PRION-1 scales analysis supports use of functional outcome measures in prion disease

S. Mead, PhD

M. Ranopa, MSc

G.S. Gopalakrishnan, PhD

A.G.B. Thompson, MRCP

P. Rudge, FRCP

S. Wroe, MD

A. Kennedy, MD

F. Hudson, BA

A. MacKay, FRCP

J.H. Darbyshire, FRCP

J. Collinge, FRS

A.S. Walker, PhD

Address correspondence and reprint requests to Prof. J. Collinge, MRC Prion Unit, Department of Neurodegenerative Disease, UCL Institute of Neurology, and National Prion Clinic, National Hospital for Neurology and Neurosurgery, UCLH NHS Trust, Queen Square, London
j.collinge@prion.ucl.ac.uk

ABSTRACT

Objectives: Human prion diseases are heterogeneous but invariably fatal neurodegenerative disorders with no known effective therapy. PRION-1, the largest clinical trial in prion disease to date, showed no effect of the potential therapeutic quinacrine on survival. Although there are several limitations to the usefulness of survival as an outcome measure, there have been no comprehensive studies of alternatives.

Methods: To address this we did comparative analyses of neurocognitive, psychiatric, global, clinician-rated, and functional scales, focusing on validity, variability, and impact on statistical power over 77 person-years follow-up in 101 symptomatic patients in PRION-1.

Results: Quinacrine had no demonstrable benefit on any of the 8 scales ($p > 0.4$). All scales had substantial numbers of patients with the worst possible score at enrollment (Glasgow Coma Scale score being least affected) and were impacted by missing data due to disease progression. These effects were more significant for cognitive/psychiatric scales than global, clinician-rated, or functional scales. The Barthel and Clinical Dementia Rating scales were the most valid and powerful in simulated clinical trials of an effective therapeutic. A combination of selected subcomponents from these 2 scales gave somewhat increased power, compared to use of survival, to detect clinically relevant effects in future clinical trials of feasible size.

Conclusions: Our findings have implications for the choice of primary outcome measure in prion disease clinical trials. Prion disease presents the unusual opportunity to follow patients with a neurodegenerative disease through their entire clinical course, and this provides insights relevant to designing outcome measures in related conditions. *Neurology*® **2011;77:1674–1683**

GLOSSARY

ADAS-cog = cognitive component of the Alzheimer's Disease Assessment Scale; **BPRS** = Brief Psychiatric Rating Scale; **CDR-SB** = Clinical Dementia Rating Scale Sum of Boxes; **GCS** = Glasgow Coma Scale; **GIC** = Global Impression of Change; **IPD** = inherited prion disease; **MMSE** = Mini-Mental State Examination; **sCJD** = sporadic Creutzfeldt-Jakob disease; **vCJD** = variant Creutzfeldt-Jakob disease.

Prion diseases are invariably fatal neurodegenerative disorders with no therapy that is known to alter the natural history in man.[1] They comprise an etiologically and clinically heterogeneous group, with genetic, acquired, or sporadic etiologies.[2] Despite promising developments in basic science and experimental therapeutics, several challenges are evident in the development of clinical trial methodology. First, prion diseases are rare, with the implication that recruitment criteria need to be broad in order to achieve acceptable trial size in a reasonable geographic area or time scale. Second, the clinical phenotype is highly variable within and between each etiologic class. Third, there are strong influences on UK family members to reject randomization in trials which include the possibility of allocation to placebo.[3,4] As a result, carefully documented historical control data may be important alternatives or adjuncts to a randomized placebo-controlled trial.

Supplemental data at
www.neurology.org

Supplemental Data

Any trial is dependent upon the relevance of its outcome measurement instruments. Mortality has been the choice of primary outcome in several prion disease trials/studies, but this is problematic for several reasons.[1,5,6] The immediate cause of death in prion disease, usually sepsis or respiratory failure, can be influenced by several factors unrelated to underlying disease progression, such as use of antibiotics to treat life-threatening infections, and supplementary nutrition. As the aim is to halt or reverse neuropsychiatric impairments with a therapeutic intervention, a more relevant trial outcome measure should consider human prion disease progression directly.

METHODS PRION-1 study design. PRION-1 was an open-label patient-preference trial in which patients were offered a choice between taking quinacrine, not taking quinacrine, or being randomized to immediate quinacrine or quinacrine deferred for 24 weeks (appendix e-1 on the *Neurology*® Web site at www.neurology.org). The objective was to obtain data on the effect of quinacrine in human prion disease, from a randomized comparison where acceptable and otherwise from observational comparisons. Full details of randomization, primary outcome, and adverse events in the pilot and main trial phases have been published.[3]

Patients were assessed at baseline (enrollment) with follow-up at 1, 2, 4, and 6 months and then 3-monthly including a standardized clinical examination, the Glasgow Coma Scale (GCS),[7] Mini-Mental State Examination (MMSE),[8] cognitive component of the Alzheimer's Disease Assessment Scale (ADAS-cog),[9] Brief Psychiatric Rating Scale (BPRS),[10] Clinical Dementia Rating Scale Sum of Boxes (CDR-SB),[11] Global Impression of Change (GIC),[12] Rankin scale,[13] and the Barthel Activities of Daily Living Index[14] (see appendix e-1).

Standard protocol approvals, registrations, and patient consents. The PRION-1 trial (ISRCTN 06722585) was approved by the Eastern Multicenter Research Ethics Committee. Written informed consent was obtained from all patients (or guardians/carers according to the Mental Capacity Act 2005, UK) participating in the study.

Statistical analysis. Changes in neurologic rating scales over time to March 30, 2007 (end of PRION-1), death, or loss to follow-up were estimated using mixed models, allowing individual random effects for intercept (enrollment value) and slope (decline). For more detail, sensitivity analyses, and subgrouping, see appendix e-1.

RESULTS Recruitment and follow-up summary. The headline results of PRION-1 have been published.[3] Here we describe the pertinent recruitment and survival data needed to understand the scales analysis. A total of 107 patients were recruited in the pilot and main phases of the PRION-1 trial. A total of 45 had sporadic Creutzfeldt-Jakob disease (sCJD), 43 inherited prion disease (IPD), and 19 acquired prion diseases (17/19 variant CJD) (note that one fewer patient had variant Creutzfeldt-Jakob disease [vCJD] than reported in[3] due

to a late reclassification as IPD). Diagnosis was confirmed by tissue biopsy, at autopsy, or by detection of a genetic mutation in 78/107. Seventy-eight of 107 patients died during the PRION-1 trial (to March 30, 2007). Including additional data from patients still alive and in follow-up on March 30, 2007, from the National Prion Monitoring Cohort to date, in total 29/107 (27%) died within 4 weeks of enrollment, 34 (32%) within 6 weeks, 45 (42%) within 16 weeks, and 75 (70%) within 2 years. See appendix e-2 for details of follow-up and data to support lack of efficacy of quinacrine.

Cross-sectional performance of scales at recruitment of symptomatic patients. Table 1 shows completion rates for the various scales at recruitment for the 101/107 (94%) patients with symptomatic prion disease at enrollment, overall and according to Rankin. Almost universal completion rates were seen for the global clinician-rated scales (Rankin, GIC) reflecting their simplicity. Very high completion rates were also obtained for the Barthel and GCS (taking into consideration that the Barthel was not used in the pilot phase[3] and, following the protocol, GCS was not formally recorded if the MMSE was ≥10 but was rather imputed as 15 in main analyses). The CDR-SB was reasonably well completed but required the participation of a carer who may not have been available at the time. Cognitive scales (MMSE, ADAS-cog) were completed in around half and one-third, respectively, either because of aphasia or advanced stage of disease. The BPRS was only completed in 28/76 participants recruited during the main phase because of the need for a lengthy patient interview.

Figure 1A illustrates the distribution of observed scores at enrollment, categorizing scores as being at the maximum possible (ceiling, best performance), at the minimum possible (floor, worst performance), or within intermediate terciles (low scores are worst for all rating scales; see appendix e-1). Despite low completion rates, most of those able to complete the cognitive and psychiatric scales scored in the upper terciles. These effects were present to a lesser extent for CDR-SB: in contrast, although Rankin was nearly universally completed, almost 40% of participants already registered the lowest score (floor) at enrollment. Overall, therefore, proportions at this lowest level or unable to complete were similar for the Rankin and CDR-SB. The Barthel, GIC, and GCS had the smallest proportion of patients already registering the worst possible score at enrollment. The only scale to show any significant ceiling effect (i.e., several patients with the best possible score at enrollment) was the GCS.

Completeness of rating scale observations over time. In order to compare interventions in terms of changes in neurologic rating scales, it is imperative

Table 1 Characteristics of participants and completion of neurologic rating scales at enrollment into PRION-1 and subsequently

	All symptomatic participants	Rankin at enrollment[a]		
		Mild to moderate disability (2–3)	Moderately severe disability (4)	Severe disability (5)
No. of patients	101	30	31	40
Age at first symptoms, y (IQR)	55 (36–65)	37 (30–52)	55 (33–66)	62 (56–71)
Age at enrollment, y (IQR)	58 (40–66)	40 (33–54)	60 (35–67)	63 (57–51)
No. of patients (%)				
sCJD	45 (44.55)	2 (6.67)	12 (38.71)	31 (77.50)
iCJD	2 (1.98)	0 (0.00)	0 (0.00)	2 (5.00)
vCJD	17 (16.83)	7 (23.33)	9 (29.03)	1 (2.50)
IPD	37 (36.63)	21 (70.00)	10 (32.26)	6 (15.00)
MMSE: assessed at enrollment	52 (51.49)	29 (96.67)	19 (61.29)	4 (10.00)
Total occasions measured (%): 0	49 (48.51)	1 (3.33)	12 (38.71)	36 (90.00)
Total occasions measured (%): ≥1	52 (51.49)	29 (96.67)	19 (61.29)	4 (10.00)
ADAS-cog: assessed at enrollment	31 (30.69)	23 (76.67)	7 (22.58)	1 (2.50)
Total occasions measured (%): 0	65 (64.36)	4 (13.33)	22 (70.97)	39 (97.50)
Total occasions measured (%): ≥1	36 (35.64)	26 (86.67)	9 (29.03)	1 (2.50)
BPRS[b]: assessed at enrollment (denominator 78 main trial participants)	28/78 (35.90)	16/18 (88.89)	10/24 (41.67)	2/36 (5.56)
Total occasions measured (%): 0	62 (61.39)	4 (13.33)	20 (64.52)	38 (95.00)
Total occasions measured (%): ≥1	39 (38.61)	26 (86.67)	11 (35.48)	2 (5.00)
GIC: assessed at enrollment	99 (98.02)	29 (96.67)	30 (96.77)	40 (100.00)
Total occasions measured (%): 0	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Total occasions measured (%): ≥1	100 (99.01)	30 (100.00)	30 (96.77)	40 (100.00)
CDR-SB: assessed at enrollment	69 (68.32)	27 (90.00)	25 (80.65)	17 (42.50)
Total occasions measured (%): 0	30 (29.70)	1 (3.33)	6 (19.35)	23 (57.50)
Total occasions measured (%): ≥1	71 (70.30)	29 (96.67)	25 (80.65)	17 (42.50)
Rankin: assessed at enrollment	99 (98.02)	28 (93.33)	31 (100.00)	40 (100.00)
Total occasions measured (%): 0	1 (0.99)	1 (3.33)	0 (0.00)	0 (0.00)
Total occasions measured (%): ≥1	100 (99.01)	29 (96.67)	31 (100.00)	40 (100.00)
Barthel[b]: assessed at enrollment	89 (88.12)	24 (80.00)	26 (83.87)	39 (97.50)
Total occasions measured (%): 0	7 (6.93)	1 (3.33)	5 (16.13)	1 (2.50)
Total occasions measured (%): ≥1	94 (93.07)	29 (96.67)	26 (83.87)	39 (97.50)
GCS[c]: assessed at enrollment	74 (73.27)	16 (53.33)	22 (70.97)	36 (90.00)
Total occasions measured (%): 0	1 (0.99)	1 (3.33)	0 (0.00)	0 (0.00)
Total occasions measured (%): ≥1	100 (99.01)	29 (96.67)	31 (100.00)	40 (100.00)

Abbreviations: ADAS-cog = cognitive component of the Alzheimer's Disease Assessment Scale; BPRS = Brief Psychiatric Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; GCS = Glasgow Coma Scale; GIC = Global Impression of Change; IPD = inherited prion disease; IQR = interquartile range; MMSE = Mini-Mental State Examination; sCJD = sporadic Creutzfeldt-Jakob disease; vCJD = variant Creutzfeldt-Jakob disease.
[a] One patient with MMSE = 15 and moderately ill on GIC did not have Rankin formally recorded at enrollment but is included in the moderate disability Rankin group.
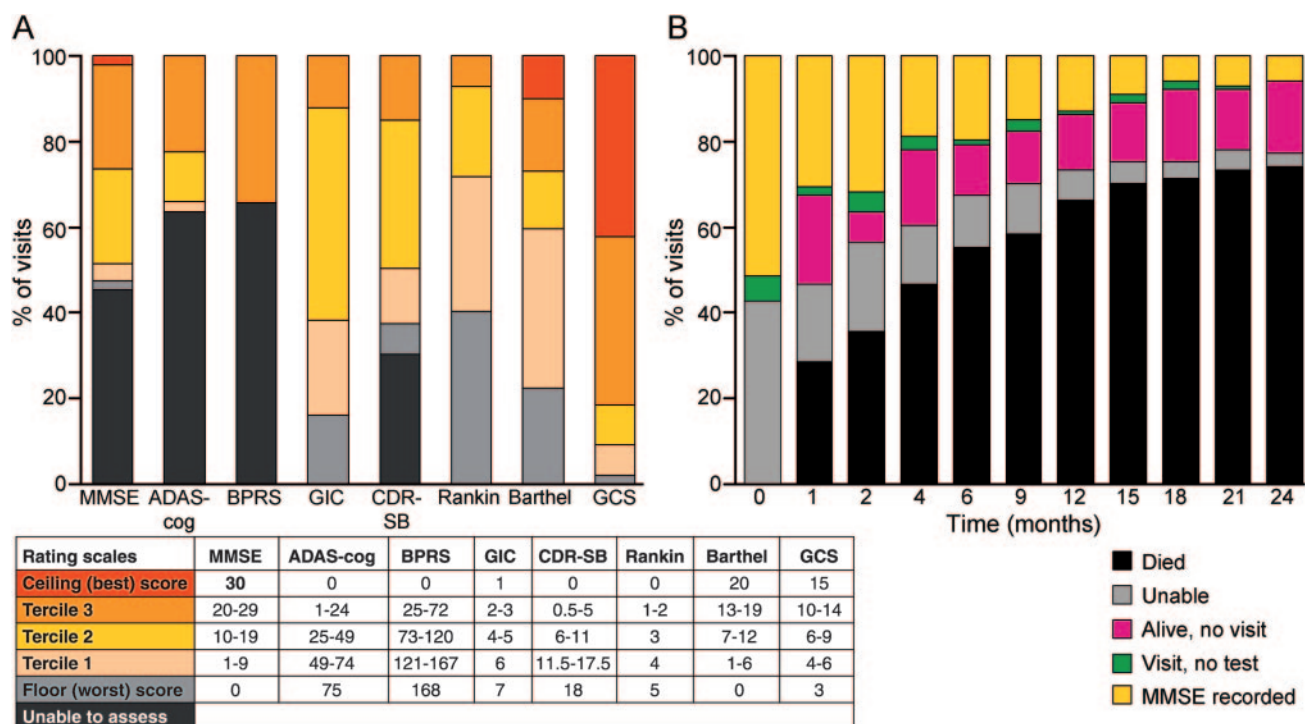[b] BPRS and Barthel were not done routinely during pilot phase of PRION-1. Barthel measurements were available for some of the 23 pilot patients.
[c] GCS imputed as 15 for all patients with enrollment MMSE ≥10 following the PRION-1 protocol (see appendix e-1).

that as many of the scheduled assessments as possible are made. This requires attendance at scheduled visits and completion of each test, which may become increasingly problematic as the battery increases in size. In PRION-1, 10%–20% of scheduled visits were not made in patients still known to be alive during the first 36 weeks, declining thereafter to about 5%.

Early in the trial most missed visits were due to patients with sporadic CJD becoming moribund. Unsurprisingly, completion of all scales was also strongly impacted by mortality (figure 1B). However, rating scales were not always completed in patients attending visits, for similar reasons as at enrollment. Table 1 shows that between 50% and 75% of the individuals enrolled had

Figure 1   Distribution of scores at enrollment and completion of MMSE over time

Rating scales	MMSE	ADAS-cog	BPRS	GIC	CDR-SB	Rankin	Barthel	GCS
Ceiling (best) score	30	0	0	1	0	0	20	15
Tercile 3	20-29	1-24	25-72	2-3	0.5-5	1-2	13-19	10-14
Tercile 2	10-19	25-49	73-120	4-5	6-11	3	7-12	6-9
Tercile 1	1-9	49-74	121-167	6	11.5-17.5	4	1-6	4-6
Floor (worst) score	0	75	168	7	18	5	0	3
Unable to assess								

(A) Distribution of observed scores at enrollment by maximum possible score (ceiling = best), minimum possible score (floor = worst), and intermediate terciles (in 101 symptomatic patients). "Unable" means this was given as a reason for noncompletion on the enrollment form. (B) Completion of MMSE, by months after enrollment (in 101 symptomatic patients). This is shown to illustrate the impact of survival and advanced clinical state on a cognitive rating scale in PRION-1. ADAS-cog = cognitive component of the Alzheimer's Disease Assessment Scale; BPRS = Brief Psychiatric Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; GCS = Glasgow Coma Scale; GIC = Global Impression of Change; MMSE = Mini-Mental State Examination.

only 0 or 1 measurement for the CDR-SB and BPRS, whereas 60%–65% of all patients, and even 32%–35% of severely affected patients, had 2 or more measurements of Barthel and GCS.

Change in rating scales over time. Some general comments can be made about the patterns of decline recorded in PRION-1, although detailed analysis of factors affecting rates of decline are beyond the scope of this manuscript. Individual patient trajectories are shown for all scales (figure 2). First, as expected, rates of decline were highly heterogeneous. Second, patients tend to follow linear patterns of decline, at least until the most advanced stages of disease. At this point several patients appear to reach a preterminal plateau, although it is unclear whether this observation is related to the insensitivity of scales at the advanced stages of disease, or reflects a true disease phenomenon. In figure 2 (see parts D, E, F, I), this phenomenon is illustrated by connecting the time of death (red point) denoted at the worst possible score on each scale with the measured value at last visit (orange line). Finally, for a reasonable proportion of patients, the last observed measurement before death is relatively far above the worst possible score on the scale (see figure 2, parts A, B, C, H)—this is a consequence of rapidity of decline in late-stage dis-

ease not covered by the visit schedule, inability to complete the scales with more severe disease (hence absence of further measurements), and the fact that death from human prion disease is often not specifically related to underlying neuropsychiatric decline (e.g., through aspiration pneumonia).

Performance of models used to describe change in rating scales over time. We estimated annual rates of decline overall, and in fast (sCJD)/medium (vCJD and some IPD, see appendix e-1)/slow progressors (Gerstmann-Straussler-Scheinker), mild to moderate disease stratified by Rankin (2–3 or 2–4), using both observed data only (table 2) and including all symptomatic patients but censoring after the first observed worst possible score for each scale and imputing death as the worst possible score (figure 3). In different disease categories we found no evidence of differences in psychiatric/cognitive scores (MMSE, ADAS-cog, BPRS) at enrollment ($p > 0.1$) although all functional and clinician-rated scales were worst in "fast" and best in "slow" progressors ($p < 0.0001$, table 2). However, we found evidence for significant variation in the annual rate of decline after enrollment both by severity and disease category for all scales ($p < 0.01$, table 2) except BPRS, which had problems with completion as noted

Figure 2   Longitudinal profiles of patients according to cognitive (MMSE, ADAS-cog, BPRS), functional (Barthel), and clinician-rated scales (Rankin, CDR-SB), GIC, and combined Barthel/CDR-SB (combination)
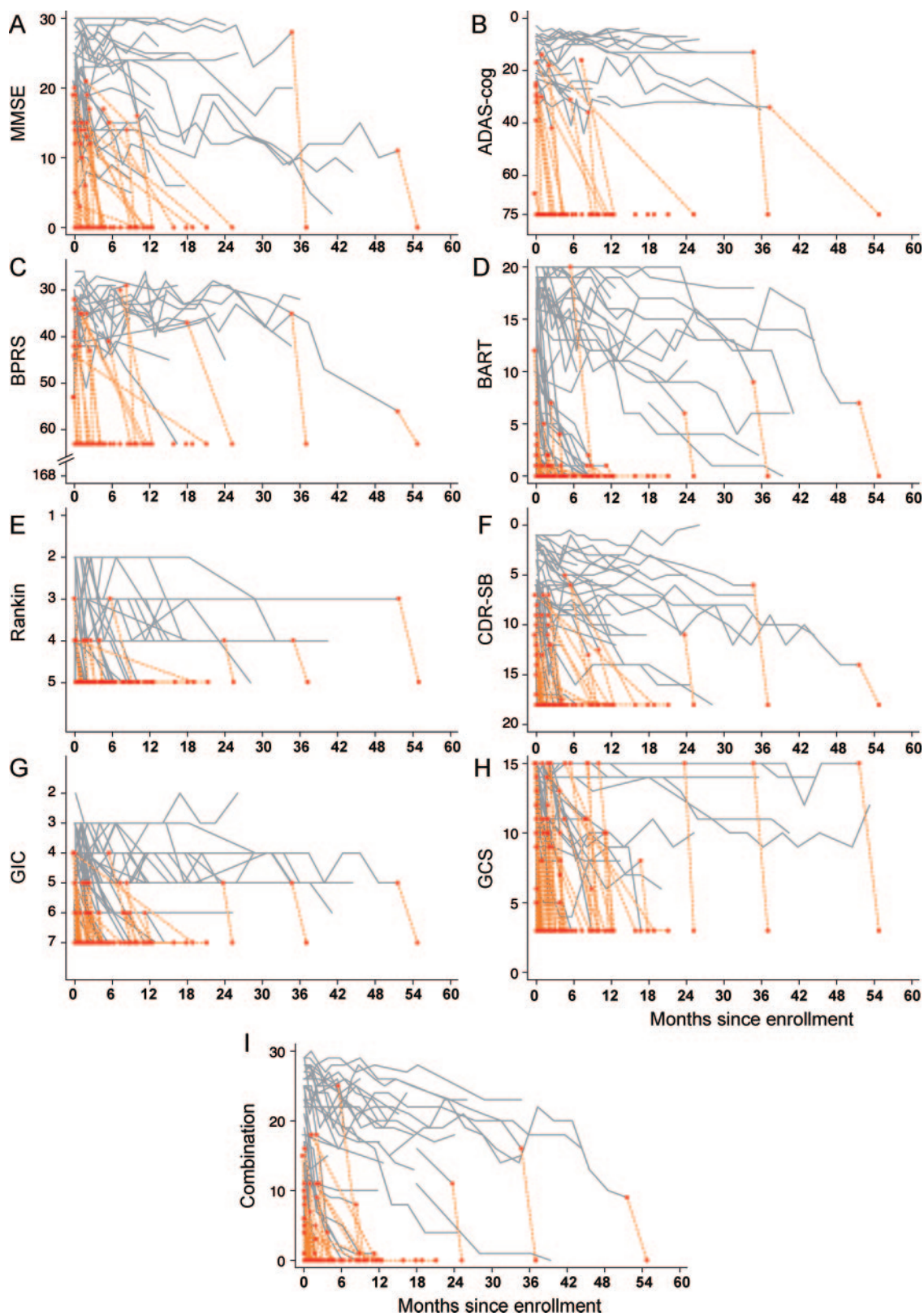
(A–I) Gray lines indicate individual patient trajectories. Red points and orange lines connect the last patient scale measurement with the time of death (allocated the worst possible score, except for BPRS, where the worst score observed during PRION-1 [64] was used rather than the worst possible score for the scale [168]; see appendix e-1). These are shown to illustrate the heterogeneity of individual patient trajectories. ADAS-cog = cognitive component of the Alzheimer's Disease Assessment Scale; BART = Barthel; BPRS = Brief Psychiatric Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; GCS = Glasgow Coma Scale; GIC = Global Impression of Change; MMSE = Mini-Mental State Examination.

Table 2 Baseline, annual rates of decline, and power in fast (sCJD)/medium (vCJD and other IPD)/slow progressors (Gerstmann-Straussler-Scheinker), stratified by Rankin (2–3)/(2–4) and imputed model

	All participants	Slow progressors	Medium progressors	Fast progressors	p^a	Rankin 2–3 at enrollment	Rankin 2–4 at enrollment	Imputation model
No.	101	34	22	45		30	61	101
MMSE: no. (%) with 1 or more values	52 (51)	29 (85)	14 (64)	9 (20)		29 (97)	48 (79)	99 (98)
Mean at enrollment	17.79	19.23	17.39	15.04	0.296	20.00	18.24	10.73
Mean annual decline (SD)	3.60 (5.41)	2.64 (3.22)	19.73 (21.56)	18.35 (10.41)	<0.0001	3.23 (4.10)	3.56 (5.20)	13.33 (17.11)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	286	188	150	42		202	270	208
ADAS-cog: no. (%) with 1 or more values	36 (36)	24 (71)	8 (37)	4 (9)		26 (87)	35 (57)	97 (96)
Mean at enrollment	23.28	21.19	28.31	23.75	0.534	22.69	22.83	49.30
Mean annual decline (SD)	7.03 (11.61)	5.16 (9.08)	64.96 (51.72)	13.73 (17.80)	0.003	3.57 (6.38)	6.96 (11.52)	35.14 (61)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	344	390	80	212		402	344	384
BPRS: no. (%) with 1 or more values	39 (39)	23 (68)	9 (41)	7 (16)		26 (87)	37 (61)	97 (96)
Mean at enrollment	35.97	34.33	39.36	37.26	0.102	34.87	35.43	50.83
Mean annual decline (SD)	1.44 (5.60)	1.78 (4.35)	3.89 (56.19)	6.43 (13.65)	0.671	2.23 (4.40)	1.62 (5.35)	6.52 (11.50)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	1912	752	>2000	568		492	1370	392
GIC: no. (%) with any value	101 (100)	34 (100)	22 (100)	45 (100)		30 (100)	61 (100)	101 (100)
Mean at enrollment	5.19	4.33	5.20	5.84	<0.0001	4.03	4.50	5.38
Mean annual decline (SD)	1.05 (1.73)	0.52 (0.99)	2.56 (1.84)	1.22 (2.71)	<0.0001	1.05 (1.28)	1.35 (1.64)	1.76 (2.06)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	346	462	66	624		190	186	172
CDR-SB: no. (%) with 1 or more values	71 (70)	32 (94)	17 (77)	22 (49)		29 (97)	54 (89)	101 (100)
Mean at enrollment	9.56	6.75	11.19	12.39	<0.0001	2.76	8.88	11.74
Mean annual decline (SD)	4.95 (6.59)	2.40 (3.02)	11.66 (7.62)	7.93 (9.37)	<0.0001	4.62 (5.75)	5.37 (6.68)	8.99 (11.97)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	224	198	54	176		196	196	224
Rankin: number with 1 or more values	100 (99)	34 (100)	21 (95)	45 (100)		29 (97)	60 (98)	101 (100)
Mean at enrollment	4.12	3.37	4.14	4.67	<0.0001	2.95	3.54	4.17
Mean annual decline (SD)	0.65 (1.06)	0.43 (0.71)	1.44 (1.31)	0.75 (1.95)	0.005	0.66 (0.67)	1.05 (1.06)	0.79 (1.20)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	334	348	106	842		132	128	290
Barthel: no. (%) with 1 or more values	94 (93)	31 (91)	19 (86)	44 (98)		29 (97)	55 (90)	101 (100)
Mean at enrollment	6.65	13.09	6.60	2.27	<0.0001	15.79	10.97	6.50
Mean annual decline (SD)	5.39 (8.07)	3.81 (5.54)	12.02 (9.81)	5.06 (14.40)	0.004	5.29 (4.83)	7.51 (7.13)	7.60 (9.63)

—Continued

	All participants	Slow progressors	Medium progressors	Fast progressors	p[a]	Rankin 2–3 at enrollment	Rankin 2–4 at enrollment	Imputation model
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	282	266	84	1018		106	114	202
GCS: no. (%) with 1 or more values	100 (99)	34 (100)	21 (95)	45 (100)		29 (97)	60 (98)	101 (100)
Mean at enrollment	12.31	14.29	13.24	10.36	<0.0001	14.90	14.21	11.41
Mean annual decline (SD)	4.49 (7.19)	0.96 (3.80)	8.84 (5.86)	5.38 (8.92)	<0.0001	1.99 (4.07)	4.55 (6.61)	14.25 (13.70)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	324	1982	56	346		528	266	118
Barthel/CDR-SB combination scale: no. (%) with 1 or more values	65 (64)	29 (85)	15 (68)	21 (47)		29 (97)	49 (80)	101 (100)
Mean at enrollment	15.04	20.67	12.61	9.02	<0.0001	22.82	17.41	10.60
Mean annual decline (SD)	10.61 (9.51)	5.18 (4.60)	23.80 (12.26)	12.72 (14.11)	<0.0001	8.09 (6.81)	10.72 (9.13)	13.45 (12.64)
Total sample size to detect 50% reduction in the annual rate of decline between 2 groups[b]	102	100	34	156		90	92	112

Abbreviations: ADAS-cog = cognitive component of the Alzheimer's Disease Assessment Scale; BPRS = Brief Psychiatric Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; GCS = Glasgow Coma Scale; GIC = Global Impression of Change; IPD = inherited prion disease; MMSE = Mini-Mental State Examination; sCJD = sporadic Creutzfeldt-Jakob disease; vCJD = variant Creutzfeldt-Jakob disease.

[a] Test for difference between slow, medium, and fast progressors. Mean baseline and annual rates of decline were estimated from mixed models including observed data only (without imputation) with random intercept and slope for each participant. SD = sqrt(n) × SE (mean annual rate of decline).

[b] Sample size estimates were based on 80% power, and a 2-sided significance level 5%. The severity stratified models were restricted to patients with baseline Rankin 2–3 or 2–4. In the imputed model the worst score was imputed for death (except for BPRS, see appendix e-1) and any subsequent scores after a patient recorded the worst score were discarded.

above. As expected, the annual rate of decline was greater in "medium" compared to "slow" progressors. However, "fast" progressors (sCJD) had smaller estimated declines than "medium" progressors as a consequence of their very low scores at enrollment, meaning that further declines were limited and therefore estimates of rates of decline are unreliable. With this exception, these data support the behavior of the rating scales as expected, and emphasize the need to diagnose sCJD at earlier stages if such rating scales are to be used as outcome measures in this population. Imputing death as the worst possible score increased the rate of decline, by a much greater extent in those scales with lowest completion rates.

Estimates of sample size in a future clinical trial. Based on the estimated annual rate of decline for each rating scale and its SD from these mixed models, we estimated the total number of participants required in a 2 parallel-group randomized trial with similar measurement frequency to PRION-1 that would provide at least 80% power to detect a 50% relative reduction in annual rate of decline for that scale (2-sided $\alpha = 0.05$) (table 2). In terms of subgroups, restricting enrollment through inclusion criteria in future trial designs to only "medium" progressors, or to those with Rankin 2–3 or 2–4 at enrollment, provides some advantages in terms of smaller sample sizes to detect reduced declines (table 2), with "fast" progressors and those at most advanced disease stages clearly able to contribute less information (and potentially also with less ability to benefit from any intervention if substantial neurologic damage has already occurred). However, restricting eligibility clearly comes at the expense of slower recruitment, given the restricted patient population. If enrollment is unrestricted, imputing deaths as worst possible score when this has not previously been observed provides greater power than using observed data only, although it is a heuristic approach to the problems described above. For comparison, assuming 25% of patients survive 2 years,[3] then a total of 106 participants would be needed to detect an increase in survival from 25% to at least 50% (i.e., a relative 33% mortality reduction) with at least 80% power and 2-sided $\alpha = 0.05$. This is about the same as the smallest total number needed to detect 50% reduction in declines in the different rating scales.

Proposed combination of Barthel with subcomponents of CDR-SB. Based on the validity and good performance of Barthel and CDR-SB in the analyses above (table e-1), we explored the utility of combinations of subcomponents of these scales. A proposed 30-point scale based on the entire Barthel with addition of modified memory, orientation, and judgment sub-

Figure 3 Linear mixed model fitted to cognitive (MMSE, ADAS-cog, BPRS), functional (Barthel), and clinician-rated scales (Rankin, CDR-SB), GIC, and combined Barthel/CDR-SB (combination)
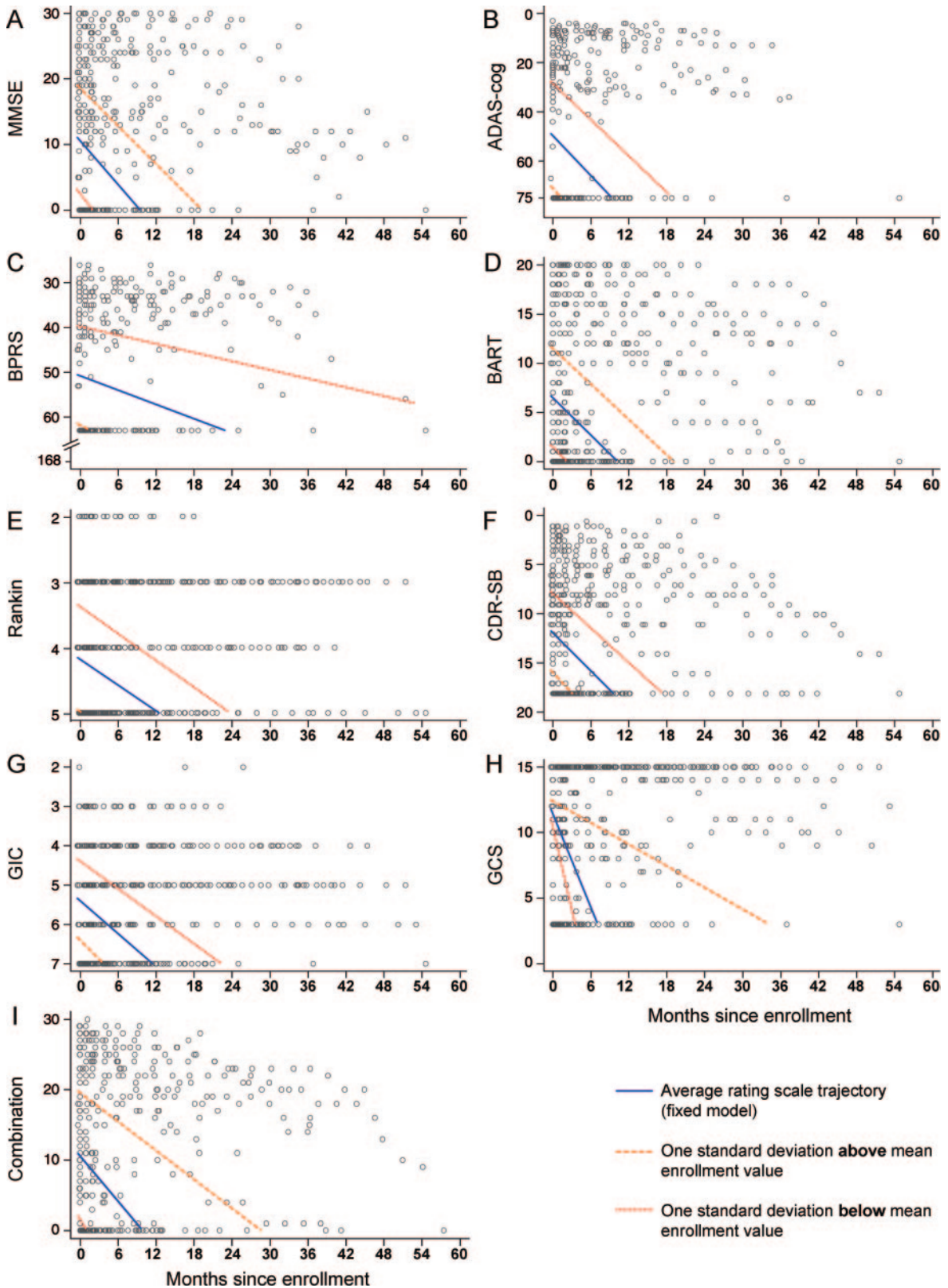
(A–I) These graphs illustrate the mean declines and observed values for each of the outcome measures that could be used in a clinical trial setting. Data are from the imputation model, i.e., censoring an individual's trajectory after observing the worst score and, if the worst score was not previously observed, imputing the worst possible score at the point of death (see appendix e-1). ADAS-cog = cognitive component of the Alzheimer's Disease Assessment Scale; BART = Barthel; BPRS = Brief Psychiatric Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; GCS = Glasgow Coma Scale; GIC = Global Impression of Change; MMSE = Mini-Mental State Examination.

components from the CDR-SB has been included in the power calculations for comparison (table 2). The CDR-SB subcomponents were selected on the basis of ease of use and validity in a prion disease patient population, confirmed by visualization of individual patient trajectories for each CDR-SB subcomponent. Further analyses to justify the choice of subcomponents and refine questions will be conducted in the National Prion Monitoring Cohort study.

DISCUSSION The desired requirements of a rating scale in prion disease are clear. The scale should be related to the disease manifestation and progression, and should be important to patients and their carers. The scale should be able to measure one or both of these accurately by minimizing disease-unrelated variations, such as those due to the particular examiner, the degree of encouragement, or patient fatigue. To facilitate patient assessment, the scale should not take long to complete and the need for specialist training of the examiner should be minimal. A hallmark of prion diseases is their rapid evolution, which necessitates a scale that can perform across a wide severity range. Given that prion diseases are rare, a scale covering a wide range of severity, which is as simple, easy to measure, and reproducible as possible, and not subject to variation in implementation over calendar time, may be important for studies with broad entry criteria, those recruiting internationally, and those with prolonged recruitment periods. In table e-1, we assess the 8 scales used in PRION-1 against these requirements.

Floor effects for cognitive and psychiatric scales were very strong in PRION-1. Serial cognitive/psychiatric measures were rarely obtained in sporadic CJD. Language output impairments have long been recognized as a classic clinical feature of sporadic CJD.[15] The frequency of this defect may overemphasize deterioration measured using the 3 cognitive/psychiatric scales which rely heavily on spoken responses by the patient. The functional and global scales were also impacted to some extent by insensitivity in the advanced stages of disease. The GCS was most resistant to this problem, but suffered from a marked ceiling effect with many patients never scoring below the best possible score, and questionable relevance to progression of the disease. In general, the floor effects were partially ameliorated by imputation of the worst possible score for patients who had died. This also allowed for the inclusion of patients in models for decline who were otherwise excluded because no measurements were obtained. A disadvantage of the imputed model is that it is only based on a single imputation, and the strong impact of survival limits the advance of the outcome measure beyond the use of mortality alone.

Across the disease-stratified, severity-stratified, and imputed models, the more functionally orientated scales (CDR-SB, Barthel) resulted in patient trajectories that appeared to match clinical experience (figure 2), they differentiated well between patients expected to progress slowly or moderately fast, and gave reasonable power in severity and imputed models to detect clinically relevant changes with moderate sample sizes. Decline in prion disease was generally linear in the measurements from these scales. These scales are quick and easy to apply with little training required, validated for use by several health care professions, can be used over the telephone, measure things fundamentally important to carers, and have been validated for use in settings of patient improvement and deterioration (see appendix e-1). Perhaps the most important advantage of the Barthel scale is that the activities it documents require a combination of mental health, cognitive, visual, balance, and other sensory or motor skills, each of which may be focally compromised in the early stages of prion disease.[14] A more specific cognitive or neurologic scale might not appropriately emphasize the impact of severe focal impairment in one of these domains. The Barthel scale has emphasis on mobility rather than cognition, whereas the latter is the most important overall clinical feature of prion disease. Subcomponents of the CDR-SB may be added to the Barthel to increase the emphasis on cognitive skills.

There are some limitations: we did not consider quality of life measures (not included as outcome measures in PRION-1), the effects of symptomatic therapies, or stratification by codon 129 genotype or prion protein type. By using models with patient-specific trajectories, our findings assume that PRION-1 patients are similar to future patients with regards to these additional factors, an assumption which is plausible due to our high recruitment of eligible patients.[3]

On the basis of these considerations we propose a scale adapted from the Barthel and CDR-SB for use as a primary outcome in future prion disease clinical trials. There are several advantages of such a scale; it appears to have comparable power to using either survival alone or any individual scale; its applicability for telephone use will allow an intensity of data collection that might lead to further improvements in power (although initial simulations suggest that these may be relatively small); and it documents issues of importance for carers. Clearly it would be advantageous to reach an international consensus on outcome measures in this disease. Further development of the scale including subgroup analysis, prediction

of survival, interrater reliability, and telephone use, among other issues, will be considered by the ongoing National Prion Monitoring Cohort study. The unusual opportunity afforded by PRION-1, to follow up patients through most of the clinical course to death, makes our conclusions relevant to other neurodegenerative diseases.

REFERENCES

1. Stewart LA, Rydzewska LH, Keogh GF, Knight RS. Systematic review of therapeutic interventions in human prion disease. Neurology 2008;70:1272–1281.
2. Collinge J. Prion diseases of humans and animals: their causes and molecular basis. Annu Rev Neurosci 2001;24:519–550.
3. Collinge J, Gorham M, Hudson F, et al. Safety and efficacy of quinacrine in human prion disease (PRION-1 study): a patient-preference trial. Lancet Neurol 2009;2009:334–344.
4. Consumer workshop on clinical trials for CJD. Department of Health, Medical Research Council; 2002.
5. Bone I, Belton L, Walker AS, Darbyshire J. Intraventricular pentosan polysulphate in human prion disease: an observational study in the UK. Eur J Neurol 2008;15:458–464.
6. Tsuboi Y, Doh-ura K, Yamada T. Continuous intraventricular infusion of pentosan polysulfate: clinical trial against prion diseases. Neuropathology 2009;29:632–636.
7. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: practical scale. Lancet 1974;2:81–84.
8. 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;12:189–198.
9. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimers disease. Am J Psychiatry 1984;141:1356–1364.
10. Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. Psychol Rep 1962;10:799–812.
11. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993;43:2412–2414.
12. Schneider LS, Olin JT, Doody RS, et al. Validity and reliability of the Alzheimer's Disease cooperative study–clinical global impression of change. Alzheimer Dis Assoc Disord 1997;11:S22–S32.
13. Rankin J. Cerebral vascular accidents in patients over the age of 60. Scott Med J 1957;2:127–136.
14. Mahoney FI, Barthel D. Functional evaluation: the Barthel Index. Md State Medical Journal 1965;14:56–61.
15. Shuttleworth EC, Yates AJ, Paltan Ortiz JD. Creutzfeldt-Jakob disease presenting as progressive aphasia. J Natl Med Assoc 1985;77:649–650, 652, 656.