Supplementary Tables

Difference between cohorts of least squares mean change from baseline (95% confidence intervals) over 18 and 36 months in patient clinical measures

1) **Sensitivity analysis* – propensity score stratum, patients with non-overlapping scores included**

Outcome	Time point	EXP-placebo vs GERAS		EXP-active vs GERAS	
		Estimate (95% CI)	p-value	Estimate (95% CI)	p-value
MMSE†	18 months	-0.68 (-1.44–0.07)	0.075	0.93 (0.17–1.69)	**0.016**
	36 months	N/A		0.35 (-0.79–1.49)	0.547
ADAS-Cog14‡	18 months	0.81 (-1.31–2.93)	0.45	-2.00 (-4.15–0.16)	0.070
	36 months	N/A		1.84 (-1.31–4.99)	0.252
ADCS-bADL§	18 months	0.23 (-0.47–0.93)	0.523	0.34 (-0.38–1.06)	0.350
	36 months	N/A		0.98 (-0.05–2.01)	0.061
ADCS-iADL§	18 months	0.95 (-0.99–2.89)	0.336	2.50 (0.41–4.59)	**0.019**
	36 months	N/A		3.59 (0.86–6.31)	**0.010**
NPI-12‖	18 months	-0.50 (-2.49–1.50)	0.625	-0.91 (-3.00–1.18)	0.394
	36 months	N/A		1.07 (-1.37–3.51)	0.390

*Mixed-effects models for repeated measures included baseline value of outcome measure, study group (GERAS or EXP), visit, visit-by-treatment group interaction and propensity score stratum; patients with non-overlapping propensity scores were included. P-values in bold represent a significant difference between EXPEDITION and GERAS.

2) Sensitivity analysis* – propensity score model, patients with non-overlapping scores excluded

Outcome	Time point	EXP-placebo vs GERAS		EXP-active vs GERAS	
		Estimate (95% CI)	p-value	Estimate (95% CI)	p-value
MMSE†	18 months	-0.65 (-1.44–0.13)	0.103	0.93 (0.14–1.72)	**0.021**
	36 months	N/A		0.30 (-0.95–1.54)	0.637
ADAS-Cog14‡	18 months	0.72 (-1.31–2.76)	0.486	-1.79 (-3.87–0.29)	0.092
	36 months	N/A		1.20 (-2.12–4.52)	0.480
ADCS-bADL§	18 months	0.26 (-0.36–0.87)	0.417	0.28 (-0.41–0.97)	0.426
	36 months	N/A		1.10 (0.06–2.14)	**0.038**
ADCS-iADL§	18 months	1.27 (-0.63–3.17)	0.189	2.41 (0.34–4.49)	**0.023**
	36 months	N/A		4.06 (1.21–6.91)	**0.005**
NPI-12||	18 months	-0.66 (-2.53–1.22)	0.492	-1.53 (-3.61–0.56)	0.151
	36 months	N/A		0.28 (-2.33–2.88)	0.835

*Mixed-effects models for repeated measures included baseline value of outcome measure, study group (GERAS or EXP), visit, visit-by-treatment group interaction and propensity score model; patients with non-overlapping propensity scores were excluded. P-values in bold represent a significant difference between EXPEDITION and GERAS.

3) Sensitivity analysis* – propensity score model, patients with non-overlapping scores included

Outcome	Time point	EXP-placebo vs GERAS		EXP-active vs GERAS	
		Estimate (95% CI)	p-value	Estimate (95% CI)	p-value
MMSE†	18 months	-0.64 (-1.40–0.11)	0.096	0.92 (0.15–1.68)	**0.019**
	36 months	N/A		0.34 (-0.80–1.47)	0.562
ADAS-Cog14‡	18 months	0.62 (-1.51–2.76)	0.567	-1.81 (-3.97–0.34)	0.099
	36 months	N/A		2.04 (-1.10–5.18)	0.202
ADCS-bADL§	18 months	0.26 (-0.45–0.97)	0.471	0.27(-0.45–0.99)	0.466
	36 months	N/A		0.89 (-0.13–1.92)	0.088
ADCS-iADL§	18 months	1.30 (-0.68–3.27)	0.197	2.45 (0.36–4.54)	**0.022**
	36 months	N/A		3.53 (0.79–6.27)	**0.012**
NPI-12‖	18 months	-0.53 (-2.54–1.48)	0.605	-0.96 (-3.06–1.14)	0.368
	36 months	N/A		0.58 (-1.43–3.44)	0.639

*Mixed-effects models for repeated measures included baseline value of outcome measure, study group (GERAS or EXP), visit, visit-by-treatment group interaction and propensity score model; patients with non-overlapping propensity scores were included. P-values in bold represent a significant difference between EXPEDITION and GERAS.

Abbreviations: ADAS-Cog14, Alzheimer's Disease Assessment Scale-cognitive subscale; ADCS, Alzheimer's Disease Cooperative Study; bADL, basic activities of daily living; CI, confidence interval; EXP, EXPEDITION study; iADL, instrumental activities of daily living; MMSE, Mini-Mental State Examination; N/A, not assessed; NPI, Neuropsychiatric Inventory.

†MMSE score range 0–30; higher score indicates better functioning.

‡ADAS-Cog14 score range 0–90; higher score indicates greater severity.

§ADCS-bADL score range 0–22 and ADCS-iADL score range 0–56; higher scores indicate better functioning.

‖NPI-12 score range 0–144; higher score indicates greater severity.