomnia | 26.1% (*n* = 46) | 32.1% (*n* = 28) | 0.58 |
| Daytime somnolence | 28.6% (*n* = 133) | 40.4% (*n* = 114) | 0.051 |
| Restless legs syndrome | 21.1% (*n* = 95) | 11.3% (*n* = 62) | 0.11 |
| Apnoea (AHI ⩾ 15) | 26.8% (*n* = 158) | 31.9% (*n* = 94) | 0.98 |
| REM %: above mean | 57.4% | 64.3% | 0.47 |
| Tonic REM % (MTL) | 50.2 ± 28.1 (*n* = 60) | 56.3 ± 31.6 (*n* = 39) | 0.33 |
| Phasic REM % (MTL) | 29.8 ± 19.9 (*n* = 42) | 35.8 ± 16.6 (*n* = 34) | 0.16 |
| % Any (SINBAR) | 66.4 ± 19.9 (*n* = 13) | 61.2 ± 26.0 (*n* = 5) | 0.70 |
| Constipation | 56.8% (*n* = 111) | 57.5% (*n* = 80) | 0.92 |
| Urinary dysfunction | 29.4% (*n* = 85) | 39.6% (*n* = 53) | 0.22 |
| Erectile dysfunction | 52.8% (*n* = 36) | 75.0% (*n* = 28) | 0.069 |
| Orthostatic symptoms | 28.4% (*n* = 67) | 39.1% (*n* = 46) | 0.23 |
| Systolic blood pressure drop | 12.7 ± 15.7 (*n* = 44) | 17.0 ± 21.9 (*n* = 37) | 0.32 |
| Abnormal office: cognitive test (regardless of complaint) | 43.2% | 65.2% | 0.003 |
| MoCA | 25.8 ± 2.6 (*n* = 49) | 22.6 ± 3.5 (*n* = 30) | <0.001 |
| MMSE | 27.8 ± 1.7 (*n* = 57) | 26.4 ± 3.3 (*n* = 70) | 0.002 |
| Neuropsychological abnormal (regardless of complaint) | 29.8% (*n* = 57) | 86.8% (*n* = 76) | <0.001 |
| Mild cognitive impairment | | | |
| Neuropsychological testing | 25.9% (*n* = 54) | 84.1% (*n* = 63) | <0.001 |
| MoCA/MMSE | 30.1% (*n* = 73) | 56.9% (*n* = 72) | 0.001 |
| Depression | 28.6% (*n* = 119) | 32.6% (*n* = 92) | 0.53 |
| Anxiety | 22.5% (*n* = 71) | 28.2% (*n* = 39) | 0.52 |
| Substantia nigra ultrasound | 60.0% (*n* = 10) | 66.7% (*n* = 3) | 0.84 |
| DAT scan (putamen) abnormal | 70.3% (*n* = 37) | 71.4% (*n* = 14) | 0.94 |

*P*-values are calculated with student t-test for continuous variables and $\chi^2$ test for categorical variables. Note that seven patients from Barcelona who converted to MCI but not yet to parkinsonism or dementia are not included in this analysis.
AHI = Apnoea–Hypopnea Index; MoCA = Montreal Cognitive Assessment; MMSE = Mini Mental State Examination; MTL = Montreal; SINBAR = Sleep Innsbruck Barcelona.

On the other hand, other stratification strategies allowed more inclusions; selecting for abnormal olfaction allowed 67% eligibility with sample size of 247–262 per group, and selecting those who met the MDS prodromal criteria allowed 77% eligibility with sample size of 282–301 per group. Among two-factor combinations, the combination of olfaction and UPDRS retained 29% eligibility, and resulted in an estimated 15.7% annual conversion rate, translating into 157 patients per group.

# Discussion

In this large multicentre study, we have confirmed the very high risk of Parkinson's disease, dementia with Lewy bodies, and MSA in 'idiopathic' RBD, and have confirmed numerous predictors of outcome. These findings have implications for potential prevention/early treatment of the neurodegenerative synucleinopathies.

## Risk of disease

As this is the largest study ever performed in iRBD, it has potentially the most precise estimates of phenoconversion rates. Overall, we found phenoconversion rates of 6.25% per year. This is broadly similar, although slightly lower than some previous estimates, including that of the only previous multicentre study (which found an 8% annual conversion; Postuma *et al.*, 2015*d*). The reason for this

**Table 4 Sample size calculations for neuroprotective trials**

| Population | Proportion of sample abnormal, % | Observed conversion rate, % | Adjusted conversion rate (adjusted for centre), % | Sample size per group - observed/ adjusted |
|---|---|---|---|---|
| All RBD | 100 | 6.25 | 6.25 | 366/366 |
| Age at least 55 | 92 | 6.32 | 6.32 | 363/363 |
| UPDRS III (combined 1987 and MDS) | 38 | 12.7 | 11.1 | 190/214 |
| Quantitative motor test (majority abnormal) | 34 | 14.7 | 12.2 | 166/197 |
| Olfaction | 67 | 9.52 | 8.93 | 247/262 |
| Colour vision | 38 | 11.9 | 8.47 | 201/275 |
| MCI (office-based) | 23 | 13.1 | 9.09 | 184/258 |
| MCI (neuropsychology) | 28 | 16.3 | 11.4 | 152/210 |
| DAT scan | 44 | 11.5 | 10.9 | 208/219 |
| Constipation | 56 | 8.33 | 8.07 | 279/288 |
| Either elevated UPDRS or MCI on neuropsychology | 53 | 10.2 | 10.2 | 232/232 |
| Elevated UPDRS and MCI (neuropsych. only) | 13 | 17.5 | 14.8 | 143/166 |
| Elevated UPDRS and MCI any[a] | 14 | 16.3 | 14.3 | 152/171 |
| Either elevated UPDRS or MCI any[a] | 55 | 11.8 | 10.3 | 203/230 |
| UPDRS and olfaction abnormal | 29 | 15.7 | 15.7 | 157/157 |
| UPDRS and constipation | 18 | 15.4 | 14.3 | 160/171 |
| Olfaction and constipation | 29 | 10.5 | 9.15 | 226/257 |
| Olfaction and MCI[a] | 14 | 15.7 | 13.3 | 157/183 |
| Olfaction and either UPDRS or MCI[a] | 39 | 14.7 | 12.4 | 166/195 |
| Meets MDS Prodromal Criteria | 77 | 8.24 | 7.69 | 282/301 |

Sample size is calculated using a time-to-event analysis for a disease phenoconversion to either dementia or parkinsonism as the primary outcome. The calculation is for a 2-year trial, with accrual set at 0 (i.e. all patients are followed for exactly 2 years). The assumption is for a disease-modifying agent that reduces HR to 0.5 (65 phenoconversion events), with power = 80% at $<0.05$ (two-tailed). For stratification, patients are included if they are abnormal for that test (or combination of tests). Note that the observed rate includes only those centres that performed the evaluation, at the time that the marker was first evaluated. The adjusted rate was calculated by dividing the overall observed rate in all centres that performed the marker (both normal and abnormal tests) by the median rate in all centres (thereby estimating the rate that would have been seen if all centres performed the test). Although the observed rate is not adjusted for centre effects, it may better reflect experience in clinical trials, in which follow-up is performed more intensively (see 'Discussion' section).

[a]In these cases, MCI can be defined as either an abnormal neuropsychological test or office-based test, plus cognitive complaint. If both were performed and contradict, the neuropsychological test result takes precedence.

slightly lower estimate is unclear. One explanation could have been secular change; as a disease becomes increasingly recognized, milder/earlier cases (with lower conversion rates) come to attention. However, we found no clear evidence for this; those diagnosed after 2010 had a 19.2% 3-year risk of disease, compared to 16.9% among those before. It could also be possible that newer centres in the RBD Study Group might have different (i.e. more permissive) diagnostic procedures, which would imply increasing proportions of patients without true synucleinopathy. However, centres who participated in the original multicentre study did not have a higher risk than those without (e.g. original centres' 3-year risk = 16.8%, versus 21.6% for new centres). We did note that annualized disease risk appeared to be lower from Years 0 to 2 than for subsequent years. This may indicate a potential selection bias; if examiners were reluctant to recruit patients who appeared on the threshold of parkinsonism or dementia, risk would be systematically underestimated (since patients would have to first develop mild signs, then full disease).

Another potentially key factor for phenoconversion may be the frequency and intensity of follow-up. Many patients do not recognize symptoms of parkinsonism/cognitive impairment, and are diagnosed only on in-person systematic examination. A striking illustration of the importance of follow-up intensity is the Montreal experience. In their 2009 report, which included patients followed clinically/*ad hoc*, conversion risk at 5 years was 18% (Postuma *et al.*, 2009). However, 6 years later, a study from the same centre, this time concentrating exclusively upon patients followed systematically by a movement disorders specialist and neuropsychologist, found a 5-year risk of 47% (Postuma *et al.*, 2015*c*). Moreover, we may see evidence of this in our cohort, as conversion estimates were higher when they were calculated starting from the first date of intensive in-person examination of Parkinson's disease/dementia risk factors (olfaction, UPDRS, cognitive exam, etc). For example, if conversion risk is tracked from performance of the first UPDRS Part III neurologist examination (a potential sign that more intensive follow-up has commenced), the estimated annual risk of conversion rises from 6.3% to 7.1%; see Table 4 for the potential effects of this on observed versus estimated sample size calculations. This might imply that a clinical trial with intensive

periodic evaluations may find a higher conversion risk than observed in this study. Regardless of the conversion rate, it is clear that the large majority of idiopathic RBD patients in fact have prodromal synucleinopathy. So, while the term 'idiopathic' RBD is used here, we recognize that few patients are truly 'idiopathic' in the original sense of the term (i.e. unclear cause), and other terms such as 'clinically isolated' RBD may be more appropriate (Hogl *et al.*, 2018).

## Predictive markers

Although comparisons of hazard ratios across different predictors should be made with caution (because centres measured different variables), it nonetheless suggests numerous findings of interest. When analysed as a binary diagnostic test, there was no clear advantage of DAT-SPECT over either the UPDRS or quantitative motor testing (note that both DAT-SPECT and quantitative motor tests were defined by each centre as normal/abnormal with no harmonization procedures; harmonization might increase the hazard ratio). Note that this finding may be unique to iRBD patients, who have an extremely high prevalence of underlying synucleinopathy; in the general population, non-specific causes of motor slowing on quantitative motor tests (e.g. arthritis) may influence estimates more (Keezer *et al.*, 2016; Jennings *et al.*, 2017). Regardless, these quantitative motor tests were simple office-based tests that required $<5$ min to administer. Clearly these are strong candidates for selecting patients for future neuroprotective trials, and could even obviate the need for sophisticated imaging techniques if simpler trial design is required. This finding illustrates both the need to improve imaging techniques for prodromal disease and the considerable future potential for more precise quantitative motor markers (e.g. wearable or smartphone-based sensors).

It is not surprising that the highest hazard ratios were for motor and cognitive measures, since these are the primary means by which parkinsonism and dementia are defined; however, the high performance of olfactory testing as a predictor is notable, as it is also easily tested in office settings. Finally, no test appeared to be able to 'rule out' phenoconversion; many of those with normal testing still went on to develop parkinsonism and dementia. For example, the highest negative predictive value was seen for the MDS prodromal criteria, but even among those negative for criteria, 5% phenoconverted at 3 years, 13% at 5 years, and 27% at 8 years (note that analysis is at baseline only, and presumably many of these patients would have developed abnormal markers before phenoconversion).

## Dementia-first versus parkinsonism-first

The comparison between dementia-first and parkinsonism-first phenoconvertors was notable for the similarity in predictive value between markers. Motor variables were highly predictive of dementia as well as parkinsonism (and for quantitative motor assessment, even more predictive of dementia than parkinsonism). This finding is consistent with previous studies which documented a longer/slower-progressing motor prodromal interval in dementia-first than parkinsonism-first convertors (Postuma *et al.*, 2012); if their motor prodromal interval is longer in prodromal dementia patients, they would be more likely to be abnormal on a cross-sectional test. Overall, the only clear differentiating variable between dementia and parkinsonism was cognition itself. It is unclear whether the conversion to dementia versus parkinsonism first is related to a different 'top-down' synuclein spread upwards to cortex before the substantia nigra (Adler and Beach, 2016), or to effects of co-morbid pathology [i.e. if a person with RBD has co-morbid amyloid cortical pathology, even modest cortical deposition of synuclein could trigger rapid cortical neurodegeneration resulting in a dementia-first phenotype (Chetelat *et al.*, 2013)].

## Sample size

We calculated the sample size requirements for a definitive neuroprotective trial, using phenoconversion as a categorical endpoint. Overall, sample sizes for a 2-year trial with HR = 0.5 ranged from 150 to 360 patients per group. In general, stratification strategies could decrease sample sizes, at the cost of reduced generalizability and less efficient recruitment. Of the selection strategies, the two most efficient appeared to be olfaction, which reduced sample size by 28.5% while retaining 67% of the sample as potential trial candidates, and the MDS prodromal criteria, which reduced sample size by 17.8% while retaining 77% of the sample. Of course, exact sample size calculations will depend on the specifics of a clinical trial; nevertheless, the fact that 24 centres combined to produce these estimates can provide some confidence for trial planners that sample sizes will be representative of the global experience. Notably, the total sample size for a future neuroprotective trial is less than the number of participants who were recruited to this study. So, it appears that a complete trial-ready population already exists in the centres of the International RBD Study Group.

## Limitations and strengths

Some limitations of this study should be pointed out. First, this study is an amalgam of the research experience of 24 different centres; there was not a single protocol for testing predictors of disease, and protocols differed greatly between centres in terms of depth, follow-up intensity, predictors assessed, and methods/cut-offs for assessing them. Therefore, the predictive data will not be fully comparable to a single clinical trial setting, which would have a single testing protocol. Second, protocols for recruiting MCI varied; 23 of 24 centres recruited patients at baseline with MCI, but the largest centre (Barcelona) did not. There is no perfect way to harmonize these completely; for the primary analysis we elected to allow the Barcelona group to define disease conversion as *de novo* MCI, to prevent underestimation of

disease risk (i.e. if patients with MCI were systematically excluded at baseline, then patients developing dementia would have to cycle from normal cognition through MCI to dementia, artificially prolonging disease-free time). However, if conversion from normal cognition to MCI were faster than from MCI to dementia/parkinsonism, disease risk might be overestimated. Regardless, excluding Barcelona data had almost no effect on risk estimates (median conversion time = 8.01 years with and 8.00 without). Third, hazard ratio comparisons between the different markers should be made with caution, as different centres (with potentially different conversion rates) tested different markers using different techniques (note that the results are adjusted for centre, which helps mitigate centre effects). Fourth, the amplitude of the hazard ratio observed in this study should not be extrapolated to the general population. When using RBD patients, the baseline risk of disease is so high that ceiling effects on hazard ratios occur [for illustration of this effect, see supplemental methods of Berg *et al.* (2015)]. Similarly, the effect of very long latency prodromal markers (e.g. autonomic dysfunction, olfaction, substantia nigra ultrasound) may be masked by floor effects; if a marker preceded RBD in almost all cases, and almost all RBD patients have prodromal synucleinopathy, there would be little apparent predictive value of that marker in this population. Fifth, RBD in Parkinson's disease marks a 'diffuse-malignant' subtype of Parkinson's disease (Fereshtehnejad *et al.*, 2015), implying that our hazard ratio findings will not completely generalize to those Parkinson's disease and dementia with Lewy bodies cases who have no RBD. Sixth, note that our markers were tested at baseline only; repeated marker testing would allow assessment of evolution of prodromal markers over time. Seventh, although sample size is large in this trial, some markers were assessed by only a few centres, and so their corresponding confidence intervals can be wide. Eighth, the final neurodegenerative disease diagnosis of all patients in this study was clinical, according to best impression of the treating neurologist; it is likely that some patients diagnosed with Parkinson's disease will eventually be discovered to have MSA, and vice versa. Finally, the number of patients with very long duration follow-up remains limited (e.g. 28 still-disease-free patients have been followed for $>12$ years); therefore, we cannot determine whether disease risk changes over very long disease durations.

In conclusion, we confirmed a high risk of phenoconversion to overt neurodegenerative disease in RBD and found numerous predictors of phenoconversion. As new disease-modifying treatments are being developed for neurodegenerative synucleinopathies, RBD patients are ideal candidates for neuroprotective trials.

# Funding

This study was funded by a grant from the Canadian Institute of Health Research (Grant #286641) the Weston Foundation (grant #N/A) and the Fonds de la Recherche en Sante Quebec (#28915). The funding sources had no role in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication

# Competing interests

The authors report no competing interests.

# Supplementary material

Supplementary material is available at *Brain* online.

# References

Adler CH, Beach TG. Neuropathological basis of nonmotor manifestations of Parkinson's disease. Mov Disord 2016; 31: 1114–9.

American Academy of Sleep Medicine TFC, Hauri PJC. The international classification of sleep disorders: diagnostic and coding manual. 2nd edn. Westchester, IL: American Academy of Sleep Medicine; 2007.

Arnulf I, Neutel D, Herlin B, Golmard JL, Leu-Semenescu S, Cochen de Cock V, et al. Sleepiness in idiopathic REM sleep behavior disorder and Parkinson disease. Sleep 2015; 38: 1529–35.

Barber TR, Lawton M, Rolinski M, Evetts S, Baig F, Ruffmann C, et al. Prodromal Parkinsonism and neurodegenerative risk stratification in REM sleep behaviour disorder. Sleep 2017; 40. doi: 10.1093/sleep/zsx071.

Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001; 2: 297–307.

Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988; 56: 893–7.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961; 4: 561–71.

Berg D, Postuma RB, Adler CH, Bloem BR, Chan P, Dubois B, et al. MDS research criteria for prodromal Parkinson's disease. Mov Disord 2015; 30: 1600–11.

Bertrand JA, Bedetti C, Postuma RB, Monchi O, Genier Marchand D, Jubault T, et al. Color discrimination deficits in Parkinson's disease are related to cognitive impairment and white-matter alterations. Mov Disord 2012; 27: 1781–8.

Chaudhuri KR, Martinez-Martin P, Schapira AH, Stocchi F, Sethi K, Odin P, et al. International multicenter pilot study of the first comprehensive self-completed nonmotor symptoms questionnaire for Parkinson's disease: the NMSQuest study. Mov Disord 2006; 21: 916–23.

Chetelat G, La Joie R, Villain N, Perrotin A, de La Sayette V, Eustache F, et al. Amyloid imaging in cognitively normal individuals, at-risk populations and preclinical Alzheimer's disease. Neuroimage Clin 2013; 2: 356–65.

Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994; 44: 2308–14.

Desrosiers J, Hebert R, Bravo G, Dutil E. The Purdue Pegboard Test: normative data for people aged 60 and over. Disabil Rehabil 1995; 17: 217–24.

Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. Physiol Behav 1984; 32: 489–502.

Dubois B, Burn D, Goetz C, Aarsland D, Brown RG, Broe GA, et al. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. Mov Disord 2007; 22: 2314–24.

Fahn S, Elton R; Members of the UDC. The Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Calne D, Goldstein M, editors. Recent developments in Parkinson's disease. Florham Park, NJ: MacMillan HealthCare Information; 1987. p. 153–63.

Farnsworth D. The Farnsworth 100-hue test and dichotomous tests for color vision. J Opt Soc Am 1943; 33: 568–78.

Fereshtehnejad SM, Romenets SR, Anang JB, Latreille V, Gagnon JF, Postuma RB. New clinical subtypes of Parkinson disease and their longitudinal progression: a prospective cohort comparison with other phenotypes. JAMA Neurol 2015; 72: 863–73.

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: 189–98.

Frauscher B, Iranzo A, Gaig C, Gschliesser V, Guaita M, Raffelseder V, et al. Normative EMG values during REM sleep for the diagnosis of REM sleep behavior disorder. Sleep 2012; 35: 835–47.

Gaudry E, Vagg P, Spielberger CD. Validation of the state-trait distinction in anxiety research. Multivariate Behav Res 1975; 10: 331–41.

Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008; 23: 2129–70.

Goldman JG, Postuma R. Premotor and nonmotor features of Parkinson's disease. Curr Opin Neurol 2014; 27: 434–41.

Haba-Rubio J, Frauscher B, Marques-Vidal P, Toriel J, Tobback N, Andries D, et al. Prevalence and determinants of REM sleep behavior disorder in the general population. Sleep 2018; 41. doi: 10.1093/sleep/zsx197.

Higgins PD, Johanson JF. Epidemiology of constipation in North America: a systematic review. Am J Gastroenterol 2004; 99: 750–9.

Hogl B, Stefani A, Videnovic A. Idiopathic REM sleep behaviour disorder and neurodegeneration—an update. Nat Rev Neurol 2018; 14: 40–55.

Hummel T, Sekinger B, Wolf SR, Pauli E, Kobal G. 'Sniffin' Sticks': olfactory performance assessed by the combined testing of odour identification, odor discrimination and olfactory threshold. Chem Senses 1997; 22: 39–52.

Iranzo A, Fernandez-Arcos A, Tolosa E, Serradell M, Molinuevo JL, Valldeoriola F, et al. Neurodegenerative disorder risk in idiopathic REM sleep behavior disorder: study in 174 patients. PLoS One 2014; 9: e89741.

Jennings D, Siderowf A, Stern M, Seibyl J, Eberly S, Oakes D, et al. Conversion to Parkinson's disease in the PARS Prodromal Cohort. JAMA Neurol 2017; 74: 933–40.

Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991; 14: 540–5.

Kang SH, Yoon IY, Lee SD, Han JW, Kim TH, Kim KW. REM sleep behavior disorder in the Korean elderly population: prevalence and clinical characteristics. Sleep 2013; 36: 1147–52.

Keezer MR, Wolfson C, Postuma RB. Age, gender, comorbidity, and the MDS-UPDRS: results from a population-based study. Neuroepidemiology 2016; 46: 222–7.

Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001; 16: 606–13.

Li Y, Kang WK, Yang Q, Zhang L, Zhang L, Dong F, et al. Predictive markers for early conversion of IRBD to neurodegenerative synucleinopathy diseases. Neurology 2017; 88: 1493–500.

Mahlknecht P, Iranzo A, Hogl B, Frauscher B, Muller C, Santamaria J, et al. Olfactory dysfunction predicts early transition to a Lewy body disease in idiopathic RBD. Neurology 2015; 84: 654–8.

McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor JP, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies: fourth consensus report of the dementia with Lewy bodies Consortium. Neurology 2017; 89: 88–100.

Montplaisir J, Gagnon JF, Fantini ML, Postuma RB, Dauvilliers Y, Desautels A, et al. Polysomnographic diagnosis of idiopathic REM sleep behavior disorder. Mov Disord 2010; 25: 2044–51.

Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005; 53: 695–9.

Nutt JG, Lea ES, Van HL, Schuff RA, Sexton GJ. Determinants of tapping speed in normal control subjects and subjects with Parkinson's disease: differing effects of brief and continued practice. Mov Disord 2000; 15: 843–9.

Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39: 142–8.

Postuma RB, Adler CH, Dugger BN, Hentz JG, Shill HA, Driver-Dunckley E, et al. REM sleep behavior disorder and neuropathology in Parkinson's disease. Mov Disord 2015a; 30: 1413–7.

Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord 2015b; 30: 1591–600.

Postuma RB, Gagnon JF, Bertrand JA, Genier Marchand D, Montplaisir JY. Parkinson risk in idiopathic REM sleep behavior disorder: preparing for neuroprotective trials. Neurology 2015c; 84: 1104–13.

Postuma RB, Gagnon JF, Vendette M, Desjardins C, Montplaisir J. Olfaction and color vision identify impending neurodegeneration in REM behavior disorder. Ann Neurol 2011; 69: 811–8.

Postuma RB, Gagnon JF, Vendette M, Fantini ML, Massicotte-Marquez J, Montplaisir J. Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder. Neurology 2009; 72: 1296–300.

Postuma RB, Iranzo A, Hogl B, Arnulf I, Ferini-Strambi L, Manni R, et al. Risk factors for neurodegeneration in idiopathic rapid eye movement sleep behavior disorder: a multicenter study. Ann Neurol 2015d; 77: 830–9.

Postuma RB, Lang AE, Gagnon JF, Pelletier A, Montplaisir JY. How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. Brain 2012; 135(Pt 6): 1860–70.

Pujol M, Pujol J, Alonso T, Fuentes A, Pallerola M, Freixenet J, et al. Idiopathic REM sleep behavior disorder in the elderly Spanish community: a primary care center study with a two-stage design using video-polysomnography. Sleep Med 2017; 40: 116–21.

Schenck CH, Boeve BF, Mahowald MW. Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. Sleep Med 2013a; 14: 744–8.

Schenck CH, Montplaisir JY, Frauscher B, Hogl B, Gagnon JF, Postuma R, et al. Rapid eye movement sleep behavior disorder: devising controlled active treatment studies for symptomatic and neuroprotective therapy—a consensus statement from the International Rapid Eye Movement Sleep Behavior Disorder Study Group. Sleep Med 2013b; 14: 795–806.

Snaith RP, Bridge GW, Hamilton M. The Leeds scales for the self-assessment of anxiety and depression. Br J Psychiatry 1976; 128: 156–65.

Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. J Psychosom Res 2000; 48: 555–60.

Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006; 166: 1092–7.

Valencia Garcia S, Brischoux F, Clement O, Libourel PA, Arthaud S, Lazarus M, et al. Ventromedial medulla inhibitory neuron inactivation induces REM sleep without atonia and REM sleep behavior disorder. Nat Commun 2018; 9: 504.

Visser M, Marinus J, Stiggelbout AM, van Hilten JJ. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. Mov Disord 2004; 19: 1306–12.

Wenning GK, Tison F, Seppi K, Sampaio C, Diem A, Yekhlef F, et al. Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS). Mov Disord 2004; 19: 1391–402.

Wing YK, Li SX, Mok V, Lam SP, Tsoh J, Chan A, et al. Prospective outcome of rapid eye movement sleep behaviour disorder: psychiatric disorders as a potential early marker of Parkinson's disease. J Neurol Neurosurg Psychiatry 2012; 83: 470–2.

Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. Journal of psychiatric research 1982; 17: 37–49.

Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67: 361–70.