



OPEN ACCESS

RESEARCH PAPER

Using 'dead or dependent' as an outcome measure in clinical trials in Parkinson's disease

David McGhee,¹ Alexander Parker,¹ Shona Fielding,¹ John Zajicek,² Carl Counsell¹

¹Division of Applied Health Sciences, University of Aberdeen, Aberdeen, Scotland, UK

²Clinical Neurology Research Group, University of Plymouth, Plymouth, Derriford, UK

Correspondence to

Dr C Counsell, Division of Applied Health Sciences, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK; carl.counsell@abdn.ac.uk

Received 23 January 2014

Revised 4 April 2014

Accepted 7 April 2014

Published Online First

22 May 2014

ABSTRACT

Background Simple, robust, sensitive and clinically meaningful outcome measures are required for neuroprotective trials in Parkinson's disease (PD). We explored the feasibility of a composite binary outcome measure, 'dead or dependent', in such trials using data from a prospective follow-up study of an incident cohort of PD patients.

Methods Two hundred incident patients had an annual follow-up, including assessment of the Hoehn-Yahr stage (H-Y) and Schwab and England Activities of Daily Living Scale (S&E). Annual scores were converted into binary variables (H-Y <3 vs H-Y ≥3, and S&E ≥80% vs S&E <80%). A new outcome of 'dead or dependent' was also created, with dependence in activities of daily living defined as S&E <80%. Using these data, sample sizes were calculated for a hypothetical three-year randomised trial in which the trial outcome was defined by a binary clinical variable, all-cause mortality, or PD-related mortality.

Results At 3 years, 18.0% of patients were dead and 38.4% were dead or dependent. At 80% power, large sample sizes were required if PD-related mortality (n=1938 per study arm) or all-cause mortality (n=734) were used as the outcome, even for large treatment effects (30% reduction in relative risk). The new outcome of 'death or dependency' required the smallest sample sizes of all the outcome measures (n=277 for 30% reduction in relative risk, 627 for a 20% reduction).

Conclusions 'Death or dependency' is a feasible and potentially useful outcome measure in PD trials of neuroprotective agents, but further work is required to validate its use and define dependency.

more likely to be of direct relevance to patients, but are perhaps more susceptible to the confounding effects of factors other than disease progression (eg, mood, finances, comorbidity) than measures towards the bottom of the hierarchy.

In PD neuroprotective clinical trials, disability outcome measures may offer the best compromise between a relevant outcome to patients and as pure as possible a measure of disease progression. Furthermore, disability outcome measures may well be less susceptible to the effects of symptomatic treatment than measures of motor impairment, which are clearly influenced by dopamine replacement therapy.

Loss of independence in either basic ADL (eg, personal hygiene, feeding, ambulation), or instrumental ADL (eg, housework, finances, shopping), are important disability outcomes which should be delayed by neuroprotective treatment. Dependence can be robustly defined, and several valid measures of it exist.^{3–5–7} Data relating to dependence also has the advantage of being simple enough to collect via the telephone, helping to minimise missing data in large randomised controlled trials (RCT). However, as with other currently used clinical rating scales, the failure of dependence outcome measures to include 'death' as a scale interval necessitates that patients dying during a study must be censored in subsequent analyses, potentially reducing the power of the study to detect a statistically significant improvement in the measure in question.

Censoring in neuroprotective trials would be avoided by using mortality alone as the trial outcome, and this should be reduced/delayed by neuroprotective treatment. However, in diseases such as PD, where patients live for many years after diagnosis, this would require prolonged periods of follow-up and/or very large sample sizes, making such trials largely impractical. A potential solution to these issues may be to use a combined outcome measure which includes mortality and loss of independence in ADL. This approach has already been used in other neurological diseases, such as stroke.⁸ We, therefore, aimed to explore the feasibility of using a new composite binary outcome measure of 'dead or dependent' as opposed to 'alive and independent' in PD clinical trials using real-life data from an ongoing prospective follow-up study of a community-based incident cohort of patients with PD.

METHODS

The Parkinsonism Incidence in North-East Scotland (PINE) study is a prospective two-phase incidence

INTRODUCTION

Simple, robust and clinically meaningful outcome measures are required for neuroprotective trials in Parkinson's disease (PD). Clinical outcomes can be classified hierarchically (figure 1)¹ into those which measure impairment (signs of underlying disease, eg, the motor subsection of the old or new Unified Parkinson's Disease Rating Scale (UPDRS)²), disability (the functional result of impairment, eg, the Schwab and England Activities of Daily Living (ADL) Scale (S&E)³), handicap (the social and societal impact of the disease) and health-related quality of life (eg, the 39-item Parkinson's disease questionnaire ([PDQ-39])⁴). More recently, WHO refined these concepts in terms of impairment in bodily structures/functions, limitations in ADL and restrictions in participation in life situations.⁵ Measures towards the top of this hierarchy are



Open Access
Scan to access more
free content



CrossMark

To cite: McGhee D, Parker A, Fielding S, et al. *J Neurol Neurosurg Psychiatry* 2015;**86**: 180–185.

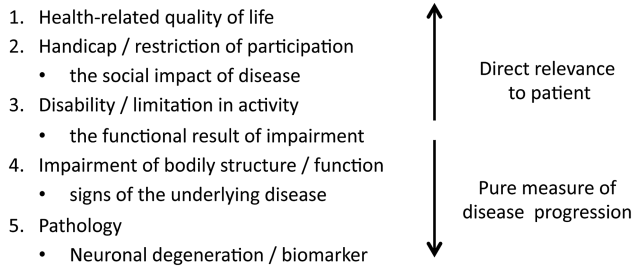


Figure 1 Hierarchy of outcomes measures.

study of parkinsonism in Aberdeen, UK, with preplanned life-long follow-up of incident patients to establish prognosis.⁹ Patients with new-onset possible or probable parkinsonism (defined as two or more of bradykinesia, rest tremor, rigidity or otherwise unexplained postural instability) were identified by a research team from 37 general (primary care) practices (baseline population 315 000) over a four-and-a-half-year period (2002–2004, 2006–2009) using multiple overlapping strategies. Detailed methods and incidence results have been previously published.⁹ The study was approved by the Scotland A Multi-centre Research Ethics Committee. All included patients gave their informed consent prior to inclusion in the study.

Ongoing annual follow-up of consenting individuals included assessment of the Hoehn-Yahr (H-Y) stage, and the S&E score. At each follow-up visit, the patient’s clinical diagnosis (PD or an alternative parkinsonian syndrome) was reviewed by a consultant neurologist with an interest in movement disorders, based on factors including the emergence of atypical features (eg, early dementia, falls, marked autonomic features, ataxia, myoclonus, supranuclear gaze palsy), the rate of disease progression, the response to dopamine replacement therapy, the development of motor complications and imaging. The UK PD society brain bank criteria¹⁰ were used to guide the clinical diagnosis of idiopathic PD. During follow-up, patients were invited to consent to postmortem and, if performed, information from these examinations informed the final diagnosis.

Patients could be seen between yearly visits if there was a clinical need, such as to start treatment. Initiation of therapy occurred when a shared decision was made by the patient and doctor. Factors including clinical presentation (tremor less likely to respond than bradykinesia), the impact of symptoms on the patient’s lifestyle and patient preference were taken into consideration, but no specific scores were used to guide initiation of treatment. There was no standard first-line dopaminergic medication: the treatment choice was usually made after discussion of the advantages and disadvantages of each option.

In this report, only consenting patients with a latest or final (if they had died) clinical diagnosis of idiopathic PD were included. Data were extracted from the PINE database on 23 November 2012. For each included patient, data from their baseline appointment, and up to 10 years of annual follow-up data were extracted. No patient had more than 10 years’ follow-up. The date of death for deceased patients was also extracted. Baseline data extracted for each patient included age, gender and the baseline motor UPDRS. For the baseline appointment and each annual follow-up, the H-Y stage and S&E were extracted and converted into binary variables, where people who had died had missing values: H-Y <3 versus H-Y ≥3 (as H-Y ≥3 is the point at which balance disturbance is present); S&E <80% versus S&E ≥80% (as S&E <80% is the point at which there is loss of independence in ADL, which in

PINE was defined as loss of independence in either washing, dressing, feeding, toileting or mobility). A new binary variable of ‘dead or dependent’ versus ‘alive and independent’ was also computed for each patient at each annual follow-up appointment where, again, dependence was defined by S&E <80%. We also performed a sensitivity analysis where dependence was defined in more severe terms (S&E ≤50% ie, ‘requiring help with 50% or more of chores’).

The PINE study case-notes of included patients were reviewed to determine the type and dose of dopamine replacement therapy taken at each annual follow-up, which were converted into levodopa-equivalent daily doses (LEDD).¹¹ Study notes, hospital and GP records, and death certificates were reviewed to determine whether those who had died did so as a result of their PD. We considered a ‘PD-related death’ to have occurred when a patient with advanced PD died as the result of a complication (eg, pneumonia, aspiration, infected bed sores, falls, dehydration, or general frailty due to PD). Patients dying suddenly with mild PD, or those with deaths obviously due to another cause (eg, myocardial infarction, stroke, or malignancy), were considered to have had a ‘non-PD related death’.

Analysis of survival rate

Using the Kaplan–Meier method, the length of survival from diagnosis (baseline visit) until death was analysed for all included patients. The length of survival in those who had died by the time of data extraction was taken as the difference between the date of death and their baseline (diagnostic) visit, and for those who were still alive, as the difference between the date of data extraction and the baseline visit. A reanalysis of the length of survival, censoring patients with a ‘non-PD related death’ was also undertaken, as was an analysis of time to ‘death or dependency’, whichever occurred first, where dependency was again defined by a S&E <80%.

Analysis of the change in binary categorical variables over 3 years

In order to assess whether the binary outcomes changed significantly over time, binary logistic generalised estimating analysis was performed using 3-years of follow-up data from those patients who had completed at least 3 years of follow-up, or would have done so had they not died before the date of data extraction. This time period was chosen, as all except one patient would have completed 3 years’ follow-up when the data were extracted. The change in the binary variables (dead or dependent, S&E <80%, and H-Y ≥3) over 3 years was analysed using data from patients who were free of the respective outcome at baseline. Patients who died were excluded from the H-Y and S&E analyses after the date of death, and the analyses assumed the data were missing completely at random. We, therefore, also did a sensitivity analysis in which those who died were classified as having a poor outcome on the H-Y and S&E. Year of follow-up and gender were included as fixed factors in the model, which was also adjusted for baseline age and motor UPDRS, LEDD at each follow-up and an interaction between LEDD and year of follow-up.

Sample size calculations for a three-year RCT

Sample sizes were calculated for a hypothetical three-year parallel-group RCT of a neuroprotective agent in which the trial outcome was either defined by one of the binary variables, all-cause mortality, or PD-related mortality. In these trials, newly diagnosed patients in active or placebo arms would be allowed to commence symptomatic treatment at any time, if required.

Movement disorders

However, where one of the binary variables defined the trial outcome, only patients who did not have the respective outcome at baseline would be eligible for inclusion.

Sample sizes were calculated with 80% and 90% power at the 5% level of significance. The proportions of control patients with each outcome were mostly taken from the PINE three-year data to reflect real life. However, a lower level of death or dependency (20%) at three years was also used to reflect what might be seen in milder/younger patients. Although initially the expected treatment effect was taken as a 30% relative risk reduction, this being the level set by the PD-NET collaborative group to detect neuroprotection,¹² a smaller effect (20% relative risk reduction) was also assessed to reflect a more modest/realistic treatment effect. Sample sizes (Fisher's exact method) were calculated using StatsDirect software (V2.7.9). Other statistical analyses were performed using the International Business Machines (IBM) Statistical Package for the Social Sciences (SPSS) Statistics V20.0.

RESULTS

Of an incident cohort of 377 patients with a baseline diagnosis of possible or probable parkinsonism, 210 have a latest clinical diagnosis of PD; 10 did not consent to clinical follow-up and were therefore excluded. Of the remaining 200 patients, only one had not completed three years of follow-up by the date of data extraction.

Seventy-one of the 200 included patients died prior to the date of data extraction. Thirty-three died from a PD-related cause (nine from pneumonia; four from dementia; two after a fall; two, who were extremely immobile due to PD, from pulmonary emboli; two with advanced PD from dehydration; two from aspiration; and twelve from general frailty with clearly documented end-stage PD).

Baseline characteristics

The baseline characteristics of the 200 included patients are presented in table 1. While the majority of patients had mild PD at baseline, around a quarter were already dependent on the help of others for ADL (S&E <80%) at diagnosis, which reflects the elderly nature of the cohort.

Table 1 Baseline characteristics of the included patients with Parkinson's disease

Characteristic	n	Median value/number with characteristic
Mean age in years (SD)	200	72.1 (10.3, range 37–90)
Male gender (%)	200	118 (59)
Symptom duration (months)	199	14 (9 to 25, range 2–121)
UPDRS (I)	166	1 (0 to 2)
UPDRS (II)	167	10 (6 to 14)
UPDRS (III)	200	24 (16 to 32)
Total UPDRS	165	37 (25 to 49)
H-Y stage ≥3.0 (%)	199	52 (26)
S&E <80% (%)	200	50 (25)
GDS-15 ≥5 (%)	156	40 (26)
MMSE	166	29 (27 to 29)
Barthel ADL index	158	20 (18 to 20)

Data are medians with IQRs unless otherwise stated. The 'n' column indicates the number of patients who contributed data in each case.

GDS-15, 15-item Geriatric Depression Scale; H-Y, Hoehn-Yahr stage; MMSE, Mini-mental state examination; S&E, Schwab and England Activities of Daily Living Scale; UPDRS, Unified Parkinson's Disease Rating Scale.

Treatment analysis

The amount of dopamine replacement therapy taken by patients increased over the 3-year follow-up: median LEDD 0 mg (IQR 0–0) at diagnosis, 240 mg (IQR 0–400) at the first annual follow-up, 300 mg (IQR 100–450) at the second, and 320 mg (IQR 230–500) at the third annual follow-up. Of those who successfully completed the third annual follow up (n=162), 136 (83%) were on dopamine replacement therapy at that time, of whom 24 were on dopamine agonist monotherapy and the remainder levodopa alone or in combination with an agonist or MAO-B inhibitor. Patients not on replacement therapy mostly had milder tremor-dominant disease, and decided against early treatment.

Survival analysis

Figure 2 illustrates the Kaplan–Meier survival curves for survival from all-cause mortality, PD-related mortality and time to 'death or dependence'. The mean survival time from all-cause mortality was shorter (6.7 (95% CI 6.1 to 7.3) years) than that from PD-related mortality (8.1 (95% CI 7.5 to 8.7) years). However, the time until patients reached 'death or dependence' was even shorter, with a mean survival time of only 3.6 (95% CI 3.1 to 4.2) years. The stepwise decline seen in the 'death or dependence' curve reflects that the S&E, which defined the development of dependency, was mostly collected at each annual follow-up.

Change in binary categorical variables over 3 years

Table 2 shows there was a statistically significant increase in the percentage of patients rated negatively in each binary categorical variable over 3 years. The H-Y change became non-significant (p=0.24) in the missing data sensitivity analysis. However, statistically, the most significant change (p=0.027), and the change of greatest magnitude (38.4%), was seen in the composite variable of 'dead or dependent' as defined by S&E <80%.

Sample size calculations

Table 3 shows that using either all-cause mortality or PD-related mortality as the trial outcome for the three-year RCT would

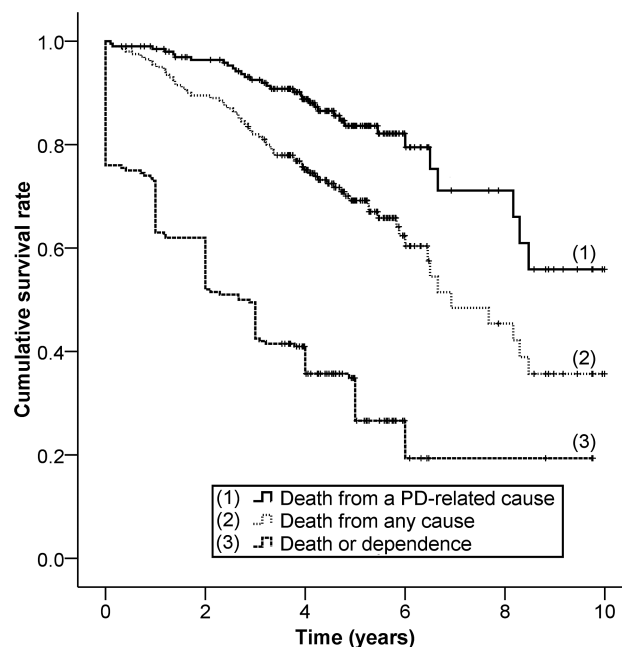


Figure 2 Kaplan–Meier curves for different outcomes in PINE study patients with Parkinson's disease.

Table 2 Percentage of patients with each binary outcome at each annual follow-up over three years

Binary outcome category	Annual follow-up appointment				Overall p Value
	Baseline n (%)	Year 1 n (%)	Year 2 n (%)	Year 3 n (%)	
H-Y stage ≥ 3	0/147 (0)	16/132 (12.1)	30/134 (22.4)	28/127 (22.0)	0.047
S&E <80%	0/151 (0)	21/144 (14.6)	34/136 (25.0)	37/130 (28.5)	0.045
'Dead or dependent'	0/151 (0)	28/151 (18.5)	48/150 (32.0)	58/151 (38.4)	0.027

The 'n' column indicates the number of patients in the stated binary category compared with the total number of patients who contributed data relevant to that binary variable at each time point. The % column indicates the actual (not modelled) percentage of patients in the stated binary category at each annual follow-up. The overall P values, derived from generalised estimating equations, relate to the significance of the change in each variable over three years. H-Y, Hoehn-Yahr scale; S&E, Schwab and England Activities of Daily Living Scale.

require large sample sizes, even when examining for a large (30%) relative risk reduction. The three other outcome variables required smaller sample sizes, with the smallest being for the composite variable of 'dead or dependent' (277–364 in each arm depending on power). However, the sample size using this outcome increased significantly when more modest treatment effects (20% relative risk reduction) and lower rates of outcomes events (20%) were expected (table 3). Given that it might be argued that combining death and mild degrees of dependency is inappropriate, the sample size for detecting death or more major dependency (S&E $\leq 50\%$) was also calculated using the number who were dead or had S&E $\leq 50\%$ at 3 years. Significantly higher sample sizes were again required than when the higher cut-off value (S&E <80%) was used (table 3).

DISCUSSION

Our data show that a composite binary measure of 'death or dependency' as opposed to 'alive and independent' is potentially a simple and feasible new outcome measure for clinical trials in PD.

Neuroprotective trials in PD have struggled to separate out symptomatic effects of potential therapeutic agents from true disease-modifying effects. Limitations in trial designs aimed at detecting a neuroprotective effect (eg, the measurement of clinical outcomes following a wash-out period,^{13 14} delayed-start trial designs,¹⁵ and only including de novo patients who are not on dopamine replacement therapy) have restricted their use. At present, the most robust way of detecting a neuroprotective effect of a novel therapeutic agent would be to conduct a long-term follow-up study in which sustained divergence in clinically important outcomes between placebo and actively treated patients would suggest a neuroprotective rather than a symptomatic effect. In a purely symptomatic effect, the improvement in outcomes would be expected to be stable or diminish over time.

However, long-term follow-up studies have several disadvantages. They are expensive and may take many years until results are available. Furthermore, high drop-out rates may be encountered, particularly if the coadministration of symptomatic treatment is prohibited, which may make interpretation of results difficult or impossible. For example, after 7 years of follow-up in a study comparing clinical outcomes in patients treated with selegiline versus those treated with placebo, only 35% of participants initially randomised to treatment were still in the study.¹⁶

Several of these problems would be overcome in a three-year trial that used 'death or dependency' as the primary outcome. First, symptomatic treatment would be permitted in addition to the study drug as, even on such treatment, we have shown significant numbers of patients develop the outcome of interest. This will also limit drop-outs. Secondly, by limiting follow-up to three years, we have hopefully struck a balance between a significantly long enough period of follow-up to detect a neuroprotective effect, and a realistic period of follow-up, financially, and in terms of patient compliance.

Our new composite measure of 'death or dependency' using a S&E cut-off of 80% has the advantage of being more clinically relevant to patients than a change in impairment measured by the motor UPDRS, and would be expected to change in response to an effective neuroprotective drug. It also avoids the need to censor data from patients who die during a study, thereby maximising the power of that study. As a clinical outcome measure, we have demonstrated that it is sensitive to disease progression even with the coadministration of symptomatic dopamine replacement therapy. Our sample size calculations demonstrated that a three-year RCT using this composite measure would be feasible, requiring about 600–800 patients per arm, depending on power, assuming a more realistic but still clinically valuable 20% relative risk reduction, and the same rate of death or dependency (38%) as we found at 3 years. Even if

Table 3 Sample sizes required to appropriately power RCTs of three years duration in patients with idiopathic PD

Primary outcome measure	Event rate in control arm (%)	Relative risk reduction (%)	Absolute risk reduction (%)	Sample size in each study arm (80% power)	Sample size in each study arm (90% power)
PD-related mortality	7.5	30	2.25	1938	2564
All-cause mortality	18.0	30	5.40	734	969
H-Y stage ≥ 3	22.0	30	6.60	577	762
S&E <80%	28.5	30	8.55	417	549
Dead or dependent (S&E <80%)	38.4	30	11.52	277	364
Dead or dependent (S&E <80%)	38.4	20	7.68	627	830
Dead or dependent (S&E <80%)	20.0	20	4.00	1497	1987
Dead or dependent (S&E $\leq 50\%$)	22.6	30	6.78	559	737

Sample sizes were calculated with a 5% level of significance. H-Y, Hoehn-Yahr scale; S&E, Schwab and England Activities of Daily Living Scale.

very large trials were required (eg, 1500–2000 per arm to detect, to a reduction in poor outcome from 20% to 16%), the simplicity and ease of collection of the measure would also allow it to be used in simple pragmatic trials, perhaps using large prospective patient registries.¹⁷

There are a number of potential limitations in our trial design. This was a small study and needs validation in larger cohorts. Patients also needed to be independent in ADL at baseline, which meant 25% were excluded (this proportion would be lower in younger cohorts). However, this is already the case in neuroprotection trials where patients usually have early stage disease, implying a reasonable proportion of intact neurons remain which may be salvageable. However, if required, an adaptation in the outcome would also allow dependent patients to be included, that is, including deterioration to a higher grade of dependency as an outcome in those already dependent. This has been used in stroke trials.

Our data come from a community-based incident cohort of PD. While this makes it more generalisable to the totality of PD patients in the community, it will be less applicable to hospital-based cohorts of people with early PD, who are the types of patient recruited into RCTs. In particular, our patients were older than most patients recruited into RCTs.¹⁸ The proportion of patients dead or dependent at 3 years will vary by age of onset, and so much larger (or longer) trials would be required if younger patients were recruited. For example, we showed that reducing the outcome event rate to 20% at 3 years would require very large trials to detect 20% relative risk reductions. However, recruitment of older people into PD trials should be encouraged to improve generalisability and clinical relevance, given that PD is predominantly a disease of the elderly.

Some of the outcomes (deaths and loss of independence in ADL) were the result of comorbid disease, not PD, and so would not be expected to be prevented by neuroprotective treatments. Restricting outcomes to those thought to be due solely to PD might seem sensible, but would add a degree of subjectivity (particularly when deciding about loss of independence) and would also dramatically increase sample sizes as shown by the data for mortality alone. Moreover, trials using impairment scores, such as the motor UPDRS, are also faced with a similar problem as it remains unclear how to score these in the presence of significant comorbidity.¹⁹

Finally, it might be argued that combining death and any dependency in ADL is inappropriate because the two are quite different. However, this outcome is already widely used in major stroke trials⁸ and meta-analyses,²⁰ and other fields of medicine frequently use endpoints composed of fatal and non-fatal outcomes, for example, cardiovascular trials often combine cardiovascular death along with non-fatal myocardial infarction and stroke. Any effective neuroprotective agent would be expected to reduce PD mortality and delay progression to dependency in survivors, so clinically it makes sense to combine the two and, as highlighted already, it allows all patients randomised to be included in the analyses at every time point. Using more severe dependency (eg, S&E ≤ 50) in combination with death can be done (and again is used in stroke trials to measure ‘devastating outcome’),²¹ but we showed it does increase the sample size significantly, even for large treatment effects (table 3).

Although ‘death or dependency’ looks promising as an outcome, further work is required before it can be widely adopted. In particular, although we used the S&E scale and clearly defined what we regarded as becoming dependent (<80%) in the PINE study, the overall validity and reproducibility of the S&E scale has been poorly studied. There are many

other disability scales that could be used, which may be better. However, we would caution against the use of the H-Y scale because it mixes impairment and disability, and there is no clear-cut point where someone becomes dependent. It would also be crucial to define dependency more clearly. For example, dependency in instrumental ADL is different to dependency in basic ADL, and may be more sensitive to early change in PD.

In conclusion, we believe that ‘death or dependency’ is a potentially useful new simple outcome measure for PD trials, particularly neuroprotective trials, but more work needs to be done to confirm this.

Acknowledgements We thank Susan Kilpatrick for secretarial support, Clare Harris and Hazel Forbes for helping with patient assessments, and Katie Wilde for maintaining the PINE study database. This article presents independent research partially funded by the NIHR under its Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-0707–10124). The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Contributors CC is chief investigator for the PINE study and conceived and designed this study. DM and AP extracted and checked the data, performed the analysis and wrote the first draft. SF helped with the design and statistical analysis. JZ is the chief investigator of the NIHR grant which funded DM. This grant aims to improve trial methods in neurodegenerative diseases including improving outcome measures. All authors commented on the draft manuscript and agreed on the final manuscript.

Funding DM was funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-0707–10124). The PINE study was funded by Parkinson’s UK (grant numbers G0502 and G0914), the BMA Doris Hillier Award, NHS Grampian Endowments, RS MacDonald Trust and SPRING (Special Parkinson’s Research Interest Group).

Competing interests None.

Ethics approval Multi-centre Research Ethics Committee for Scotland, Committee A.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The PINE data are available for sharing by approaching Dr Counsell.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- 1 World Health Organization. International classification of impairments, disabilities and handicaps. Geneva, Switzerland: World Health Organization, 1980.
- 2 Fahn S, Eton RL, UPDRS Development Committee. The Unified Parkinson’s Disease Rating Scale. In: Fahn S, Marsden CD, Goldstein M, Calne DB, eds. Recent Developments in Parkinson’s Disease. Volume 2. New Jersey: Florham Park, Macmillan Healthcare Information, 1987:153–63.
- 3 Schwab R, England A. Projection technique for evaluating surgery in Parkinson’s disease. In: Gillingham F, Donaldson I, eds. Third Symposium on Parkinson’s Disease. Edinburgh: E&S Livingstone, 1969:152–7.
- 4 Jenkinson C, Fitzpatrick R, Peto V, et al. The Parkinson’s Disease Questionnaire (PDQ-39): development and validation of a Parkinson’s disease summary index score. *Age Ageing* 1997;26:353–7.
- 5 ICF. Towards a common language for functioning, disability and health. The International Classification of Functioning, Disability and Health. Geneva: World Health Organisation, 2002. <http://www.who.int/classifications/icf/training/icfbeginnersguide.pdf> (accessed 12 Jan 2014).
- 6 Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke* 2007;38:1091–6.
- 7 Mahoney F, Barthel D. Functional evaluation: the Barthel index. *Md State Med J* 1965;14:61–5.
- 8 International Stroke Trial Collaborative Group. The International Stroke Trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19435 patients with acute ischaemic stroke. *Lancet* 1997;349:1569–81.
- 9 Caslake R, Taylor K, Scott N, et al. Age-, gender-, and socioeconomic status-specific incidence of Parkinson’s disease and parkinsonism in North East Scotland: The PINE study. *Park Rel Dis* 2013;19:515–21.

- 10 Hughes AJ, Daniel SE, Kilford L, *et al.* Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992;55:181–4.
- 11 Tomlinson CL, Stowe R, Patel S, *et al.* Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov Disord* 2010;25:2649–53.
- 12 Elm JJ, Goetz CG, Ravina B, *et al.* A responsive outcome for Parkinson's disease neuroprotection futility studies. *Ann Neurol* 2007;57:197–203.
- 13 Clarke CE. A "cure" for Parkinson's disease: can neuroprotection be proven with current trial designs? *Mov Disord* 2004;19:491–8.
- 14 Rascol O. "Disease-modification" trials in Parkinson disease: target populations, endpoints and study design. *Neurology* 2009;72:S51–8.
- 15 Clarke CE. Are delayed-start design trials to show neuroprotection in Parkinson's disease fundamentally flawed? *Mov Disord* 2008;23:784–9.
- 16 Pålhagen S, Heinonen E, Häggglund J, *et al.* Selegiline slows the progression of the symptoms of Parkinson disease. *Neurology* 2006;66:1200–6.
- 17 Lauer MS, D'Agostino RB Sr. The randomized registry trial—the next disruptive technology in clinical research? *N Engl J Med* 2013;369:1579–81.
- 18 Mitchell SL, Sullivan EA, Lipsitz LA. Exclusion of elderly subjects from clinical trials for Parkinson disease. *Arch Neurol* 1997;54:1393–8.
- 19 Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Mov Disord* 2003;18:738–50.
- 20 Wardlaw JM, Murray V, Berge E, *et al.* Thrombolysis for acute ischaemic stroke. *Cochrane Database Syst Rev* 2009;4: CD000213. doi:0.1002/14651858.CD000213.pub2.
- 21 Hofmeijer J, Kappelle LJ, Algra A, *et al.* Surgical decompression for space-occupying cerebral infarction (the Hemicraniectomy After Middle Cerebral Artery infarction with Life-threatening Edema Trial [HAMLET]): a multicentre, open, randomised trial. *Lancet Neurol* 2009;8:326–33.