Change in timed walk as primary outcome measure of treatment response in HAMLET-P: HAM/TSP MuLticentre Efficacy Trial-Prednisolone

Fabiola Martin [1], Eisuke Inoue[2], Ramon Kruschewsky [3] , Raya Massoud [4] , Irene Cortese[4], Marcus T Silva [5] , Maria Fernanda Rios Grassi [3], Bernardo Galvao-Castro[3] , Steven Jacobson[4], Yoshi Yamano[6], Graham P Taylor[7] , Martin Bland [1]

1 University of York, UK; 2 Kitasato University, Japan; 3 NIH/NINDS, USA ; 4 Fundação Oswaldo Cruz, Fiocruz, Brazil; 5 Escola Bahiana de Medicinae Saúde Pública, Brazil; 6 St. Marianna University School of Medicine, Japan; 7 Imperial College London, UK

Figure1: Measured 10m walk times from four countries, n=7.

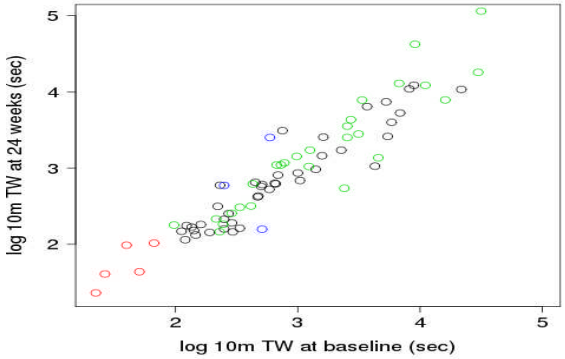
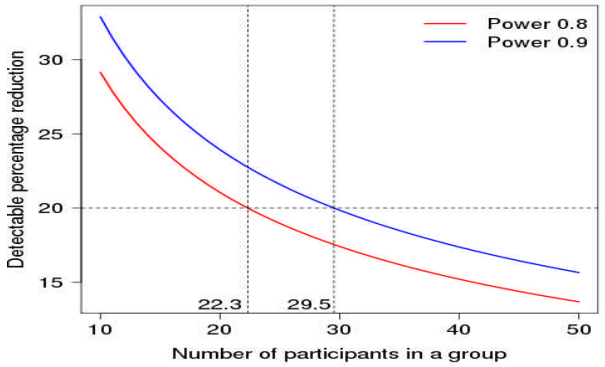


Table1: Detectable change calculations given fixed sample size for the 10mTW at 24 weeks.

Detectable change calculations given fixed sample size for the 10mTW at 24 weeks					
		Power = 0.8		Power = 0.9	
SD	# participant in one arm	Detectable change (log)	Detectable change (%)	Detectable change (log)	Detectable change (%)
0.2	20	0.18	19.94	0.21	23.42
	25	0.16	17.56	0.19	20.58
	30	0.15	15.85	0.17	18.56
	35	0.14	14.55	0.16	17.02
	40	0.13	13.53	0.15	15.81
	45	0.12	12.69	0.14	14.82
	50	0.11	11.99	0.13	13.99
0.26	20	0.24	26.66	0.27	31.46
	25	0.21	23.40	0.24	27.55
	30	0.19	21.08	0.22	24.77
	35	0.18	19.32	0.20	22.67
	40	0.16	17.93	0.19	21.02
	45	0.16	16.80	0.18	19.68
	50	0.15	15.85	0.17	18.56
0.3	20	0.27	31.35	0.32	37.11
	25	0.24	27.46	0.28	32.41
	30	0.22	24.69	0.26	29.09
	35	0.20	22.61	0.24	26.59
	40	0.19	20.96	0.22	24.63
	45	0.18	19.62	0.21	23.03
	50	0.17	18.50	0.20	21.70

Figure2: Detectable change (increase) against sample size for two chosen powers



Background: The HAM/TSP clinical trial study group sought a primary outcome measure for an international double-blind randomised placebo-controlled prednisolone efficacy trial based in Brazil, Japan, UK and USA. What matters to patients with HAM/TSP most is improvement of their gait when treated for HAM/TSP. In the absence of an internationally recognised biomarker of treatment response in HAM/TSP, the group chose improvement of 10 meter time walk (TW) at 24 weeks compared to baseline. Inter- and intra-patient variabilities of TW over time need to be taken into account in order to be able to calculate the sample size required to have a high probability to refute the null hypothesis if the treatment really is effective and so justify the power of the trial.

Aim: To define the minimum change in TW required between baseline and 24 weeks for an observed treatment effect to be of clinical importance and detectable in practice. To calculate the minimum size of the trial cohort required for 90% power to detect this difference.

Methods: Prospectively collected TW (seconds/10meter) of HAM/TSP patients from the four trial countries were analyzed. To interpret a detectable difference as a percentage change, log transformed TW were used. Analysis of covariance was used to analyze log TW.

Results: Matched TW data at baseline and at 6 months were available for a total of 76 patients. Baseline mean (SD), median TW were 23.46 (8.9), 16. 32 sec/10 m and 24.85 (23.89), 16.38 sec/10m. Baseline mean(SD), median $\log_e$ TW were 2.89 (0.72), 2.79 and 2.91 (0.74), 2.80 at 6 months. We estimated that the standard deviation of log 10m timed walk (TW) after adjustment for the baseline measurement was 0.26, based on available patient data. With 30 participants per group, we would have 90% power to detect a difference of ±0.21 . This corresponds to a ratio of 0.81 or 1.23, so we could detect a decrease in time of 19% or an increase of 23%. With power 80% we could detect a difference of −15% or +18%. To power the trial at 90%, a minimum of 30 patients are needed in each arm, to be increased by 3-5 patients/arm to cover for 10-15% estimated trial dropout rate.

Conclusions: For the first time ever we have been able to calculate a sample size for a HAM/TSP clinical trial based on a clinical outcome measure as a result of a multi-centre, international, collaborative clinical trial network. Prospectively collected longitudinal data on TW is useful in measuring inter- and intra-patient variability of this clinical efficacy marker. Treatment efficacy trials should aim for a sample size large enough to have 90% power to show a treatment effect of > 20% improvement in TW.

Contact: Fabiola.martin@hyms.ac.uk

#2242