

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Bensimon, G;Ludolph, A;Agid, Y;Vidailhet, M;Payan, C;Leigh, PN;Dib, M;Lacomblez, L;Burn, D;Landwehrmeyer, GB;Asselain, B;Allain, H;Chadwick, D;Perret, JE;Warlow, C;Ben-Shlomo, Y;Bonnet, AM;Fermanian, J;Verny, M;Moore, P;Tranchant, C;Borg, M;McCrone, P;Durif, F;Evans, A;Viallet, F;Verin, M;Deasy, N;Jarosz, J;Hauser, TK;Kraft, E;Broussolle, E;Dormont, D;Marsault, C;Tourbah, A;Defebvre, L;Delmaire, L;Roland, Y;Hauw, JJ;Duyckaerts, C;Seilhean, D;Al-Sarraj, S;Revesz, T;Kretzschmar, HA;Brown, R;Bak, T;Danek, A;Dubois, B;Marié, RM;Uttner, I;Dürr, A;Al-Chalabi, A;Wood, N;Brice, A;Camu, W;Morrison, K;Graf, M;Paillasseur, P;Pham, HP;Dedise, N;Hermine, C;Sagnes, S;Poître, B;Foucart, C;Dougherty, A;Murphy, C;Mason, H;Hermann, T;Klempp, K;Niess, A;Stange, V;Ouslimani, A;Lehmann, B;Tibi, A;Couratier, C;Arguillère, S;Payan-Cassin, H;Vassault, GM;Henon, P;Gimeno, S;Delfini, M;Dubas, F;Fressinaud, C;Tanguy, JY;Legall, D;Rumbach, L;Vidry, E;Kraehenbuhl, J;Bonneville, JF;Chopard, G;LeDoze, F;Defer, G;Viader, F;Huet, H;Daniel, F;Lalevée, C;Debilly, B;Derost, P

Title:

Riluzole treatment, survival and diagnostic criteria in Parkinson plus disorders: The NNIPPS Study

Date:

2009-01-01

Citation:

Bensimon, G., Ludolph, A., Agid, Y., Vidailhet, M., Payan, C., Leigh, P. N., Dib, M., Lacomblez, L., Burn, D., Landwehrmeyer, G. B., Asselain, B., Allain, H., Chadwick, D., Perret, J. E., Warlow, C., Ben-Shlomo, Y., Bonnet, A. M., Fermanian, J., Verny, M. Derost, P. (2009). Riluzole treatment, survival and diagnostic criteria in Parkinson plus disorders: The NNIPPS Study. *Brain*, 132 (1), pp.156-171. <https://doi.org/10.1093/brain/awn291>.

Persistent Link:

<https://hdl.handle.net/11343/256961>

License:

CC BY-NC

Riluzole treatment, survival and diagnostic criteria in Parkinson plus disorders: The NNIPPS Study

Gilbert Bensimon,¹ Albert Ludolph,² Yves Agid,³ Marie Vidailhet,³ Christine Payan,¹
P. Nigel Leigh⁴ and for the NNIPPS Study Group*

1 Département de Pharmacologie Clinique, Hôpital de la Pitié-Salpêtrière, Assistance Publique Hôpitaux de Paris & UPMC University Paris 06, UMR 7087, Paris, France

2 Department of Neurology, University of Ulm, Oberer Eselsberg, Ulm, Germany

3 Fédération de Neurologie, Hôpital de la Pitié-Salpêtrière, Assistance Publique Hôpitaux de Paris & UPMC University Paris 06, Paris, France

4 MRC Centre for Neurodegeneration Research, King's College London, Institute of Psychiatry, PO 41, Department of Clinical Neuroscience, London, UK

*See appendix for the details on NNIPPS Study Group

Correspondence to: Prof. P. Nigel Leigh,
MRC Centre for Neurodegeneration Research,
King's College London, PO 41, Institute of Psychiatry,
Camberwell, London SE5 8AF, UK
E-mail: n.leigh@iop.kcl.ac.uk

Parkinson plus diseases, comprising mainly progressive supranuclear palsy (PSP) and multiple system atrophy (MSA) are rare neurodegenerative conditions. We designed a double-blind randomized placebo-controlled trial of riluzole as a potential disease-modifying agent in Parkinson plus disorders (NNIPPS: Neuroprotection and Natural History in Parkinson Plus Syndromes). We analysed the accuracy of our clinical diagnostic criteria, and studied prognostic factors for survival. Patients with an akinetic-rigid syndrome diagnosed as having PSP or MSA according to modified consensus diagnostic criteria were considered for inclusion. The psychometric validity (convergent and predictive) of the NNIPPS diagnostic criteria were tested prospectively by clinical and pathological assessments. The study was powered to detect a 40% decrease in relative risk of death within PSP or MSA strata. Patients were randomized to riluzole or matched placebo daily and followed up to 36 months. The primary endpoint was survival. Secondary efficacy outcomes were rates of disease progression assessed by functional measures. A total of 767 patients were randomized and 760 qualified for the Intent to Treat (ITT) analysis, stratified at entry as PSP (362 patients) or MSA (398 patients). Median follow-up was 1095 days (range 249–1095). During the study, 342 patients died and 112 brains were examined for pathology. NNIPPS diagnostic criteria showed for both PSP and MSA excellent convergent validity with the investigators' assessment of diagnostic probability (point-biserial correlation: MSA $r_{pb}=0.93$, $P < 0.0001$; PSP, $r_{pb} = 0.95$, $P < 0.0001$), and excellent predictive validity against histopathology [sensitivity and specificity (95% CI) for PSP 0.95 (0.88–0.98) and 0.84 (0.77–0.87); and for MSA 0.96 (0.88–0.99) and 0.91 (0.86–0.93)]. There was no evidence of a drug effect on survival in the PSP or MSA strata (3 year Kaplan–Meier estimates PSP-riluzole: 0.51, PSP-placebo: 0.50; MSA-riluzole: 0.53, MSA-placebo: 0.58; $P=0.66$ and $P=0.48$ by the log-rank test, respectively), or in the population as a whole ($P=0.42$, by the stratified-log-rank test). Likewise, rate of progression was similar in both treatment groups. There were no unexpected adverse effects of riluzole, and no significant safety concerns. Riluzole did not have a significant effect on survival or rate of functional deterioration in PSP or MSA, although the study reached over 80% power to detect the hypothesized drug effect within strata. The NNIPPS diagnostic criteria were consistent and valid. They can be used to distinguish between PSP and MSA with high accuracy, and should facilitate research into these conditions relatively early in their evolution.

Keywords: progressive supranuclear palsy; multiple system atrophy; randomized controlled trial; riluzole; natural history

Abbreviations: ADL = activity of daily living; ALS = amyotrophic lateral sclerosis; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CBD = corticobasal degeneration; CGI-ds = clinical global impression for disease severity; CGI-dysautonomia = clinical global impression for autonomic dysfunction; GCP = good clinical practice; IDMSC = independent data monitoring and safety committee; IPD = idiopathic Parkinson's disease; IRB = institutional review board; ITT = Intent-to-treat; MMSE = Mini-Mental State Examination; MSA-C = multiple system atrophy, Cerebellar form; MSA-P = multiple system atrophy, parkinsonian form; NINDS-SPSP = National Institute of Neurological Disorders and the Society for Progressive Supranuclear Palsy; PPPT = per protocol per treatment; SAE = serious adverse event; SEADL = Schwab and England activities of daily living scale; SMDS = Short Motor Disability Scale; VAS = Visual Analogue Scale

Introduction

Progressive supranuclear palsy (PSP) and multiple system atrophy (MSA) are disabling and fatal neurodegenerative disorders for which no disease-modifying treatment is available. For the majority of patients with PSP and MSA, many of whom present with an atypical Parkinsonian or akinetic-rigid syndrome ('Parkinson plus disorder') the course is one of relentless progression, increasing disability and death with a median survival of 5–10 years from onset of symptoms (Litvan *et al.*, 1996a; Testa *et al.*, 1996, 2001; Ben-Shlomo *et al.*, 1997; Schrag *et al.*, 1999, 2008; Litvan, 2003; Golbe and Ohman-Strickland, 2007).

PSP and MSA have similar prevalence rates estimated at 2–7 per 100 000 person years (Golbe *et al.*, 1988; Ben-Shlomo *et al.*, 1997; Bower *et al.*, 1997; Schrag *et al.*, 1999; Nath and Burn, 2000; Nath *et al.*, 2001; Vanacore *et al.*, 2001a, b; Watanabe *et al.*, 2002). These are probably underestimates, however, because current diagnostic criteria are based on retrospective clinicopathological studies (Litvan *et al.*, 1996a, b, d, 2003; Litvan, 2003) and both delayed diagnosis and mis-diagnosis are common. Although published consensus diagnostic criteria are highly specific they are relatively insensitive and a definite diagnosis can only be made through histopathology (Litvan *et al.*, 2003). Since it is likely that neuroprotective strategies are best tested at a relatively early stage of disease, more sensitive diagnostic criteria are required for trials of potential disease-modifying agents. Although PSP and MSA often present as akinetic-rigid syndromes, each has distinctive pathological features. In MSA, a key feature is glial cytoplasmic inclusions with accumulation of α -synuclein in oligodendrocytes and neurons (Papp *et al.*, 1989; Lantos and Papp, 1994; Spillantini and Goedert, 2000) whilst in PSP the hallmark is accumulation of abnormally phosphorylated microtubule-associated tau protein (τ) in neurons and glia (Hauw *et al.*, 1994; Dickson *et al.*, 2007). Although the pathogenic mechanisms underlying MSA and PSP are unknown, there is evidence that glutamate toxicity may contribute to neuronal damage in these and other neurodegenerative diseases (Albin and Greenamyre, 1992; Albers and Augood, 2001; Mattson, 2003; Przedborski, 2005). The benzothiazole drug riluzole has a number of pharmacological effects that contribute to neuroprotection in experimental paradigms of neurodegenerative diseases including anti-excitotoxic activity, blocking of voltage dependent sodium-channels, free-radical scavenging, anti-apoptotic and neurotrophic effects and inhibition of protein aggregation (Doble, 1999; Heiser *et al.*, 2002; Yoo *et al.*, 2005; Caumont *et al.*, 2006; Shortland *et al.*, 2006). In a rodent model of MSA, riluzole improved some measures

of neuronal damage (Diguët *et al.*, 2005; Scherfler *et al.*, 2005). Riluzole (up to 200 mg daily) is well tolerated and prolongs survival in amyotrophic lateral sclerosis (Bensimon *et al.*, 1994; Lacomblez *et al.*, 1996; Miller *et al.*, 2007). Thus far, riluzole remains the only agent shown to modify disease progression in a human neurodegenerative disorder.

In order to test the hypothesis that riluzole may slow disease progression in PSP and MSA, we carried out a phase-III, randomized double-blind placebo controlled trial in 44 centres in France, Germany and the UK. The study design incorporated ancillary objectives including natural history, development and validation of more sensitive diagnostic criteria and functional measures of disease severity and progression, cognition, quality of life, health economics, MRI changes, pathology and the establishment of brain and DNA banks. We describe here the design and main outcomes of the NNIPPS trial in terms of the efficacy and safety of riluzole, the psychometric validity of the NNIPPS diagnostic criteria in relation to clinic and pathology and the major factors influencing prognosis for PSP and MSA.

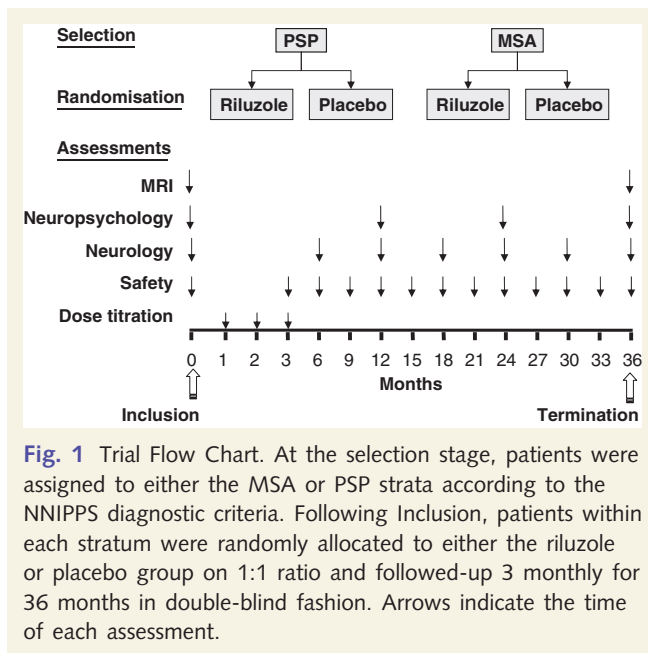
Methods

NNIPPS was designed as a double-blind placebo-controlled, stratified (by diagnosis of MSA or PSP, and by centre), parallel group, European (France, Germany, United Kingdom) trial assessing the efficacy and safety of riluzole at flexible dose (50–200 mg/day) (Fig. 1). The primary objective of NNIPPS was to demonstrate the efficacy of riluzole on survival (primary end-point) and rate of decline in motor function (secondary end-points).

The number of patients required was determined for the primary end-point survival and for each stratum (PSP, MSA). Assumptions included mean disease duration of 3 years prior to entry and no loss to follow-up. Assuming a 41% death rate at 3 years in the placebo group (Litvan *et al.*, 1996a; Testa *et al.*, 1996, 2001; Ben-Shlomo *et al.*, 1997; Litvan, 2003; Golbe and Ohman-Strickland, 2007; Schrag *et al.*, 2007), 400 patients provide over 80% power to detect a 40% decrease in the relative risk of death in the treated group using the log-rank test with two-sided α risk set at 0.05. With both strata combined (800 patients), the power reached 98% to detect a 40% decrease in the risk of death with active treatment compared to placebo, assuming that a minimum of 272 events would be observed over the 36-month trial period.

Patients and treatment

From consensus criteria (Litvan *et al.*, 1996a, d, 2003) we derived simplified operational diagnostic criteria suitable for large-scale clinical



trials, with the aim of maximizing sensitivity for recruitment. Strata were defined according to these operational diagnostic criteria (Table 1). The history of the condition in each patient was evaluated at entry using systematic questionnaires recording the initial (presenting) symptom, and current syndromes (entry and last visit). Response to levodopa therapy was evaluated at entry. Additional assessments related to ancillary studies including a new functional scale (the Parkinson Plus scale), magnetic resonance imaging, neuropsychology, health economics and quality of life, genetic testing and neuropathology, will be presented separately.

Patients were allocated to treatment according to a computer generated randomization list, stratified for diseases (MSA or PSP) and clinical centre, with a riluzole to placebo ratio of 1:1. Riluzole (50 mg) or placebo was prepared as identical tablets (Sanofi-Aventis, Antony-France). Packaging and labelling (LC², Lentilly-France) and treatment management (Cardinale, Corby-UK; Clindata, Weilerswist-Germany; AGEPS-AP HP, Paris-France) was performed so as to safeguard blinding to treatment allocation throughout the trial duration.

Following randomization, a monthly dose-titration over 3 months was used with increasing dosage from one, two and four tablets daily during which time tolerance was assessed with monthly laboratory tests for haematology (full blood count) and liver function (including ALT and AST) and patient's reporting of adverse events. Dose flexibility (one to four tablets per day) according to tolerance was allowed throughout the 36-month study period, with each dose-adjustment recorded. Study treatment withdrawal was not considered as an 'end of study' and the protocol required that patients should be followed for the ITT analysis to the end of the planned double-blind period. Treatments were delivered to patients every 3 months and the tablets returned were counted to evaluate compliance.

Assessments

To assess the psychometric validity (see Statistical analysis section) of the NNIPPS operational diagnostic criteria, investigators were required to assign at entry a diagnostic probability (PSP or MSA) for

each patient using two 100 mm-visual analogue scales (VAS). These clinical diagnostic assessments were completed at entry, every 12 months thereafter, and at the last visit. When possible, clinical diagnosis was compared to neuropathological diagnosis, which was assessed blind to the clinical diagnosis at all stages of processing and analysis of donated brains. The latter were processed according to a standard protocol incorporating formalin fixation of one hemisphere (randomly allocated by the UK, French or German coordinating centres) with freezing (at -80°C) of the other hemisphere for banking. Tissue sections were assessed against standard diagnostic criteria (Lantos and Papp, 1994; Litvan *et al.*, 1996d; Dickson, 1999) with cross-examination and consensus scoring for each case.

The primary criterion of efficacy was defined as survival during the 36-month double-blind period of the study (1095 days included) or until the administrative cut-off date (November 30, 2004 included) whichever came first. The primary end-point was defined as death from any cause. All dates of death were documented with death certificates. For all surviving patients, the date of last contact was also documented.

Secondary end-points included standard assessments used in idiopathic Parkinson's disease completed at entry and 6 monthly, using the Hoehn and Yahr staging scale (Hoehn and Yahr, 1967), the Schwab and England Activities of Daily Living scale (SEADL, Schwab and England, 1969) as well as generic health scale assessments using the Clinical Global Impression for disease severity completed by investigators (CGI-ds; Streiner and Norman, 2003) and a CGI adapted to autonomic function assessment (CGI-dysautonomia). The Mini-Mental State Examination (MMSE; Folstein *et al.*, 1975) was completed by the clinical observer at entry and every 12 months. In addition, a specific functional measure assessing ambulation was developed (Short Motor Disability Scale, SMDS) and performed at entry and 3 monthly (Supplementary Table 1). Anticipated non-serious adverse events of riluzole included dizziness, gastrointestinal symptoms, and fatigue (Lacomblez *et al.*, 1996). The main anticipated serious adverse event (SAE) for riluzole was serious abnormality of liver function tests defined as ALT $>5\times$ the upper limit of normal. Safety was evaluated through clinical examination, vital signs, routine laboratory tests including haematology and transaminases (AST, ALT), concomitant medication and patient reports of adverse events (at entry and 3 monthly). Adverse events, serious and non-serious were coded using MedDRA[®] version 6.0 (MedDRA MSSO, Northrop Grumman Corp., Reston, VA, USA). Weight and electrocardiogram were assessed at entry and at the final visit.

Statistical analysis

An Independent Data Monitoring and Safety Committee (IDMSC) was established for unblinded review of all SAEs during the trial and to advise the Steering Committee at regular intervals regarding continuation of the trial according to predefined stopping rules. No interim analysis for efficacy was planned but four safety analyses were performed throughout the study to ensure that mortality in the treated group was not in excess compared to the placebo group.

The detailed statistical plan was submitted to the French IRB prior to unblinding. The primary analysis was conducted following the ITT principles. The ITT population was defined as all randomized subjects who received at least one dose of study medication and for whom there were no major violations of GCP (ICH-6). Sensitivity analyses population definition is described in Supplementary Text 1.

For the primary analysis, the diagnosis at inclusion was used to define the PSP and MSA strata. In addition, the statistical plan defined

Table 1 NNIPPS Inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
BOTH STRATA	All of the following: -Akinetic-rigid syndrome; -Age at disease onset \geq 30 years; -Disease duration (12 months to 8 years); -Signed informed consent.	Any of the following: -Idiopathic Parkinson's disease; -Evidence of any other neurological disease that could explain signs; -History of repeated strokes with stepwise progression of parkinsonian features; -History of major stroke; -Any history of severe or repeated head injury; -A history of encephalitis; -A history of neuroleptic use for a prolonged period of time or within the past 6 months; -Street-drug related parkinsonism; -Significant other neurological disease on CT-scan/MRI; -Oculogyric crises; -Signs of corticobasal degeneration; -Signs of lewy body disease; -Other life-threatening disease likely to interfere with the main outcome measure; -Any clinically significant laboratory abnormality, with the exception of cholesterol, triglyceride and glucose; -Renal failure (serum creatinine $>$ 300 μ M/l); -Transaminase elevation $>$ 2 time upper limit of normal; -Presence of contra-indicated treatments; -Any previous participation in a therapeutic trial within 3 months prior to entry; -Patient likely to be non-compliant or not easily reached in case of emergency; -Patient under legal guardianship (France only).
PSP	All of the following: -Supranuclear ophthalmoplegia; -Postural instability or falls (within 3 years from disease onset).	Any of the following: -Cerebellar ataxia; -Symptomatic autonomic dysfunction; -Tremor at rest.
MSA	One or more of the following: -Symptomatic autonomic dysfunction; -Cerebellar ataxia; -Postural instability or falls (within 3 years from disease onset); -Pyramidal signs.	Any of the following: -Supranuclear ophthalmoplegia; -Signs of severe dementia.

According to the NNIPPS standard operating procedures, for inclusion into the PSP stratum, supranuclear ophthalmoplegia required 'definite slowness and/or moderate to definite limitation of downward gaze'. For MSA, cerebellar ataxia required a moderate to severe ataxia of trunk and/or limbs. Less marked signs which the investigator nonetheless considered clinically significant were not considered as inclusion criteria but allowed investigators to report the presence or absence of an oculomotor or cerebellar syndrome. The akinetic-rigid syndrome was defined as mild to severe rigidity or slowness of neck or limbs. Significant symptomatic autonomic dysfunction (not treatment induced) was defined as moderate to severe CGI-dysautonomia. A MMSE score of \leq 20 was regarded as evidence of severe dementia. Contraindicated treatments included glutamatergic drugs (e.g. amantadine, lamotrigine, dextrometorphan, gabapentin, glutamate containing drugs), free radical scavengers (selegiline, vit-E/ or C at very high dose) or any drug given to treat the disease and not the symptoms; potentially hepatotoxic drugs (e.g. dantrolene); drugs interacting with riluzole metabolism (CYP1A2 inhibitors or inducers); and ropirinoles (due to decreased levels of the drug induced by riluzole).

sensitivity analyses by diagnosis at the end of the study to allow for the possibility of misdiagnosis.

The *Safety Population* comprised all randomized subjects who received at least one dose of study medication. All analyses were performed using SAS Software version 11. According to the guidelines for the standard for educational and psychological testing (American Psychological Association, 1985) two facets of the psychometric validity of the diagnostic criteria were evaluated, the convergent and the predictive validity. The convergent validity was assessed by the degree of correlation between diagnostic classification according to inclusion criteria and the investigator's assessment of diagnostic probability on the VAS. As inclusion criteria represent a nominal variable with two modalities (MSA, PSP), we used the point biserial coefficient which is the relevant method in this case (point-biserial correlation, r_{pb} ; Nunnally and Bernstein, 1994). Predictive

criterion-related validity of the clinical diagnostic criteria was assessed by calculating the specificity, sensitivity, percent correct classification and positive likelihood (Attia, 2003), using pathological diagnosis as the gold standard.

Descriptive analyses summarized the overall population, and the population sub-divided by treatment group, by strata (PSP versus MSA) and by treatment within strata. Categorical data were summarized by frequency and percentage. The log-linear model was used to compare distributions between factors including treatment, strata, country and all interactions (chi-square tests of partial association) (Bishop *et al.*, 1975). Continuous data were summarized by mean and standard deviation. Between treatment groups comparison at entry was carried out using three-factor variance analyses including treatment, strata, country and all interaction factors (Winer, 1971).

For the primary end-point survival, between groups survival curves were compared using the Mantel-Cox (log-rank) test (Mantel and Haenszel, 1958). Treatment effect was assessed with the stratified log-rank test (p_{α} two-tailed test <0.05). The Cox model (Cox, 1972) including treatment, strata and interaction factors, was used to check for treatment by strata (MSA, PSP) interaction.

The influence of demographic and clinical variables at entry on survival was tested with univariate and multivariate Cox proportional-hazard analysis. Multivariate analysis used an automatic up and down stepwise selection of variables (Allison, 1995).

Secondary end-point variables used for assessing disease progression included the SMDS, the SEADL, the Hoehn and Yahr staging and the CGI-ds. For each patient, repeated measurements were summarized by slope of change over time using linear regression methods (unweighted least square estimate; Wu, 1988). Comparisons of slopes between the treatment groups were performed using three-factor variance analyses, including strata, country, treatment and interaction factors. Serious and non-serious adverse events were compared between treatment groups with the Pearson's chi-square test or Fisher exact test where appropriate.

Ethics

Prior to inclusion, patients gave their informed written consent to participate in the study. Separate consent was obtained both for DNA sampling and for post-mortem brain tissue donation. The protocol and subsequent amendments were approved by Ethics Committees/Institutional Review Boards of each coordinating centre in the three participating countries. The trial was conducted according to International standards of Good Clinical Practice-(ICH guidelines

and the Helsinki Declaration). An internal audit of the study was carried out at the end of the trial before unblinding by an independent auditor (*Qualilab*, Olivet, France) in the nine largest centres (three in France, three in Germany and three in UK) accounting for 43% of the overall trial population, and the primary criteria survival was audited and appropriate documentation certified as 100% complete over the whole study population.

Results

During the study period, April 2000 to December 2004, the Independent Data Monitoring and Safety Committee performed four safety analyses (last review, June 2004) with advice to continue the trial on each occasion.

Study populations

From April 2000 to July 2002, a total of 767 patients (PSP, $n=363$; MSA, $n=404$) were recruited in 44 centres with 19 centres in France ($n=320$), 12 in Germany ($n=243$) and 13 in the UK ($n=204$). Two patients never received treatment, and therefore were excluded from all analyses. For five additional patients, signed informed consent ($n=3$) or code break envelopes ($n=2$) were missing and therefore these patients were excluded from the ITT population (All, $n=760$; PSP, $n=362$; MSA, $n=398$) but remained in the safety population analysis ($n=765$) (Fig. 2). The mean daily dose on treatment in ITT population was 3.6 tablets (180 mg riluzole or placebo), and overall mean compliance was

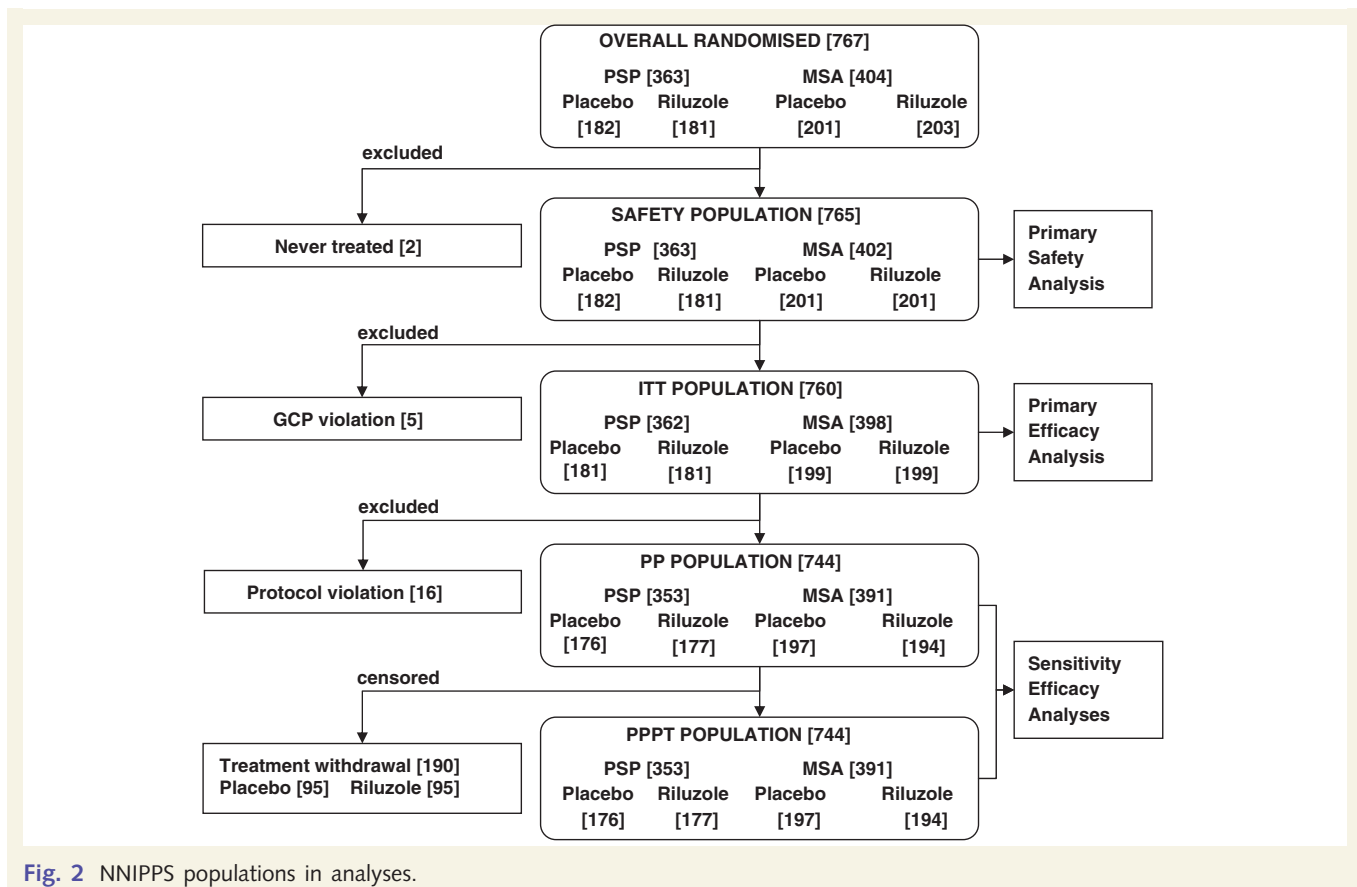


Fig. 2 NNIPPS populations in analyses.

Table 2 ITT population characteristics at entry

	Placebo		Riluzole	
	PSP, <i>n</i> = 181	MSA, <i>n</i> = 199	PSP, <i>n</i> = 181	MSA, <i>n</i> = 199
Sex				
Female (%)	40.3	44.2	44.8	46.2
Age at entry (years)*	67.7 ± 7.0	62.6 ± 8.1	68.0 ± 6.6	61.9 ± 8.5
Age at onset (years)*	63.8 ± 7.0	58.2 ± 8.3	63.9 ± 7.0	57.3 ± 8.6
Disease duration (years)*	3.9 ± 1.9	4.4 ± 2	4.1 ± 1.9	4.5 ± 1.9
Short Motor Disability Scale (0–17)	6.4 ± 3.7	6.1 ± 3.9	6.7 ± 3.6	6.1 ± 3.6
Frontal Assessment Battery (0–18)*	11.1 ± 4.3	14.6 ± 3.2	11.3 ± 4.1	14.3 ± 3.3
Mini-Mental Status Examination (0–30)*	25.4 ± 4.4	27.8 ± 2.3	25.2 ± 4.4	27.6 ± 2.5
Schwab & England Activity Daily Living (0–100)*	50.2 ± 24.5 ^a	53.8 ± 24.8	48.3 ± 23.6	53.3 ± 24.3
Hoehn and Yahr Staging (0–5)*	3.6 ± 1.0	3.4 ± 1.0	3.6 ± 0.9	3.5 ± 1.0
Clinician Global Impression Disease severity (0–6)*	3.6 ± 1.0	3.6 ± 1.0	3.7 ± 1.0	3.6 ± 0.9
Clinician Global Impression Dysautonomia (0–3)*	0.6 ± 0.6	1.8 ± 0.8	0.6 ± 0.6	1.8 ± 0.8

Quantitative variables were analysed using variance analysis and categorical data using a log-linear model. Factors included in models were strata, treatment, country and strata by treatment interactions. Significance was set at $P < 0.05$ (two-tailed test) for each factor or interaction. No differences related to treatment group were detected at entry for the whole population or within strata.

All values are mean ± SD.

^a $n = 180$ due to one missing value.

* Differences between PSP and MSA strata ($P < 0.05$; two-tailed test) were found for all variables except for sex and the Short Motor Disability Scale score.

81.2% ± 31.5. In the riluzole group, the maximum tolerated dose of riluzole for the 284 (75%) patients who did not stop study treatment until death or study cut-off was 200 mg in 237 patients (83.4%), 150 mg in 7 patients (2.5%), 100 mg in 25 patients (8.8%) and 50 mg in 15 patients (5.3%).

The sensitivity analysis population is described in Supplementary Text 1.

Effect of riluzole in patients with PSP and MSA

Treatments were well balanced in the overall population and within strata. There was no significant difference between treatment groups in terms of demographic features, disease severity, previous medical history or concomitant disease at entry in the overall population or within strata (Table 2).

Follow-up at cut-off date was complete and documented in all patients regardless of treatment compliance except for three patients, two of whom withdrew consent in order to be included in another trial and one who underwent medically assisted suicide. At the cut-off date, the surviving patients had a mean (±SD) follow-up time from randomization of 1055 ± 88 days with 81.1% having completed 3 years double-blind follow-up (1095 days). Overall 140 patients (placebo $n = 71$, riluzole $n = 72$) had less than 3 years follow-up at the administrative cut-off date, and three had to be censored at the time they withdrew from the study, as mentioned above. The mean time in study for these 143 patients (18.8%) was 979.4 ± 117.4 days (placebo: 970.5 ± 127.3 days; riluzole: 988.2 ± 106.9 days).

Overall 342 patients (45.0%) died during the double-blind period, with no difference between the PSP and the MSA strata [PSP, $n = 171$ (47.2%); MSA, $n = 171$ (43.0%); $P = 0.21$ by the log-rank test]. With the strata combined, Kaplan–Meier survival

estimates at 36 months for riluzole and placebo groups were 52.6% and 54.9%, respectively. Comparison of survival curves in treatment groups showed no statistically significant difference either in the overall population ($P = 0.42$ by the stratified log-rank test) or within strata (Kaplan–Meier estimates for PSP–riluzole: 0.51, PSP–placebo: 0.50; and for MSA–riluzole: 0.53, MSA–placebo: 0.58; $P = 0.66$ and $P = 0.48$ by the log-rank test, respectively) (Fig. 3). Accordingly there was no statistically significant treatment by strata interaction ($P = 0.85$, by a Cox model analysis).

For secondary efficacy end-points of disease progression, 651 patients (86%) had at least two assessments allowing calculation of a slope of change in scores for Hoehn and Yahr Staging, SEADL and CGI for severity. For SMDS 714 (94%) patients had usable data for calculating slope of change. All scales showed high sensitivity to change with time ($P < 0.0001$). Strata differences were evident with the SEADL, CGI for severity and SMDS ($P = 0.003$, $P = 0.0003$, $P = 0.03$, respectively; Table 3), with the PSP group showing more rapid progression compared to the MSA group.

As with the primary end-point, there was no statistically significant difference between treatment groups in progression rate with any of the scales, either in the overall population or within strata. The PPPT analysis showed identical results for primary and secondary end-points (data not shown).

To test our initial assumptions, using the data acquired during the study, we calculated the detectable difference in survival and functional change. With power greater than 0.8, and α risk at 0.05 (two-tailed), the actual number of patients recruited and events observed would have allowed us to detect a 40% decrease in relative risk of death for PSP and 35% for MSA, consistent with our initial hypothesis. For functional change, using the SEADL as the most sensitive scale, we would have been able to detect differences in annual rate of progression of 32% for PSP and 36% for MSA.

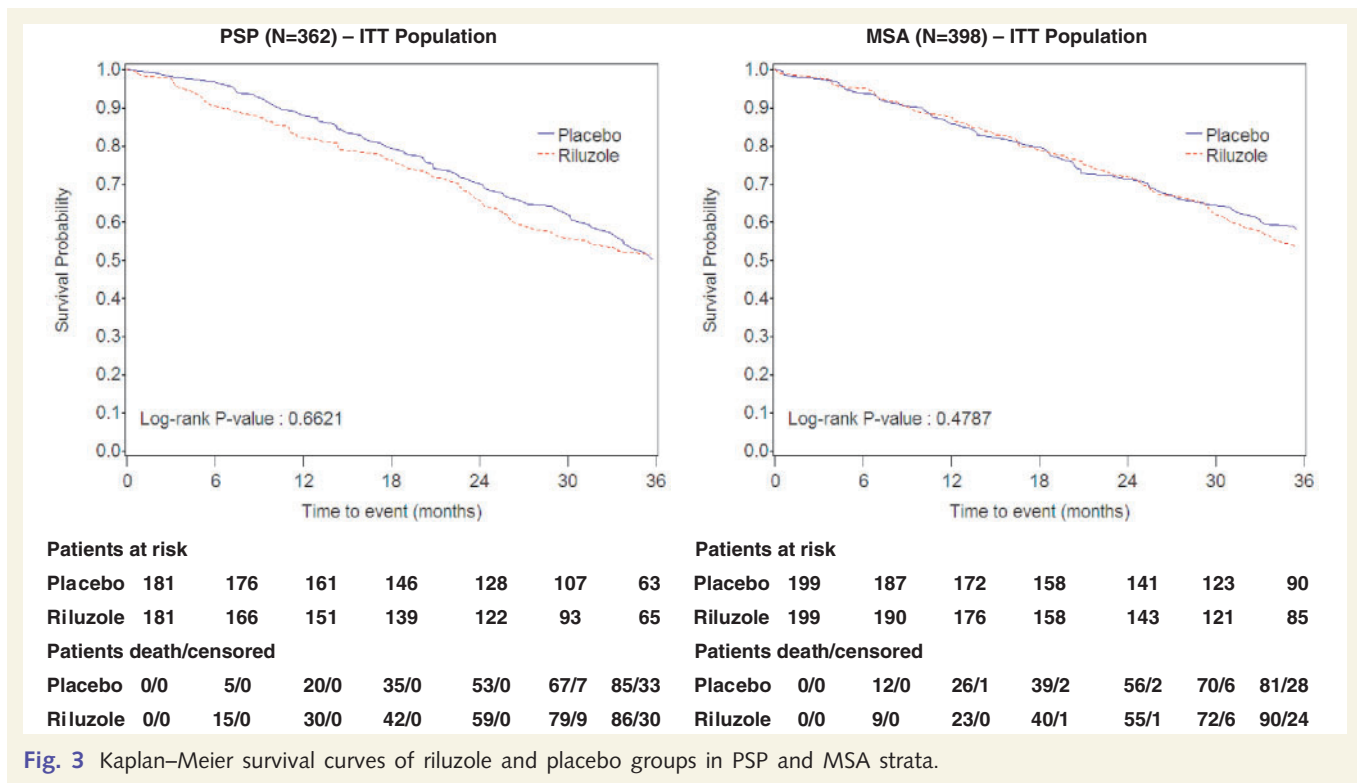


Fig. 3 Kaplan–Meier survival curves of riluzole and placebo groups in PSP and MSA strata.

Table 3 Slope of change in functional measures

	Placebo		Riluzole	
	PSP, n = 181	MSA, n = 199	PSP, n = 181	MSA, n = 199
Short Motor Disability Scale*	3.1 ± 4.4 (171)	2.3 ± 2.7 (190)	3.0 ± 4.0 (168)	2.7 ± 2.9 (185)
Schwab and England Activity Daily Living*	−16.3 ± 17.5 (162)	−11.5 ± 12.4 (172)	−15.2 ± 14.1 (149)	−12.9 ± 16.5 (168)
Hoehn and Yahr Staging	0.5 ± 0.8 (162)	0.5 ± 0.6 (172)	0.6 ± 0.7 (149)	0.5 ± 0.8 (168)
Clinical Global Impression of disease severity*	0.7 ± 0.7 (162)	0.5 ± 0.6 (171)	0.7 ± 0.8 (148)	0.5 ± 0.8 (168)

Slope of change for each patient was calculated using simple linear regression method of dependent variable with time since randomization. Patient with at least two measures in study were included in analyses. Slope of change in functional scale scores are shown as mean points/year ± SD with numbers of patients in each analysis in parentheses. No statistically significant difference between treatment groups could be evidenced in either scale.

* Strata differences were observed with the SEADL ($P=0.003$), CGI disease severity ($P=0.0003$) and SMDS ($P=0.03$).

Safety

Analysis of all SAEs revealed no significant differences between riluzole and placebo groups for the frequency of unexpected SAEs. However, for expected SAEs (other than death) with frequency of greater than 5%, gastrointestinal events were more common with riluzole treatment (Table 4). Borderline significant differences were observed in SAEs with frequency below 5%. Cardiac disorders were more common in the riluzole group (2.6% versus 0.8% placebo) and urinary disorders more frequent in the placebo group (4.4% versus 1.8% riluzole). With all treatment groups combined, there were significant differences between PSP and MSA, with injuries due to falls more common in PSP (18% versus 8% in MSA) and gastrointestinal disorders related to dysphagia more frequent in PSP (11% versus 6%). Events related to death are given in Table 4. The main causes were related to respiratory disorders, general condition (terminal state) and

infections (urinary or pulmonary). Gastrointestinal complications were more common in PSP (4% versus 1% for MSA). Among the non-serious events, gastrointestinal disorders were again more common with riluzole (34% versus 27% with placebo, $P=0.03$). The frequency of vascular disorders (mainly haematoma related to falls) was also slightly higher in the riluzole group (13% versus 8% with placebo, $P=0.02$). Elevation of transaminases was rare, with values above three times normal reported for only 11 patients on riluzole (3%) and eight patients on placebo (2%).

Validity of NNIPPS diagnostic criteria

At entry, the investigators' assessments of diagnostic probability for PSP or MSA were in close agreement with the inclusion category (point-biserial correlation for MSA $r_{pb}=0.93$, $P<0.0001$, and for PSP $r_{pb}=0.95$, $P<0.0001$) demonstrating an excellent convergent validity. For a few patients, the VAS probability was equal

for PSP and MSA (Fig. 4). Overall, diagnostic confidence (VAS probability of each diagnosis) was significantly higher (Student's *t*-test 2.89, $P=0.004$) for PSP patients [mean 81.2, SD \pm 12.8 (range 0.40–1)] than for MSA patients [mean 78.4, SD \pm 14.1 (range 0.21–1)].

A total of 210 patients (27%) consented to brain donation. Analysis of 112 brains was completed at the time of writing. Comparisons of patients with a pathological diagnosis with those dying without pathological studies showed no difference in demographic characteristics, disease severity at entry or diagnostic probability (data not shown).

Histopathology showed that the NNIPPS clinical diagnostic criteria had correctly identified Parkinson plus syndromes in 94% of cases. Two patients with pathologically confirmed MSA were mis-stratified at entry as PSP, and three with pathologically confirmed PSP had been mis-stratified as MSA (overall 4.9% mis-stratification). For the PSP stratum, seven cases were found to have other conditions, including two with Lewy body disease, one with amyotrophic lateral sclerosis (ALS), one with basophilic inclusion body disease, and three with corticobasal degeneration (CBD). In the MSA stratum, there were three misdiagnoses including one case with Lewy body disease, one with ALS and one with non-specific lesions. The sensitivity (95% CI) and specificity (95% CI) of the NNIPPS clinical diagnostic criteria for MSA were 0.96 (0.88–0.99) and 0.91 (0.86–0.93) respectively, with a correct clinical diagnosis in 0.93 (0.87–0.96) of cases and positive likelihood of 10.67 (6.28–14.39) (Table 5). For patients diagnosed clinically as PSP, but excluding cases diagnosed pathologically as

CBD, sensitivity (95% CI) and specificity (95% CI) of clinical diagnostic criteria were 0.95 (0.89–0.98) and 0.84 (0.77–0.87) respectively, with 0.89 (0.83–0.93) having a correct clinical diagnosis and positive likelihood of 5.79 (3.79–7.58). These results are evidence of an excellent predictive validity of the NNIPPS diagnostic criteria. This was not explained by the higher than expected disease severity and thus the large number of late stage patients in the overall population. We conducted the same analysis on the population broken down by severity according to the median of the CGI-ds at baseline, which was identical in the overall population and in the sub-group with neuropathology diagnosis. Mean \pm SD CGI-ds was 2.6 ± 0.5 and 2.9 ± 0.3 in early patients with PSP and MSA, respectively, and 4.5 ± 0.6 and 4.5 ± 0.7 in late patients with PSP and MSA, respectively. As shown in Table 5, results in early patients were not significantly different from those in late patients demonstrating a good consistency of the parameters with regard to disease progression. All three CBD cases fell within the late stage category. When CBD was included in the PSP neuropathology cases, there was slightly but not significantly increased diagnostic specificity, overall fraction correct, and positive likelihood for the overall sample and the late disease patients subgroup (data not shown).

Reassessment of the clinical diagnosis was achieved at least once during the trial period in 554 (73%) patients who did not have a neuropathological diagnosis (PSP, 71%; MSA, 75%). The overall rate of change in clinical diagnosis after entry was 7%, (Table 6), consistent with the neuropathological findings.

Table 4 Serious Adverse Events—MedDRA classification (By System Organ Class)

System Organ Class, n (%)	Related to death		Related to hospitalisation	
	Placebo N=383	Riluzole N=382	Placebo N=383	Riluzole N=382
Respiratory, thoracic and mediastinal disorders	60 (25)	65 (29)	33 (14)	40 (12)
General disorders and administration site conditions	56 (24)	56 (25)	27 (12)	39 (12)
Infections and infestations	37 (16)	49 (22)	38 (16)	50 (15)
Cardiac disorders	25 (11)	14 (6)	3 ^a (1)	10 (3)
Surgical and medical procedures	13 (6)	8 (4)	13 (6)	19 (6)
Gastrointestinal disorders	13 (6)	7 (3)	23 ^a (10)	41 (13)
Nervous system disorders	12 (5)	7 (3)	22 (9)	25 (8)
Metabolism and nutrition disorders	3 (1)	6 (3)	10 (4)	8 (2)
Injury, poisoning and procedural complications	6 (3)	2 (1)	48 (20)	49 (15)
Neoplasms benign, malignant and unspecified	5 (2)	2 (1)	2 (1)	2 (1)
Psychiatric disorders	2 (1)	4 (2)	13 (6)	19 (6)
Skin and subcutaneous tissue disorders	3 (1)	1 (0.4)	3 (1)	1 (0.3)
Vascular disorders	–	3 (1)	7 (3)	8 (2)
Renal and urinary disorders	2 (1)	–	17 ^a (7)	7 (2)
Reproductive system and breast disorders	1 (0.4)	–	3 (1)	1 (0.3)
Hepatobiliary disorders	–	–	3 (1)	3 (1)
Investigations	–	–	5 (2)	1 (0.3)
Musculoskeletal and connective tissue disorders	–	–	3 (1)	3 (1)
Blood and lymphatic system disorders	–	–	3 (1)	–
Eye disorders	–	–	2 (1)	–
Ear and labyrinth disorders	–	–	–	1 (0.3)
Total events	238	224	235	327
Total patients, Number of events (percentage of patients)	169 (44)	176 (46)	145 (38)	164 (43)

^a Statistically significant difference between treatment by the Fischer exact test.

Clinical features and natural history of the NNIPPS cohort

Analysis of the systematic questionnaire on initial clinical signs showed that the akinetic-rigid syndrome (with or without falls) was a frequent presenting syndrome in the PSP stratum (70.2%) and in the MSA stratum (61.8%). In the PSP stratum, oculomotor abnormalities were an initial feature in only 7.7% of patients, while 11.0% had presented with a behavioural or cognitive syndrome, and 5.8% had bulbar or pseudo-bulbar features. In MSA, second to akinetic-rigid syndrome, the most common presenting clinical features were cerebellar (22.1%) and genito-urinary (9.1%). The date of onset of gait instability or falls was

documented in 93.3% of the population (92.5% PSP, 94.0% MSA). Falls within the first year of disease onset, as incorporated in the NINDS-SPSP criteria (Litvan *et al.*, 2003) were present in only 49.6% of the PSP patients, and were also present in 21.9% of the MSA group. Similar results were observed in neuropathologically confirmed cases (PSP 53.6%, MSA 27.9%). A similar proportion of patients in the PSP and MSA strata had levodopa therapy at entry [PSP, $n=307$ (85%); MSA, $n=334$ (84%)], although the mean daily dose of levodopa was higher in the MSA group (636 mg/day, range 50–2100 mg) compared to the PSP group (509 mg/day, range 50–1600 mg) ($P<0.0001$) (Supplementary Table 2). Overall, most patients had a very poor response to levodopa therapy. A greater than 50% response to levodopa was reported for only 1.5% of MSA patients, and none of the PSP group. A best-ever response to levodopa therapy $>50\%$ was reported more frequently for MSA patients (9.1%) than for PSP patients (2.6%) ($P=0.0002$). In the majority of those who had experienced a good response to levodopa, the duration of response was <1 year.

At time of entry in the study, the distribution of the various syndromes is given for both diagnostic categories (Fig. 5). Oculomotor abnormalities were present in all PSP patients except one, but were also noted in 19% of MSA patients. PSP patients showed a higher rate of cognitive and behavioural syndromes. Dysautonomia was present in the majority of MSA patients, with urinary symptoms (87%) more common than cardiovascular symptoms (57%). Urinary symptoms were also present in 48% of PSP patients. A cerebellar syndrome was reported in 50% of MSA cases, but also a small number of PSP patients (6%). Pyramidal signs were present in approximately half the cases, slightly more in the MSA strata. A high frequency of bulbar/pseudobulbar features was reported in the MSA patients group (63%) and the PSP group (76%).

There were no differences between disease strata in terms of gender, weight or height. MSA patients were younger than the PSP patients at entry, younger at disease onset and had longer disease duration prior to entry (Table 2). Assessments of disease severity by the CGI-ds or the modified Hoehn and Yahr staging, showed there were significantly fewer MSA patients in the most severe stage ($P=0.024$ and $P=0.048$). Likewise, strata difference in scores for the SEADL indicated less dependency for MSA than for the PSP patients ($P=0.017$). As expected, PSP patients scored worse on cognitive functions as assessed with the MMSE ($P=0.0001$) (Table 2).

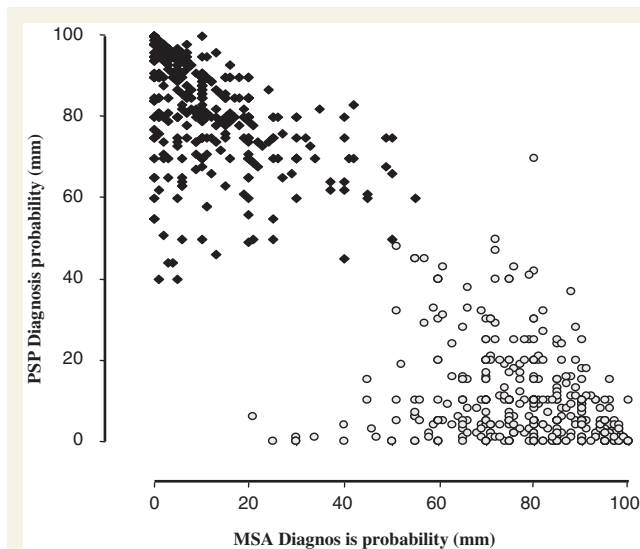


Fig. 4 Convergent validity of NNIPPS Diagnostic Criteria with Investigators' Diagnostic Probability (VAS). At the inclusion visit, following patients' assignment to strata using the NNIPPS diagnostic criteria, investigators were asked to evaluate the probability of each diagnosis (PSP, MSA), using a 100 mm VAS. All 760 patients are plotted on the graph according to the probability score on each VAS (PSP-vertical axis, MSA-horizontal axis). Solid diamonds represent patients included in the PSP stratum; White circles represent patients included in the MSA stratum. Convergent validity of the NNIPPS inclusion criteria with the investigators' assessment of diagnostic probability was tested with the point-biserial correlation. MSA, $r_{pb}=0.93$ ($P<0.0001$), PSP, $r_{pb}=0.95$, ($P<0.0001$).

Table 5 Predictive validity of NNIPPS clinical diagnostic criteria in deceased patients with neuropathological diagnosis

	CGI<4 $n=39$		CGI \geq 4 $n=73$		ALL $n=112$	
	PSP	MSA	PSP	MSA	PSP	MSA
Sensitivity (95% CI)	0.95 (0.82–0.99)	0.88 (0.73–0.95)	0.95 (0.86–0.98)	1.0 (0.91–1)	0.95 (0.88–0.98)	0.96 (0.88–0.99)
Specificity (95% CI)	0.84 (0.7–0.89)	0.91 (0.79–0.96)	0.83 (0.74–0.87)	0.91 (0.86–0.91)	0.84 (0.77–0.87)	0.91 (0.86–0.93)
Overall fraction correct (95% CI)	0.90 (0.77–0.94)	0.90 (0.77–0.96)	0.89 (0.80–0.93)	0.95 (0.88–0.95)	0.89 (0.83–0.93)	0.93 (0.87–0.96)
Positive likelihood (95% CI)	6.01 (2.83–8.60)	9.71 (3.50–26.48)	5.68 (3.35–7.71)	11.25 (6.32–11.25)	5.79 (3.79–7.58)	10.67 (6.28–14.39)

Sensitivity, specificity, overall fraction correct and positive likelihood (95% CI) of NNIPPS diagnostic criteria in the overall population with neuropathology diagnosis, and broken down by disease severity as defined by the CGI. Early patients were defined as those below the median (CGI, range 1–3); late patients were those equal or over the median (CGI, range 4–6).

Table 6 Change in clinical diagnosis during the trial showing predictive validity of NNIPPS clinical diagnostic criteria in surviving patients

Clinical diagnosis at inclusion	Final clinical diagnosis			All
	PSP	MSA	Other*	
PSP	246 (96%)	2 (1%)	8 (3%)	256
MSA	13 (4%)	270 (91%)	15 (5%)	298
All	259	272	23	554

* Other diagnosis: MSA ($N=15$): IPD $n=7$, ALS $n=2$, CBD $n=3$, LBD $n=1$, Carbon dioxide intox $n=1$, MPI multilacunar $n=1$; PSP ($N=8$): Mixed PSP/MSA $n=1$, Lower body Parkinsonism – pseudo PSP $n=1$, CBD $n=1$, FTDP or Motor neuron disease $n=1$, ophthalmoplegia $n=1$, FTD $n=1$, multiple cerebral infarct $n=1$, unidentified $n=1$.

Demographics and clinical factors influencing survival were assessed with univariate and multivariate Cox model analysis (Supplementary material and Supplementary Table 7). In the univariate analyses, only variables scoring disease severity were predictive of survival. Accordingly, the stepwise multivariate analysis selected first the SEADL, followed by the CGI scales for disease severity and for dysautonomia, as strong predictors of survival.

However, following adjustment on these variables, disease duration was selected in the model (RR = 0.923, $P=0.007$) indicating that at constant severity patients with a shorter disease duration at entry had a worse prognosis ('fast progressors'). To visualise the discriminating accuracy of the combined variables, we constructed a prognostic score for each patient using a Cox coefficient weighted linear combination of the selected prognostic variables (Supplementary Fig. 1A). Following adjustment on the prognostic score as a variable, the strata factor became significant (RR = 0.657, $P=0.0002$) indicating that PSP patients on the whole had a worse prognosis compared to MSA patients at constant disease severity and disease duration at entry (Supplementary Fig. 1B). Testing for a treatment effect on survival following adjustment on the prognostic variables did not change the results of the log-rank analysis nor was there any significant interaction between treatment and prognostic variables (Supplementary Table 3).

Discussion

Our results show that riluzole (up to 200 mg/day) is unlikely to be effective as a disease-modifying agent in PSP or MSA. Riluzole has a definite effect in ALS with an estimated 30–40% decrease in relative risk at 12–18 months (Bensimon *et al.*, 1994; Lacomblez *et al.*, 1996; Miller *et al.*, 2007). NNIPPS was therefore designed to detect an effect of similar size. Power calculation assumptions for the placebo group were based largely on retrospective data (Litvan *et al.*, 1996a, c, 2003; Testa *et al.*, 1996, 2001; Ben-Shlomo *et al.*, 1997), which underestimated the number of events observed in NNIPPS by 2–6%. Hence the NNIPPS trial achieved the power necessary to detect the hypothesized drug effect. The failure of NNIPPS to detect that effect could be due to heterogeneity in the trial population (loss of power), uneven distribution of demographic and clinical prognostic factors or comorbidity across treatment groups (randomization bias), poor compliance, or simply a smaller pharmacological effect in these conditions than the one used at the planning stage (effect size).

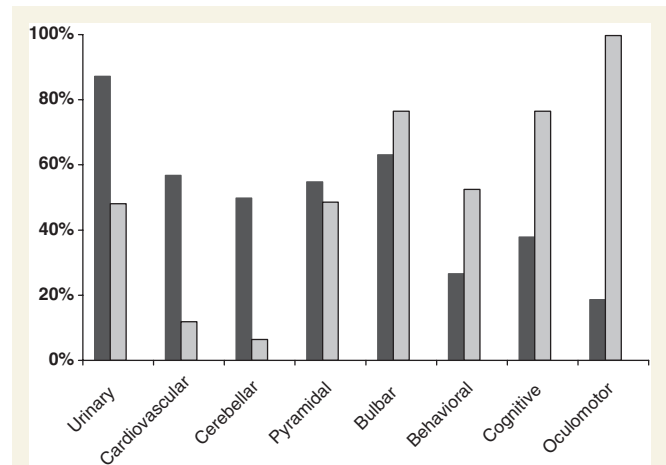


Fig. 5 Syndrome profile of patients in PSP and MSA Strata. At the inclusion visit, and based on the clinical neurological assessments, investigators were asked to describe the syndrome profile of the patients using a systematic (yes/no) questionnaire. Each bar represents the percentage of patients within each stratum positive for a given syndrome. Black bars represent the MSA stratum, grey bars the PSP stratum. The akinetic rigid syndrome was a mandatory inclusion criterion for both strata and therefore is not represented.

With regard to disease heterogeneity, a major difficulty in designing trials of disease-modifying therapy in neurodegenerative diseases is lack of diagnostic confidence, particularly at a stage of disease evolution when intervention is likely to be most effective. Indeed, there are no prospective studies of clinical diagnostic criteria assessed against the definitive criterion of histopathology in PSP or MSA (Litvan *et al.*, 2003). Such studies are not easy to achieve, as long term follow-up and a high autopsy rate are required. Until validated biomarkers are available for use in trials, clinical criteria prospectively tested against pathology are essential to understand heterogeneity, which is a potential source of bias in estimating drug effects in Parkinson plus syndromes, as in idiopathic Parkinson's disease. Furthermore, criteria for large scale clinical trials in rare diseases such as PSP and MSA should be robust and simple, so that they can be applied in international multi-centre studies and not only in highly specialized centres. The NNIPPS diagnostic criteria represent a simplification of existing diagnostic criteria and we have shown them to have excellent convergent and predictive validities in this study population,

despite variations in practice across three European countries and 44 centres. Although pathological examination was only possible in 112 cases (15% of recruited patients), the pathological sample is representative of those patients who died (45%) and, most likely, of the population as a whole. Further analysis of these diagnostic criteria in relation to MRI, DNA and neuropsychology is ongoing. At this stage it can be concluded that the estimated rate of misdiagnosis (about 6% overall, representing at most 46 patients in the ITT) is acceptably low and is unlikely to have biased the estimate of the drug effect. We could not detect imbalance in entry parameters with relevant prognostic significance for survival or disease progression which could account for the lack of treatment efficacy.

It is likely that a neuroprotective agent should be started in the earliest phase of the disease in order to demonstrate efficacy. Hence the impact of a neuroprotective agent is likely to be an inverse function of the stage of disease progression. In support of this assumption, riluzole did not show efficacy in ALS when tested in a trial including late stage disease patients (Besimon *et al.*, 2002). In the present trial, disease severity in the population included was greater than anticipated and the death rate was also higher than expected. Overall, about 50% of patients were classified as severely disabled or wheelchair bound (i.e. in the latest stages) by the Hoehn and Yahr staging and over 50% similarly classified as markedly, severely or extremely ill by the CGI scale for disease severity. However, adjustment for prognostic factors with the Cox model did not show any trend for a treatment effect, nor was there a significant treatment by prognostic factor interaction, meaning that the treatment effect is constant across the levels of the prognostic variables.

Even though treatment withdrawal rate was relatively high (25%), and might have had an impact on the size of a treatment effect, the overall mean dose was close to the maximum, and compliance rate was good (81.2%) considering the length of the trial, with a mean follow-up time close to the maximum planned (3 years). Finally, the PPPT sensitivity analysis was consistent with the ITT. Hence, the observed lack of treatment effect in this Parkinson plus population seems relatively robust. Our results are similar to those observed in Huntington's disease (Landwehrmeyer *et al.*, 2007) and contrast with the findings in ALS. This may indicate that the disease pathways targeted by riluzole in ALS are more specific than previously thought.

At the time of designing the NNIPPS study there were no validated instruments for measuring change in function in these disorders. Since the priority was to detect a disease-modifying effect of riluzole, we chose to use survival as a robust and unambiguous endpoint for the primary analysis. There are several reasons why prospectively acquired information on survival provides a unique tool for understanding these diseases and for developing new assessment instruments for clinical trials. First, it is possible to achieve complete ascertainment of the endpoint, as we have shown. Secondly, the high rate of death in PSP and MSA limits the use of functional scales as measurements of disease progression. Inevitably, the extent of missing functional data is not random, so that in case of drug toxicity increasing death rate, the remaining functional data in the active drug group is misleading as regards efficacy (Wu, 1988; Wu and Bailey, 1988)

because they originate from selection of a biased sub-group. In early patients, when the likelihood of informative censoring is least, slope of change in functional measures might be appropriate as an end point in a definitive phase III trial where using survival as an endpoint would result in substantially increasing the number of patients and/or length of the trial. The parameters provided in this article allow power calculations for either context since linearity is a basic assumption for slope of change and should hold true early or late in the course of disease.

Third, survival data provide a unique insight into the natural history of these disorders and help to validate functional assessments in relation to prognosis. In this study, the median survival from onset of symptoms is in keeping with our prior assumptions based on retrospective (Litvan *et al.*, 1996d; Testa *et al.*, 1996, 2001; Vanacore *et al.*, 2001a, b; Litvan *et al.*, 2003) and prospective studies (Schrag *et al.*, 2008). Following a step-wise Cox model adjustment on prognostic factors which selected disease severity scores (SEADL, CGI-ds, CGI-dysautonomia) and disease duration, there was a significant difference between strata consistent with the difference in the rate of disease progression between the two strata. Finally, our data show that it is feasible to use survival as a primary endpoint in phase III studies and that our strategy of recruiting patients with PSP and MSA into a single stratified trial is methodologically valid where a generic neuroprotective effect is hypothesised. On the other hand, where the agent to be tested is thought to act on disease-specific pathways (e.g. processing of tau in PSP, or alpha-synuclein in MSA) a different strategy is likely to be more appropriate. The demographic features of the NNIPPS cohort clearly indicate that patients presenting with an akinetic-rigid syndrome and categorized as having PSP or MSA with reasonably high confidence have very poor prognosis, and no significant response to levodopa therapy. Unfortunately, there are no comparable prospective studies with which to compare our cohort. The European MSA Study Group (EMSA-SG) has reported preliminary data on a functional scale developed for MSA based on assessments of 50 patients diagnosed on the basis of the Gilman criteria (Gilman *et al.*, 1998). A total of 412 patients with presumed MSA were included in a European registry (Geser *et al.*, 2005, 2006), but no information on survival is yet available from the EMSA database. Nevertheless, demographic features of the 50 patients reported from the EMSA are broadly comparable to those in the NNIPPS MSA stratum (Geser *et al.*, 2006). In a cohort study of 162 PSP patients the median survival since disease onset was 7.3 years, compared to 7.8 years in the NNIPPS cohort (Golbe and Ohman-Strickland, 2007). As with NNIPPS, functional scores predicted survival. However, multivariate analysis of prognostic factors revealed that age of onset and gender were predictors of survival. For the former, the apparent discrepancy is likely to be related to differences in the method of survival analysis since when survival is calculated from disease onset in our cohort, age of onset becomes a significant predictor of survival as well (data not shown). In relation to a gender effect, the sex ratio in most studies, as in the NNIPPS cohort, shows an equal proportion of men and women or a slight excess of men (Golbe *et al.*, 1988; Santacruz *et al.*, 1998; Schrag *et al.*, 1999; Nath *et al.*, 2001, 2003), whereas in the study of Golbe and Ohman-Strickland (2007) the sex ratio was reversed, indicating

a potential sample bias. However, in MSA, demographic features such as age of onset and sex ratio are similar in the NNIPPS cohort to those recorded in previous natural history studies (Bower *et al.*, 1997; Schrag *et al.*, 1999; Vanacore *et al.*, 2001a, b; Schrag *et al.*, 2008).

In order to achieve a relatively homogeneous population of patients in the PSP strata, oculomotor abnormalities were mandatory at inclusion. This increased the accuracy of diagnosis (as confirmed pathologically) but did not take into account 'PSP-Parkinsonism' patients without oculomotor abnormalities (Williams *et al.*, 2005, 2007) who may be confused with patients with PD or with MSA-P. Likewise, in order to recruit MSA patients with relatively early disease, we took a pragmatic approach so as to avoid, as far as possible, recruiting patients with disorders such as olivopontocerebellar atrophy (Berciano *et al.*, 2006) and idiopathic late-onset cerebellar ataxia, (Gilman and Quinn, 1996; Gilman *et al.*, 2000). We expected this to reduce the proportion of patients with the cerebellar form of MSA (MSA-C) in our cohort, but cerebellar features were noted at presentation in 20% of our patients, similar to the findings of a prospective study (Schrag *et al.*, 2008). Thus the NNIPPS population is reasonably representative of the whole (European and North American) MSA population. In this context, we have clearly shown that clinical features that have been regarded as characteristic of each disorder, and which are of key importance in the different consensus diagnostic criteria (Litvan *et al.*, 1996a, b, d, 1997, 2003; Gilman *et al.*, 1998) are common to both disorders. Thus at entry urinary symptoms were present in over 85% of patients in the MSA stratum but also in 50% in the PSP stratum (Fig. 5) and oculomotor signs, required for inclusion in the PSP stratum, were recorded in about 20% of patients in the MSA stratum. However, this included any evidence of supranuclear ophthalmoplegia, whereas allocation to the PSP stratum required significant supranuclear impairment of downward gaze. Likewise, behavioural and cognitive abnormalities were judged as being common in both PSP and MSA, in keeping with studies on cognitive function in these disorders (Robbins *et al.*, 1992, 1994; Burk *et al.*, 2006; Lyoo *et al.*, 2008). Despite this overlap, we have shown that PSP and MSA can be differentiated by neurologists at a relatively early stage in the disease process using the NNIPPS diagnostic criteria.

A striking feature of the NNIPPS diagnostic criteria is the high sensitivity at inclusion, although a direct comparison with published diagnostic criteria is not possible as the timing of assessments was different (Osaki *et al.*, 2002, 2004; Litvan *et al.*, 2003). The NNIPPS diagnostic assessments were carried out on average 2 years before death in the pathologically confirmed cases whereas in the published criteria the assessments were made at the first ('diagnostic') visit when sensitivity is low, and at the end of the disease (i.e., on the last visit before death) when sensitivity is high. Studies of the sensitivity of the NNIPPS criteria at the earliest stages of disease process are now needed, though we provide evidence that the criteria still display acceptable quality in the sub-group below the mid-range of disease progression. Indeed, only because we studied both conditions in one cohort could we discern the close similarities and the diagnostically important differences between PSP and MSA.

We have shown that a large scale randomized trial with survival as the primary endpoint is feasible using the Parkinson plus concept to allow relatively early diagnosis and recruitment of a surprisingly homogeneous population, as shown by the performance of the NNIPPS diagnostic criteria informed by pathology. The rate of death was higher than anticipated, allowing us to shorten the trial slightly. We have learnt from clinical trials in ALS that survival does not necessarily equate with functional change (Lacomblez *et al.*, 1996; Meininger, 2005). This has resulted in recommendations by the regulatory authorities (e.g. European Medicines Agency, Committee for Proprietary Medicinal Products, CPMP/EWP/565/98) and by the World Federation of Neurology Committee on Research (Miller *et al.*, 1999) that survival should be the primary endpoint for definitive demonstration of a neuroprotective drug effect. Future phase III studies on potential neuroprotective agents in these disorders should consider using survival as a primary endpoint, and combining patients with presumed PSP and MSA in the same trial in order to improve our understanding of both.

Supplementary material

Supplementary material is available at *Brain* online.

Funding

NNIPPS was an academic-led study with core funding from the European Union 5th Framework Programme (QLG1-CT-2000-01262); support from French Health Ministry, Programme Hospitalier de Recherche Clinique (AOM97073, AOM01125) and from Sanofi-Aventis affiliates in the UK, France and Germany providing an unconditional research grant and drug supply throughout the study. Three academic institutions (Institute of Psychiatry, King's College London; Assistance Publique-Hôpitaux de Paris; and University of Ulm) were sponsors of the study in each country, and jointly own the data. The protocol and amendments were reviewed and approved by the Comité de Protection des Personnes of Pitié-Salpêtrière Hospital (France), the UK Multicentre Research Ethics Committee (MREC), (UK), Ethikkommission of the University of Ulm, (Germany), and by local Institutional Review Boards (Ethics Committees) where appropriate (UK, Germany).

Acknowledgements

We thank the patients and their families for their commitment and altruism, and The French and UK PSP Associations and the UK Parkinson's Disease Research Group for their help and support. We are grateful to the many colleagues who were not formally part of the NNIPPS consortium but whose support contributed to the success of the study. The study protocol was filed in the open clinical trial registry (www.clinicaltrials.gov) with ID number NCT00211224.

References

- Albers DS, Augood SJ. New insights into progressive supranuclear palsy. *Trends Neurosci* 2001; 24: 347–53.
- Albin RL, Greenamyre JT. Alternative excitotoxic hypotheses. *Neurology* 1992; 42: 733–8.
- Allison P. *Survival Analysis using the SAS System – A Practical Guide*. SAS Institute Inc.; 1995.
- American Psychological Association. *Standard for educational and psychological testing*. Washington DC: American Psychological Association ed.; 1985.
- Attia J. Moving beyond sensitivity and specificity: Using likelihood ratios to help interpret diagnostic tests. *Australian Prescriber* 2003; 26: 111–13.
- Ben-Shlomo Y, Wenning GK, Tison F, Quinn NP. Survival of patients with pathologically proven multiple system atrophy: a meta-analysis. *Neurology* 1997; 48: 384–93.
- Bensimon G, Lacomblez L, Delumeau JC, Bejuit R, Truffinet P, Meininger V. A study of riluzole in the treatment of advanced stage or elderly patients with amyotrophic lateral sclerosis. *J Neurol* 2002; 249: 609–15.
- Bensimon G, Lacomblez L, Meininger V. A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group. *N Engl J Med* 1994; 330: 585–91.
- Berciano J, Boesch S, Pérez-Ramos JM, Wenning GK. Olivopontocerebellar atrophy: toward a better nosological definition. *Mov Disord* 2006; 21: 1607–13.
- Bishop Y, Fienberg S, Holland P. *Discrete multivariate analysis: theory and practice*. Cambridge, MA: MIT Press; 1975.
- Bower JH, Maraganore DM, McDonnell SK, Rocca WA. Incidence of progressive supranuclear palsy and multiple system atrophy in Olmsted County, Minnesota, 1976 to 1990. *Neurology* 1997; 49: 1284–8.
- Burk K, Daum I, Rub U. Cognitive function in multiple system atrophy of the cerebellar type. *Mov Disord* 2006; 21: 772–6.
- Burn DJ, Lees AJ. Progressive supranuclear palsy: where are we now? *Lancet Neurol* 2002; 1: 359–69.
- Caumont AS, Octave JN, Hermans E. Specific regulation of rat glial cell line-derived neurotrophic factor gene expression by riluzole in C6 glioma cells. *J Neurochem* 2006; 97: 128–39.
- Cox D. Regression models and life tables. *J R Stat Soc* 1972; B34: 187–220.
- Dickson DW. Neuropathologic differentiation of progressive supranuclear palsy and corticobasal degeneration. *J Neurol* 1999; 246 (Suppl. 2): 116–15.
- Dickson DW, Rademakers R, Hutton ML. Progressive supranuclear palsy: pathology and genetics. *Brain Pathol* 2007; 17: 74–82.
- Diguët E, Fernagut PO, Scherfler C, Wenning G, Tison F. Effects of riluzole on combined MPTP + 3-nitropropionic acid-induced mild to moderate striatonigral degeneration in mice. *J Neural Transm* 2005; 112: 613–31.
- Doble A. The role of excitotoxicity in neurodegenerative disease: implications for therapy. *Pharmacol Ther* 1999; 81: 163–221.
- 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.
- Geser F, Seppi K, Stampfer-Kountchev M, Kollensperger M, Diem A, Ndayisaba JP, et al. The European Multiple System Atrophy-Study Group (EMSA-SG). *J Neural Transm* 2005; 112: 1677–86.
- Geser F, Wenning GK, Seppi K, Stampfer-Kountchev M, Scherfler C, Sawires M, et al. Progression of multiple system atrophy (MSA): a prospective natural history study by the European MSA Study Group (EMSA SG). *Mov Disord* 2006; 21: 179–86.
- Gilman S, Little R, Johanns J, Heumann M, Kluin KJ, Junck L, et al. Evolution of sporadic olivopontocerebellar atrophy into multiple system atrophy. *Neurology* 2000; 55: 527–32.
- Gilman S, Low P, Quinn N, Albanese A, Ben-Shlomo Y, Fowler C, et al. Consensus statement on the diagnosis of multiple system atrophy. American Autonomic Society and American Academy of Neurology. *Clin Auton Res* 1998; 8: 359–62.
- Gilman S, Quinn NP. The relationship of multiple system atrophy to sporadic olivopontocerebellar atrophy and other forms of idiopathic late-onset cerebellar atrophy. *Neurology* 1996; 46: 1197–9.
- Golbe LI, Davis PH, Schoenberg BS, Duvoisin RC. Prevalence and natural history of progressive supranuclear palsy. *Neurology* 1988; 38: 1031–4.
- Golbe LI, Ohman-Strickland PA. A clinical rating scale for progressive supranuclear palsy. *Brain* 2007; 130: 1552–65.
- Hauw JJ, Daniel SE, Dickson D, Horoupian DS, Jellinger K, Lantos PL, et al. Preliminary NINDS neuropathologic criteria for Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). *Neurology* 1994; 44: 2015–19.
- Heiser V, Engemann S, Brocker W, Dunkel I, Boeddrich A, Waelter S, et al. Identification of benzothiazoles as potential polyglutamine aggregation inhibitors of Huntington's disease by using an automated filter retardation assay. *Proc Natl Acad Sci USA* 2002; 99 (Suppl. 4): 16400–6.
- Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967; 17: 427–42.
- Lacomblez L, Bensimon G, Leigh PN, Guillet P, Meininger V. Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II. *Lancet* 1996; 347: 1425–31.
- Landwehrmeyer GB, Dubois B, de Yebenes JG, Kremer B, Gaus W, Kraus PH, et al. Riluzole in Huntington's disease: a 3-year, randomized controlled study. *Ann Neurol* 2007; 62: 262–72.
- Lantos PL, Papp MI. Cellular pathology of multiple system atrophy: a review. *J Neurol Neurosurg Psychiatry* 1994; 57: 129–33.
- Litvan I. Update on epidemiological aspects of progressive supranuclear palsy. *Mov Disord* 2003; 18 (Suppl. 6): S43–50.
- Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, et al. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology* 1996a; 47: 1–9.
- Litvan I, Agid Y, Jankovic J, Goetz C, Brandel JP, Lai EC, et al. Accuracy of clinical criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome). *Neurology* 1996b; 46: 922–30.
- Litvan I, Bhatia KP, Burn DJ, Goetz CG, Lang AE, McKeith I, et al. Movement Disorders Society Scientific Issues Committee report: SIC Task Force appraisal of clinical diagnostic criteria for Parkinsonian disorders. *Mov Disord* 2003; 18: 467–86.
- Litvan I, Goetz CG, Jankovic J, Wenning GK, Booth V, Bartko JJ, et al. What is the accuracy of the clinical diagnosis of multiple system atrophy? A clinicopathologic study. *Arch Neurol* 1997; 54: 937–44.
- Litvan I, Hauw JJ, Bartko JJ, Lantos PL, Daniel SE, Horoupian DS, et al. Validity and reliability of the preliminary NINDS neuropathologic criteria for progressive supranuclear palsy and related disorders. *J Neuropathol Exp Neurol* 1996c; 55: 97–105.
- Litvan I, Hutton M. Clinical and genetic aspects of progressive supranuclear palsy. *J Geriatr Psychiatry Neurol* 1998; 11: 107–14.
- Litvan I, Lees AJ. Progressive supranuclear palsy. *Adv Neurol* 1999; 80: 341–5.
- Litvan I, Mangone CA, McKee A, Verny M, Parsa A, Jellinger K, et al. Natural history of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome) and clinical predictors of survival: a clinicopathological study. *J Neurol Neurosurg Psychiatry* 1996d; 60: 615–20.
- Lyo CH, Jeong Y, Ryu YH, Lee SY, Song TJ, Lee JH, et al. Effects of disease duration on the clinical features and brain glucose metabolism in patients with mixed type multiple system atrophy. *Brain* 2008; 131: 438–46.
- Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1958; 22: 719–48.
- Mattson MP. Excitotoxic and excitoprotective mechanisms: Abundant targets for the prevention and treatment of neurodegenerative disorders. *Neuromolecular Med* 2003; 3: 65–94.
- Meininger V. Clinical trials in ALS: what did we learn from recent trials in humans? *Neurodegener Dis* 2005; 2: 208–14.

- Miller RG, Mitchell JD, Lyon M, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev* 2007; CD001447.
- Miller RG, Munsat TL, Swash M, Brooks BR. Consensus guidelines for the design and implementation of clinical trials in ALS. *World Federation of Neurology committee on Research. J Neurol Sci* 1999; 169: 2–12.
- Nath U, Ben-Shlomo Y, Thomson RG, Lees AJ, Burn DJ. Clinical features and natural history of progressive supranuclear palsy: a clinical cohort study. *Neurology* 2003; 60: 910–16.
- Nath U, Ben-Shlomo Y, Thomson RG, Morris HR, Wood NW, Lees AJ, et al. The prevalence of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome) in the UK. *Brain* 2001; 124: 1438–49.
- Nath U, Burn DJ. The epidemiology of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome). *Parkinsonism Relat Disord* 2000; 6: 145–53.
- Nunnally J, Bernstein I. *Psychometric theory*. New York: McGraw-Hill, Inc; 1994.
- Osaki Y, Ben-Shlomo Y, Lees AJ, Daniel SE, Colosimo C, Wenning G, et al. Accuracy of clinical diagnosis of progressive supranuclear palsy. *Mov Disord* 2004; 19: 181–9.
- Osaki Y, Ben-Shlomo Y, Wenning GK, Daniel SE, Hughes A, Lees AJ, et al. Do published criteria improve clinical diagnostic accuracy in multiple system atrophy? *Neurology* 2002; 59: 1486–91.
- Papp MI, Kahn JE, Lantos PL. Glial cytoplasmic inclusions in the CNS of patients with multiple system atrophy (striatonigral degeneration, olivopontocerebellar atrophy and Shy-Drager syndrome). *J Neurol Sci* 1989; 94: 79–100.
- Przedborski S. Pathogenesis of nigral cell death in Parkinson's disease. *Parkinsonism Relat Disord* 2005; 11 (Suppl. 1): S3–7.
- Robbins TW, James M, Lange KW, Owen AM, Quinn NP, Marsden CD. Cognitive performance in multiple system atrophy. *Brain* 1992; 115: 271–91.
- Robbins TW, James M, Owen AM, Lange KW, Lees AJ, Leigh PN, et al. Cognitive deficits in progressive supranuclear palsy, Parkinson's disease, and multiple system atrophy in tests sensitive to frontal lobe dysfunction. *J Neurol Neurosurg Psychiatry* 1994; 57: 79–88.
- Santacruz P, Uttl B, Litvan I, Grafman J. Progressive supranuclear palsy: a survey of the disease course. *Neurology* 1998; 50: 1637–47.
- Scherfler C, Sather T, Diguët E, Stefanova N, Puschban Z, Tison F, et al. Riluzole improves motor deficits and attenuates loss of striatal neurons in a sequential double lesion rat model of striatonigral degeneration (parkinson variant of multiple system atrophy). *J Neural Transm* 2005; 112: 1025–33.
- Schrag A, Ben-Shlomo Y, Quinn NP. Prevalence of progressive supranuclear palsy and multiple system atrophy: a cross-sectional study. *Lancet* 1999; 354: 1771–5.
- Schrag A, Wenning GK, Quinn N, Ben-Shlomo Y. Survival in multiple system atrophy. *Mov Disord* 2008; 23: 294–6.
- Shortland PJ, Leinster VH, White W, Robson LG. Riluzole promotes cell survival and neurite outgrowth in rat sensory neurones in vitro. *Eur J Neurosci* 2006; 24: 3343–53.
- Spillantini MG, Goedert M. The alpha-synucleinopathies: Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy. *Ann N Y Acad Sci* 2000; 920: 16–27.
- Streiner D, Norman G. *Health Measurement Scales: a practical guide to their development and use*. Oxford: Oxford University Press; 2003.
- Testa D, Filippini G, Farinotti M, Palazzini E, Caraceni T. Survival in multiple system atrophy: a study of prognostic factors in 59 cases. *J Neurol* 1996; 243: 401–4.
- Testa D, Monza D, Ferrarini M, Soliveri P, Girotti F, Filippini G. Comparison of natural histories of progressive supranuclear palsy and multiple system atrophy. *Neurol Sci* 2001; 22: 247–51.
- Vanacore N, Bonifati V, Colosimo C, Fabbrini G, De Michele G, Marconi R, et al. Epidemiology of progressive supranuclear palsy. ESGAP Consortium. European Study Group on Atypical Parkinsonisms. *Neurol Sci* 2001b; 22: 101–3.
- Vanacore N, Bonifati V, Fabbrini G, Colosimo C, De Michele G, Marconi R, et al. Epidemiology of multiple system atrophy. ESGAP Consortium. European Study Group on Atypical Parkinsonisms. *Neurol Sci* 2001a; 22: 97–9.
- Watanabe H, Saito Y, Terao S, Ando T, Kachi T, Mukai E, et al. Progression and prognosis in multiple system atrophy: an analysis of 230 Japanese patients. *Brain* 2002; 125: 1070–83.
- Williams DR, de Silva R, Paviour DC, Pittman A, Watt HC, Kilford L, et al. Characteristics of two distinct clinical phenotypes in pathologically proven progressive supranuclear palsy: Richardson's syndrome and PSP-parkinsonism. *Brain* 2005; 128: 1247–58.
- Williams DR, Holton JL, Strand C, Pittman A, de Silva R, Lees AJ, et al. Pathological tau burden and distribution distinguishes progressive supranuclear palsy-parkinsonism from Richardson's syndrome. *Brain* 2007; 130: 1566–76.
- Winer B. *Design and analysis of factorial experiments in Statistical principles in experimental design*. New York: McGraw-Hill Inc; 1971. p. 309–430.
- Wu MC. Sample size for comparison of changes in the presence of right censoring caused by death, withdrawal, and staggered entry. *Control Clin Trials* 1988; 9: 32–46.
- Wu MC, Bailey K. Analysing changes in the presence of informative right censoring caused by death and withdrawal. *Stat Med* 1988; 7: 337–46.
- Yoo MH, Hyun HJ, Koh JY, Yoon YH. Riluzole inhibits VEGF-induced endothelial cell proliferation in vitro and hyperoxia-induced abnormal vessel formation in vivo. *Invest Ophthalmol Vis Sci* 2005; 46: 4780–7.

Appendix

The NNIPPS STUDY GROUP

Principal Investigator: P.N. Leigh (London, UK)

Co-ordination: European and UK: P.N. Leigh (London, UK), France: G. Bensimon (Paris, France), Germany: A.C. Ludolph (Ulm, Germany)

Steering Committee: Chair: P.N. Leigh (London, UK), Members: Y. Agid, G. Bensimon, M. Dib, L. Lacomblez, M. Vidailhet (Paris, France), D. Burn (Newcastle, UK); B. Landwehrmeyer, A.C. Ludolph (Ulm, Germany)

Independent Data Monitoring and Safety Committee: Chair: B. Asselain (Paris-France), Members: H. Allain (Rennes, France), D. Chadwick (Liverpool, UK), JE. Perret (Grenoble, France), C. Warlow (Glasgow, UK)

Technical Committees

Clinical diagnostic criteria: Chair: D. Burn (Newcastle, UK), Members: Y. Ben-Shlomo (Bristol, UK), AM. Bonnet, J. Fermanian, C. Payan, M. Verny, M. Vidailhet (Paris, France), P. Moore (Liverpool, UK), C. Tranchant (Strasbourg, France)

Motor Function, QoL & Health Service Research: Chair: C. Payan (Paris, France), Members: M. Borg (Nice, France), P. McCrone (London, UK), F. Durif (Clermont-Ferrand, France), A. Evans (London, UK), J. Fermanian (Paris, France), F. Viallet (Aix en Provence, France)

Neuro-Imaging: Chair: M. Verin (Rennes, France), Members: N. Deasy, J. Jarosz (London, UK), T.K. Hauser (Tübingen, Germany), E. Kraft (Ulm, Germany), E. Broussole (Lyon, France), D. Dormont, C. Marsault, A. Tourbah (Paris, France), L. Defebvre, L. Delmaire (Lille, France), Y. Roland (Rennes, France)

Neuro-Pathology: Chair: J.J. Hauw (Paris, France), Members: C. Duyckaerts, D. Seilhean (Paris, France), S. Al-Sarraj, T. Revesz (London, UK), B. Landwehrmeyer (Ulm, Germany), H. A. Kretzschmar (Munich, Germany)

Neuropsychology: Chair: R. Brown (London, UK), Members: T. Bak (Cambridge, UK), A. Danek (Munich, Germany), B. Dubois, L. Lacomblez (Paris, France), R.M. Marié (Caen, France), I. Uttner (Ulm, Germany)

Genetics: Chair: A. Dürr (Paris, France), Members: A. Al-Chalabi, N. Wood (London, UK), A. Brice (Paris, France), W. Camu (Montpellier, France), K. Morrison (Birmingham, UK)

Logistics, Treatments, Monitoring, Data Management & Statistical analysis

Chair: G. Bensimon (Paris, France), **European Project Manager:** M. Graf (Paris, France), **Data Manager:** C. Payan (Paris, France), **Data entry:** P. Paillasseur (Theramis – St Maur des Fossés, France), **Senior Statistician:** C. Payan (Paris, France), Assistant Statistician: H.P. Pham (Paris, France), Functional scales development: J. Fermanian (Paris, France), Neuropsychology: R. Brown (London, UK), Health economics: P. Mc Crone (London, UK), **Clinical Research Assistants:** N. Dedise, C. Hermine, S. Sagnes, B. Poître, C. Foucart (Paris, France), A. Dougherty, C. Murphy, H. Mason (London, UK), T. Hermann, K. Klempf, A. Niess, V. Stange (Ulm, Germany), Regulatory affairs France: A. Ouslimani (Paris, France), **Treatment Manufacturing:** Sanofi-aventis (Antony, France), LC² (Lentilly, France), **Treatment management:** B. Lehmann, A. Tibi, Fabreguette (Paris, France), Cardinal (UK), Clindata (Germany)

Investigators within Countries

Principal Investigator (France/Germany/UK)

Centres (number of patients), Principal Investigators, Co-investigators (clinician/radiologist/psychologist)

France

Principal Investigator France: Y. Agid (Paris, France)

Aix en Provence (n=20): F. Viallet, C. Couratier, S. Arguillère (clinicians), H. Payan-Cassin, G.M. Vassault (radiologists), P. Henon, S. Gimeno (psychologists); Angers (5): F. Dubas, C. Fressinaud (Clinicians), J.Y. Tanguy (radiologist), D. Legall (psychologist); Besançon (n=7): L. Rumbach, E. Vidry (clinicians), J. Kraehenbuhl, J.F. Bonneville (radiologists), G. Chopard (psychologist); Caen (n=13): F. LeDoze, G. Defer, F. Viader, R-M. Marié (clinicians), H. Huet (radiologist), F. Daniel, C. Lalevée (psychologists); Clermont-Ferrand (n=21): F. Durif, B. Debilly, Ph. Derost, C. Tilignac (clinicians), J. Gabrillargues (radiologist), D. Lauvergne

Crégu, C. Rosière (psychologists); Grenoble (n=7): G. Besson, C. Mallaret (clinicians), S. Grand (radiologist), H. Klinger, A. Funkiewiez (psychologists); Lille (n=8): A. Destée, L. Defebvre (clinicians), C. Delmaire (radiologist), K. Dujardin (psychologist); Limoges (n=12): P. Couratier (clinician), MP Boncoeur-Martel (radiologist), M. Chazot-Balcon (psychologists); Lyon (n=11): E. Broussole, H. Mollion (clinicians), M. Hermier (radiologist), M. Bouvard (psychologists); Marseille (n=15): JP Azulay, T. Witjas (clinicians), (radiologist cf Aix en Provence), M. Delfini (psychologist); Montpellier (n=16): W. Camu, F. Portet, J. Khoris, N. Pageot, G. Garrigues (clinicians), B. Viaud (radiologist), K. Martin, J. Bernard (psychologists); Nice (n=8): M. Borg (clinician), S. Chanalet (radiologist), B. Bailet (psychologist); Paris patient clinical selection: M. Vidailhet, S. Sangla (Hôpital St Antoine), D. Ranoux (Hôpital St Anne), J.P. Brandel (Hôpital Leopold Belland), T. De Broucker (Hôpital St Denis), Y. Agid, B. Dubois, Meininger, Verny (Hopital Pitié-Salpêtrière), P. Cesaro (Hôpital Henri Mondor), G. Fenelon (Hôpital Tenon); Paris CIC Pitié-Salpêtrière-inclusion and follow-up (n=111): Y. Agid, F. Bloch, A.M. Bonnet, L. Lacomblez, D. Maltête, A. Memin, Torni, M.L. Welter, J. Worbe (Clinicians), T. Lalam, A. Tourbah, C. Marsault, Pr. D. Dormont (Radiologists), B. Pillon, V. Czernecki, A. Picard, B. Passaquet (psychologists); Pointe à Pitre (n=5) D. Caparros-Lefebvre, A. Lannuzel (clinicians), (no radiologist), F. Verlut (psychologist); Poitiers (n=12): R. Gil, M. Bailbé, S. Venisse, H. Moumy, V. Mesnage, J.L. Houeto, F. Petit (clinicians), P. Vandermarcq (radiologist), V. Bonnaud, C. Ornon (psychologists); Rennes (n=13): M. Verin (clinician), Y. Rolland (radiologist), P. Trébon, G. Salicé (psychologists); Strasbourg (n=16): C. Tranchant, G. Steinmetz (clinicians), JL Dietemann (radiologist), Crémel (psychologist); Toulouse (n=8): O. Rascol, M. Galitzky, C. Thalamas (clinicians), P. Manelfe (radiologist), S. Lemoal, M.C. Deneuille, H. Delabaere (psychologists); Tours (n=11): C. Prunier, A. Autret, P. Corsia (clinicians), P. Cottier, S. Gallas (radiologists), D. Beauchamp (psychologist).

Germany

Principal Investigator Germany: A. Ludolph (Ulm, Germany)

Aachen (n=24): J. Noth, C. Kosinski, C. Geyer, M. Kronenbürger, C. Schlangen (clinicians), M. Doenges, S. Kémeny (radiologists), M. Kronenbürger (psychologist); Berlin (n=37): K. Einhaeupl, PD G. Arnold, B. Hauptmann, A. Lipp (clinicians), A. Villringer, A. Lipp (radiologists), K. Fassdorf (psychologist); Bochum (n=7): H. Przuntek, T. Müller, G. Gagel-Schweibold, M. Siepmann, S. Benz (clinicians), G. Schmid (radiologist), M. Finger, P. Klotz (psychologists); Dresden (n=20): H. Reichmann, B. Herting (clinicians), R. von Kummer, D. Mucha (radiologists), E. Reissner (psychologist); Freiburg (n=14): C. H. Lücking, I. Bötöfür, S. Braune, C. Magerkurth, V. Mylius (clinicians), M. Schumacher, J. Spreer, S. Ziyeh (radiologists), C. Magerkurth, V. Mylius (psychologists); Halle (n=7): S. Zierz, M. Kornhuber, T. Mueller, S. Neudecker, U. Seifert (clinicians), C. Behrmann, A. Schlueter (radiologists), A. Rockahr (psychologist); Hannover (n=28): R. Dengler, A. Hauswedell, H. Kolbe, T. Peschel, C. Schrader, S. Siggelkow, J. Stewen, H.-H. Kapels, C. Winkler (clinicians), H. Heinze, G.

Kauffmann, M. Rotte (radiologists), C. Schrader (psychologist); Magdeburg (n=7): C. W. Wallesch, C. Bartels, M. Fork (clinicians), S. Reissberg (radiologist), M. Fork (psychologist); München (n=21): T. Brandt, F. Asmus, M. Bauer, T. Gasser, S. Maass, J. Velden, A. Viehöver, D. Wassilowsky, K. Bötzel (clinicians), T. Youssri, T. Wesemann, H. Brückmann, R. Brüning (radiologists), B. Baur, C. von Schlippenbach, G. Stenglein-Krapf (psychologists); Regensburg (n=11): U. Bogdahn, J. Klucken, Z. Kohl, M. Lange, C. Thun, J. Winkler, B. Winner (clinicians), G. Schuierer (radiologist), M. Lange (psychologist); Rostock (n=12): R. Benecke, D. Dressler, A. Wolters, G. Zegowitz (clinicians), G. Grau (radiologist), A. Fister (psychologist); Tübingen (n=23): J. Dichgans, O. Eberhardt, K. Gröschel, T. K. Hauser, J. B. Schulz (clinicians), M. Skalej, T. K. Hauser (radiologists), T. K. Hauser (psychologist); Ulm (n=31): A. C. Ludolph, D. Ecker, A. Jung, B. Kramer, G. B. Landwehrmeyer, A. Storch, S. D. Sussmuth (clinicians), E. Kraft, J. Kassubek (radiologists), I. Uttner (psychologist).

United Kingdom

Principal Investigator UK: PN Leigh (London, UK)

Belfast (n=9): M. Gibson, R. Forbes (clinicians), C. Reynolds, C. S. McKinstry (radiologists), A. Dick (psychologist); Birmingham – City Hospital (n=5): C. Clarke (clinician), S. Chavda (radiologist), S. Dhariwal, R. Hornabrook, S. Colhoun, V. Barnfield (psychologists); Birmingham – Queen Elizabeth Hospital (n=9): H. Pall,

D. Nicholls (clinicians), S. Chavda (radiologist), D. Nicholl (psychologist); Cambridge (n=17): J. Hodges, T. Bak (clinicians), A. Carpenter (radiologist), T. Bak, V. Hearn, L. Donald (psychologists); Liverpool (n=17): P. Moore (clinician), T. Dixon (radiologist), G. Baker, L. Owen, I. O'Brien (psychologists); London, King's College London (n=55): P.N. Leigh, K. R. Chaudhuri, D. Heaney, C. Blain, S. Azam, V. Williams, J. Isaacs, C. Smallman, B. Stanton (clinicians), J. Jarosz, N. Deasey (radiologists), R. Brown, A. Dittner, C. Donnellan, D. Secker, A. Langman, C. Lomax (psychologists); C. Simeon (research nurse) London NHNN & Queen Square Hospital (n=29): A. Lees, N. Quinn, A. Evans, T. Scaravilli, N. Russo, E. Trikouli, D. Paviour, Luke Massey (clinicians), C. Andrews, J. Stevens (radiologists), M. Jahanshahi, D. Winterburn (psychologists); Middlesbrough (n=7): P. Newman, Bathgate (clinicians), N. Bradey (radiologist), Z. Cowen (psychologist); Newcastle upon Tyne (n=20): D. Burn, A. Zermansky, N. Warren (clinicians), P. English, A. Gholkar (radiologists), J. Welch, P. Welch (psychologists); Stafford (n=9): B. Summers (clinician), D. Steventon (radiologist), B. Summers, L. Silver (psychologists); Aberdeen (n=4): C. Counsell (clinician), A. Murray (radiologist), J. Gordon, C. Harris, K. Perkins (psychologists); Guernsey (n=9) S. Bhaumick, S. Evans, G. Turner (clinicians), (no radiologist), S. Adam (psychologist); Swansea (14): R. Weiser, C. Lawthorn, A. Lowman (clinicians), (no radiologist), G. Forwood, M. Moran, L. Bastin (psychologists).