

G.9.1.8 Exercise

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
15	Not serious	Serious ¹	Not serious	Not serious	1148	MD 1.30 (0.49, 2.11)	Moderate
Cognition: MMSE (post-intervention, excluding multimodal interventions) – higher numbers favour intervention							
12	Not serious	Serious ¹	Not serious	Not serious	987	MD 1.55 (0.56, 2.55)	Moderate
Cognition: MMSE (follow-up) – higher numbers favour intervention							
2	Very serious ²	Serious ¹	Not serious	Serious ³	156	MD 1.21 (-3.51, 5.93)	Very low
Cognition: all measures (post-intervention) – higher numbers favour intervention							
16	Not serious	Serious ¹	Not serious	Serious ⁴	1179	SMD 0.36 (0.14, 0.58)	Low
Cognition: all measures (post-intervention, excluding multimodal interventions) – higher numbers favour intervention							
13	Not serious	Serious ¹	Not serious	Serious ⁴	1,018	SMD 0.41 (0.16, 0.66)	Low
Cognition: all measures (follow-up) – higher numbers favour intervention							
2	Very serious ²	Serious ¹	Not serious	Very serious ⁵	156	SMD 0.20 (-0.83, 1.23)	Very low
ADL: ADCS-ADL (post-intervention) – higher numbers favour intervention							
1 (Hoffman 2015)	Not serious	N/A	Not serious	Serious ³	190	MD -0.70 (-3.54, 2.14)	Moderate
ADL: all measures (post-intervention) – higher numbers favour intervention							
13	Not serious	Serious ¹	Not serious	Serious ⁴	1474	SMD 0.26 (0.09, 0.43)	Low
ADL: all measures (post-intervention, excluding multimodal interventions) – higher numbers favour intervention							
11	Not serious	Serious ¹	Not serious	Serious ⁴	1,264	SMD 0.32 (0.15, 0.50)	Low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
ADL: all measures (follow-up) – higher numbers favour intervention							
1 (Littbrand 2009)	Serious ⁶	N/A	Not serious	Serious ⁴	91	SMD 0.23 (-0.18, 0.64)	Low
Behavioural and psychological symptoms: NPI (post-intervention) – lower numbers favour intervention							
6	Not serious	Not serious	Not serious	Not serious	729	MD -1.58 (-2.76, -0.41)	High
Behavioural and psychological symptoms: all measures (post-intervention) – higher numbers favour intervention							
6	Not serious	Not serious	Not serious	Serious ⁴	729	SMD -0.26 (-0.41, -0.11)	Moderate
Global assessment (post-intervention) – higher numbers favour intervention							
1 (Luttenberger 2012)	Very serious ²	N/A	Not serious	Not serious	119	SMD 0.80 (0.42, 1.17)	Low
Depression: Cornell scale for depression in dementia (post-intervention) – higher numbers favour intervention							
3	Not serious	Serious ¹	Not serious	Serious ³	379	MD 1.50 (-0.15, 3.16)	Low
Depression: all measures (post-intervention) – higher numbers favour intervention							
7	Not serious	Serious ¹	Not serious	Serious ⁴	762	SMD 0.11 (-0.19, 0.40)	Low
Depression: all measures (post-intervention, excluding multimodal interventions) – higher numbers favour intervention							
6	Not serious	Serious ¹	Not serious	Serious ⁴	719	SMD 0.14 (-0.18, 0.46)	Low
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
1 (Yang 2015)	Serious ⁷	N/A	Not serious	Serious ³	50	MD 2.16 (-0.44, 4.76)	Low
Quality of life: EQ-5D (post-intervention) – higher numbers favour intervention							
1 (Hoffman 2015)	Not serious	N/A	Not serious	Serious ³	190	MD 0.00 (-0.03, 0.03)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
5	Not serious	Not serious	Not serious	Serious ⁴	459	SMD -0.01 (-0.20, 0.17)	Moderate
Carer burden: ZBI (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious ³	69	MD -4.12 (-11.44, 3.20)	Moderate
Carer burden: all measures (post-intervention) – higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Very serious ⁵	96	SMD -0.12 (-0.52, 0.29)	Low

1. $i^2 > 40\%$

2. Evidence of selective outcome reporting

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
3. Non-significant result							
4. 95% CI crosses 1 line of a defined MID interval							
5. 95% CI crosses 2 lines of a defined MID interval							
6. Assessors not blinded to group allocation							
7. No details of randomisation method or assessor blinding reported							