Supplementary Appendix

Supplement to:

Jesus S. Mora, Angela Genge, Adriano Chio, Conrado J. Estol, Delia Chaverri, et al.

Masitinib as an add-on therapy to riluzole in patients with amyotrophic lateral sclerosis: a randomized clinical trial.

Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2019

DOI:10.1080/21678421.2019.1632346

This appendix has been provided by the authors to give readers additional information about their work.

LIST OF SUPPLEMENTARY TABLES

LIST OF SUPPLEMENTARY FIGURES

SUPPLEMENTARY DISCUSSION AND DATA

DOI:10.1080/21678421.2019.1632346

eTable 1: List of AB10015 Study Group collaborators (non-author investigators).

Argentina:	Nogues M, de Ambrosi B, Nofal P, Rey R, Mendez C, Mancuso M, Mainella C, Gonzalez L, Gargiulo G, Garcia-Mena MC, Ellenberg A, de Navarrete A, Carreño S, Bohorquez N.
Canada:	Zinman L, Shoesmith C, Larue S.
Greece:	Tavernarakis A, Doskas T, Verantioti A, Siatouni A, Gkatzonis S, Alexoudi A.
Italy:	Comi G, Silani V, Sabatelli M, Messina S, Giannini F, Caponnetto C, Zanolini A, Vita G, Scialò C, Russo M, Patanella AK, Moglia C, La Rosa M, Insana L, Ilardi A, Fuda G, Fini N, Fasano A, Falzone Y, Falzone F, Di Stefano M-G, Conte A, Carone M, Cammarosano S, Calvo A, Cabona C, Battistini S, Barcellona C.
Mexico:	Meza J, Ticozzi N, Lerario A.
Netherlands:	van den Berg LH, Westeneng H-J, Van es M.
Portugal:	de Carvalho M, Viana P, Oliveira M.
Slovakia:	TurčániP, Šutovský S, Kurča E, Krastev G, Gurcik L, Cuchran P, Tomášová A, Puzderová L, Števková Z, Sivák S.
Spain:	Carbajo P, Morán Y, Salas T, Marey Lopez J, Essanhaji A, Diez L, Rubio MA.

DOI:10.1080/21678421.2019.1632346

eTable 2: Distribution of patients with slowly progressive disease at baseline (various ΔFS cutoffs) according to treatment-arm and ΔFS-tiered cohort

	PLACEBO	**M4.5**	**M3.0**
'Normal Progressor' dataset‡	n = 113	n = 105	n = 110
ΔFS < 0.1 points/month	2 (1.8%)	3 (2.9%)	2 (1.8%)
ΔFS < 0.2 points/month	13 (11.5%)	13 (12.4%)	13 (11.8%)
ΔFS < 0.3 points/month	28 (24.8%)	23 (21.9%)	31 (28.2%)
'Normal and Fast Progressor' dataset†	n = 132	n = 128	n = 131
ΔFS < 0.1 points/month	2 (1.5%)	3 (2.3%)	2 (1.5%)
ΔFS < 0.2 points/month	13 (9.8%)	13 (10.2%)	13 (9.9%)
ΔFS < 0.3 points/month	28 (21.2%)	23 (18.0%)	31 (23.7%)

‡ 'Normal Progressor' dataset defined as patients with a post-onset ΔFS of less than 1.1 points/month. †'Normal and Fast Progressor' dataset includes all patients, regardless of the post-onset ΔFS selection criterion. ΔFS = ALSFRS-R progression rate from disease-onset to baseline. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. M3.0 = Masitinib 3.0 mg/kg/day plus riluzole.

DOI:10.1080/21678421.2019.1632346

eTable 3: Summary of predefined sensitivity analyses (rules 2–7) on the primary endpoint of 'Normal Progressor'* patients receiving masitinib 4.5 mg/kg/day versus placebo (primary efficacy population). Several sensitivity analyses were performed to test robustness of the primary analysis result (rule 1), censoring by reason for discontinuation (rules 2–5) and full analysis dataset imputation (rules 6–7) (see Supplementary eDiscussion Section B for detailed description of rules).

		n	LSM (ALSFRS-R)	ΔLSM [95%CI]	*P* value
† **Rule 1 (LOCF)**					
	PBO	102	-12.6	3.4 [0.7;6.1]	0.016
	M4.5	99	-9.2		
‡ **Rule 2 (LOCF)**					
	PBO	103	-12.5	3.3 [0.6;6.0]	0.019
	M4.5	99	-9.3		
‡ **Rule 3 (LOCF)**					
	PBO	107	-12.1	3.1 [0.4;5.7]	0.025
	M4.5	102	-9.0		
‡ **Rule 4 (LOCF)**					
	PBO	108	-12.0	3.0 [0.3;5.6]	0.029
	M4.5	102	-9.0		
‡ **Rule 5 (LOCF)**					
	PBO	111	-11.8	2.9 [0.3;5.4]	0.029
	M4.5	104	-9.0		
‡ **Rule 6 (Full analysis dataset)**					
	PBO	111	-14.0	3.0 [0.5;5.5]	0.018
	M4.5	104	-11.0		
‡ **Rule 7 (Full analysis dataset)**					
	PBO	111	-14.4	3.0 [0.5;5.5]	0.018
	M4.5	104	-11.4		

†Primary endpoint analysis. ‡Sensitivity analyses on the primary endpoint. *'Normal Progressor' dataset defined as post-onset ΔFS <1.1 points/month. ΔFS = ALSFRS-R progression rate from disease-onset. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. LSM = Least-squares means difference from baseline. ΔLSM = Between treatment-arm difference of LSM. 95% two-sided confidence intervals [95%CI].

DOI:10.1080/21678421.2019.1632346

eTable 4: Summary for the secondary efficacy populations including the 'Normal and Fast Progressor' masitinib 4.5 mg/kg/day cohort and ΔFS-tiered low-dose (masitinib 3.0 mg/kg/day) cohorts.

Cohort	Endpoint	Arm	n	LSM	ΔLSM [95%CI]	§Effect	*P* value
'Normal and Fast Progressor' dataset receiving masitinib 4.5 mg/kg/day†							
	ΔALSFRS-R*	**PBO**	119	-13.0	2.09 [-0.5 ; 4.7]	16%	0.12
		M4.5	120	-10.9			
	ALSAQ-40$	**PBO**	119	28.2	-6.6 [-11.9 ; -1.3]	23%	0.015
		M4.5	119	21.6			
	FVC$	**PBO**	119	-36.4	5.6 [-0.9 ; 12.2]	15%	0.09
		M4.5	118	-30.8			
'Normal Progressor' patients receiving masitinib 3.0 mg/kg/day‡							
	ΔALSFRS-R*	**PBO**	102	-11.3	2.7 [-0.2 ; 5.6]	24%	0.066
		M3.0	106	-8.6			
	ALSAQ-40$	**PBO**	102	23.6	-8.0 [-13.7 ; -2.4]	34%	0.006
		M3.0	106	15.6			
	FVC$	**PBO**	102	-28.9	5.0 [-1.9 ; 11.9]	17%	0.16
		M3.0	106	-23.9			
'Normal and Fast Progressor' dataset receiving masitinib 3.0 mg/kg/day†							
	ΔALSFRS-R*	**PBO**	119	-12.1	1.8[-0.9 ;4.5]	15%	0.19
		M3.0	126	-10.3			
	ALSAQ-40$	**PBO**	119	25.5	-7.2 [-12.4 ; -1.9]	28%	0.008
		M3.0	126	18.3			
	FVC$	**PBO**	119	-32.4	3.4 [-3.3; 10.0]	10%	0.32
		M3.0	126	-29.1			

† 'Normal and Fast Progressor' dataset includes all patients from a given treatment-arm, regardless of the post-onset ΔFS selection criterion. ‡ 'Normal Progressor' dataset defined as patients with a post-onset ΔFS of less than 1.1 points/month. ΔFS = ALSFRS-R progression rate from disease-onset to baseline. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. § Treatment-effect defined as slowing in the rate of decline for masitinib treatment-arm relative to placebo. *Parameter used for primary endpoint in the primary efficacy analysis. $Secondary endpoint. £All analyses shown performed on assessable patient dataset as determined by the predefined rule 1 for missing data imputation (see Supplementary eDiscussion Section B). LSM = Least-squares means difference from baseline. ΔLSM = Between treatment-arm difference of LSM. 95% two-sided confidence intervals [95%CI]. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. ALSAQ-40 = ALS Assessment Questionnaire. FVC = Forced Vital Capacity.

DOI:10.1080/21678421.2019.1632346

eTable 5: Subgroup analysis exploring effect of baseline disease severity, as measured by the individual component scores of ALSFRS-R. Summary results for the 'Normal and Fast Progressor' masitinib 4.5 mg/kg/day cohort ‡.

ALSFRS-R ANALYSIS[*] ('Normal and Fast Progressor' masitinib 4.5 cohort)

	n	LSM	ΔLSM [95%CI]	ΔEffect [£]	*P* value
ALSFRS-R item ≥1[†]					
PBO	105	-13.1	3.27 [0.4 ; 6.1]	25%	0.0266
M4.5	92	-9.83			
ALSFRS-R item ≥2[†]					
PBO	57	-11.51	4.81 [0.9 ; 8.7]	42%	0.0152
M4.5	48	-6.70			
ALSFRS-R item ≥3[†]					
PBO	27	-15.1	10.8 [3.2 ; 18.4]	72%	0.0064
M4.5	20	-4.3			

TIME-TO-EVENT[$] ANALYSIS ('Normal and Fast Progressor' masitinib 4.5 cohort)

	n	Median[95%CI]	ΔMedian	*P*-value[§]
ALSFRS-R item ≥1[†]				
PBO	115	16 [11; 19]	4 months	0.022
M4.5	96	20 [14; 30]		
ALSFRS-R item ≥2[†]				
PBO	63	17 [11; 33]	13 months	0.1502
M4.5	50	30 [15; NR]		
ALSFRS-R item ≥3[†]				
PBO	28	11 [8.3; 19]	19 months	0.0071
M4.5	22	30 [22; NR]		

‡ 'Normal and Fast Progressor' dataset includes all patients from a given treatment-arm, regardless of the post-onset ΔFS selection criterion. ΔFS = ALSFRS-R progression rate from disease-onset to baseline. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. [†] Patients having a score above a given threshold value for each ALSFRS-R item (a higher threshold indicates less severe disease). [*]ALSFRS-R analysis according to rule 1 for missing data imputation (see Supplementary eDiscussion Section B for detailed description of rules). LSM = Least-squares means difference from baseline. ΔLSM = Between treatment-arm difference of LSM. 95% two-sided confidence intervals [95%CI]. [£]Treatment-effect defined as slowing in the rate of decline for masitinib treatment-arm relative to placebo. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. [$]Time-to-event analysis defined as time interval (months) for ALSFRS-R deterioration of 9 points from baseline or death. ΔMedian = Between treatment-arm difference of time-to-event median. [§]P-value calculated using the Wilcoxon test. NR = not reached.

DOI:10.1080/21678421.2019.1632346

eTable 6: All non-fatal serious adverse events during 48-week treatment period with at least one event in the masitinib treatment-arms (safety dataset, regardless of causality, listed as per MedDRA Preferred Terms[*]).

	PBO (n=133)	M4.5 (n=129)	Δ[M4.5] (%)	M3.0 (n=131)	Δ[M3.0] (%)
Respiratory Failure	2 (1.5%)	8 (6.2%)	4.7	5 (3.8%)	2.3
Transaminases Increased	0 (0.0%)	2 (1.6%)	1.6	0 (0.0%)	0.0
Dysphagia	9 (6.8%)	10 (7.8%)	1.0	15 (11.5%)	4.7
Normochromic Normocytic Anaemia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Microcytic Anaemia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Myocardial Infarction	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Diarrhoea	0 (0.0%)	1 (0.8%)	0.8	1 (0.8%)	0.8
Pneumoperitoneum	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Abdominal Pain Upper	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Chest Pain	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Bronchitis	1 (0.8%)	2 (1.6%)	0.8	1 (0.8%)	0.0
Pneumonia Bacterial	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Oropharyngeal Candidiasis	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Oral Candidiasis	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Upper Respiratory Tract Infection	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Skin Infection	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Gastroenteritis	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Bronchitis Haemophilus	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Femur Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Fall	0 (0.0%)	1 (0.8%)	0.8	1 (0.8%)	0.8
Wrist Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Scapula Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Rib Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Ligament Sprain	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Subarachnoid Haemorrhage	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Hip Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Weight Decreased	1 (0.8%)	2 (1.6%)	0.8	3 (2.3%)	1.5
Troponin T Increased	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Blood Bilirubin Increased	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Intervertebral Disc Protrusion	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Transitional Cell Carcinoma	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Cauda Equina Syndrome	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Brain Injury	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Anxiety	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Dyspnoea	2 (1.5%)	3 (2.3%)	0.8	2 (1.5%)	0.0

DOI:10.1080/21678421.2019.1632346

	PBO (n=133)	M4.5 (n=129)	Δ[M4.5] (%)	M3.0 (n=131)	Δ[M3.0] (%)
Pulmonary Embolism	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Pneumonia Aspiration	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Urticaria	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Skin Toxicity	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Seborrhoeic Dermatitis	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Rash Generalised	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Pruritus Generalised	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Eczema	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Dry Skin	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Wisdom Teeth Removal	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Deep Vein Thrombosis	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Neutropenia	0 (0.0%)	0 (0.0%)	0.0	2 (1.5%)	1.5
Cardio Respiratory Arrest	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Pancreatitis	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Cholecystitis	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Pyelonephritis Acute	0 (0.0%)	0 (0.0%)	0.0	0 (0.0%)	0.0
Face Injury	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Neutrophil Count Decreased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Haemoglobin Decreased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Muscle Spasticity	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Generalised Tonic Clonic Seizure	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Brain Oedema	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Amnesia	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Panic Attack	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Nephrolithiasis	0 (0.0%)	0 (0.0%)	0.0	0 (0.0%)	0.0
Ureterolithiasis	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Respiratory Distress	0 (0.0%)	0 (0.0%)	0.0	0 (0.0%)	0.0
Drug Reaction Eosinophilia Systemic Symptoms	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Lower Respiratory Tract Infection	2 (1.5%)	1 (0.8%)	-0.7	2 (1.5%)	0.0
Gastrostomy	2 (1.5%)	0 (0.0%)	-1.5	2 (1.5%)	0.0

* Adverse events described using MedDRA Preferred Terms. Any given AE can be listed under multiple MedDRA preferred terms, which are not therefore cumulative. Safety dataset excluded 1 patient from ITT because of no intake of study drug. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. Δ[M4.5] = difference between M4.5 and placebo treatment-arms (M4.5 minus PBO). M3.0 = Masitinib 3.0 mg/kg/day plus riluzole. Δ[M3.0] = difference between M3.0 and placebo treatment-arms (M3.0 minus PBO). AEs were recorded until 28 days after treatment interruption.

DOI:10.1080/21678421.2019.1632346

eTable 7: All severe (grade 3/4) adverse events during 48-week treatment period with at least one event in the masitinib treatment-arms (safety dataset, regardless of causality, listed as per MedDRA Preferred Terms*).

	PBO (n=133)	M4.5 (n=129)	Δ[M4.5] (%)	M3.0 (n=131)	Δ[M3.0] (%)
Respiratory Failure	4 (3.0%)	9 (7.0%)	4.0	8 (6.1%)	3.1
Blood Phosphorus Decreased	0 (0.0%)	3 (2.3%)	2.3	0 (0.0%)	0.0
Dysphagia	3 (2.3%)	5 (3.9%)	1.6	6 (4.6%)	2.3
Gamma Glutamyltransferase Increased	0 (0.0%)	2 (1.6%)	1.6	0 (0.0%)	0.0
Dyspnoea	1 (0.8%)	3 (2.3%)	1.6	0 (0.0%)	-0.8
Anaemia Vitamin B12 Deficiency	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Normochromic Normocytic Anaemia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Microcytic Anaemia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Myocardial Infarction	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Eye Irritation	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Pneumoperitoneum	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Liver Disorder	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Lower Respiratory Tract Infection	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Scapula Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Rib Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Femur Fracture	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Subarachnoid Haemorrhage	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Weight Decreased	0 (0.0%)	1 (0.8%)	0.8	1 (0.8%)	0.8
Transaminases Increased	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Aspiration Bronchial	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Hypophosphataemia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Hypokalaemia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Muscular Weakness	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Amyotrophic Lateral Sclerosis	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Occipital Neuralgia	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Dysarthria	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Brain Injury	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Pulmonary Embolism	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Rash Maculo Papular	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Rash Generalised	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Pruritus Generalised	0 (0.0%)	1 (0.8%)	0.8	0 (0.0%)	0.0
Neutropenia	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Lymphopenia	0 (0.0%)	0 (0.0%)	0.0	2 (1.5%)	1.5
Iron Deficiency Anaemia	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Cardiopulmonary Failure	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8

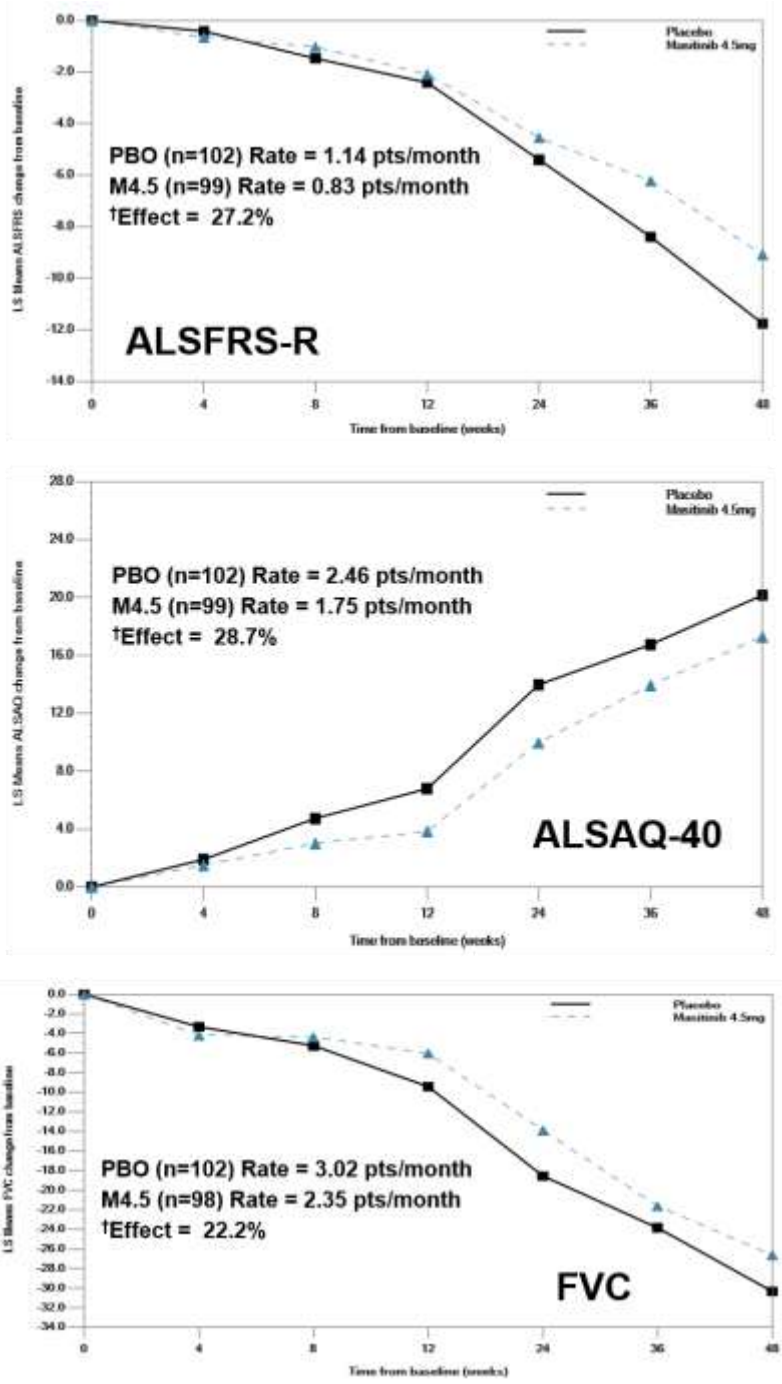
DOI:10.1080/21678421.2019.1632346

	PBO (n=133)	M4.5 (n=129)	Δ[M4.5] (%)	M3.0 (n=131)	Δ[M3.0] (%)
Bronchitis	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Laceration	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Fall	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Face Injury	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Haemoglobin Decreased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Blood Triglycerides Increased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Blood Glucose Increased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Blood Calcium Decreased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Aspartate Aminotransferase Increased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Alanine Aminotransferase Increased	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Hyponatraemia	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Myoclonic Epilepsy	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Muscle Spasticity	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Brain Oedema	0 (0.0%)	0 (0.0%)	0.0	2 (1.5%)	1.5
Anxiety	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Agitation	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Respiratory Arrest	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Acute Respiratory Failure	1 (0.8%)	1 (0.8%)	0.0	0 (0.0%)	-0.8
Pulmonary Oedema	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Obstructive Airways Disorder	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Drug Reaction Eosinophilia Systemic Symptoms	0 (0.0%)	0 (0.0%)	0.0	1 (0.8%)	0.8
Cardio Respiratory Arrest	2 (1.5%)	1 (0.8%)	-0.7	4 (3.1%)	1.5
Neutrophil Count Decreased	1 (0.8%)	0 (0.0%)	-0.8	1 (0.8%)	0.0
Hypertriglyceridaemia	1 (0.8%)	0 (0.0%)	-0.8	1 (0.8%)	0.0

* Adverse events described using MedDRA Preferred Terms. Any given AE can be listed under multiple MedDRA preferred terms, which are not therefore cumulative. Safety dataset excluded 1 patient from ITT because of no intake of study drug. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. Δ[M4.5] = difference between M4.5 and placebo treatment-arms (M4.5 minus PBO). M3.0 = Masitinib 3.0 mg/kg/day plus riluzole. Δ[M3.0] = difference between M3.0 and placebo treatment-arms (M3.0 minus PBO). AEs were recorded until 28 days after treatment interruption.

DOI:10.1080/21678421.2019.1632346

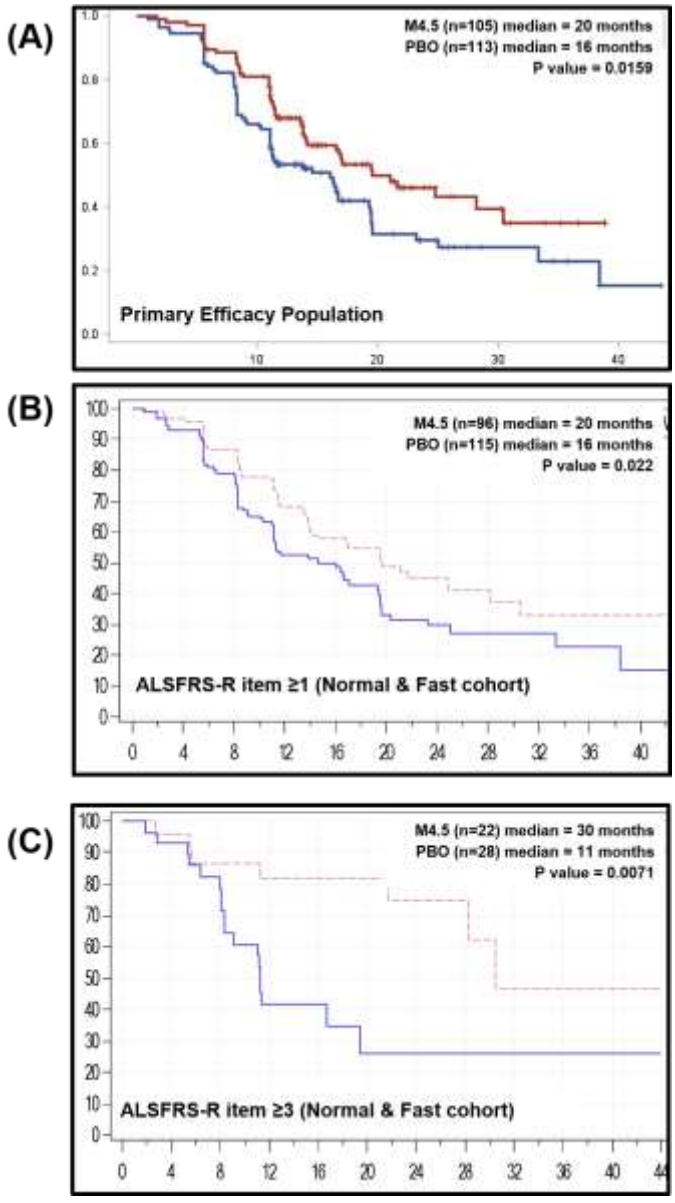
e**Figure 1:** Least-squared mean scores for the efficacy measures of ALSFRS-R[$], FVC[‡], and ALSAQ-40[‡] during 48-week treatment period in 'Normal Progressor'[*] patients receiving masitinib 4.5 mg/kg/day versus placebo (primary efficacy population).



[*] 'Normal Progressor' dataset defined as post-onset ΔFS <1.1 points/month. ΔFS = ALSFRS-R progression rate from disease-onset. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. [$]Primary endpoint. [‡]Secondary endpoint. [†]Treatment-effect defined as slowing in the rate of decline for masitinib treatment-arm relative to placebo. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. ALSAQ-40 = ALS Assessment Questionnaire. FVC = Forced Vital Capacity.

DOI:10.1080/21678421.2019.1632346

eFigure 2: Kaplan-Meier plot of time-to-event (defined as ALSFRS-R deterioration of 9 points from baseline or death) of masitinib (red line) versus placebo (blue line). (A) Primary efficacy population ('Normal Progressor' patients‡ receiving masitinib 4.5 mg/kg/day versus placebo). (B) and (C) Subgroup analyses exploring effect of baseline disease severity (as measured by the individual component scores of ALSFRS-R with a higher threshold indicating less severe disease) on treatment-effect in the 'Normal and Fast Progressor'† masitinib 4.5 mg/kg/day cohort.



‡ 'Normal Progressor' dataset defined as patients with a post-onset ΔFS of less than 1.1 points/month. †'Normal and Fast Progressor' dataset includes all patients, regardless of the post-onset ΔFS selection criterion. ΔFS = ALSFRS-R progression rate from disease-onset to baseline. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. PBO = Placebo plus riluzole. M4.5 = Masitinib 4.5 mg/kg/day plus riluzole. P-values calculated using Wilcoxon test.

DOI:10.1080/21678421.2019.1632346

eDiscussion Section A

Extended discussion on design of study AB10015, including rationale and validation of post-onset ΔFS as a robust instrument to reduce sample heterogeneity

- **Rationale for the dichotomization of the study population based on ΔFS**

The assumption of a homogenous treatment-effect across patients with ALS seems no longer tenable; moreover, the intrinsic heterogeneity of this population (in terms of phenotype and genotype) means that subgroups of patients may modify and confound a drug's treatment effect. Indeed, it is well-recognized that design and methodological shortcomings related to disease diversity or heterogeneity, and overestimation of expected effect size are potential reasons for negative results from ALS randomized controlled trials [Mitsumoto, 2014]. To address such issues, it is recommended to implement innovative design strategies with predictive cohort-enrichment being one such approach.

It is becoming increasingly clear that ALS is a complex multisystem neurodegenerative syndrome with marked heterogeneity at the level of clinical expression and also etiologically [Strong, 2017; Bäumer, 2014;]. The decision to dichotomize the ALS population for study AB10015 was in part motivated by emerging evidence of there being distinct forms of ALS, with dysregulation of the immune system being one possibly factor [Henkel, 2013], and also in view of repeated historical failure in trials with protocols that included all ALS phenotypes. Disease aggressiveness, i.e. the rate of disease progression, was considered one of the most fundamental ways in which such heterogeneity manifests itself clinically. Although heterogeneity in ALS survival and rate of disease progression is a well-established clinical observation, at the time study AB10015 was implemented there is a lack of published research regarding biomarkers of such variability. This has recently been addressed through gene expression and proteomic profiling studies but even so there are still no viable molecular biomarkers for identification of ALS subgroups according to disease aggressiveness.

In the absence of practicable molecular biomarkers for direct identification of ALS subgroups according to disease aggressiveness, an indirect measure of this phenomenon was required. The rate of disease progression (ΔFS), as measured by decline in ALSFRS-R, therefore represented the only viable tool for effective patient randomization and selection of a more homogeneous study population. At the time of implementation, ΔFS was considered a clinically relevant tool by merit of its close relationship with clinical manifestations of the disease (i.e. individual components of the ALSFRS-R score) and documented correlation with disease deterioration and patient survival. Moreover, clinical importance of ALSFRS-R, and therefore ΔFS, is further evidenced in its regular use by neurologists in the management of ALS. Indeed, the sanctioned use of ALSFRS-R was clearly described in the EMA guidance document EMA/CHMP/40105/2013 [EMA, 2013], wherein it is stated that "… *ALSFRS-R is the most widely used instrument to measure function in ALS clinical trials. It is a validated disease-specific questionnaire [Kaufmann 2007; Maier 215 2012; Leigh 2004; Cedarbaum 1999]. … Other scales that measure functional disability … may also be used, however the ALSFRS-R should be the preferred scale.*"

ΔFS was thus incorporated into the design of study AB10015 in two ways:

- ΔFS had previously been reported to be a sensitive and independent clinical prognostic parameter in ALS [Kimura, 2006; Gordon, 2006; Kollewe, 2008]. It was therefore important that randomization of patients in study AB10015 included stratification on this factor, thereby protecting against its potential confounding influence on any measured treatment-effect (e.g. through unbalanced treatment-arms). Gordon and colleagues commenting on the article of Kimura [Gordon, 2006; Kimura, 2006] stated that, "*Stratified enrollment lowers variability by reducing heterogeneity in the treatment arms. While site of onset and riluzole treatment may impart modest effects, the person's rate of progression is the most important predictor of outcome. It is theoretically possible to assign strata using historical information on progression at the baseline visit of a trial using the DeltaFS.*"
- Moreover, because of the large variability observed in ALS patient survival, which in turn is driven by divergent disease aggressiveness that presumably arises from differing and/or additional disease mechanisms, ΔFS can to some extent reflect underlying pathophysiological-based differences within the overall ALS population. Accordingly, patient categorization in terms of relative disease aggressiveness via ΔFS in effect groups patients with respect to underlying pathophysiology. In this manner ΔFS can be used as a possible

DOI:10.1080/21678421.2019.1632346

predictive tool for patient selection, permitting definition of a more homogeneous study population based on likely susceptibility to a given drug's mechanism of action.

When incorporating ΔFS into the design of study AB10015, two decisions were necessary: (a) setting a single cut-off that effectively isolated those patients with faster progressing (more aggressive) disease from the overall ALS population, and (b) choosing which of the two resulting groups would represent the primary efficacy population (thereby retaining the option of also assessing treatment-effect in the overall ALS population or complementary subgroup and also increasing the pool of patients available for safety assessment).

a) The ΔFS cut-off was based on available information from the scientific literature [Kollewe, 2008; Kimura, 2006; Gordon, 2006] and expert advice, and while no claim can be made that the optimal cut-off in terms of specificity and sensitivity was chosen, there existed no other practicable way for making such a distinction (a situation that is equally true as of today).

b) The decision to exclude Fast Progressor patients from the primary efficacy population was based on the common observation from clinical practice that the ΔFS histogram distribution has a right-skewed (i.e. positive-skew or tail) characteristic [Proudfoot, 2016]. This showed that patients with rapidly progressive disease represented the largest source of variability. Moreover, this group represents a relatively small proportion of the ALS population, with retention of the more populous group being preferable in terms of sample size requirements, especially given that ALS is a rare orphan disease.

Two subgroups were thus predefined according to ΔFS calculated from disease-onset to baseline, using the cut-off < 1.1 points/month to define 'Normal Progressors' and ≥ 1.1 points/month to define 'Fast Progressors'. The patient group of 'Normal Progressors' was prospectively declared as being the study's primary efficacy population. The design of study AB10015 therefore effectively represents a prospectively defined two-tiered design, in which a ΔFS-based cohort (i.e. 'Normal Progressors') can be compared with the broader population (i.e. 'Normal plus Fast Progressors'). This design feature defined a more homogenous target population while concurrently permitting assessment of predictive value in the proposed selection criterion. It follows that the prospectively declared primary efficacy population will be smaller than the overall study population.

Hence, it should be apparent that the predefined patient selection criteria introduced to study AB10015, wherein the primary efficacy analysis population is restricted to patients with ΔFS <1.1 points/month (i.e. via exclusion of patients with ΔFS ≥1.1 points/month), was implemented with the objective of managing anticipated response heterogeneity, thereby improving signal-to-noise. This design aspect of study AB10015 helped minimize patient numbers, a strategy consistent with published guidance [ALS Association, 2016] and is an approach also reflected in the draft and final EMA guidance document (*circa* 2013 and 2015, respectively), which specifies that "*Study participants should be stratified according to known prognostic factors*" and that "… *prognostic models may be used to stratify the study population by predicted rates of progression*" [EMA, 2015; EMA, 2013].

- **Validation of post-onset ΔFS as a robust instrument to reduce sample heterogeneity**

Verification that ΔFS represents an appropriate and robust tool for selection of the primary efficacy population in study AB10015, is provided by sensitivity analyses to determine the margin of error associated with the ΔFS cut-off and potential impact of misclassification from the requisite estimation of time to first symptom. As described in further detail below, results showed that the ΔFS cut-off value was well-judged and associated with a sizeable margin of error, while the possibility and impact of misclassification is small. This indicates that it is the action of dichotomization that is of key importance and not optimization of a specific cut-off.

Notably, the predefined 1.1 points/month cut-off of ΔFS has external validation through evidence published prior to initiation of study AB10015 [Kollewe, 2008], as well as after study initiation [Labra, 2016], the latter of which independently identify this rate of ALSFRS-R decline as having significant prognostic value. Briefly, Kollewe and colleagues defined a dichotomizing threshold for distinguishing between rapid and non-rapid progressing ALS over the duration of disease (i.e. post-onset) as a ΔFS of 1.185 points per month (corresponding to the median ΔFS from a study cohort of 479 patients) [Kollewe, 2008]. Patients with ΔFS above 1.185 points per month had a significantly shorter median survival time compared with patients with ΔFS below 1.185 and higher risk of death (HR 3.6, 95% CI 1.6–7.9). In another study, Labra and colleagues assessed the utility of rate of disease progression (measured by post-onset ΔFS) as a prognostic biomarker in ALS. They found that the ΔFS score at initial visit was a significant predictor

DOI:10.1080/21678421.2019.1632346

of survival in ALS (p<0.001) and remained significant when adjusted for age and site of onset. Statistically derived prognostic subgroups emerged with the threshold of ΔFS >1.11 points/month delineating patients more rapid disease progression and poorest prognosis, to those with non-rapid progression. The authors concluded that ΔFS and the proposed thresholds could be potentially utilized as prognostic guides in patient management and future clinical trials. The authors also discussed external validation of these findings by citing the work of Kimura and colleagues [Kimura, 2006], with both studies deriving similar prognostic cut-off values. This led Labra to conclude that the ΔFS score prognostic values appear to be similar irrespective of the origins of the ALS cohorts, thereby supporting the generalization and utility of the ΔFS score as a prognostic biomarker in ALS.

- *The predefined ΔFS cut-off of 1.1 was well-judged and is associated with a sizeable margin of error*

The distinction between 'Normal Progressors' and 'Fast Progressors' according to ΔFS was implemented for the reasons described above, although the actual threshold of ΔFS=1.1 was based on accumulate experience rather than prospective, data-driven evidence. To test the robustness of this cut-off threshold, sensitivity analyses were performed based on the primary endpoint (rule 1 for handling of missing data) and using a ΔFS cut-off ranging from 0.8 (corresponding to a predefined sensitivity analysis cut-off from the Statistical Analysis Plan of study AB10015) until the cut-off at which treatment-effect became non-significant. The table below presents results from sensitivity analyses to determine the margin of error associated with the ΔFS cut-off.

ΔFS cut-off	**ΔLSM (ALSFRS-R)†**	**95%[CI]**	**P-value**	**Proportion of randomized patients n, (%)**
0.8	3.4951	[0.62;6.37]	0.0174	191 (73.5%)
0.9	3.3355	[0.56;6.11]	0.0188	204 (78.5%)
1.0	3.1453	[0.40;5.89]	0.0251	210 (80.8%)
1.1*	3.3878	[0.65;6.13]	0.0157	218 (83.9%)
1.2	3.2707	[0.55;5.99]	0.0186	223 (85.8%)
1.3	3.2972	[0.60;5.99]	0.0168	226 (86.6%)
1.4	2.6897	[0.01;5.37]	0.0495	233 (89.9%)
1.5	2.5620	[-0.08;5.21]	0.0576	241 (92.7%)

†Primary endpoint = change in ALSFRS-R from baseline to week-48. Assessable patients for primary endpoint according to rule 1 for missing data imputation, i.e. last observation carried forward methodology for those patients discontinuing because of toxicity or lack of efficacy before week 48 (see Supplementary eDiscussion Section B for detailed description of rules). *Primary efficacy population prospectively defined as 'Normal Progressor' patients receiving masitinib at 4.5 mg/kg/day versus placebo. 'Normal Progressor' dataset defined as patients with a post-onset ΔFS of less than 1.1 points/month. ΔFS = ALSFRS-R progression rate from disease-onset to baseline. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised. LSM = Least-squares means difference from baseline. ΔLSM = Between treatment-arm difference of LSM. 95% two-sided confidence intervals [95%CI].

It is seen that treatment-effect remains in favor of masitinib until a ΔFS threshold of 1.4 (inclusive), accounting for 90% of patients. From this we can conclude that the positive treatment-effect seen for masitinib in study AB10015 cannot be dismissed as being a statistical anomaly of the dataset and its interaction with the ΔFS=1.1 cut-off, nor that this threshold makes for a highly volatile, changeable outcome. On the contrary, this analysis shows the categorization of patients based on our estimate of ΔFS=1.1 to be robust with a demonstrable 'buffer zone' (i.e. ΔFS cut-offs from 0.8 to 1.3 or 1.4 points/month, ΔFS = 1.1 ± 0.3) for maintained positive treatment-effect.

- *Sensitivity analyses showed ΔFS-categorization of patients was relatively insensitive to recollection of date of first symptom*

When considering the ΔFS-tiered design used in study AB10015, it is important to note that there is a fundamental difference in long-term estimates of 'post-onset' ΔFS and shorter-term 'in-treatment' ΔFS. The former measure was used for the categorization of patients in study AB10015 and is calculated from first symptom or duration of disease as described in the literature [Labra, 2016; Kimura, 2006]. The latter measure is used for calculating the rate of decline over the treatment period, also referred to as the slope or gradient of the ALSFRS-R curve, which is typically calculated therefore at a more advanced disease stage and with a likelihood of greater variability.

DOI:10.1080/21678421.2019.1632346

The calculation of 'post-onset' ΔFS depends on time from first symptom to baseline and therefore the patient's recollection of first symptom date. As per the El Escorial DC (EEDC), first symptom of disease is any progressive muscle weakness or atrophy, including symptoms such as slurring of a syllable or clumsiness in hand or foot, but excluding fasciculations or cramps or fatigue. An investigator's questioning of patient and caregiver, including review of previous records from previous or referring physicians, can accurately determine the month of onset if they know the EEDC. Nevertheless, this prerequisite of the ΔFS-tiered does introduce a risk that patients could be misclassified due to an error in the definition of the date of first symptom. We have therefore assessed the potential impact of errors associated with the time to first symptom. This analysis confirms that although there is theoretically a risk of misclassification, in particular in patients with baseline ALSFRS score ranging from 41 to 47, in practice the risk is very limited.

Considering a theoretical model, for a given baseline ALSFRS-R score we simulated the number of days between date of first symptom and baseline corresponding to an ALSFRS-R progression rate (ΔFS) of 1.1 points/month. That is to say, starting at a baseline ALSFRS-R score of 48 points, a decline of 1 point would occur after 27 days. Likewise, a decline of 10-points would occur over about 270 days, assuming a linear rate of decline over time. This process was repeated for upper and lower boundaries of ΔFS = 0.8 and 1.3 points/month, which define the aforementioned 'buffer zone' for maintained positive treatment-effect. From this is was possible to determine the margin of error associated with date of first symptom that would place the patient within the buffer zone described. For example, a decline of 10-points at ΔFS = 0.8 and 1.3 points/month would occur over about 375 and 231 days, respectively (a range of 144 days). Hence, an error in recollection of first symptom date by 102 and 42 days, respectively, would still place the patient within the 'buffer zone' for maintained positive treatment-effect. This model can thus be used to estimate the permissible margin of error in date of first symptom according to baseline ALSFRS-R score, thereby indicating the potential risk of patient misclassification in study AB10015; see table below, showing margin of error on time to first symptom (TTFS).

It can be seen from this analysis that the most 'recall sensitive' cohort are patients presenting with relatively rapid progressive disease and a baseline ALSFRS score ranging from 41 to 47, as evident by a deviation in TTFS of less than 30 days impacting strongly on their estimated 'post-onset' ΔFS. However, in the clinical setting there is typically some inertia between symptom onset and confirmed diagnosis; indeed, total diagnostic time has been reported to range from 8 to 15 months [Paganoni, 2014]. In a study by Paganoni and colleagues (n = 304) it was shown that the median total diagnostic time was 11.5 months (interquartile range, IQR, 7–20 months) [Paganoni, 2014]. For a patient with ΔFS of 1.3 these times translate into baseline ALSFRS-R scores of median 33.1 points (IQR 38.9–22.0 points). Likewise, for a patient with ΔFS of 1.1 these times translate into baseline ALSFRS-R scores of median 35.4 points (IQR 40.3–26.0 points). Hence, a large majority of patients with ΔFS of ≥1.1 are likely to have a baseline ALSFRS-R score of less than 40 based on the typical diagnostic timelines in ALS.

This assumption is supported by considering the actual distribution of patients presenting with baseline ALSFRS-R scores of 41 to 47 from study AB10015 (n=127/391, 32.5%). For this cohort, median ΔFS was 0.27 points/month with just 5 patients (1.25% of the ITT population) having a score in the vicinity of the 1.1 cut-off. Specifically, 2 patients with baseline ΔFS >1.1 (none exceeded a ΔFS of 1.3) and 3 patients with baseline ΔFS between 0.9 to 1.1. Hence, there is a strong preponderance towards patients in this 'recall sensitive' cohort having baseline ΔFS well-below the predefined threshold of 1.1. Consequently, tolerances on accuracy are relaxed, lowering the risk of misclassification.

Additionally, we assessed the risk for study AB10015 that an error in time to first symptom could shift a patient out of, or into, the aforementioned buffer zone of ΔFS = from 0.8 to 1.3 or 1.4 points/month. The table below, showing shift in classification depending on error in TTFS, considers two scenarios:

- Normal Progressors (ΔFS<1.1) that shift to ΔFS>1.3 or ΔFS>1.4 (upper limit of ΔFS buffer zone) due to recall inaccuracy and would therefore be considered as a Fast Progressor (i.e. false negative). In this scenario, only a shortening of the time to first symptom can generate misclassification.
- The complement scenario is when recall inaccuracy leads patients with baseline ΔFS of 1.3 or 1.4 to shift to ΔFS<1.1, therefore being considered as a Normal Progressor (i.e. false positive). In this scenario, only a lengthening of the time to first symptom can generate misclassification.

Results show that the risk of error in ΔFS classification for study AB10015 was very small, with a recall errors of 30 days having practically no impact, while a recall error of 60 days would have theoretically affected just 1% of patients.

DOI:10.1080/21678421.2019.1632346

Margin of Error on Time to First Symptoms (TTFS)

Baseline score	**21**	**23**	**25**	**27**	**29**	**31**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**
TTFS for slope =-1.1 (days)	736	682	627	573	518	464	409	382	355	327	300	273	245	218	191	164	136	109	82	55	27
TTFS for slope -1.3 (days)	623	577	531	485	438	392	346	323	300	277	254	231	208	185	162	138	115	92	69	46	23
TTFS for slope -0.8 (days)	1013	938	863	788	713	638	563	525	488	450	413	375	338	300	263	225	188	150	113	75	38
Range 1.1 to 1.3 (days)	113	105	96	88	80	72	63	59	55	50	46	42	37	33	29	26	21	17	13	9	4
Range 0.8 to 1.1 (days)	277	256	236	215	195	174	154	143	133	123	113	102	93	82	72	61	52	41	31	20	11
Overall range (buffer zone) (days)	390	361	332	303	275	246	217	202	188	173	159	144	130	115	101	87	73	58	44	29	15

Shift in classification depending on error in TTFS

Shift (days)		**-30**	**+30**	**-45**	**+45**	**-60**	**+60**	**-75**	**+75**	**-90**	**+90**
Upper threshold for ΔFS buffer zone >1.3 (i.e. recall error in ΔFS of 0.2)											
From Normal (<1.1) to Fast (>1.3)	n	1	-	2	-	4	-	10	-	15	-
	%	0.3%	-	0.5%	-	1.0%	-	2.6%	-	3.8%	-
From Fast (>1.3) to Normal (<1.1)	n	-	0	-	0	-	5	-	8	-	16
	%	-	0.0%	-	0.0%	-	1.3%	-	2.0%	-	4.1%
Upper threshold for ΔFS buffer zone >1.4 (i.e. recall error in ΔFS of 0.3)											
From Normal (<1.1) to Fast (>1.4)	n	0	-	1	-	3	-	6	-	9	-
	%	0.0%	-	0.3%	-	0.8%	-	1.5%	-	2.3%	-
From Fast (>1.4) to Normal (<1.1)	n	-	0	-	0	-	1	-	2	-	8
	%	-	0.0%	-	0.0%	-	0.3%	-	0.5%	-	2.0%

DOI:10.1080/21678421.2019.1632346

eDiscussion Section B

Extended discussion on sensitivity analyses

The primary endpoint was decline in ALSFRS-R from baseline to week-48 (ΔALSFRS-R), with missing data imputed via last observation carried forward (LOCF) methodology for those patients discontinuing because of toxicity or lack of efficacy before week-48. The procedure was prespecified in the study's statistical analysis plan prior to unblinding and referred to as 'rule 1' for handling of missing data. Several predefined sensitivity analyses were performed to test robustness of the primary analysis result, in part because LOCF methodology can generate biased results. Six predefined sensitivity analyses were conducted on the primary analysis, including four variations on LOCF via censoring on reason for discontinuation (rules 2–5) and two full analysis dataset (non-LOCF) imputation methods (rules 6–7). Additionally, sensitivity analyses based on multiple imputation and tipping-point analyses were done to further challenge robustness of the primary analysis findings.

- *Predefined sensitivity analyses as per protocol and statistical analysis plan*

Rules for handling missing data at week 48 in ALSFRS score are detailed below, with the different imputation methods used in each of the predefined sensitivity analysis summarized in tabular form according to reason of discontinuation (which is consistent with EMA guideline: EMA/CPMP/EWP/1776/99 Rev.1).

Predefined methods of missing data imputation for primary and sensitivity analyses

	Primary	**Rule 2**	**Rule 3**	**Rule 4**	**Rule 5**	**Rule 6**	**Rule 7**
Reason of discontinuation							
Lack of Efficacy	LOCF	LOCF	LOCF	LOCF	LOCF	Imput.	Imput. with penalty
Toxicity	LOCF	LOCF	LOCF	LOCF	LOCF	Imput.	Imput.
Procedure	OC	LOCF	OC	LOCF	LOCF	Imput.	Imput.
Travel	OC	OC	LOCF	LOCF	LOCF	Imput.	Imput.
Lost to follow up	OC	OC	OC	OC	LOCF	Imput.	Imput.
Protocol deviation	OC	OC	OC	OC	OC	Imput.	Imput.
Other	OC	OC	OC	OC	LOCF	Imput.	Imput.
Non-compliance	OC	OC	OC	OC	LOCF	Imput.	Imput.

LOCF: Last Observation Carried Forward. OC: Observed Case (patients not included in analysis). Imput: Full analysis dataset retained via single imputation method.

Data already presented in eTable 3 shows that all these predefined sensitivity analyses on the primary endpoint, including those based on the full analysis dataset, were statistically significant.

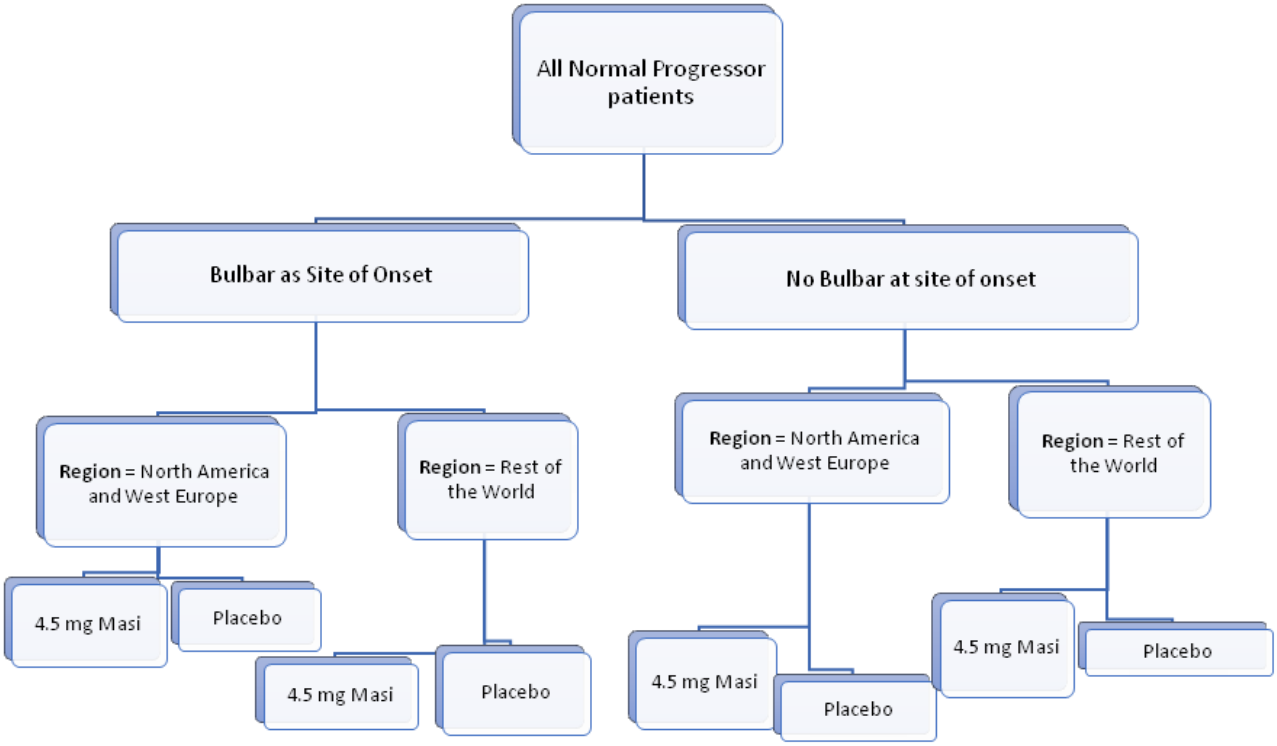
- Rule 1 was used for the primary analysis. Missing data was imputed via LOCF when patients discontinued before week-48 for documented reasons of toxicity or lack of efficacy. If patient died before week-54 (included) after randomization, ALSFRS-R score was replaced by zero (0). Patients discontinuing prematurely for the following documented reasons were not imputed (lost to follow-up, non-compliance, travel, procedure, protocol deviation, any other reason not mentioned above).
- Rule 2 was the same as rule 1 with the exception that imputation via LOCF was also done in case of premature discontinuation due to withdrawal of consent related to study procedure.
- Rule 3 was the same as rule 1 with the exception that imputation via LOCF was also done in case of premature discontinuation due to any travel issue.
- Rule 4 was the same as rule 1 with the exception that imputation via LOCF was also done in case of premature discontinuation due to withdrawal of consent related to study procedure or due to any travel issue.
- Rule 5 imputed data via LOCF for all patients with the exception of those patients who were non-compliant after the date of non-compliance. For these patients, the week-48 ALSFRS score was imputed using the last available score before non-compliance, where non-compliance was defined as per site clinical judgement, '*patient who could and should have continued to use study treatment but did not due to different reasons as mentioned in statement signed by the investigators'*.
- Rule 6 used single imputation methodology, i.e. non-LOCF, copying increment from similar patients, i.e. imputation was done by clustering patients by a given prognostic factor, then using the average increment within groups.

DOI:10.1080/21678421.2019.1632346

- Rule 7 was the same as rule 6 with the exception that a penalty of 50% was applied to those patients who discontinued early due to lack of efficacy, an imputation method based on recommendations from The National Research Council Panel on Handling Missing Data in Clinical Trials [Permutt, 2016].

Considering further the full analysis dataset methods of rules 6 and 7, and accompanying tipping analyses, these sensitivity analyses estimate ALS disease progression for similarly clustered patients, with clustering being done on the variables of: site of onset at baseline, region, and treatment group. This method decreases possible bias introduced by LOCF methods because it tries to identify the trend within similar patients and then imputes the missing value using this average trend.

The flow-chart presented below details the formation of clusters using the primary efficacy population of 'Normal Progressors' receiving masitinib at 4.5 mg/kg/day as an example. It is seen that altogether eight groups are formed in this manner.



Prior to performing imputation, the ALSFRS-R score of any patient dying up to 54 weeks (included) after randomization was replaced by zero (0). The sensitivity analysis for rule 6 defined that all remaining patients who discontinued during the 48-week treatment period had their week-48 ALSFRS-R score imputed via the following formula:

> *Imputed score at week 48 = Last non-missing ALSFRS-R score plus average (negative) increment from week 'X' to week 48 (derived from all patients in the same cluster with non-missing compliant data over that period).*

The sensitivity analysis for rule 7 defined that all remaining patients who discontinued during the 48-week treatment period had their week-48 ALSFRS-R score imputed via the following formula, with inclusion of a penalty for those patients who discontinued early due to lack of efficacy:

> *Imputed score at week 48 = Last non-missing ALSFRS-R score plus average (negative) increment from week 'X' to week 48 (derived from all patients in the same cluster with non-missing compliant data over that period) plus 50% of the average increment from week 'X' to week 48 for those patients discontinued due to lack of efficacy.*

The rationale for adding a 50% penalty on patients who discontinue due to lack of efficacy is that for study AB10015 it can be assumed the masitinib-treated patients discontinuing for lack of efficacy will have continued to decline at a faster rate than other patients if they had continued the trial. This assumption is consistent with Permutt publication in which it is stated that, "*Most missing data in clinical trials are not missing completely at random. Patients have reasons for discontinuing study medication, study participation, or both. The reasons are most commonly toxicity and lack of efficacy, even when they are not carefully ascertained and therefore are*

DOI:10.1080/21678421.2019.1632346

recorded as withdrawal of consent. Patients who withdraw for lack of efficacy have systematically different efficacy from other patients. Patients who withdraw for toxicity might be thought to have the same (de jure) efficacy if they had been able to continue." Hence, a penalty of 50% of the average increment of the cluster has been imposed on these patients' imputed score [Permutt, 2016].

As presented in Supplementary eTable 3, the result for sensitivity analysis rule 7 showed a positive trend ($p < 0.05$) and was similar to the primary endpoint analysis (rule 1).

- *Multiple Imputation FCS REGPMM (Additional sensitivity analyses as per guidelines)*

We referred to the following guidelines and references used widely across the industry:

- EMA Guideline on Missing Data in Confirmatory Clinical Trials (02 July 2010)
- Prevention and Treatment of Missing Data in Clinical Trials, National Research Council (2010)
- ICH E9: Statistical Principles for Clinical Trials
- E9(R1) Estimands and Sensitivity Analysis in Clinical Trials, (16 June 2017)
- Preventing and Treating Missing Data in Longitudinal Clinical Trials, Craig Mallinckrodt

All these guidelines recommend use of multiple imputation methodologies.

Multiple imputation (MI) is the most widely used sensitivity analysis technique and is highly recommended by all health authorities [EMA Guideline, 2010; NRC Report, 2010].

The Fully Conditional Specification (FCS) Regression Predictive Mean Matching (PMM) method was used for imputing missing values. A total of 2000 simulated datasets were generated via PROC MI.

The FCS REGPMM [Smith, 2017] method was chosen because it allows imputation of an arbitrary pattern and allows classification variables as covariates.

The imputation was carried out as follows:

- A total of 2000 data sets were generated using PROC MI
- Change from baseline at week-48 was analyzed using PROC MIXED with appropriate factors
- PROC MIANALYZE combined all the results from every MI repetition and provided valid Statistical Inference

The following three models were evaluated. Akiake's Information Criteria (AICC), a standard model diagnostic technique [Akaike, 1974] was used to assess the models (b) and (c).

a) All stratification factors for imputation and analysis. Analysis similar to primary analysis as specified in the protocol.

b) The factors explaining maximum variability in ALSFRS-R used for imputation and analysis. Treatment, Bulbar, Region and ΔFS best explain the variability for the model's dependent variable (i.e. change from baseline in ALSFRS at week-48). Of all the models assessed, AICC was at a minimum for this model (AICC = 9764.8).

c) The second-best set of factors having maximum impact on the variability in ALSFRS-R used for imputation and analysis. Treatment, Bulbar, Region, ΔFS and Baseline ALSFRS-R best explain the variability for this model's dependent variable. AICC for this model was AICC = 9769.9.

Results consistently show that there is a statistically significant difference between masitinib and placebo for each of the FCS REGPMM MI scenarios described (see table below).

DOI:10.1080/21678421.2019.1632346

Summary of FCS REGPMM MI analyses on the primary endpoint of 'Normal Progressor' patients receiving masitinib 4.5 mg/kg/day versus placebo (primary efficacy population).

Imputation Factors	Included	Analysis Factors	Included	Estimate (95%CI)
Model (1): All factors for imputation and analysis				
Treatment	Yes	Treatment	Yes	
Bulbar	Yes	Bulbar	Yes	
Region	Yes	Region	Yes	**2.95 (0.03-5.86)**
Age	Yes	Age	Yes	**P = 0.0476**
ΔFS	Yes	ΔFS	Yes	
Baseline ALSFRS-R	Yes	Baseline ALSFRS-R	Yes	
Model (2): According to AICC the factors explaining maximum variability in ALSFRS-R used for imputation and analysis				
Treatment	Yes	Treatment	Yes	
Bulbar	Yes	Bulbar	Yes	
Region	Yes	Region	Yes	**3.44 (0.53-6.33)**
Age	No	Age	No	**P = 0.020**
ΔFS	Yes	ΔFS	Yes	
Baseline ALSFRS-R	No	Baseline ALSFRS-R	No	
Model (3): According to AICC the second-best set of factors explaining maximum variability in ALSFRS-R used for imputation and analysis				
Treatment	Yes	Treatment	Yes	
Bulbar	Yes	Bulbar	Yes	
Region	Yes	Region	Yes	**3.18 (0.28-6.07)**
Age	No	Age	No	**P = 0.0317**
ΔFS	No	ΔFS	Yes	
Baseline ALSFRS-R	No	Baseline ALSFRS-R	Yes	

- *Multiple Imputation Jump to Reference (Additional sensitivity analyses as per guidelines)*

This Jump to Reference (J2R) approach imputed missing data for reason of discontinuation due to lack of efficacy or toxicity, using estimates from the control group. This is justifiable scientifically under the assumption that patients who stop taking the therapy for lack of efficacy will no longer benefit from it in the future, and thus will tend to have outcomes similar to those in the control group.

This approach is highly recommended in the literature for sensitivity analyses and an Expert Working Group of highly recognized statisticians from the London School of Hygiene and Tropical Medicine developed set of macros for this approach. These macros were used for our analysis.

Results consistently show that there is a statistically significant difference between masitinib and placebo for each of the Jump to Reference MI scenarios described (see table below).

Summary of Jump to Reference MI analyses on the primary endpoint of 'Normal Progressor' patients receiving masitinib 4.5 mg/kg/day versus placebo (primary efficacy population).

J2R analysis	Estimate	95% confidence interval	p-value
Lack of efficacy 100 and toxicity 100	2.80	[0.1462; 5.45716]	0.0386
Lack of efficacy 100 and toxicity 50	3.81	[1.0493; 6.56790]	0.0068
Lack of efficacy 100 and toxicity 0	2.73	[0.1036; 5.36051]	0.0416
Lack of efficacy 50 and toxicity 50	4.57	[1.6136; 7.52656]	0.0024
Lack of efficacy 50 and toxicity 0	3.49	[0.6243; 6.36269]	0.0170

DOI:10.1080/21678421.2019.1632346

- *Tipping-point (on Rule 6) analysis*

Tipping point analysis consists of applying Jump to Reference on all reasons of discontinuation for masitinib-treated patients. This method determines what proportion of discontinued masitinib patients would need to have their scores imputed at the placebo treatment effect in order to overturn conclusions from the primary analysis; i.e. the tipping-point approach serves as form of stress-testing, with a high penalty (range 100─0) indicating greater robustness.

The following method was used for tipping-point analysis:

- It is assumed that the patients who discontinue in the masitinib 4.5 mg/kg/day treatment-arm (M) do not respond as well as those who do not discontinue, the worst-case scenario being that the treatment effect for those patients is similar to that observed in the placebo arm (P). Hence, the imputed values for these patients will be shifted gradually towards the observed placebo treatment effect by a certain additive shift and the point at which the analysis p-value changes from ≤ 0.05 to > 0.05 will be noted.
- The formula is: $P \rightarrow P$

  $M \rightarrow P + (M-P) \times (1-\text{penalty})$

Hence, if the penalty is 100% (i.e. 1) then M = P, and if the penalty is 0% (i.e. 0) then M = M.

With a 100% penalty (i.e. imputing the placebo average for masitinib 4.5 mg/kg/day patients)

Treatment group	**n**	**LSM (ALSFRS-R)**	**ΔLSM [95%CI]**	***P* value**
PBO	111	-14.09	2.31 [-0.17;4.80]	0.0678
Masitinib	104	-11.79		

With a 76% penalty

Treatment group	**n**	**LSM (ALSFRS-R)**	**ΔLSM [95%CI]**	***P* value**
PBO	111	-14.06	2.48 [0.01;4.96]	0.0498
Masitinib	104	-11.58		

With a 77% penalty

Treatment group	**n**	**LSM (ALSFRS-R)**	**ΔLSM [95%CI]**	***P* value**
PBO	111	-14.06	2.47 [-0.01;4.95]	0.0504
Masitinib	104	-11.59		

Hence, the tipping point was achieved with a penalty of 77%. This is equivalent to saying that the primary analysis remains positive even after applying the conservative assumption that those patients discontinuing from masitinib would experience only 24% of the average masitinib treatment-effect had they continued treatment until week 48.

- *Conclusion from sensitivity analyses*

Taken together, these positive results from sensitivity analyses, including recommended multiple imputation methods, corroborate the robustness of the primary endpoint data and indicate that the observed treatment-effect cannot be dismissed as possible LOCF bias.

DOI:10.1080/21678421.2019.1632346

eDiscussion Section C

Extended discussion on exploratory subgroup analyses showing that initiation of masitinib at a less severe stage of disease produced greater treatment-effect

Subgroup analyses explored whether patient susceptibility to masitinib was influenced by baseline disease severity, as measured by ALSFRS-R individual component scores. Results showed that initiation of masitinib at a less severe stage of disease produced greater treatment-effect for both ΔFS-tiered high-dose (masitinib 4.5 mg/kg/day) cohorts.

Notably, this minor adjustment in patient selection criteria revealed a significant benefit for masitinib over placebo in the broader ('Normal and Fast Progressor') masitinib 4.5 mg/kg/day cohort (Supplementary eTable 5). For patients in this cohort with a baseline score of ≥1 on each ALSFRS-R item (92 and 105 patients in the masitinib and placebo treatment-arms, respectively), ΔALSFRS-R was -9.8 versus -13.1 (P=0.0266), corresponding to a 25% slower rate of decline.

Likewise, for patients with a baseline score of ≥3 on each ALSFRS-R item (20 and 27 patients in the masitinib and placebo treatment-arms, respectively), ΔALSFRS-R was -4.3 versus -15.1 (P=0.0064), corresponding to a 72% slower rate of decline. Considering the time-to-event analysis in this subgroup (22 and 28 patients in the masitinib and placebo treatment-arms, respectively), there was a significant 19-month difference between treatment-arms in favour of masitinib (P=0.0071) (Supplementary eFigure 2).

Hence, these subgroup analyses indicated further improvement is possible when initiating treatment at a less severe stage of disease, e.g. exclusion of patients with zero-point ALSFRS-R items. This represents a realistic clinical scenario in which treatment is initiated prior to severe symptom onset. Indeed, this minor adjustment in patient selection criteria, which in hindsight is a logical design feature considering that any ALSFRS-R component scoring zero at baseline will be insensitive to treatment-effect, revealed significant benefit in ΔALSFRS-R for the broader ('Normal and Fast Progressor') masitinib 4.5 mg/kg/day cohort.

DOI:10.1080/21678421.2019.1632346

References for Supplementary Appendix

- Akaike, H., 1974, 19 (6): 716– 23, doi:10.1109/TAC.1974.1100705, MR 0423716
- ALS Association. Guidance for Industry Drug Development for Amyotrophic Lateral Sclerosis. 2016. [Internet] http://www.alsa.org/advocacy/fda/assets/als-drug-development-guidance-for-public-comment-5-2-16.pdf. Accessed 06 March 2018.
- Bäumer D, Talbot K, Turner MR. Advances in motor neurone disease. J R Soc Med. 2014 Jan;107(1):14-21.
- EMA, 2010. EMA Guideline on Missing Data in Confirmatory Clinical Trials (02-July-2010)
- EMA, 2013 (EMA/CHMP/40105/2013) Committee for Medicinal Product for Human Use (CHMP). Guideline on clinical investigation of medicinal products for the treatment of amyotrophic lateral sclerosis (ALS). Accessed 06 February 2018
- EMA, 2015 (EMA/531686/2015, Corr.1.1) Committee for Medicinal Product for Human Use (CHMP). Guideline on clinical investigation of medicinal products for the treatment of amyotrophic lateral sclerosis (ALS). Accessed 06 February 2018
- Gordon, P. H., Doorish, C., Montes, J., Mosley, R. L., Mosely, R. L., Diamond, B., et al. (2006). Randomized controlled phase II trial of glatiramer acetate in ALS. Neurology 66, 1117–9.
- Henkel JS, Beers DR, Wen S, et al. Regulatory T-lymphocytes mediate amyotrophic lateral sclerosis progression and survival. EMBO Mol Med. 2013 Jan;5(1):64-79.
- Kimura F, Fujimura C, Ishida S, et al. Progression rate of ALSFRS-R at time of diagnosis predicts survival time in ALS. Neurology 2006;66:265–7.
- Kollewe K, Mauss U, Krampfl K, et al. ALSFRS-R score and its ratio: a useful predictor for ALS-progression. J Neurol Sci 2008;275:69–73.
- Labra J, Menon P, Byth K, Morrison S, Vucic S. Rate of disease progression: a prognostic biomarker in ALS. J Neurol Neurosurg Psychiatry. 2016 Jun;87(6):628-32.
- Mitsumoto H, Brooks BR, Silani V. Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? Lancet Neurol. 2014 Nov;13(11):1127-1138.
- NRC Report, 2010. Prevention and Treatment of Missing Data in Clinical Trials, National Research Council report.
- Paganoni S, Macklin EA, Lee A, et al. Diagnostic timelines and delays in diagnosing amyotrophic lateral sclerosis (ALS). Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration. 2014;15(0):453-456.
- Permutt. Statistics in medicine. 2016:35(17): 2876-2879
- Proudfoot M, Jones A, Talbot K, Al-Chalabi A, Turner MR. The ALSFRS as an outcome measure in therapeutic trials and its relationship to symptom onset. Amyotroph Lateral Scler Frontotemporal Degener. 2016 Jul-Aug;17(5-6):414-25.
- Smith C, Kosten S. Multiple Imputation: A Statistical Programming Story, PharmaSUG 2017 - Paper SP01.
- Strong MJ. Revisiting the concept of amyotrophic lateral sclerosis as a multisystems disorder of limited phenotypic expression. Curr Opin Neurol. 2017 Dec;30(6):599-607.

DOI:10.1080/21678421.2019.1632346